Stanford Technology Law Review

Fall 2012

CONTENTS

ARTICLES

NOTES

Information About the *Stanford Technology Law Review*

About: The *Stanford Technology Law Review* is published three times a year (January, March, and June) by students of the Stanford Law School, Crown Quadrangle, 559 Nathan Abbott Way, Stanford, California 94305-8610. The *Stanford Technology Law Review* is available free of charge online at http://stlr.stanford.edu/.

Manuscripts: The *Stanford Technology Law Review* accepts the submission of unsolicited manuscripts through *ExpressO*, a service of the Berkeley Electronic Press. Additional information about submissions is available online at http://stlr.stanford.edu/submissions/. Questions may be directed to stlr-submissions@lists.stanford.edu.

Citations: The text and citations of the *Stanford Technology Law Review* generally conform to *The Bluebook: A Uniform System of Citation* (19th ed. 2010), copyright by The Columbia Law Review Association, The Harvard Law Review Association, the *University of Pennsylvania Law Review*, and the *Yale Law Journal*.

Internet Address: The *Stanford Technology Law Review* homepage is located at http://stlr.stanford.edu/. Follow the *Stanford Technology Law Review* on Twitter at StanTechLRev.

EDITORIAL BOARD
VOLUME 16

The Stanford University School of Law

OFFICERS OF ADMINISTRATION

John Hennessy, B.E., M.S., Ph.D., President of the University
John Etchemendy, B.A., M.A., Ph.D., Provost of the University
M. Elizabeth Magill, B.A., J.D., Dean and Richard E. Lang Professor of Law
Mark G. Kelman, B.A., J.D., James C. Gaither Professor of Law and Vice Dean
Deborah R. Hensler, B.A., Ph.D., Judge John W. Ford Professor of Dispute Resolution, and Associate Dean for Graduate Studies
Lawrence C. Marshall, B.A., J.D., Professor of Law, Associate Dean for Clinical Education and David &Stephanie Mills Director of the Mills Legal Clinic
Jane Schacter, A.B., J.D., William Nelson Cromwell Professor of Law and Associate Dean for Curriculum
F. Daniel Siciliano, B.A., J.D., Professor of the Practice of Law and Associate Dean for Executive Education and Special Programs
Frank F. Brucato, B.A., Senior Associate Dean of Administration and Chief Financial Officer
Diane T. Chin, B.A., J.D., Associate Dean for Public Service and Public Interest Law and Lecturer in Law
Faye Deal, A.B., Associate Dean for Admissions and Financial Aid
Julia Erwin-Weiner, B.A., M.A., Associate Dean for External Relations
Catherine Glaze, B.A., J.D., Associate Dean for Student Affairs
Sabrina Johnson, B.A., Associate Dean for Communications and Public Relations and Chief Communications Officer
Susan C. Robinson, B.A., J.D., Associate Dean for Career Services and Lecturer in Law

FACULTY EMERITI

Barbara Babcock, B.A., LL.B., LL.D. (hon.), Judge John Crown Professor of Law, Emerita
Paul Brest, B.A., LL.B., Professor of Law, Emeritus and Former Dean
Gerhard Casper, Referendar, LL.M., Dr. iur. utr., LL.D. (hon.), President Emeritus and Peter and Helen Bing Professor in Undergraduate Education, Professor of Law, Senior Fellow, Institute for International Studies, and Professor (by courtesy) of Political Science
William Cohen, B.A., LL.B., C. Wendell and Edith M. Carlsmith Professor of Law, Emeritus
Lance E. Dickson, B.A., LL.B., B.Com., M.L.S., Professor of Law, Emeritus, and Former Director of Robert Crown Law Library
Marc A. Franklin, A.B., LL.B., Frederick I. Richman Professor of Law, Emeritus
William B. Gould IV, B.A., LL.B., LL.D. (hon.), Charles A. Beardsley Professor of Law, Emeritus
Thomas C. Grey, B.A., B.A., LL.B., LL.D. (hon.), Nelson Bowman Sweitzer and Marie B. Sweitzer Professor of Law, Emeritus
Thomas C. Heller, B.A., LL.B., Lewis Talbot and Nadine Hearn Shelton Professor of International Legal Studies, Emeritus
Miguel A. Méndez, A.A., B.A., J.D., Adelbert H. Sweet Professor of Law, Emeritus
John Henry Merryman, B.S., M.S., J.D., LL.M., J.S.D., Nelson Bowman Sweitzer and Marie B. Sweitzer Professor of Law, Emeritus, and Affiliated Professor, Department of Art, Emeritus
Kenneth E. Scott, B.A., M.A., LL.B., Ralph M. Parsons Professor of Law and Business, Emeritus
Michael S. Wald, B.A., M.A., LL.B., Jackson Eli Reynolds Professor of Law, Emeritus

PROFESSORS

Janet Cooper Alexander, B.A., M.A., J.D., Frederick I. Richman Professor of Law
Joseph M. Bankman, B.A., J.D., Ralph M. Parsons Professor of Law and Business
R. Richard Banks, B.A., M.A., J.D., Jackson Eli Reynolds Professor of Law
Juliet M. Brodie, A.B., J.D., Professor of Law and Director, Stanford Community Law Clinic
James Cavallaro, A.B., J.D., Professor of Law and Director, Stanford International Human Rights and Conflict Resolution Clinic
Joshua Cohen, B.A., M.A., Ph.D., Professor of Political Science, Philosophy and Law (on leave)
G. Marcus Cole, B.S., J.D., Wm. Benjamin Scott and Luna M. Scott Professor of Law
Richard Craswell, B.A., J.D., William F. Baxter-Visa International Professor of Law
Mariano-Florentino Cuéllar, A.B., M.A., J.D., Ph.D., Stanley Morrison Professor of Law, Professor of Political Science (by courtesy), and Senior Fellow, Freeman Spogli Institute for International Studies
Robert M. Daines, B.S., B.A., J.D., Pritzker Professor of Law and Business and Professor (by courtesy) of Finance
Michele Landis Dauber, B.S.W., J.D., Ph.D., Professor of Law, Bernard D. Bergreen Faculty Scholar, and Professor (by courtesy) of Sociology
John J. Donohue III, B.A., J.D., Ph.D., C. Wendell and Edith M. Carlsmith Professor of Law
David Freeman Engstrom, A.B., M.Sc., J.D., Ph.D., Associate Professor of Law
Nora Freeman Engstrom, B.A., J.D., Associate Professor of Law
George Fisher, B.A., J.D., Judge John Crown Professor of Law and Faculty Co-Director, Criminal Prosecution Clinic
Jeffrey L. Fisher, A.B., J.D., Professor of Law and Co-Director, Supreme Court Litigation Clinic
Richard Thompson Ford, B.A., J.D., George E. Osborne Professor of Law
Barbara H. Fried, B.A., M.A., J.D., William W. and Gertrude H. Saunders Professor of Law

Lawrence M. Friedman, B.A., J.D., LL.M., LL.D. (hon.), Marion Rice Kirkwood Professor of Law, Professor (by courtesy) of History, and Professor (by courtesy) of Political Science
Ronald J. Gilson, B.A., J.D., Charles J. Meyers Professor of Law and Business
Paul Goldstein, B.A., LL.B., Stella W. and Ira S. Lillick Professor of Law
Robert W. Gordon, B.A., J.D., Professor of Law
Henry T. Greely, B.A., J.D., Deane F. and Kate Edelman Johnson Professor of Law and Professor (by courtesy) of Genetics
Joseph A. Grundfest, B.A., M.Sc., J.D., W.A. Franke Professor of Law and Business
Deborah R. Hensler, B.A., Ph.D., Judge John W. Ford Professor of Dispute Resolution and Associate Dean for Graduate Studies
Daniel E. Ho, B.A., A.M., Ph.D., J.D., Professor of Law and Robert E. Paradise Faculty Fellow for Excellence in Teaching and Research
Erik G. Jensen, B.A., J.D., Professor of the Practice of Law
Pamela S. Karlan, B.A., M.A., J.D., Kenneth and Harle Montgomery Professor of Public Interest Law and Co-Director, Supreme Court Litigation Clinic
Mark G. Kelman, B.A., J.D., James C. Gaither Professor of Law and Vice Dean
Amalia D. Kessler, B.A., M.A., J.D., Ph.D., Lewis Talbot and Nadine Hearn Shelton Professor of International Legal Studies
Daniel P. Kessler, B.A., J.D., Ph.D., Professor of Law, Senior Fellow, Hoover Institution, and Professor of Health Research and Policy (by courtesy) of School of Medicine
Michael Klausner, B.A., M.A., J.D., Nancy and Charles Munger Professor of Business and Professor of Law
William S. Koski, B.B.A., J.D., Ph.D., Eric and Nancy Wright Professor of Clinical Education and Director, Youth and Education Law Project
Larry D. Kramer, B.A., J.D., Former Dean and Professor of Law (on leave)
Mark A. Lemley, B.A, J.D., William H. Neukom Professor of Law
M. Elizabeth Magill, B.A., J.D., Dean and Richard E. Lang Professor of Law
Lawrence C. Marshall, B.A., J.D., Professor of Law, Associate Dean for Clinical Education and David & Stephanie Mills Director of the Mills Legal Clinic
Jenny S. Martinez, B.A., J.D., Professor of Law and Warren Christopher Professor in the Practice of International Law and Diplomacy
Michael W. McConnell, B.A., J.D., Richard and Frances Mallery Professor of Law
David W. Mills, B.A., J.D., Professor of the Practice of Law and Senior Lecturer in Law
Jay Mitchell, B.A., J.D., Associate Professor of Law and Director, Organizations and Transactions Clinic
Alison D. Morantz, B.A., M.Sc., J.D., Ph.D., Professor of Law and John A. Wilson Distinguished Faculty Scholar
Joan Petersilia, B.A., M.A., Ph.D., Adelbert H. Sweet Professor of Law
A. Mitchell Polinsky, B.A., Ph.D., M.S.L., Josephine Scott Crocker Professor of Law and Economics
Robert L. Rabin, B.S., J.D., Ph.D., A. Calder Mackay Professor of Law
Dan Reicher, B.A., J.D., Professor of the Practice of Law
Deborah L. Rhode, B.A., J.D., Ernest W. McFarland Professor of Law
Jane Schacter, A.B., J.D., William Nelson Cromwell Professor of Law and Associate Dean for Curriculum
F. Daniel Siciliano, B.A., J.D., Professor of the Practice of Law and Associate Dean for Executive Education and Special Programs
Deborah A. Sivas, B.A., M.S., J.D., Luke W. Cole Professor of Environmental Law and Director, Environmental Law Clinic
Norman W. Spaulding, B.A., J.D., Nelson Bowman Sweitzer and Marie B. Sweitzer Professor of Law
Jayashri Srikantiah, B.S., J.D., Professor of Law and Director, Immigrants' Rights Clinic
James Frank Strnad II, B.A., J.D., Ph.D., Charles A. Beardsley Professor of Law
Barton H. Thompson, Jr., B.A., J.D., M.B.A., Robert E. Paradise Professor of Natural Resources Law and Perry L. McCarty and Director, Woods Institute for the Environment
George Triantis, J.S.D., James and Patricia Kowal Professor of Law
Ronald Tyler, B.S., J.D., Associate Professor of Law and Director, Criminal Defense Clinic
Barbara van Schewick, Ph.D., Associate Professor of Law and Helen L. Crocker Faculty Scholar
Shirin Sinnar, A.B., J.D., Assistant Professor of Law
Michael Wara, B.A., Ph.D., J.D., Associate Professor of Law
Robert Weisberg, B.A., M.A., Ph.D., J.D., Edwin E. Huddleson, Jr. Professor of Law

SENIOR LECTURERS

Margaret R. Caldwell, B.S., J.D., Senior Lecturer in Law and Executive Director for Oceans Solutions and Senior Lecturer, Woods Institute for the Environment
Janet Martinez, B.S., J.D., M.P.A., Senior Lecturer in Law
Allen S. Weiner, B.A., J.D., Senior Lecturer in Law

VISITING PROFESSORS & AFFILIATED FACULTY

Michelle Anderson, M. Sc., J.D., Visiting Assistant Professor of Law
Michael Asimow, J.D., Visiting Professor of Law
Elizabeth Blankespoor, M.Acc., Ph.D., Affiliated Faculty
Alexandria (Ali) Boehm, M.S., Ph.D., Affiliated Faculty
Kate Bundorf, M.B.A., M. Ph., Ph.D., Affiliated Faculty
John-Paul Ferguson, M.A., Ph. D., Affiliated Faculty

Siegfried Fina, J.D., J.S.D., Visiting Associate Professor of Law
Joy Ishii, A.M., Ph.D., Affiliated Faculty
Michael Karayanni, LL.B., LL.M., LL.D., S.J.D., Visiting Professor of Law
David F. Larcker, M.S., Ph.D., Professor (by courtesy) of Law
Jose Maldonado, M.D., Professor (by courtesy) of Law
Robert Merges, J.D., LL.M., J.S.D., Visiting Professor of Law
Clifford Nass, M.A., Ph.D., Professor (by courtesy) of Law
Burt Neuborne, LL.B., Visiting Professor of Law
Joost Pauwelyn, Ph.D., Visiting Professor of Law
Rogello Perez-Perdomo, Ph.D., Visiting Professor of Law
Paul C. Pfleiderer, M.Phil., Ph.D., Professor (by courtesy) of Law
Madhav Rajan, M.S., M.B.A., Ph.D., Professor (by courtesy) of Law
Jack Rakove, Ph.D., Professor (by courtesy) of Law
Stefan Reichelstein, M.S., Ph.D., Affiliated Faculty
Helen Stacy, LL.B., Ph.D., Affiliated Faculty
Barry Weingast, Ph.D., Affiliated Faculty

LECTURERS AND TEACHING FELLOWS

Julie Ahrens, J.D.
Alvin Attles, M.A.
Simao Avila, J.D.
Daniel Barton, J.D.
William Baude, J.D.
Marilyn Bautista, J.D.
Jeanine Becker, J.D.
Tanya Beder, M.B.A.
Byron Bland, M.A., M.Div.
Viola Canales, J.D.
Diane Chin, J.D.
Daniel Cooperman, J.D., M.B.A.
Maggie Crosby, J.D.
Betsy de la Vega, J.D.
Michael Dickstein, J.D.
Ron Dolin, Ph.D.
Lisa Douglass, J.D.
Jared Ellias, J.D.
Bonnie Eskenazi, J.D.
Russ Feingold, J.D.
Randee Fenner, J.D.
Bertram Fields, LL.B.
Jay Finkelstein, J.D.
David Forst, J.D.
Simon Frankel, M.Phil, J.D.
Laurence Franklin, J.D., M.B.A.
Michelle Galloway, J.D.
Mei Gechlik, J.S.D., M.B.A.
Tracy Genesen, J.D.
Marshall Goldberg, J.D.
Alex Gould, J.D.
Jonathan Greenberg, J.D.
Lucas Guttentag, J.D.
Tim Hallahan, J.D.
Brad Handler, J.D.
Laurie Hane, J.D.
Keith Hennessey, M.A.
Brooke Heymach, J.D., M.S.W.
Ivan Humphreys, J.D.
David Johnson, J.D., J.S.M.
Danielle Jones, J.D.
Julie Matlof Kennedy, J.D.
Jason Kipnis, M.S., J.D.
Suzanne McKechnie Klahr, J.D.
Jeffrey Kobrick, J.D.
Charles Koob, J.D.
Phillip Levine, J.D.
May Liang, J.D.
Goodwin Liu, J.D.
Mina Titi Liu, J.D.
Paul Lomio, J.D., LL.M., M.L.I.S.
Beth McLellan, J.D.
Jeanne Merino, J.D.
Roberta Morris, J.D., Ph.D.
Carly Munson, J.D.
Linda Netsch, J.D.
Thomas Nolan, J.D.
Jessica Notini, J.D.
B. Howard Pearson, J.D.
Lisa Pearson, J.D., J.S.M.
Sergio Puig, LL.B., J.S.M., J.S.D.
John Quigley, J.D., M.B.A., M.Phil.
Sarah Rajec, J.D.
Stephan Ray, J.D.
Susan Robinson, J.D.
John Rodkin, M. Eng., J.D., M.B.A.
Michael Romano, J.D.
Andrew Roper, M.A., Ph.D.
Stephen Rosenbaum, J.D., M.P.P.
Matthew Rossiter, J.D.
Kevin Russell, J.D.
Leah Russin, J.D.
Ticien Sassoubre, Ph.D.
William (Billy) Schwartz, J.D.
Rachelle Silverberg, J.D.
Bryant Walker Smith, J.D., LL.M.
Stephanie Smith, J.D.
Steven Smith, M.Phil., LL.B., J.D.
James (Jim) Sonne, J.D.
Stephan Sonnenberg, J.D.
Michelle Sonu, S.B., J.D.
Kimberly Summe, J.D.
Alicia Thesing, J.D.
Dan Torres, J.D.
Adine Varah, J.D.
Roland Vogl, J.D., J.S.M., J.S.D.
Nisha Vyas, J.D.
Vaughn Walker, J.D.
Erika Wayne, J.D., M.S.
Dana Weintraub, M.D.
Robert Wexler, J.D.
Katherine Wright, J.D.
James Yoon, J.D.

PROFESSIONAL LIBRARY STAFF

Annie Chen, B.A., M.L.S.
Alexandra Lee Delgado, J.D., M.L.I.S.
Alba Holgado, B.A.
J. Paul Lomio, B.A., J.D., LL.M., M.L.I.S.
Rachael Samberg, J.D.
Sergio Stone, J.D., M.L.I.S.
Erika V. Wayne, A.B., J.D., M.S.
George D. Wilson, B.A., J.D., M.L.I.S.
Sarah Wilson, B.A., M.L.I.S.
Kathleen M. Winzer, B.A., M.L.S.
Naheed Zaheer, B.S., M.S.C., M.L.I.S.

STANFORD TECHNOLOGY LAW REVIEW
VOLUME 16, NUMBER 1 FALL 2012

WHY THE "SCOPE OF THE PATENT" TEST CANNOT SOLVE THE DRUG PATENT SETTLEMENT PROBLEM

Michael A. Carrier*

CITE AS: 16 STAN. TECH. L. REV. 1 (2012)
http://stlr.stanford.edu/pdf/scopeofthepatent.pdf

INTRODUCTION

One of the most difficult legal issues today involves settlements by which brand-name drug companies pay generic firms to delay entering the market. Such conduct requires courts to consider not only patent and antitrust law, but also the Hatch-Waxman Act, the complex regime governing behavior in the pharmaceutical industry.

Courts have analyzed these agreements by relying on a test that asks if the settlement falls within the "scope of the patent." They have found, in nearly all

* Professor of Law, Rutgers University School of Law-Camden.

of these cases, that it does. And, as a result, they have concluded that the agreements do not violate the antitrust laws.

This Article shows why the scope test is not appropriate in determining the antitrust treatment of drug patent settlements. It recounts the history of the test, showing its increasing deference over time. And it demonstrates the three primary problems with the test: (1) it involves a transformation that has left the test toothless, (2) it assumes that the patent at issue is valid, and (3) it neglects the issue of infringement.

I. HISTORY OF THE SCOPE TEST

A. Cardizem – *Outside the Scope*

The scope test can be traced to the Sixth Circuit's decision in *In re Cardizem CD Antitrust Litigation*.[1] In *Cardizem*, the generic company agreed not to market a generic version of the brand firm's patented blood-pressure drug until it obtained a final determination that the patent was not infringed.[2] Of concern to the court, the agreement prevented the marketing of generic versions of not only the patented drug, but also drugs "not at issue in the pending litigation."[3]

The court found that the brand paid "the only potential competitor $40 million per year to stay out of the market."[4] And it concluded that the settlement was "a horizontal agreement to eliminate competition" and was "a classic example of a *per se* illegal restraint of trade."[5] The court's punishment of conduct outside the patent's scope was adopted by later courts that used the scope test for different purposes.

B. Valley Drug – *A Calibrated Test*

The Eleventh Circuit took a calibrated approach to the scope issue in *Valley Drug Co. v. Geneva Pharmaceuticals, Inc*.[6] In that case, the court reversed the district court's determination that a settlement involving a blood-pressure drug was *per se* illegal.[7] It found that a full analysis of the agreement, which provided "restrictions on infringing products"[8] and the marketing of "any" generic product covering the relevant active ingredient,[9] required

1. 332 F.3d 896 (6th Cir. 2003).
2. *Id.* at 902.
3. *Id.* at 908 n.13.
4. *Id.* at 908.
5. *Id.*
6. 344 F.3d 1294 (11th Cir. 2003).
7. *Id.* at 1306.
8. *Id*. at 1311.
9. *Id*.

"consideration of the scope of the exclusionary potential of the patent, the extent to which these provisions of the [a]greements exceed that scope, and the anticompetitive effects thereof."[10]

In determining whether the settlement provisions resembled a preliminary injunction and stay, the Eleventh Circuit explained that courts must analyze "the likelihood of [the patentee's] obtaining such protections."[11] The court remanded for the lower court to determine whether the settlement was a "reasonable implementation" of the "protection afforded by the patents."[12]

C. Taxmoxifen – *A Shrinking Test*

Courts then imperceptibly shifted from punishing conduct "outside the scope" of the patent to immunizing conduct "within the scope" of the patent. In doing so, the test took a dramatic turn toward deference.

In *In re Tamoxifen Citrate Antitrust Litigation*,[13] the Second Circuit upheld a grant of the defendants' motion to dismiss regarding a settlement on a breast-cancer-treatment drug. It concluded that as long as "the patent litigation is neither a sham nor otherwise baseless" or beyond the patent's scope, the patentee can enter into a settlement "to protect that to which it is presumably entitled: a lawful monopoly over the manufacture and distribution of the patented product."[14]

The court concluded that the settlement did not "unlawfully extend the reach" of the patent.[15] Because the brand's patent "preclude[d] all generic versions of [the drug]," any competing version "would . . . necessarily infringe the patent."[16] The court also noted that the agreement did not "restrain[] the introduction or marketing of unrelated or non-infringing products," in contrast to the settlement in *Cardizem*, which "included not only a substantial reverse payment but also an agreement that the generic manufacturer would not market non-infringing products."[17]

D. Cipro – *The Noose Tightens*

The Federal Circuit in *In re Ciprofloxacin Hydrochloride Antitrust*

---

10. *Id.* at 1312; *see also* Schering-Plough Corp. v. FTC, 402 F.3d 1056, 1066 (11th Cir. 2005) (focusing on "(1) the scope of the exclusionary potential of the patent; (2) the extent to which the agreements exceed that scope; and (3) the resulting anticompetitive effects").

11. *Valley Drug*, 344 F.3d at 1312.

12. *Id.*

13. 466 F.3d 187 (2d Cir. 2006).

14. *Id.* at 208-09, 213.

15. *Id.* at 213.

16. *Id.* at 214.

17. *Id.* at 213-14.

Litigation[18] continued the trend toward deference in affirming a motion to dismiss on an agreement concerning an antibiotic. Its analysis focused on the patent system's right to exclude and the presumption that patents are valid.[19] The court concluded that, in the absence of evidence of fraud before the PTO or sham litigation, the court "need not consider the validity of the patent."[20]

The court found that the agreements at issue only "exclude[d] the defendants from profiting from the patented invention," thus falling "well within Bayer's rights as the patentee."[21] It found that "a patent is presumed to be valid," with patent law bestowing "the right to exclude others from profiting by the patented invention."[22] And it explained that the "essence of the inquiry" was "whether the agreements restrict competition beyond the exclusionary zone of the patent."[23] The court concluded that "all anticompetitive effects of the settlement agreement are within the exclusionary power of the patent."[24]

E. Androgel – *The Ambiguity Disappears*

Even though its initial version of the scope test appeared nuanced in its focus on the patent's "exclusionary potential" and "likelihood" of obtaining an injunction, the Eleventh Circuit dispensed with any ambiguity in *FTC v. Watson Pharmaceuticals*,[25] making clear, in upholding a settlement concerning a testosterone drug, that it was lining up behind the version articulated by the Second and Federal Circuits.

The court stated that "[a] patent holder and any of its challengers cannot enter into an agreement that excludes more competition than the patent has the potential to exclude."[26] And it clarified that its use in an earlier case of the phrase "strength of the patent" referred to "the potential exclusionary scope of the patent," which means "the exclusionary rights appearing on the patent's face and not the underlying merits of the infringement claim."[27]

F. K-Dur – *A Turn Toward Scrutiny*

Bucking the march toward deference, in 2012 the Third Circuit in *In re K-Dur Antitrust Litigation* criticized the scope test in reversing the district court's grant of summary judgment.[28] It recognized that the test "assumes away the

18. 544 F.3d 1323 (Fed. Cir. 2008).
19. *Id*. at 1333, 1337.
20. *Id.* at 1336.
21. *Id.* at 1333.
22. *Id.* at 1337.
23. *Id.* at 1336.
24. *Id.*
25. 677 F.3d 1298 (11th Cir. 2012).
26. *Id.* at 1308.
27. *Id*. at 1311 n.8.
28. 686 F.3d 197 (3d Cir. 2012).

question being litigated in the underlying patent suit, enforcing a presumption that the patent holder would have prevailed."[29] And it observed that "the scope of the patent test does not subject reverse payment agreements to any antitrust scrutiny," explaining that "no court applying the scope of the patent test has ever permitted a reverse payment antitrust case to go to trial."[30] The court concluded by adopting a test by which "the finder of fact must treat any payment from a patent holder to a generic patent challenger who agrees to delay entry into the market as *prima facie* evidence of an unreasonable restraint of trade."[31]

II. CRITIQUES OF THE SCOPE TEST

There are three primary problems with the scope test. First, the version used today has shed any potential nuance in morphing into a test granting automatic legality. Second, the test is based on the crucial assumption that the relevant patent is valid. Third, it cannot address the issue of infringement.

A. *Transformed Scope*

Although each of the decisions discussed above used the concept of patent scope, the meaning of the test has shifted dramatically. The framework was first articulated in *Cardizem* with the court punishing conduct lying outside the coverage of the patent. The Eleventh Circuit then applied the test by using language that left open the possibility that it would consider whether the patent at issue actually allowed the brand to exclude the generic. In *Valley Drug*, for example, the court explored the likelihood that a patentee would have been successful in obtaining an injunction against infringement.

But the test then shifted toward deference. Such a move was a subtle – and until now unnoticed – shift based on an improper inference drawn from *Cardizem*. The court in *Cardizem* made clear that an agreement covering a product outside the scope of the patent was per se illegal. In that case, the agreement applied not only to products covered by the patent but also to unpatented products.

The fact, however, that a settlement reaching a product *outside* the scope of the patent violates the antitrust laws does not mean that one falling *within* the facial scope of the patent is automatically valid. The Second, Eleventh, and Federal Circuits thus used the test for a new and dramatically different purpose. They did not employ the framework to address the easy cases where the settling parties blocked competition on products not covered by the patent.

Instead, they imported the test into the more complex setting of agreements

---

29. *Id.* at 214.

30. *Id.*

31. *Id.* at 218 (providing that presumption could be rebutted by showing that payment (1) was for purpose other than delayed entry or (2) offered some pro-competitive benefit).

that do not reach beyond the facial scope of the patent. These agreements cannot be so easily dealt with. For they might or might not violate the antitrust laws. That depends on whether the patent is valid. But that cannot be determined by the mere existence of the patent.

B. *Assumption of Validity*

The fundamental problem with the court's transformed, simplistic scope test is that it assumes the validity issues that are central to the determination of antitrust analysis. The overriding question in cases analyzing pharmaceutical settlements is whether the patent is valid.

If the patent is valid, then an agreement by which the brand pays the generic to drop its challenge and delay entering the market could fall within the patent's scope and not present antitrust concerns.[32] After all, the brand could rely on the patent itself to exclude competitors before the end of the term.

But if, in contrast, the patent is not valid, then it does not have any scope at all. The patentee is not entitled to pay the generic to drop its patent challenge since, by definition, the patent is not valid. In this setting, the behavior resembles market allocation, one of the most severe anticompetitive harms, with two competitors dividing the market and eliminating competition.

The problem with courts that rely on the scope test today is that they unwittingly assume that the patent is valid. These courts ignore potential indications of patent validity – such as judicial findings of invalidity or substantial payments to generics – in relying on the mere issuance of the patent.

Not every patent issued by the U.S. Patent and Trademark Office, however, is valid. Empirical studies have consistently shown that at least 40% of granted patents that are litigated to decision are invalid.[33] The rate of invalidity is even higher in the present context, with the FTC finding that generics prevailed in 73% of challenges between 1992 and 2000.[34]

Courts that have applied the scope test often address the validity issue by relying on the procedural presumption of Section 282 of the Patent Act, which

---

32. Michael A. Carrier, *Unsettling Drug Patent Settlements: A Framework for Presumptive Illegality*, 108 MICH. L. REV. 37, 66 (2009).

33. *See* John R. Allison & Mark A. Lemley, *Empirical Evidence on the Validity of Litigated Patents*, 26 AIPLA L.Q. 185, 205 (1998) (finding that courts invalidated 46% of patents between 1989 and 1996); Kimberly A. Moore, *Judges, Juries, and Patent Cases – An Empirical Peek Inside the Black Box*, 99 MICH. L. REV. 365, 385 (2000) (demonstrating that alleged infringer prevailed in 42% of patent cases that reached trial between 1983 and 1999); University of Houston Law Center, Decisions for 2000-2004, Issue Codes 01-16, 23, 24, http://www.patstats.org/2000-04.htm (stating that, in patent cases between 2000 and 2004, courts found 43% of patents invalid and 75% not infringed).

34. FED. TRADE COMM'N, *Generic Drug Entry Prior to Patent Expiration: An FTC Study* 16 (2002), http://www.ftc.gov/os/2002/07/genericdrugstudy.pdf (providing results for "Paragraph IV" challenges by which generic firms certify that brand firms' patents are invalid or not infringed, seeking to enter before the end of the patent term).

states that patents "shall be presumed valid."[35] But a presumption of validity is only a procedural presumption governing the order in which proof is presented. It is not substantive evidence of validity.[36] In addition, such a presumption should be entitled to the least amount of deference where parties "enter agreements that prevent validity from even being challenged," which is especially problematic given the Hatch-Waxman Act's emphasis on challenges to invalidity and infringement.[37]

C. *Inapplicability to Infringement*

A final problem with the scope test is that it ignores the issue of infringement. A brand firm must show not only that the patent is valid but also that the generic's drug infringes its patent. The simplistic version of the scope test cannot resolve that question.

One fundamental difference between validity and infringement is that the patentee bears the burden of demonstrating infringement. For validity, the patentee at least can point to an initial presumption that the patent is valid. In contrast, the Federal Circuit has made clear that "[t]he patentee bears the ultimate burden of proof to demonstrate infringement by a preponderance of the evidence."[38] For this reason, a court cannot dispose of the issue of infringement by observing the mere existence of the patent.

The *K-Dur* case is instructive in this regard. In that case, one generic claimed that its product did not infringe the brand's patent because its product lay "outside the range limited by claim 1 of the [] patent."[39] The other generic claimed that its product did not infringe since it lacked the "coating material with different ingredients" covered by the patent.[40]

These claims were plausible since the brand's patent did not cover the active ingredient in the patented supplement, but applied only to a weaker formulation that covered a certain type of tablet with a certain percentage of potassium chloride crystals and coating material.[41] So even though the district court "declined to discount the exclusionary power of [the brand's] patent

---

35. 35 U.S.C. § 282; *see Tamoxifen*, 466 F.3d at 211 (finding that presumption of validity allows parties to settle "weak patent cases" even though "such settlements will inevitably protect patent monopolies that are, perhaps, undeserved"); *Schering-Plough*, 402 F.3d at 1066-67 (relying on presumption in concluding that brand would not suffer antitrust liability for exclusionary activity unless generics were able to prove patent's invalidity or noninfringement); *Ciprofloxacin*, 544 F.3d at 1337 (asserting that "analysis of patent validity" is not "appropriate in the absence of fraud or sham litigation" since "a patent is presumed to be valid").

36. *See* Stratoflex, Inc. v. Aeroquip Corp., 713 F.2d 1530, 1534 (Fed. Cir. 1983).

37. Carrier, *supra* note 32, at 64.

38. *E.g.,* Egyptian Goddess, Inc. v. Swisa, Inc., 543 F.3d 665, 679 (Fed. Cir. 2008).

39. *In re* K-Dur Antitrust Litig., No. 01-1652, 2009 WL 508869, at *6 (D.N.J. Feb. 6, 2009), *rev'd*, 686 F.3d 197 (3d Cir. 2012).

40. *Id.* at *8.

41. *Id.* at *4.

based on the *possibility* that it was not infringed by the [generic] products," the issue could not really be resolved by relying on the scope of the patent.[42]

CONCLUSION

The scope test applied by courts today cannot resolve the issue of whether drug patent settlements violate the antitrust laws. The test has ventured far beyond its initial version employed for the narrow purpose of punishing conduct reaching products clearly outside the scope of the patent.

The simplistic version used today is employed to give automatic immunity to conduct that might – or might not – be justified. The test assumes issues of validity and infringement that cannot possibly be determined from the mere issuance of the patent. With all potential nuance stripped out of the scope test, courts today are relegated to the role of traffic cops shooing agreements through an antitrust light always flashing green. The simplistic scope test is not appropriate for analyzing the complex issues presented by drug patent settlements.

42. *Id.* at *25 (emphasis in original).

STANFORD TECHNOLOGY LAW REVIEW
VOLUME 16, NUMBER 1 FALL 2012

PATENT INFRINGEMENT IN THE CONTEXT OF FOLLOW-ON BIOLOGICS

Janet Freilich*

CITE AS: 16 STAN. TECH. L. REV. 9 (2012)
http://stlr.stanford.edu/pdf/follow-onbiologics.pdf

ABSTRACT

This Article fills a gap in the literature by conducting a comprehensive analysis of patent infringement in the context of follow-on biologics. Patent infringement is an important topic because, like small molecule generic drugs, follow-on biologics are likely to begin their life facing infringement suits. Because it is tremendously expensive to develop a follow-on biologic, it is vital that there be consistency in how they are treated in the courts once the inevitable patent infringement suits arrive. If follow-on biologics companies cannot predict how their product will be received in court, they may decide it is not worth the risk to develop the product. This Article looks at types of strategies industry is likely to use to avoid infringement and how courts are likely to respond to these strategies. This Article focuses predominantly on the doctrine of equivalents, both because it will be particularly important in suits concerning follow-on biologics (it is nearly impossible to make a follow-on biologic identical to the reference drug) and because it represents the outer limits of the scope of a patent, and thus the most difficult cases. The Article is important for courts that must create a coherent body of law where no precedent yet exists, for industry members trying to predict how their products will be received and for policy makers who seek to understand the nature of infringement suits and shape this body of law in a

* Associate, Covington & Burling, LLP, Washington, D.C. This Article contains the views of the author only and does not reflect the views of Covington & Burling or any of its clients. I am a John M. Olin Fellow in Law and Economics at Harvard Law School and I acknowledge support from the School's John M. Olin Center for Law, Economics and Business. The Article won the Irving Oberman Memorial Prize for Intellectual Property. For their edits and advice, I thank Julie Cohen, Rochelle Cooper Dreyfuss, Robin Feldman, Peter Barton Hutt, Bryan Laulicht, Benjamin Roin, Steven Shavell, Robert Sitkoff and Henry Smith.

direction that makes sense for all parties involved.

INTRODUCTION

Follow-on biologics[1] have attracted a great deal of attention, first as a conceptual matter, and then as a practical matter in the wake of the 2009 Biologics Price Competition and Innovation Act (BPCIA) which created a pathway for "generic" biologics.[2] The literature contains in-depth coverage of questions of proper legislative design,[3] whether follow-on biologics will be safe and effective,[4] and how a pathway for follow-on biologics will affect

1. Alternatively called follow-on protein products and subsequent entry biologics.

2. 42 U.S.C. § 262(i)(2) (West 2012).

3. Brian R. Bouggy, *Follow-On Biologics Legislation: Striking a Balance Between Innovation and Affordability*, 7 IND. HEALTH L. REV. 367 (2010); Tam Q. Dinh, *Potential Pathways for Abbreviated Approval of Generic Biologics under Existing Law and Proposed Reforms to the Law*, 62 FOOD & DRUG L.J. 77, 81 (2007); Michael P. Dougherty, *The New Follow-On-Biologics Law: A Section by Section Analysis of the Patent Litigation Provisions in the Biologics Price Competition and Innovation Act of 2009*, 65 FOOD & DRUG L.J. 231 (2010); Donna M. Gitter, *Innovators and Imitators: An Analysis of Proposed Legislation Implementing an Abbreviated Approval Pathway for Follow-on Biologics in the United States*, 35 FLA. ST. U. L. REV. 555 (2008); Ingrid Kaldre, *The Future of Generic Biologics: Should the United States "Follow-on" the European Pathway?*, 2008 DUKE L. & TECH. REV. 1 (2008); Kathleen Kelleher, Note, *FDA Approval of Generic Biologics: Finding a Regulatory Pathway*, 14 MICH. TELECOMM. & TECH. L. REV. 245, 261-63 (2007); Alana Montas, *Cheaper Clinical Trials: The Real Solution to the Biologic Industry's Gordian Knot*, 37 AM. J.L. & MED. 172 (2011); Jordan Paradise, Foreword, *Follow-On Biologics: Implementation Challenges and Opportunities*, 41 SETON HALL L. REV. 501 (2011); Sarah Sorscher, *A Longer Monopoly for Biologics?: Considering the Implications of Data Exclusivity as a Tool for Innovation Policy*, 23 HARV. J.L. & TECH. 285 (2009-2010); Joyce Wing Yan Tam, *Biologics Revolution: The Intersection of Biotechnology, Patent Law, and Pharmaceutical Regulation*, 98 GEO. L.J. 535, 558-62 (2010); Linfong Tzeng, *Follow-on Biologics, Data Exclusivity, and the FDA*, 25 BERKELEY TECH. L.J. 135 (2010); Dawn Willow, *The Regulation of Biologic Medicine: Innovators' Rights and Access to Healthcare*, 6 J. INTELL. PROP. 32, 34 (2006).

4. Lisa D. DiMartino et al., *Using Medicare Administrative Data to Conduct*

brand-name incentives to innovate.[5] However, the literature contains no comprehensive treatment of patent infringement in the context of follow-on biologics.

Patent infringement is an important topic because, like small molecule generic drugs, follow-on biologics are likely to begin their life facing infringement suits. The BPCIA sets up complex procedures for resolving patent disputes prior to entry.[6] Although follow-on biologics will not enter the market until after expiration of the core (new biological entity) patent covering the reference drug, the reference drug will still be covered by a variety of weaker patents protecting matters such as manufacturing processes, formulation or packaging.[7] Because the BPCIA requires follow-on biologics to be "highly similar to the reference product," there is the potential for patent conflict every time a follow-on biologic enters the market.

Once the first follow-on biologic infringement suit is filed, courts will have the grueling task of sorting through the science to apply it to a body of law invented long before the elemental discoveries of biotechnology even happened, much less understood. Unfortunately, courts do not have the luxury of muddling through early cases and creating conflicting standards before eventually settling into a more coherent body of law with the help of the Federal Circuit. Courts need to create a coherent body of law right from the beginning. It is tremendously expensive to develop a follow-on biologic, so it is

---

Postmarketing Surveillance of Follow-On Biologics: Issues and Opportunities, 63 FOOD & DRUG L.J. 891 (2008); Elysa B. Goldberg, *Fixing a Hole: Will Generic Biologics Find a Niche Within the Hatch-Waxman Act*, 20 FORDHAM INTELL. PROP. MEDIA & ENT. L.J. 327, 331 (2010); Megan Thisse, *Working the Bugs out of Biologics: A Look at the Access to Life-Saving Medicines Act and Follow-On Biologics*, 18 ALB. L.J. SCI. & TECH. 543 (2008); Jeanne Yang, *A Pathway to Follow-On Biologics*, 3 HASTINGS SCI. & TECH. L.J. 217, 220 (2011); Joshua Boger, *Follow-on Biologics: Balancing Innovation and Cost Savings*, Health Care Cost Monitor (Nov.12, 2009), http://healthcarecostmonitor.thehastingscenter.org/joshuaboger/follow-on-biologics-balancing-innovation-and-cost-savings.

5. Katherine N. Addison, *The Impact of the Biosimilars Provision of the Health Care Reform Bill on Innovation Investments*, 10 J. MARSHALL REV. INTELL. PROP. L. 553 (2011); Yaniv Heled, *Patents vs. Statutory Exclusivities in Biological Pharmaceuticals—Do We Really Need Both?*, 18 MICH. TELECOMM. & TECH. L. REV. 419 (2012); Jeremiah J. Kelly, *Follow-on Biologics: Legal, Scientific, and Policy Considerations*, 13 J. HEALTH CARE L. & POL'Y 257, 257 (2010); Jeremiah J. Kelly & Michael David, *No Longer "If," But "When": The Coming Abbreviated Approval Pathway for Follow-on Biologics*, 64 FOOD & DRUG L.J. 115, 138-40 (2009); Maxwell R. Morgan, *Regulation of Innovation Under Follow-On Biologics Legislation: FDA Exclusivity As An Efficient Incentive Mechanism*, 11 COLUM. SCI. & TECH. L. REV. 93 (2010); John A. Vernon et al., *Exploration of Potential Economics of Follow-on Biologics and Implications for Data Exclusivity Periods for Biologics*, 16 B.U. J. SCI. & TECH. L. 55 (2010).

6. *See generally* 42 U.S.C. § 262 (2010).

7. Although the BPCIA contains anti-evergreening provisions intended to curb some of the strategic patenting seen in generic drugs, biologics are still likely to be covered by a broad patent portfolio to give them maximum protection against follow-on biologics. *See infra* Part II.

vital that there be consistency in how they are treated in the courts once the inevitable patent infringement suits arrive. If follow-on biologics companies cannot predict how their product will be received in court, they may decide it is not worth the risk to create it.

A product may infringe either literally, meaning that the accused product copies every detail of the patent, or by equivalents, meaning that there are "insubstantial differences" between the products.[8] A product infringes by equivalents if it does "the same work in substantially the same way[s] and accomplish[es] substantially the same result" even if it "differ[s] in name, form, or shape."[9]

Relative to other types of products, literal infringement is likely to be somewhat less important in the context of follow-on biologics. This is because it is incredibly difficult—perhaps impossible—for the follow-on biologic to be identical to the reference drug.[10] While arguing against the creation of a follow-on biologic pathway, the brand-name industry itself stated that "[t]o achieve identical composition between biologics produced by unrelated manufacturers is virtually impossible because of the nature of biological manufacturing."[11] While this does not mean that a follow-on biologic cannot literally infringe, it does suggest that literal infringement will be a more challenging argument. Therefore the doctrine of equivalents will likely be of outsized importance in infringement litigation concerning follow-on biologics.[12]

---

8. Warner-Jenkinson Co., Inc. v. Hilton Davis Chem. Co., 520 U.S. 17, 19 (1997).

9. Union Paper-Bag Mach. Co. v. Murphy, 97 U.S. 120, 125 (1877).

10. This has led to several articles suggesting that the difficulty in replicating biologics might mean that the patents covering biologics are not properly enabled and thus invalid. Dmitry Karshtedt, *Limits on Hard-to-Reproduce Inventions: Process Elements and Biotechnology's Compliance with the Enablement Requirement*, 3 HASTINGS SCI. & TECH. L.J. 109 (2011); Gregory N. Mandel, *The Generic Biologics Debate: Industry's Unintended Admission that Biotech Patents Fail Enablement*, 11 VA. J.L. & TECH. 11, 1 (2006); Joyce Wing Yan Tam, *Biologics Revolution: The Intersection of Biotechnology, Patent Law, and Pharmaceutical Regulation*, 98 GEO. L.J. 536, 544-47 (2010).

11. Memorandum from the Pharm. Research and Mfrs. of Am. to the Food and Drug Admin. (Nov. 12, 2004), http://www.fda.gov/ohrms/dockets/dockets/04n0355/04n-0355-c000004-01-vol1.pdf.

12. Courts and scholars have recognized that the science of biotechnology makes it uniquely challenging to apply the doctrine of equivalents to biotechnology. *See, e.g.*, Lawrence S. Graham, *Equitable Equivalents: Biotechnology and the Doctrine of Equivalents After* Warner-Jenkinson Co. v. Hilton Davis Chemical Co., 6 J.L. & POL'Y 741, 777-85 (1997-1998) (arguing that the doctrine of equivalent is not readily applied to biotechnology and providing several examples of cases where it is inappropriate); Qing Lin, *A Proposed Test for Applying the Doctrine of Equivalents to Biotechnology Inventions: The Nonobviousness Test*, 74 WASH. L. REV. 885, 900 (1999) (arguing that the doctrine of equivalents is difficult to apply to biotechnology because scientists often do not understand the "way" in which a biotechnology product works, and therefore cannot provide enough evidence to fulfill the "way" requirement). *See also Genentech, Inc. v. Wellcome Found. Ltd.*, 29 F.3d 1555, 1570 (Fed. Cir. 1994) (Lourie, J., concurring) (explaining some of the difficulties in applying the doctrine of equivalents to the case at hand). However, the

This Article fills a gap in the literature by conducting a comprehensive analysis of infringement in the context of follow-on biologics. I look both at the types of strategies that follow-on biologics companies are likely to use to avoid infringement, and how courts are likely to respond to these strategies. I focus predominantly on the doctrine of equivalents, both because it represents the outer limits of the scope of a patent, and thus the most difficult cases, and because it will be particularly important in suits concerning follow-on biologics. I find that it will be easiest for follow-on biologics to make changes at certain stages of the manufacturing process where the BPCIA and FDA regulations give them more latitude to stray from the precise form of the brand-name product. These less regulated areas give follow-on biologics companies greater scope to make changes that will bring them outside the range of equivalents for the brand-name product.

The Article is important for courts that must create a coherent body of law where no precedent yet exists. It is important for policy makers and scholars who seek to understand the nature of follow-on biologics infringement suits and how to shape this body of law in a direction that makes sense for the ultimate stakeholder: the patient.

In Part I, I give a brief explanation of patent infringement, with an emphasis on the doctrine of equivalents, because it defines the outer border of patent protection. In Part II, I define "biologics," explain how they differ from small molecule drugs, and describe the history of biologics regulation world-wide, in particular the history of the BPCIA. I next summarize the follow-on biologics that have been approved in Europe, and what changes those follow-on biologics have made from the innovator drug. I summarize the FDA regulations governing follow-on biologics and what they mean for types of work-arounds that follow-on biologics will be permitted to attempt.

In Part III, I explore how the BPCIA and patent law will shape infringement suits. I draw my conclusions from the interaction between patent law and the BPCIA, which make certain types of infringement more likely, from infringement cases involving biotechnology (not follow-on biologics, as none have been brought—yet) and from doctrine of equivalents suits that have been brought for generic small molecule drugs. I then make policy suggests for how courts should treat these cases when they inevitably begin arriving on dockets across the country.

---

doctrine of equivalents has been applied to numerous biotechnology cases, suggesting that courts will continue attempting to apply it in the context of follow-on biologics. *See, e.g.*, Amgen Inc. v. Hoffmann-La Roche Ltd., 580 F.3d 1340 (Fed. Cir. 2009); Carnegie Mellon University v. Hoffmann-La Roche Inc., 541 F.3d 1115 (Fed. Cir. 2008); Boehringer Ingelheim Vetmetica, Inc. v. Schering-Plough Corp., 320 F.3d 1339 (Fed. Cir. 2003); Genentech, Inc. v. Boehringer Mannheim GmbH, 47 F. Supp. 2d 91 (D. Mass. 1999).

I. PATENT LAW AND THE DOCTRINE OF EQUIVALENTS

Patents "promote the Progress of Science and useful Arts"[13] by granting property rights in information in exchange for full disclosure of the invention.[14] However, patent law exists in a careful balance. If inventors receive too little reward for their invention, innovation will decrease. If inventors receive too much reward for their invention, their monopoly rights prevent secondary innovation[15] and may prevent optimal public use.[16]

---

13. U.S. CONST. art. I, § 8, cl. 8.

14. *See, e.g.*, Graham v. John Deere Co. of Kan. City, 383 U.S. 1, 9 (1966) ("The patent monopoly was not designed to secure to the inventor his natural right in his discoveries. Rather, it was a reward, an inducement, to bring forth new knowledge."); Mazer v. Stein, 347 U.S. 201, 219 (1954) ("The economic philosophy behind the clause empowering Congress to grant patents and copyrights is the conviction that encouragement of individual effort by personal gain is the best way to advance public welfare through the talents of authors and inventors in 'Science and useful Arts.'"). *See also* Dan L. Burk & Mark A. Lemley, *Policy Levers in Patent Law*, 89 VA. L. REV. 1575, 1576 (2003) ("Patent law is our primary policy tool to promote innovation, encourage the development of new technologies, and increase the fund of human knowledge."); Mark A. Lemley & David McGowan, *Legal Implications of Network Economic Effects*, 86 CALIF. L. REV. 479, 526 (1998) ("Indeed, the principle behind intellectual property law is that competition should be sacrificed to some extent in order to give sufficient incentive for innovation."); Lawrence Lessig, *Intellectual Property and Code*, 11 ST. JOHN'S J. LEGAL COMMENT. 635, 638 (1996) ("while we protect real property to protect the owner from harm, we protect intellectual property to provide the owner sufficient incentive to produce such property.").

15. *See, e.g.*, Jonathan M. Barnett, *Cultivating the Genetic Commons: Imperfect Patent Protection and the Network Model of Innovation*, 37 SAN DIEGO L. REV. 987, 1000 (2000) ("Today academic and industrial researchers commonly lament the ballooning costs of navigating around proliferating clusters of patent claims, and some commentators contend that patent claims ultimately will result in upstream strangleholds on basic-research discoveries that will significantly impede downstream technological applications."); Michael A. Carrier, *Resolving the Patent-Antitrust Paradox Through Tripartite Innovation*, 56 VAND. L. REV. 1047, 1081-85 (2003) (explaining how patents can interfere with cumulative innovation); Michael A. Heller & Rebecca S. Eisenberg, *Can Patents Deter Innovation? The Anticommons in Biomedical Research*, 280 SCI. 698, 698 (1998) (arguing in the context of biomedical research that patent holders can impede downstream research); Lisa Mandrusiak, *Balancing Open Source Paradigms and Traditional Intellectual Property Models to Optimize Innovation*, 63 ME. L. REV. 303, 310-11 (2010) (providing an overview of the anticommons patent problem); Robert P. Merges & Richard R. Nelson, *On the Complex Economics of Patent Scope*, 90 COLUM. L. REV. 839, 843 (1990). However, some scholars have argued that the original inventor is in the best position to develop and coordinate downstream innovation. *See* Edmund W. Kitch, *The Nature and Function of the Patent System*, 20 J.L. & ECON. 265, 276-77 (1977).

16. Optimal public use is demonstrated by charging inflated prices for the product, which increases prices for consumers. *See* Note, *Limiting the Anticompetitive Prerogative of Patent Owners: Predatory Standards in Patent Licensing*, 92 YALE L.J. 831, 836 (1983) ("The patent system . . . reflects a tradeoff between dynamic and static efficiency.").

A related problem is the recent rise of patent trolls. For more information on the problem of patent trolls, *see, e.g.,* Einer Elhauge, *Do Patent Holdup and Royalty Stacking Lead to Systematically Excessive Royalties?*, 4 J. COMPETITION L. & ECON. 535, 537 (2008); Damien Geradin et al., *The Complements Problem Within Standard Setting: Assessing the*

A patent is made up of two main parts: the specification and the claims.[17] The specification is a narrative description of the invention, while the claims define the boundaries of the patent. Historically, claims were not required in a patent, which consisted only of a description of the invention in the specification.[18] Courts interpreting the patent looked at the specification to determine the "essence" of the patent[19] in order to answer fuzzy questions about the "similarity" of the inventions.[20] The specification-only system worked poorly.[21] Because the specification did not clearly define the bounds of the patent, it was nearly impossible for either the patentee or the public to determine exactly where those boundaries were located.[22]

Claims were first statutorily required in the Patent Act of 1870.[23] Claims define the bounds of the patent's scope.[24] Claims also serve a public notice

---

Evidence on Royalty Stacking, 14 B.U. J. SCI. & TECH. L. 144, 145 (2008); John M. Golden, Commentary, "*Patent Trolls" and Patent Remedies*, 85 TEX. L. REV. 2111, 2145-47 (2007); J. Gregory Sidak, *Holdup, Royalty Stacking, and the Presumption of Injunctive Relief for Patent Infringement: A Reply to Lemley and Shapiro*, 92 MINN. L. REV. 714, 714 (2008). However, note that the negative view of patent trolls is not unanimous. Some think that they provide a useful economic function. *See, e.g.*, Sannu K. Shrestha, *Trolls or Market-Makers? An Empirical Analysis of Nonpracticing Entities*, 110 COLUM. L. REV. 114, 115-16 (2010) (suggesting that patent trolls enhance innovation by serving a sort of venture capital role to capital-poor inventors by creating a market for patents and inventions). *See also* James F. McDonough III, Comment, *The Myth of the Patent Troll: An Alternative View of the Function of Patent Dealers in an Idea Economy*, 56 EMORY L.J. 189, 190 (2006) ("[T]rolls act as a market intermediary in the patent market. Patent trolls provide liquidity, market clearing, and increased efficiency to the patent markets—the same benefits securities dealers supply capital markets.").

17. 35 U.S.C. § 112 (2006).

18. Christopher A. Cotropia, *Patent Claim Interpretation Methodologies and Their Claim Scope Paradigms*, 49 WM. & MARY L. REV. 49, 63 (2005) (explaining that the 1793 patent statute did not require a claim).

19. Odiorne v. Winkley, 18 F. Cas. 581, 582 (C.C.D. Mass. 1814) (No. 10,432).

20. Keystone Bridge Co. v. Phoenix Iron Co., 95 U.S. 274, 278 (1877).

21. John F. Duffy, *The Festo Decision and the Return of the Supreme Court to the Bar of Patents*, 2002 SUP. CT. REV. 273, 309 (2002) (Noting that "lay jurors would find no infringement because they would see many superficial differences between the defendant's machine and the description of the patented invention and thus believe the two not substantially identical.").

22. Cotropia, *supra* note 18, at 63.

23. Act of July 8, 1870, § 26, 16 Stat. 201 (the patent must "particularly point out and distinctly claim" the invention).

24. *See, e.g.*, Festo Corp. v. Shoketsu Kinzoku Kogyo Kabushiki Co., 234 F.3d 558, 622 (Fed. Cir. 2000) ("In drafting an original claim of a patent application, the writer sets out the metes and bounds of the invention . . ."); Markman v. Westview Instruments, Inc., 52 F.3d 967, 1000 (Fed. Cir. 1995) ("The legal effect of the patent claim is to establish the metes and bounds of the patent right to exclude . . ."); Zenith Labs., Inc. v. Bristol-Myers Squibb Co., 19 F.3d 1418, 1424 (Fed. Cir. 1994) ("It is the claim that sets the metes and bounds of the invention entitled to the protection of the patent system."); Corning Glass Works v. Sumitomo Elec. U.S.A., Inc., 868 F.2d 1251, 1257 (Fed. Cir. 1989) ("A claim in a patent provides the metes and bounds of the right which the patent confers on the patentee to exclude others from making, using or selling the protected invention.").

function, giving third parties notice of the existence of the claim and the location of the boundaries.[25] Claims can then be used by competitors as a guide to designing around the patent.[26] The Supreme Court promotes using claims to develop work-arounds, stating, for example, that claims "inform the public during the life of the patent of the limits of the monopoly asserted, so that it may be known which features may be safely used or manufactured without a license and which may not."[27]In an infringement suit, the court begins by inquiring as to whether a claim has been literally infringed, meaning that the defendant has copied every detail of the claim.[28] However, the Supreme Court worried that restricting patent protection to cases where the defendant literally infringed would make it simple for an "unscrupulous copyist to make unimportant and insubstantial changes and substitutions in the patent[,]"[29] which would greatly diminish the value of the patent.[30]

---

25. PSC Computer Prods., Inc. v. Foxconn Int'l, Inc., 355 F.3d 1353, 1359 (Fed. Cir. 2004) ("[C]laims serve the important notice function of informing the public that anyone who makes, uses, or sells the claimed invention infringes the patent.").

26. *See, e.g.*, Read Corp. v. Porter, Inc., 970 F.2d 816, 828 (Fed. Cir. 1992) ("We have often noted that one of the benefits of the patent system is the incentive it provides for 'designing around' patented inventions, thus creating new innovations."); Slimfold Mfg. v. Kinkead Indus., 932 F.2d 1453, 1457 (Fed. Cir. 1991) ("Designing around patents is, in fact, one of the ways in which the patent system works to the advantage of the public in promoting progress in the useful arts, its constitutional purpose."); State Indus., Inc. v. A.O. Smith Corp., 751 F.2d 1226 (Fed. Cir. 1985) ("One of the benefits of a patent system is its so-called "negative incentive" to "design around" a competitor's products, even when they are patented, thus bringing a steady flow of innovation to the marketplace. It should not be discouraged . . ."). However, courts do not always regard designing-around as a benefit of the patent system. *See, e.g.*, Graver Tank & Mfg. Co. v. Linde Air Prods. Co., 339 U.S. 605, 607 (1950) (concerned that allowing too much design-around would "covert the protection of the patent grant into a hollow and useless thing."). *See also* Tun-Jen Chiang, *The Levels of Abstraction Problem in Patent Law*, 105 NW. U. L. REV. 1097, 1138 (2011) (arguing that if the patent's scope were "confined to precise replication . . . , then pirates would quickly learn to copy the principle or the heart of the patent without replicating the precise embodiment . . . [.] [P]rotection limited to literal reproduction is worthless and easily circumvented.").

27. Permutit Co. v. Graver Corp., 284 U.S. 52, 60 (1931).

28. *Graver Tank*, 339 U.S. at 607.

29. *Id*. *See also* Tun-Jen Chiang, *The Levels of Abstraction Problem in Patent Law*, 105 NW. U. L. REV. 1097, 1138 (2011) (explaining the importance of the doctrine of equivalents). *But see* Timothy R. Holbrook, *Equivalency and Patent Law's Possession Paradox*, 23 HARV. J.L. & TECH. 1, 39 (2009) (pointing out that the doctrine of equivalents may be responsible for decreasing a patentee's incentive for downstream innovation because if there was no doctrine of equivalents, patentees would have an incentive to "continue to innovate and improve upon her invention because others will have the opportunity to invent and patent improvements on it."); Lee Petherbridge, *On the Decline of the Doctrine of Equivalents*, 31 CARDOZO L. REV. 1371, 1404 (2010) (noting that courts are increasingly reluctant to rule for plaintiffs on doctrine of equivalents grounds but that "[a]mple evidence suggests that all the while the courts were killing the doctrine of equivalents, patent applicants were increasing the rate at which they filed applications for new inventions.").

30. Festo Corp. v. Shoketsu Kinzoku Kogyo Kabushiki Co., 535 U.S. 722, 731 (2002) ("If patents were always interpreted by their literal terms, their value would be greatly

The Court therefore expanded the scope of the monopoly that patentees could claim by creating the equitable doctrine of equivalents.[31] The doctrine of equivalents expands the patentee's right to exclude beyond the fence created by the literal meaning of the claims to include inventions that perform "substantially the same function in substantially the same way to obtain the same result."[32] The intent is to prevent a competitor from committing a "fraud on a patent"[33] by creating a product that is functionally identical to the patented product and thus should equitably fall within the patent's scope.

Infringement under the doctrine of equivalents can be decided using one of two tests. The "function-way-result" test asks whether the defendant's device functions substantially the same way to achieve substantially the same result.[34] The "insubstantial differences" test asks whether the defendant's device is substantially different from the patent scope.[35] The Supreme Court has expressed no preference between the tests, stating that the "particular linguistic framework used [to determine equivalency] is less important than whether the test is probative of the essential inquiry."[36]

Equivalency is a question for the jury[37] although in practice it is often decided on summary judgment.[38] An inquiry into equivalence is fact-heavy and must consider "the context of the patent, the prior art, and the particular circumstances of the case. Equivalence . . . is not the prisoner of a formula and is not an absolute to be considered in a vacuum."[39] The Supreme Court also instructs juries to consider "whether persons reasonably skilled in the art would have known of the interchangeability of an ingredient not contained in the patent with one that was."[40] In fields involving quickly developing technology, the doctrine protects patentees from "'after-arising' technology because a patent draftsman has no way to anticipate and account for later developed

---

diminished.").

31. The doctrine of equivalents first appeared in Supreme Court jurisprudence in *Winans v. Adams*, 56 U.S. 330 (1853).

32. Graver Tank & Mfg. Co. v. Linde Air Prods. Co., 339 U.S. 605, 608 (1950) (quoting Sanitary Refrigerator Co. v. Winters, 280 U.S. 30, 42 (1929)).

33. *Graver Tank*, 399 U.S. at 608.

34. *Sanitary Refrigerator Co.*, 280 U.S. at 42.

35. Warner-Jenkinson Co. v. Hilton Davis Chem. Co., 520 U.S. 17, 39 (1997).

36. *Id*. at 40.

37. *Graver Tank*, 399 U.S. at 609.

38. Allison & Lemley suggest that the doctrine of equivalents has been in decline since *Markman* made claim construction a matter of law. Because courts now resolve questions of claim construction as a matter of law, they are incentivized to resolve the entire matter on summary judgment to avoid a trial. Thus if they make a finding on literal infringement as a matter of law, they are likely to do the same for infringement under the doctrine of equivalents. John R. Allison & Mark A. Lemley, *The (Unnoticed) Demise of the Doctrine of Equivalents*, 59 STAN. L. REV. 955, 977 (2007).

39. *Graver Tank*, 339 U.S. at 609.

40. *Id*.

substitutes for a claim element."[41] Whether the accused product is patented is relevant, but not dispositive.[42]

The doctrine of equivalents is controversial.[43] Part of the controversy—both scholarly and judicial—arises because the doctrine creates an inherent tension between its goal of protecting patent rights and its unintended consequence of increasing uncertainty and reducing the clarity of patents.[44] Patents, like any property right, function best when they clearly delineate the boundaries of the property, enabling other parties to invest and invent around those boundaries with confidence that they are not infringing. On one hand, the doctrine of equivalents reflects courts' desire to ensure patent protection is broad enough that inventors have an incentive to innovate and to publically disclose their inventions.[45] However, by extending a patent's boundaries to

---

41. Al-Site Corp. v. VSI Int'l, Inc., 174 F.3d 1308, 1320 n.2 (Fed. Cir. 1999). *See also* Chiuminatta Concrete Concepts, Inc. v. Cardinal Indus., Inc., 145 F.3d 1303, 1310 (Fed. Cir. 1998) ("Due to technological advances, a variant of an invention may be developed after the patent is granted, and that variant may constitute so insubstantial a change from what is claimed in the patent that it should be held to be an infringement.").

42. Federal Circuit jurisprudence is imprecise on the importance of this factor. In *Hoechst Celanese Corp. v. BP Chems. Ltd.*, the Federal Circuit held that the accused product's "patentability presents no legal or evidentiary presumption of noninfringement." 78 F.3d 1575, 1582 (Fed. Cir. 1996). The fact that the defendant's device is patented over the plaintiff's device does not preclude a finding that the defendant's device infringes by equivalents. Roton Barrier, Inc. v. Stanley Works, 79 F.3d 1112, 1128 (Fed. Cir. 1996) (Nies, J., concurring). However, if an accused product is patented, the USPTO must have determined that the accused product did not read onto the plaintiff's patent, which would have been prior art. In *Hoganas AB v. Dresser* Indus., the Federal Circuit noted that the defendant had obtained a patent covering their product, and that the plaintiff's patent was listed as art of record for the defendant's product, but that the USPTO had nevertheless granted the defendant's patent. 9 F.3d 948, 954 (Fed. Cir. 1993).

43. Michael J. Meurer & Craig A. Nard, *Innovation, Refinement and Patent Claim Scope: A New Perspective on the Doctrine of Equivalents*, 93 GEO. L.J. 1947, 1948 (2005); Lee Petherbridge, *On the Decline of the Doctrine of Equivalents*, 31 CARDOZO L. REV. 1371, 1372 (2010). *See also* Martin J. Adelman and Gary L. Francione, *The Doctrine of Equivalents in Patent Law: Questions that Pennwalt Did Not Answer*, 137 U. PENN. L. REV. 673 (1989); Joseph S. Cianfrani, *An Economic Analysis of the Doctrine of Equivalents*, 1 VA. J.L. & TECH. 1 (1997); Paul R. Michel, *The Role and Responsibility of Patent Attorneys in Improving the Doctrine of Equivalents*, 40 IDEA 123 (2000); Joshua D. Sarnoff, *Abolishing the Doctrine of Equivalents and Claiming the Future after Festo*, 19 BERK. TECH. L.J. 1157 (2004); John R. Thomas, *Claim Re-Construction: The Doctrine of Equivalents in the Post-Markman Era*, 9 LEWIS & CLARK L. REV. 153 (2005); T. Whitley Chandler, *Prosecution History Estoppel, the Doctrine of Equivalents, and the Scope of Patents*, 13 HARV. J.L. & TECH. 465 (2000).

44. *E.g.*, Warner-Jenkinson Co. v. Hilton Davis Chem. Co., 520 U.S. 17, 29 (1996) ("[T]he doctrine of equivalents, when applied broadly, conflicts with the definitional and public-notice functions of the statutory claiming requirement."); D. Alan White, *The Doctrine of Equivalents: Fairness and Uncertainty in an Era of Biologic Pharmaceuticals*, 60 EMORY L.J. 751, 773 (2011); Petherbridge, *supra* note 43, at 1374.

45. White, *supra* note 44, at 756. Petherbridge suggests that the doctrine of equivalents bolsters the patent system because innovators might be reluctant to make new inventions if competitors could get around their patent by making minor changes to the product. *Supra*

inventions not literally within the bounds of the claims, the doctrine increases uncertainty and may deter investment and business activities.[46]

It is unclear how this problem plays out in the pharmaceutical and biotechnology industries. Industries innovate differently, thus conclusions about patents in general or about a specific industry do not always apply to a particular industry.[47] Empirical research has shown that pharmaceutical and biotechnology industries are underrepresented in doctrine of equivalent cases compared to the mechanical and electronics industries, accounting for only 9.2% of doctrine of equivalents cases compared to 11.5% of all patents.[48] Several theories attempt to explain this discrepancy. One conjecture is that the doctrine was designed for mechanical inventions and thus works less well for other industries.[49] A second hypothesis suggests that the information technology industry changes rapidly and thus its inventions are less well expressed in patent claims, making the doctrine of equivalents more important than in the life sciences industry where it is easier to express the scope of an invention in the patent claim.[50] Regardless, the success rate for plaintiffs using the doctrine of equivalents is consistent (and consistently low) across industries.[51]

However, patents remain a vital part of the pharmaceutical industry[52] and uncertainty in patent boundaries would surely make it difficult for firms to raise funds and develop products.[53] Moreover, if the theory that the doctrine was designed for mechanical inventions and thus works best in that industry is true, it follows that the doctrine would be less predictable and less well applied in the life sciences industries, leading to even more confusion and uncertainty. In addition, most pharmaceutical and biotechnology cases involve complex science, and many studies have shown that juries (and judges) struggle with

---

note 43, at 1374.

46. Donald S. Chisum, *The Scope of Protection for Patents After the Supreme Court's Warner-Jenkinson Decision: The Fair Protection-Certainty Conundrum*, 14 SANTA CLARA COMPUTER & HIGH TECH. L.J. 1, 1, 62 (1998).

47. *E.g.*, Dan L. Burk & Mark A. Lemley, *Policy Levers in Patent Law*, 89 VA. L. REV. 1575 (2003).

48. Allison & Lemley, *supra* note 38, at 972-73.

49. *Id.* at 973.

50. Julie E. Cohen and Mark A. Lemley, *Patent Scope and Innovation in the Software Industry*, 89 CALIF. L. REV. 1, 45-47 (2001).

51. Allison & Lemley, *supra* note 38, at 973.

52. Natalie M. Derzko, *The Impact of Recent Reforms of the Hatch-Waxman Scheme on Orange Book Strategic Behavior and Pharmaceutical Innovation*, 45 IDEA 165 (2005) (explaining how FDA exclusivity periods are short enough that there is almost always a period of time when a drug is covered by a patent but not by market exclusivity).

53. *E.g.*, Henry Grabowski et al., *The Market for Follow-On Biologics: How Will it Evolve?*, 25 HEALTH AFF. 1291, 1300 (2006) ("[I]ncreased uncertainty and IP litigation in biotech also would have major negative-incentive effects on capital market decisions for developing private and public biotech firms with promising pipelines.").

scientific cases.[54]

Irrespective of the challenges inherent in the doctrine, it remains influential and popular. Over the past year, 190 district court cases cited the doctrine, and the Federal Circuit heard two doctrine of equivalents cases.[55]

II. BIOLOGICS AND FOLLOW-ON BIOLOGICS

Biologics are regulated under the Public Health and Service Act (PHSA).[56] A biologic, or "biological product" is defined to mean "a virus, therapeutic serum, toxin, antitoxin, vaccine, blood, blood component or derivative, allergenic product, protein (except any chemically synthesized polypeptide), or analogous product . . . applicable to the prevention, treatment, or cure of a disease or condition of human beings."[57] Biologics are complex proteins which are bigger, more intricate, and more poorly-understood than small molecule drugs.[58] Biologics can be extracted from animal cells or tissue that naturally produce the protein or scientists can genetically modify cells or tissue to create a system that produces larger quantities of the protein. Because of the potential to scale up production, most biologic proteins are produced using the latter technique.[59]

It is much more difficult to create (and to regulate) a "generic" biological product than a generic small molecule drug. Small molecule generics usually include an identical active ingredient which is chemically identical to the brand name drug's active ingredient, and which can be synthesized in a predictable and replicable process.[60] Small molecule drugs are also generally easy to characterize. Biologics, in contrast, cannot be synthesized chemically and are instead usually produced through a recombinant cell line.[61] Compounding these challenges, the details of the production process used by the pioneer company

---

54. Alan Feigenbaum, *Special Juries: Deterring Spurious Medical Malpractice Litigation in State Courts*, 24 CARDOZO L. REV. 1361, 1389-96 (2003); Jody Weisberg Menon, *Adversarial Medical and Scientific Testimony and Lay Jurors: A Proposal for Medical Malpractice Reform*, 21 AM. J.L. & MED. 281, 281 (1995).

55. To get a rough estimate of the number of cases citing the doctrine of equivalents, I searched on Westlaw's ALLFEDS database for "doctrine of equivalents" and restricted to "year to date" (searching between 01/13/2011 and 01/13/2012).

56. 42 U.S.C. § 262(j) (2006).

57. 42 U.S.C. § 262(i)(1) (2006).

58. *See, e.g.*, Alan J. Morrison, *Biosimilars in the United States: A Brief Look at Where We Are and the Road Ahead*, 26 BIOTECHNOLOGY L. REP. 463, 465 (2007).

59. Robert N. Sahr, *The Biologics Price Competition and Innovation Act: Innovation Must Come Before Price Competition*, B.C. INTELL. PROP. & TECH. F., 2009, at 6.

60. Jeanne Yang, *A Pathway to Follow-On Biologics*, 3 HASTINGS SCI. & TECH. L.J. 217, 221 (2011).

61. *Id.* Note this inability to synthesize the biologic only extends to protein drugs. Some nucleotide products can be synthesized chemically.

are protected by various intellectual property methods.[62] The production process is thus not fully controlled (or understood), and small differences in production process—or even production by the same process but in a different facility—can result in differences in the product, which can have adverse clinical consequences.[63] Moreover, it may not even be possible, given the current state of scientific knowledge, to determine whether two biologics are, in fact, identical.[64]

Because of the challenges in reproducing biologics and the lack of sensitive assays for differences,[65] the data requirements for comparing follow-on biologics to a reference product are likely to be considerably higher than the data requirements for generic companies comparing their small molecule drug to a reference product.[66] Small molecule drug manufacturers are usually required to conduct approximately 40 to 50 clinical tests, whereas follow-on biologic manufacturers in Europe (which has had follow-on biologic legislation

---

62. Islah Ahmed et al., *Follow-on Biologics: Impact of Biologic Product Life Cycle and European Experience on the Regulatory Trajectory in the United States, Clinical Therapeutics* (forthcoming, 2012) (manuscript at 4) (on file with author),; *See also* U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES ET AL., GUIDANCE FOR INDUSTRY: SCIENTIFIC CONSIDERATIONS IN DEMONSTRATING BIOSIMILARITY TO A REFERENCE PRODUCT (DRAFT GUIDANCE), 5-6 (February 2012) [hereinafter *Scientific Considerations*] ("[T]he manufacturer of a proposed [biosimilar] product will likely have . . . . . .no direct knowledge of the manufacturing process for the reference product.").

63. Paul J. Declerck, *Biotherapeutics in the Era of Biosimilars: What Really Matters is Patient Safety*, 30 DRUG SAFETY 1087, 1088 (2007) ("Small distinctions in the cell line, the manufacturing process or in any step from the cell line stage through to administration to the patient can make a major difference in adverse effects observed during treatment . . . . Therefore, unlike chemical pharmaceuticals, substitutions between biologics, including [follow-on biologics], can have clinical consequences and create health concerns for patients.").

64. Erika Lietzan & Emily Alexander, *Biosimilars: What US Regulators Might Learn From Others*, REG. AFF. PHARMA 18, 19 (2011) (Speakers at the FDA's comment session regarding implementation of the BPCIA "disagreed sharply over whether it is even possible for a biosimilar applicant to satisfy [the BPCIA's interchangeability] standard given the current state of science."); *see also* Ahmed, *supra* note 62 ("In theory you can develop technology sensitive enough to establish clinically relevant thresholds of heretogeneity such that Hatch-Waxman type structure could be applied. In practice, this is extremely challenging because it is difficult to establish a correlation between biophysical differences and clinical effects."); Derclerck, *supra* note 63, at 1089 ("As a consequence of the complexity of both the biotechnology product and the production process . . . and the limitation of sensitivity of analytical tools (i.e. the process determines the product), no solid scientific grounds exist to guarantee safe interchangeability between any biologics . . . obtained through different manufacturers.").

65. *See, e.g.*, Ahmed, *supra* note 62 ("For many structurally complex drugs, current technology is insufficient for establishing the identical nature of the active molecule in comparison to the approved reference.").

66. *See, e.g.*, Jonathan Stroud, *The Illusion of Interchangeability: The Benefits and Dangers of Guidance-Plus Rulemaking in the FDA's Biosimilar Approval Process*, 63 ADMIN. L. REV. 599, 624 (2011) ("[T]he burden of evidence for generic biologics applicants could be far higher than it is for generic drugs . . .").

since 2003) are required to conduct over 200 tests.[67] In addition, the plain language of the BPCIA seems to require much more data than the plain language of the Hatch-Waxman Act.[68] Moreover, the BPCIA empowers the FDA to request additional data beyond the statutory requirements, whereas the Hatch-Waxman Act explicitly does not allow this.[69]

Despite these difficulties, economic pressure to lower healthcare costs led governments worldwide to attempt to develop an abbreviated approval pathway for generics. Many countries, including Canada, Japan, Korea, and the European Union, have developed such pathways.[70] In the United States, various lobbying groups and members of Congress began to push for new legislation to create such a pathway.[71]

In 2009, Congress passed the Biologics Price Competition and Innovation Act (BPCIA),[72] a subtitle within the larger Patient Protection and Affordable Care Act. The statute defines follow-on biologic to mean "(A) that the biological product is highly similar to the reference product notwithstanding minor differences in clinically inactive components; and (B) there are no clinically meaningful differences between the biological product and the reference product in terms of the safety, purity, and potency of the product."[73] The statute also describes the data requirements for a follow-on biologic application. An applicant must submit "analytical studies that demonstrate that the biological product is highly similar to the reference product notwithstanding minor differences in clinically inactive components," and animal studies and clinical studies that are "sufficient to demonstrate safety, purity, and potency in one or more appropriate conditions of use . . ."[74] The applicant must also show that both products use "the same . . . mechanisms of action" and that "the route of administration, the dosage form, and the strength of the biological product are the same as those of the reference product" and that the facility in which the follow-on biologic is produced "meets standards designed to assure that the

---

67. Ingrid Kaldre, *The Future of Generic Biologics: Should the United States 'Follow-On' the European Pathway?* 9 DUKE L. & TECH. REV. ¶14 (2008).

68. Stroud, *supra* note 66, at 625 ("[The follow-on biologics] standards outlined [in the BPCIA] will require additional studies showing that the physical chemical structures of the two biologics are highly similar.").

69. *Id.* at 627.

70. For a discussion of the differences in the regulatory schemes of the countries that have an abbreviated biologics pathway, see Noel Courage & Ainslie Parsons, *The Comparability Conundrum: Biosimilars in the United States, Europe and Canada*, 66 FOOD & DRUG L.J. 203 (2011).

71. For a description of the legislative history of the BPCIA, *see generally*, Krista Hessler Carver, Jeffrey Elikan & Erika Lietzan, *An Unofficial Legislative History of the Biologics Price Competition and Innovation Act of 2009*, 65 FOOD & DRUG L.J. 671 (2010).

72. 42 U.S.C. § 262 (2012).

73. 42 U.S.C. § 262(i)(2) (2012).

74. 42 U.S.C. § 262(k)(2)(A)(i) (2012).

biological product continues to be safe, pure, and potent."[75]

The BPCIA also includes a provision for determining when a follow-on biologic is sufficiently similar to the reference product that it may be deemed "interchangeable" with the reference product and be substituted for the brand name drug by a pharmacist even if the physician did not prescribe the follow-on biologic.[76] The standard for "interchangeability" is a product that "(i) is biosimilar to the reference product; and (ii) can be expected to produce the same clinical results as the reference product in any given patient . . ."[77]

Unlike the Hatch-Waxman Act, the BPCIA does not include a 180-day exclusivity period for the first generic company to challenge a patent, or a 30-month stay when a brand name company sues.[78] In addition, the BPCIA includes an "anti-evergreening" provision: a list of improvements in a drug that do not qualify for an exclusivity period—an effort to reduce the strategic small improvements made by producers of small molecule drugs in an attempt to extend their market monopoly.[79] The anti-evergreening provision provides that the following improvements will not receive exclusivity: (i) "a supplement for the biological product that is the reference product" or an application filed by the sponsor of the original reference product for a change "that results in a new indication, route of administration, dosing schedule, dosage form, delivery system, delivery device, or strength" or "a modification in the structure of the biological product that does not result in a change in safety, purity, or potency."[80] Further, the BPCIA gives brand-name companies twelve years of data exclusivity as compared to the Hatch-Waxman Act's five years of market exclusivity.[81]

The interpretation of many portions of the BPCIA has been left to FDA discretion.[82] The FDA must define "highly similar" and "interchangeable."[83] In addition, the FDA must determine what tests must be done and what data must be acquired in order to satisfy the similarity requirements.[84] The FDA has indicated that it will look at biologics on a case-by-case basis, rather than a one-size-fits-all approach. The FDA Commissioner, Dr. Margaret Hamburg, stated that "there will not be a 'one-size-fits-all' approach. There will, rather, be a science-driven, case-by-case decision-making process rooted in the

---

75. *Id.*

76. 42 U.S.C. § 262(i)(3) (2012).

77. 42 U.S.C. § 262(k)(4) (2012).

78. Henry Grabowski, Genia Long & Richard Mortimer, *Implementation of the Biosimilar Pathway: Economic and Policy Issues*, 41 SETON HALL L. REV. 511, 515 (2011).

79. For a comprehensive discussion of evergreening, *see* ROBIN FELDMAN, RETHINKING PATENT LAW 170-77 (2012).

80. 42 U.S.C. § 262(k)(7)(C) (2012).

81. Patient Protection and Affordable Care Act, 124 Stat. 119 (2010).

82. 42 U.S.C. § 262(k)(8) (2012).

83. Courage & Parsons, *supra* note 70, at 215.

84. 75 Fed. Reg. 61498 (Oct. 5, 2010).

regulatory studies . . .”[85] Other FDA officials, including Dr. Janet Woodcock, Director for the Center for Drug Evaluation and Research, published an article in the New England Journal of Medicine stating that given “the complex nature of biologics, it’s unlikely that a “one size fits all” systematic assessment of [follow-on biologicity] can be developed. Instead, FDA scientists will need to integrate various types of information to provide an overall assessment that a biologic is [follow-on biologic] to an approved reference product.”[86]

On February 9, 2012, the FDA issued draft guidances for industry outlining how it will define “highly similar” and what studies it will require follow-on biologic companies to submit. The documents confirmed that the FDA will determine what evidence is required on a case-by-case basis, noting that the “type and amount of analyses and testing that will be sufficient to demonstrate [follow-on “biologicity”] will be determined on a product-specific basis.”[87] In addition, the FDA indicated that its evaluation will not depend on any one piece of evidence, but it will instead “consider the *totality of the evidence* provided by a sponsor.”[88]

The FDA explained that products must be “highly similar,” citing the statutory language of “no clinically meaningful differences between [the products] in terms of safety, purity, and potency.”[89] However, the document highlights certain areas where the FDA expects a follow-on biologic product might differ from the reference product. [90] The primary amino acid sequence must remain substantially the same, but “minor modifications such as N- or C-terminal truncations that will not affect safety and effectiveness may be justified . . . “[91] The choice of cell expression system is another area of possible difference, “because the type of expression system and host cell will significantly affect the types of process- and product-related substances and impurities . . . that may be present in the protein product.”[92] Differences in post

---

85. Margaret A. Hamburg, Comm’r, Food & Drug Admin., Remarks at Generic Pharmaceutical. Ass’n Ann. Meeting (Feb. 18, 2010), available at http://www.fda.gov/NewsEvents/Speeches/ucm201833.htm.

86. Steven Kozlowski, Janet Woodcock, Karen Midthun & Rachel Behrman Sherman, *Developing the Nation’s Biosimilars Program*, 365 NEW ENG. J. MED. 385, 386 (2011).

87. *Scientific Considerations*, *supra* note 62, at 8.

88. *Id.* at 2.

89. *Scientific Considerations*, *supra* note 62, at 3.

90. *See e.g.*, U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES ET AL., GUIDANCE FOR INDUSTRY: BIOSIMILARS: QUESTIONS AND ANSWERS REGARDING IMPLEMENTATION OF THE BIOLOGICS PRICE COMPETITION AND INNOVATION ACT OF 2009 (DRAFT GUIDANCE), 4-5 (Feb. 2012) [hereinafter *Questions and Answers*] (“[D]ifferences between the formulation of a proposed product and the reference product may be acceptable . . .”).

91. *Scientific Considerations*, *supra* note 62, at 9.

92. U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES ET AL., GUIDANCE FOR INDUSTRY: QUALITY CONSIDERATIONS IN DEMONSTRATING BIOSIMILARITY TO A REFERENCE PROTEIN PRODUCT (DRAFT GUIDANCE), 9 (Feb. 2012) [hereinafter *Quality Considerations*].

translational modifications "might not preclude a finding of biosimilarity."[93] If "the manufacturing process produces different levels of impurities, the biosimilar can still be accepted by the FDA."[94] Differences "between the formulation of a proposed product and the reference product may be acceptable."[95] Finally, "some design differences in the delivery device or container closure system used with the proposed biosimilar product may be acceptable."[96]

The draft guidelines indicate that the FDA will require follow-on biologic companies to submit extensive studies demonstrating that their product is "highly similar" to the reference product if the follow-on biologic has any of the differences listed above. "The type, nature, and extent of any differences . . . introduced by design or observed from comprehensive analytical characterization of multiple manufacturing lots, should be clearly described and discussed. . . . The potential effect of the differences on safety, purity, and potency should be addressed and supported by appropriate data."[97]

The FDA has not yet issued guidance on how it will determine interchangeability (which would allow pharmacists to substitute the follow-on biologic for the brand name drug). In its draft guidance, it notes that it has the power to make a determination of interchangeability, but "[a]t this time, it would be difficult as a scientific matter for a prospective [follow-on biologic] applicant to establish interchangeability . . ."[98]

Although the FDA has not yet approved a drug under the follow-on biologic pathway, drugs that could be considered follow-on biologics have been approved in the United States though other pathways. Omnitrope™, a recombinant human-growth hormone (rhGH) produced by Sandoz was approved through the 505(b)(2) pathway.[99] The 505(b)(2) pathway is considered an "abbreviated" application pathway in the sense that applicants may rely on safety studies submitted by a pioneer drug manufacturers.[100] In its 505(b)(2) application, Omnitrope relied on studies done by Pfizer for their pioneer rhGH product Genotropin.[101] The FDA allowed Omnitrope to be approved through 505(b)(2), but emphasized that Omnitrope is a "relatively

93. *Scientific Considerations*, *supra* note 62, at 8.

94. *Quality Considerations*, *supra* note 92, at 12.

95. *Questions and Answers*, *supra* note 90, at 4.

96. *Id.* at 5.

97. *Quality Considerations*, *supra* note 92, at 8.

98. *Questions and Answers*, *supra* note 90, at 11.

99. Covington & Burling, FDA Approval of Sandoz's 505(b)(2) Application for a Follow-On Recombinant Human Growth Hormone Product, 1 (June 5, 2006), available at http://www.cov.com/files/Publication/8405cdb8-b5ca-4050-a2a7-84ea54b23aac/Presentation/PublicationAttachment/356e78e0-06fc-45e3-b0b8-9385b2b205b0/oid20985.pdf.

100. 21 C.F.R. § 314.3.

101. Covington & Burling, *supra* note 99.

simple recombinant protein [and] it is possible to determine that the end products of different manufacturing processes are highly similar . . . "[102] Other biological drugs, including GlucaGen, Hylenex, and Fortical have also been approved through 505(b)(2) applications.[103] Shortly before the BPCIA was passed, Teva applied for a BLA for its product TevaGrastim (follow-on biologic to Amgen's Neupogen).[104]

III. WORK-AROUNDS AND INFRINGEMENT

Because no follow-on biologics have been approved under the BPCIA, courts have not yet addressed the question of infringement. However, the first biologics are starting to come off patent, meaning that they will go forward protected only by the weaker drug product, method, or product patents seen in the section on small-molecule drugs. This will spawn opportunities for follow-on biologic work-arounds which will, like their generic predecessors, struggle with maintaining sufficient similarity to the reference drug to satisfy the FDA while maintaining sufficient differences from the reference drug to avoid infringing by equivalents.[105] In this section, I make predictions about how

---

102. Letter from Steven Galson, Dir., Center for Drug Evaluation and Res., to Kathleen Sanzo et al. 4 (May 30, 2005).

103. Courage & Parsons, *supra* note 70, at 213-14.

104. Press Release, Teva, Teva Announces the Submission Of A Biologics License Application (BLA) for XM02 for The Treatment Of Chemotherapy-Induced Neutropenia (Dec. 1, 2009); s*ee also* Courage & Parsons, *supra* note 70, at 214; Richard A. Epstein, *The Constitutional Protection of Trade Secrets and Patents Under the Biologics Price Competition and Innovation Act of 2009*, 66 FOOD & DRUG L.J. 285, 326 (2011). The FDA responded to Teva's BLA with a request for more information. *Id.* at n.190.

105. In the world of small molecule drugs, it is extremely common for brand name drugs to remain protected by peripheral patents and for generic companies to attempt to enter the market by working-around these patents. *See, e.g.*, Sanofi-Aventis v. Sandoz, Inc., 2009 WL 1741571 (D.N.J. 2009) (The plaintiff produces Eloxatin, an anti-cancer drug used in conjunction with chemotherapy to slow the growth of cancer cells in the body. The active ingredient is oxaliplatin, a chemical that can come in one of two orientations, called enantiomers, which are mirror images of each other. One enantiomer is toxic; therefore it must be separated out before the drug can be used. Plaintiff's patent claims "optically pure [oxaliplatin]," (meaning that the toxic enantiomer is completely separated out) purified using High Performance Liquid Chromatography (HPLC). The defendants both produce optically pure oxaliplatin using methods other than HPLC. The court found that because the generic products are not resolved using HPLC, they do not infringe by either literally or by equivalents); Astrazeneca Pharm. LP v. Mayne Pharma Inc., 2005 WL 2864666 (S.D.N.Y. 2005) (The plaintiff held a patent on Diprivan, a mixture of injectable propofol (an anesthetic) and disodium edetate (EDTA, an antimicrobial compound added to improve the shelf-life of the product). The generic company used a formulation that mixed injectable propofol with diethylenetriaminepentaacetate (DTPA), a compound similar to the EDTA used in the brand-name product. The court found that the generic did not literally infringe, but it did infringe by equivalents); Janssen Pharm. N.V. v. Eon Labs Mfg., Inc., 374 F. Supp. 2d 263 (E.D.N.Y. 2004) (The plaintiff produces brand-name Sporanox, an anti-fungal. The patent claims "Beads Having a Core Coated with an Antifungal and a Polymer" and further

follow-on biologic manufacturers and courts will handle this problem. I base my predictions off patent infringement cases for small molecule drugs, cases for biologics approved through BLAs or 505(b)(2)s, and the strategies used by follow-on biologic manufacturers who have had their products approved in other countries. I focus on how these cases will be treated under the doctrine of equivalents because the doctrine of equivalents represents the outer boundary of patent protection and thus is the most relevant to understanding how far follow-on biologics must stay from the reference drug's patent.

A. *What Will Follow-on Biologics Look Like?*

The types of work-arounds most commonly seen in small molecule drugs are changes to inactive ingredients, packaging, and chemical synthesis of the drug and stability agents.[106] Because the production process of biologics is much more complex, the approaches for biological work-arounds are likely to be somewhat different. Based on the predictions of several scientific scholars and industry experts,[107] I divide follow-on biologic work-arounds into the

---

limits the claim to beads with a "polymer layer, characterized in that the core has a diameter of about 600 to about 700 um (25-30 mesh)." The defendant developed a bioequivalent formulation using a 20-25 mesh. The court held that defendant's product did not infringe either literally or by equivalents); Glaxo Wellcome, Inc. v. Pharmadyne Corp., 32 F. Supp. 2d 265 (D. Md. 1998) (The plaintiff makes Zantac, a medication used to treat heartburn and ulcers. Zantac is made from ranitidine hydrochloride combined with ethanol, an antimicrobial put into the solution to preserve shelf-life. The generic product combined ranitidine with propylene glycol, which is also an anti-microbial agent. The court found that the generic did not literally infringe, but did infringe by equivalents).

106. Janet Freilich, *The Paradox of Legal Equivalents and Scientific Equivalence: Reconciling Patent Law's Doctrine of Equivalents with the FDA's Bioequivalence Requirements*, 66 SMU L. REV. (forthcoming 2012).

107. My categories are based on categories listed in the following articles: Islah Ahmed et al, *Biosimilars: Impact of Biological Product Life Cycle and European Experience on the Regulatory Trajectory in the United States,* 34 CLINICAL THERAPEUTICS 400, 405 (2012). (Listing the following categories in Table 1: Cloning (coding gene, plasmid), Transformation/Transfection (host cell/method), Cell Culture (temperature/media/oscillation of cells), Purification (method of purification/removal of epitopes/formulation and packaging)); Wolfgang Jelkmann, *Recombinant EPO Production—Points the Nephrologist Should Know*, 22 NEPHROL. DIAL. TRANSPLANT. 2749, 2751 (2007). (Manufacturing steps influencing the product include: "Sequence of cDNA, type of vector/plasmid, promoter and other accessory DNA elements, type of host cell, technique of transfection, propagation of host cell clones, maintenance of production cultures, composition of culture medium, type of culture vials/bottles, type of fermenter/bioreactor, extraction and purification of recombinant product from culture medium, analysis of product, formulation."); Wolfgang Jelkmann, *Biosimilar Epoetins and Other "Follow-On" Biologics: Update on the European Experience*, 85 AMERICAN JOURNAL OF HEMATOLOGY 771, 771 (2010). ("The main factors influencing the composition of recombinant medicines are: (i) the plasmid (promoter, marker genes), (ii) the host cell (origin, species, clone), (iii) the culturing process (fermenter, culture media), (iv) the purification steps, (v) posttranslational modifications (oxidation, deamidation, addition of polymers), and (vi) the formulation and packaging."); Huub Schellekens, *Biosimilar Therapeutics—What Do We Need To Consider*, 2 NEUROLOGY

following categories for ease of discussion:

- Pre-transformation (changes in promoters, enhancers, termination sequences, selection markers, genetic sequences)
- Transformation (changes in cell lines, glycosylation patterns, transfection efficiency, transcription/translation efficiency)
- Cell culture (changes in temperatures, media, reactor turnover)
- Purification (changes in method of purification, removal of epitopes, degree of impurity)
- Formulation (changes in inactive ingredients such as buffers or stabilizing solutions)

Thus far, other countries have approved follow-on biologics with changes in several of these categories. The European Medicine Agency (EMA) has approved Zarzio and Filgrastim Hexal, two follow-on biologics of Neupogen, a drug used to stimulate white blood cell growth. Both Zarzio and Filgrastim Hexal are identical to Neupogen except that the buffer used in the follow-on biologics is glutamate while Neupogen uses acetate.[108] This is a formulation switch. The EMA has also approved Tevagrastim, Ratiograstim and Biograstim, also follow-on biologics to Neupogen.[109] These three follow-on biologics are identical to the reference product except for a slightly different pH and concentration of polysorbate 80 (used to stabilize solutions intended for parenteral administration). These are also formulation changes.

Both the EMA and the FDA have approved Valtropin, a recombinant human growth hormone. Valtropin was approved as an NDA in the United States[110] and as a follow-on biologic of Humatrope in Europe. Valtropin is produced in *S. cerevisiae* (yeast) whereas the reference product is produced in *E.coli*.[111] This is a cell line (or transformation) change.

The EMA approved Abseamed, Binocrit, Epoetin alpha Hexal, Retacrit, and Silapo, all follow-on biologics of Eprex, a recombinant epoetin which stimulates the production of red blood cells. They have different glycosylation levels and lower levels of neuraminic acid as compared to the reference

---

DIALYSIS TRANSPLANTATION i27, i28 (2009). ("Changes may occur to the expression systems used for production, culture conditions (e.g. temperature and nutrients), purification and processing, formulation, storage and packaging . . . Structural differences between proteins may arise for a number of reasons, including oligomerization, modification of the primary protein sequence, glycosylation patterns or the conformational state . . . ").

108. European Medicines Agency, European public assessment report (EPAR) for Zarzio (2009); European Medicines Agency, EPAR for Filgrastim Hexal (2009).

109. European Medicines Agency, EPAR for Biograstim (2008); European Medicines Agency, EPAR for Tevagrastim (2008); European Medicines Agency, EPAR for Ratiograstim (2008).

110. Center for Drug Evaluation and Research, Application Number: 21-905 Approval Letter (2007).

111. European Medicines Agency, EPAR for Valtropin (2012). // authorization removed

drug.[112] These are all transformation changes.

In addition, epoetin follow-on biologics approved in Korea (Eporon, Espogen, and Epokine) had a purity difference compared to the reference product.[113] The follow-on biologic products contained several different isoforms (a different form of the same protein) as compared to Eprex.[114]

Epoetin products (not follow-on biologics) have also been changed using PEGylation, a formulation change that involves attaching a polyethylene glycol (PEG) molecule to the protein to increase water solubility and thus stability and shelf-life.[115] Another formulation change used with epoetin involved changing the stabilizer from human serum albumin (HSA) to polysorbate 80.[116]

B. *How Will Each Category of Change Be Treated?*

In general, courts have been reluctant to hold that a change in a biotechnology product infringes under the doctrine of equivalents. This may be because courts struggle to understand the technology, or because scientists themselves struggle to understand how the mechanics of small changes affect the function, way, and result of biotechnologies to the same extent that they understand the function, way and result of small molecule drugs. In addition, there is very little precedent in this area, so courts may be reluctant to move ahead of the development of the case law and hold that a product does infringe by equivalents. Rather, courts prefer to leave the question to a jury by refusing to grant summary judgment to the patentee.[117]

The doctrine of equivalents will become an increasingly popular litigation strategy as follow-on biologics enter the market. Although the doctrine of equivalents has been litigated in non-drug biotechnology cases, these non-drug biotechnologies are not constrained by the FDA to closely resemble a patented product—the way that a follow-on biologic would be. Therefore scientists searching for a work-around have a larger number of variations to attempt, making it less likely that their product will infringe. Once follow-on biologics enter the market, the doctrine of equivalents is likely to become as highly used in the biotechnology sphere as it is in the small molecule drug sphere.

Follow-on biologic manufacturers, like small molecule drug

---

112. European Medicines Agency, EPAR for Binocrit (2007).

113. *See* S. Park, K. Patel and J. Ko et al., *Analytical comparisons of erythropoietin products from Korea and Amgen's Epogen (epoetin alfa)*, 21 NEPHROL. DIAL. TRANSPLANT iv14 (2006).

114. Schellekens, *supra* note 107, at i29.

115. Amgen Inc. v. Hoffman-La Roche Ltd., 580 F.3d 1340, 1381 (Fed. Cir. 2009).

116. Michael Lissy et al., *Comparison of the Pharmacokinetic and Pharmacodynamic Profiles of One US-Marketed and Two European-Marketed Epoetin Alfas*, 11 DRUGS R. D. 61, 62 (2011).

117. Allison & Lemley, *supra* note 38, at 38 (remarking that in other contexts courts prefer to decide questions of infringement by equivalents on summary judgment).

manufacturers, will be forced to create a product that is similar enough to satisfy the FDA, but different enough to avoid infringing on the reference drug's patent. Of the possible categories of changes they can make—from pre-transformation changes to formulation changes—I will show that there will [be?] the broadest intellectual space for work-arounds far upstream from the final product, in the pre-transformation or transformation categories, or at the final stage, with formulation changes. This is because pre-transformation or transformation changes may not be as integral to the FDA's "highly similar" comparison, thus, there is more room to make changes as long as the end product is still "highly similar." Although formulation changes will certainly be part of the FDA's "highly similar" analysis, the statute explicitly allows changes in formulation, stating that although there can be "no clinically meaningful differences between the biological product and the reference product in terms of the safety, purity, and potency of the product," changes in excipients are allowed.[118] Companies given more leeway by the FDA to make changes will be more likely to make changes further away from the boundaries of the patent, and thus less likely to infringe by equivalents.

Conversely, the middle stages of the production process—cell culture and purification steps—will be the most difficult for follow-on biologic companies to change while remaining "highly similar" and avoiding infringement by equivalents. Because the manufacturing process is an integral part of a biologic, and because small changes may have large—and poorly understood—effects, the FDA will likely govern all stages of the manufacturing process closely and require them to resemble the brand name company's manufacturing process. Thus, the space to stray from the process used by the innovator company will be very narrow, forcing follow-on biologics to stay closer to the patented formulation. Resultantly, they will be more likely to infringe.

Whether any particular work-around will be barred on the grounds of patent infringement will, of course, depend on the precise wording of the patent protecting the innovator product. However, there are trends in what types of work-arounds are more likely to allow generic drugs to avoid infringement;[119] therefore, it stands to reason that there would also be trends for follow-on biologics. My conclusions will be useful to follow-on biologic companies because they suggest fruitful directions for research, and will be useful for brand name companies because they suggest areas that should be accounted for when developing a patent portfolio around a product.

1. *Pre-transformation changes*

Follow-on biologic companies will likely be able to make pre-transformation changes from the reference product that will allow them to

118. 42 U.S.C. § 262(i)(2).
119. Freilich, *supra* note 106.

create a product that does not infringe the brand-name patent. Pre-transformation changes include changes to the coding nucleic acid sequence and the use of different promoters, enhancers, or termination sequences. It will be easier for follow-on biologic companies to make changes to promoters, enhancers, or termination sequences than to the portion of the nucleic acid sequence that codes for the drug.

Although the FDA will restrict the range of changes a follow-on biologic company will be able to make at the pre-transformation stage, it will not do so to the same extent as [it will to?] some of the downstream steps. Therefore, follow-on biologic companies will be able to make a wider range of changes and will be more likely to be able to innovate far enough away from the boundary of the patent to avoid infringement. The FDA's requirement for a "highly similar" product applies to the finished product, not the starting material.[120] The FDA has indicated in its draft guidance documents that "minor modifications such as N- or C- terminal truncations that will not affect safety and effectiveness may be justified."[121] The draft guidances do not mention how the FDA will treat other pre-transformation changes, such as different promoters or enhancers, but presumably the FDA would allow the changes as long as the final product remained highly similar to the reference product and the change did not introduce additional impurities.[122]

The courts have favored defendants in cases involving pre-transformation changes. Although there are no cases addressing follow-on biologics, there have been a significant number of cases exploring the outer boundaries of patent protection in the context of pre-transformation techniques used in biotechnology. This is because pre-transformation technologies have been used extensively in laboratories and in biologics research. Courts have been extremely reluctant to hold that a pre-transformation change infringes under the doctrine of equivalents. Courts will not grant summary judgment to a plaintiff moving for a decision on infringement by equivalents, and will often grant summary judgment to a defendant moving for a finding of no infringement as a matter of law. However, courts deal more favorably with changes to promoters, enhancers, or termination sequences than to changes in the portion of the genetic sequence that encodes the protein.

Because there are many cases dealing with pre-transformation changes, follow-on biologic applicants can look to this law to predict how courts will treat follow-on biologic litigation. Courts have struggled to determine when a change in the DNA or amino acid sequence is small enough to infringe by

---

120. *See generally Scientific Considerations*, *supra* note 62; *Quality Considerations*, *supra* note 92.

121. *Scientific Considerations*, *supra* note 62, at 9.

122. *Quality Considerations*, *supra* note 92, at 13 (noting that the guidance mentions the possibility of such impurities, suggesting that "process-related impurities arising from cell substrates (e.g., host cell DNA, host cell proteins) . . . should be evaluated.").

equivalents.[123] The Federal Circuit has stated that "[t]he mere possibility that a single mutation could affect biological function cannot as a matter of law preclude an assertion of equivalence."[124]

A recent case where the court acknowledged the difficulty of applying the doctrine of equivalents to a defendant who used a different generic sequence is *Regents of University of California v. Monsanto Co.*[125] In *Regents*, the plaintiff held a patent on the recombinant nucleotide sequence encoding bovine growth hormone (bGH). Monsanto used a slightly different DNA sequence to encode its version of bGH, which plaintiff argued infringed by equivalents.[126] The court struggled to define the 'function' of the biotechnology product. The court was unsure whether the function of recombinant DNA is to "require expression of the bGH protein" or whether its function is "merely to provide a blueprint for bGH."[127] Because the outcome of the case turns on the definition of function, the court denied the plaintiff's motion for summary judgment, holding that whether the products were equivalent was a matter of fact.[128]

The Federal Circuit struggled with the question of how to apply the doctrine of equivalents to a change in nucleic acid sequence in *Boehringer Ingelheim Vetmedica, Inc. v. Schering-Plough Corp*.[129] In *Boehringer*, both companies made a vaccine for Porcine Reproductive Respiratory Syndrome (PRRS). The genetic sequence of the vaccines differed by at least 73 nucleotides.[130] A jury found that the defendant did infringe by equivalents. The Federal Circuit upheld the jury's verdict holding that "[a] reasonable jury could easily . . . conclude that the genetic differences between [the two vaccines] are insubstantial in the context of the claimed method."[131] Note that the patent does not actually claim the genetic sequence; it merely names the strain of virus used in the vaccine, which was the reason for the court's caution about context—the result could have been different if the patent had described the vaccine using a different method.

Although the Federal Circuit in *Boehringer* upheld a jury verdict of infringement, in another genetic sequence case, the Federal Circuit upheld a

---

123. *See* JANICE M. MUELLER, PATENT LAW 352 (3d ed. 2009) (questioning whether a change in a single nucleotide would be infringing, and concluding that the answers are likely to be case-specific); D. Alan White, *supra* note 44, at 762-3 (noting that it will be extremely difficult to apply the doctrine of equivalents to small changes in generic sequences).

124. Boehringer Ingelheim Vetmedica, v. Schering-Plough, 320 F.3d 1339, 1353 (Fed. Cir. 2003).

125. Regents of University of California v. Monsanto Co., No. C 04-0634 PJH, 2005 WL 3454107 (N.D. Cal. Dec. 16, 2005).

126. *Id.* at *8.

127. *Id.*

128. *Id.*

129. 320 F.3d 1339. (Fed. Cir. 2003).

130. *Id.* at 1352.

131. *Id.*

jury verdict of non-infringement. In *Genentech, Inc. v. Wellcome Foundation Ltd*,[132] Genentech held patents covering the protein t-PA (tissue plasminogen activator) that dissolves stroke-causing fibrin clots and restores blood flow. The patent covered the DNA sequence, the expression vector containing the sequence, the cell culture capable of expressing t-PA using the vector, and the process for producing recombinant t-PA.[133] Defendants made FE1X, a protein variant of t-PA. FE1X lacks a structure of t-PA called the finger region and has a one amino acid substitution, which changes the glycosylation pattern.[134] After a jury trial returned a verdict of infringement by equivalents, the defendants asked the court to hold that FE1X could not infringe as a matter of law.

The court applied the function-way-result test but struggled to define the 'function' prong. While the trial court found that the function of t-PA was "dissolution of fibrin clots through the cleavage of plasminogen to plasmin," the Federal Circuit worried that if this was true: it "is difficult to imagine how . . . any version of t-PA . . . would avoid infringement under the doctrine of equivalents because t-PA, or any operative variant, would by definition necessarily perform this function in the same general way with the same results."[135] Therefore the Federal Circuit defined the function of t-PA to be "catalyzing the conversion of plasminogen to plasmin, [and] bind[ing] to fibrin."[136] Based on this definition, the court held that FE1X did not function substantially the same way or achieve substantially the same results because the different structure of FE1X resulted in a significant change in binding efficiency and a roughly ten times increase in half-life.[137] Therefore the court held that there was no infringement as a matter of law.[138]

In a similar case, *Carnegie Mellon University v. Hoffman-LaRoche, Inc.*,[139] plaintiff had a patent on a "recombinant plasmid containing a cloned complete structural gene encoding . . . DNA polymerase I."[140] The court construed the term "DNA polymerase I" to mean an enzyme that, among other things, had 3'-5' exonuclease activity.[141] Defendant's product, *Taq* polymerase,[142] does not

---

132. Genentech Inc. v. Wellcome Found., 29 F.3d 1555 (Fed. Cir. 1994).

133. *Id.* at 1558.

134. *Id.* at 1559.

135. *Id.* at 1567 (internal quotation marks omitted).

136. *Id.*

137. *Id.* at 1569.

138. *Id.*

139. Carnegie Mellon Univ. v. Hoffman-LaRoche, Inc., 55 F. Supp. 2d 1024 (N.D. Cal. 1999).

140. *Id.* at 1028.

141. *Id.* at 1045.

142. *Taq* polymerase is named after the bacteria from which it was derived—*Thermus Aquaticus*.

have 3'-5' exonuclease activity.[143] Plaintiffs argued that *Taq* polymerase infringed by equivalents because DNA polymerase I's 3'-5' exonuclease activity includes a proofreading function, and *Taq* polymerase also performs a proofreading function.[144] The court, relying on *Genentech v. Wellcome*, found that *Taq* polymerase is missing the amino acids used for 3'-5' exonuclease activity and did not perform a proofreading function in the same way.[145] Thus, the defendant's product did not infringe by equivalents.

As demonstrated in the cases above, courts are unsure of how to treat changes in genetic sequence under the doctrine of equivalents. In general, changes in genetic sequence will likely run into the most trouble with the 'way' prong of the equivalents test. A change to the genetic sequence that will be more likely to produce a product that is "highly similar" to the reference drug will be less likely to change the way the protein interacts with its target. A change to the genetic sequence that changes the way that the protein interacts with its target will likely effect how the drug functions, and thus it will be harder for the drug to obtain FDA approval as a follow-on biologic.

However, depending on the brand name patent in question, follow-on biologic companies may not need to alter the genetic sequence that codes for their protein in order to make a pre-transformation change. They may also be able to use a different promoter or enhancer, or make some other modification to the pre-translational process. Courts have been very favorable to defendants in cases involving this sort of change.

In *Regents of University of California v. Dako North America, Inc.*,[146] plaintiff held a patent on a method of using complementary DNA segments to bind to DNA in a cell. Defendants had a similar binding system which used peptide nucleic acid (PNA) instead of DNA. PNA is a synthetic molecule similar to DNA except that it has a polyamide backbone and binds more tightly to complementary DNA than DNA or RNA would.[147] Plaintiffs argued that the PNA product infringed by equivalents and moved for summary judgment. The court found that the "function and result" prongs were the same as a matter of law, but could not find that they functioned the same *way* as a matter of law because PNA binds more selectively and effectively than DNA.[148] Thus, the two inventions were not equivalent as a matter of law.

In *Gen-Probe v. Vysis*,[149] Vysis held a patent on DNA probes that capture and amplify a DNA sequence. The court constructed "amplification" to mean

---

143. *Id.* at 1045.

144. *Id.*

145. *Id.* at 1046-47.

146. Regents of Univ. of Cal. v. Dako N. Am., Inc., 615 F. Supp. 2d 1087 (N.D. Cal. 2009).

147. *Id.* at 1090.

148. *Id.* at 1095-98.

149. No. 99-CV-2667 H(AJB), 2002 WL 34413199 (S.D. Cal. 2002).

non-specific amplification.[150] Vysis used non-specific random hexamers to bind to all DNA in a sample and to amplify the entire sample. Gen-Probe used specific primers to bind to a pre-determined sequence, and amplify only that sequence.[151] A jury found that Gen-Probe did not infringe by equivalents, and Vysis asked for a judgment as a matter of law.[152] The court found that the jury was given evidence that Gen-Probe's system had a different function (using specific primers, rather than non-specific random hexamers), operated a different way (by using specific primers and promoters rather than non-specific primers and promoters), and had a different result (increasing the proportion of the target sequence compared to the overall pool of nucleic acid, rather than increasing the proportion of all nucleic acid).[153] Thus, the court did not overturn the jury verdict.

The sorts of changes demonstrated in these two cases could all be applied to some aspect of the pre-transformational process of producing a follow-on biologic. The FDA has been silent on this type of change, but it is unlikely to affect the final product; therefore generic companies will be able to make a wide range of changes. In addition, courts lean towards finding no infringement for pre-transformational changes.

Why are non-coding sequence types of pre-transformation changes easier to make? First, non-coding sequences may be changed while still producing an identical product,[154] meaning that they may be changed with less concern that the FDA will reject the product for lack of biosimilarity. This means that there is much wider room for change, and it follows that work-arounds for these sequences will fall further from the boundaries of the brand name company's patent than patent work-arounds in areas where deviation from the reference product is more closely regulated by the FDA.

Another advantage for follow-on biologic companies making changes at the pre-transformation stage is that this type of change has been the most heavily litigated to date. This is simply because pre-transformation technology was developed earlier and has been used for longer. The majority of the litigation does not concern biologic medicine used in humans, which is a relatively recent phenomenon, but rather biotechnology used in the lab, which traces its origins to the 1970s.[155] With a longer history comes more litigation.

---

150. *Id.* at *11.

151. *Id.* at *13.

152. *Id.*

153. *Id.*

154. Note that while some pre-transformation changes, such as changes in nucleic acids, may result in a different product, they will not necessarily do so. A several base change in a DNA sequence may still produce the same amino acid sequence, resulting in an identical protein.

155. *See* Sally Smith Hughes, *Making Dollars out of DNA*, 1974-1980, 92 ISIS 541, 542 (2001) (describing how recombinant DNA revolutionized the biotechnology field. Recombinant DNA was also one of the first biotechnologies to be patented). For more

A high volume of litigation benefits follow-on biologic companies for several reasons. First, they will be better able to predict how courts will treat any particular change and how broadly courts will define the boundaries of a brand name patent. This predictability will allow them to design a work-around that they can be more confident will not infringe by equivalents. Moreover, if they engage in litigation, a greater number of precedential cases should allow the case to settle more quickly.[156] Finally, the greater volume of litigation and the longer history of how biotechnology works at a pre-transformation level will allow the litigants to provide more evidence to satisfy the 'function' and 'way' prongs of the test that courts have struggled with.[157] However, this longer history of litigation benefits brand name companies too. They can draft stronger patents based on how courts have treated earlier patents, and will also have a greater understanding of the technology and will be able to provide more evidence to the court. Nevertheless, pre-transformation changes will provide fruitful ground for follow-on biologic work-arounds.

2. *Transformation changes*

Transformation changes include use of different cell lines, different glycosylation patterns, and improvement in transfection efficiency. There is some overlap between the technologies placed in the transformation category and technologies placed in the pre-transformation category, and many of the

---

information on the effect of the rDNA patent on biotechnology patenting practices *see* Janet Freilich, *A Nuisance Model for Patent Law*, U. ILL. J.L. TECH. & POL'Y 329, 363 (2011).

156. *See* Steven Shavell, FOUNDATIONS OF ECONOMIC ANALYSIS OF LAW 401 (2004) (demonstrating that parties are more likely to settle if their predictions of the outcome of the case are closer together).

157. This has traditionally been a problem in small molecule drug litigation. Courts want the plaintiff to produce specific evidence of how the defendant's product operates, and because many biological functions are poorly understood, this can be challenging for the plaintiff. *See, e.g.*, Cephalon v. Watson, 769 F. Supp. 2d 729 (D. Del. 2011) (Plaintiff produces Fentora, a product used to treat breakthrough pain in cancer. The drug consists of fentanyl buccal tablets given sublingually, which evolve gas by means of an effervescent reaction to increase the rate of absorption across the oral mucosa. The plaintiff produces the effervescent reaction using sodium bicarbonate. The plaintiff alleged that the defendant infringed because their generic tablets were "bioequivalent to Fentora" but did not provide any evidence of the nature of the chemical reaction occurring or conduct experiments to determine the rate and extent of absorption across the oral mucosa. The court scolded the plaintiff for lack of evidence and held that there was insufficient evidence to find infringement.); *see also* In re Omeprazole Patent Litig., 490 F. Supp. 2d 381 (S.D.N.Y. 2007) (Plaintiff produces Prilosec, used to treat ulcers. The active ingredient is omeprazole, a proton pump inhibitor which slows gastric acid secretion. The plaintiff's product contained an inert core coated with omeprazole, talc, hydroxypropyl methylcellulose, and several other coatings. The defendant's product contained different chemicals to stabilize the omeprazole. The court found that the plaintiff had not produced sufficient evidence to show that the chemicals in the defendants' product performed substantially the same function in substantially the same way to produce substantially the same result. The court therefore held that the defendants did not infringe).

points made about pre-transformation changes also apply to transformation changes. As with pre-transformation changes, courts have resolved cases in this category favorably for the defendant, therefore transformation changes will be a useful approach for follow-on biologic companies. However, as compared to pre-transformation changes, the FDA is likely to look more closely for similarity in transformation changes because differences in the transformation process can have a significant effect on the finished product. Production using a different cell line or a glycosylation change (glycosylation changes often come from using a different cell line) can affect the structure—and therefore function—of the protein.

European companies have successfully made follow-on biologics using transformation changes. Valtropin is grown in a different cell line than its reference product Humatrope™ (yeast versus *E. coli*) and has been approved as a follow-on biologic in Europe. It has also been approved in the United States through an NDA.[158] Valtropin is "analytically comparable to the marketed European reference product Humatrope."[159] Several follow-on biologics of the erythropoietin product Eprex have different glycosylation patterns but are nevertheless approved as follow-on biologics in Europe, and "[c]omparison of the purity and in-vivo bioactivity did not reveal any remarkable difference.[160]

The FDA is likely to approve follow-on biologics that have some transformation differences. Because there is a relatively large space between the reference drug and the limits of what would be considered follow-on biologic, it will be possible for generic companies to make a variety of changes, and thus to stay further from the reference product's patent and be less likely to infringe. In its draft guidelines, the FDA has indicated certain changes that may be acceptable as long as they do not produce clinically meaningful differences in safety, purity, and potency as compared to the reference product.[161] For example, the guidelines indicate that the FDA envisions some differences in the expression systems (cell lines) used to produce follow-on biologic products: "Differences between the chosen expression system of the proposed follow-on biologic product and that of the reference product should be carefully considered. . . ."[162] The guidelines also indicate that the FDA envisions differences in amino acid modifications that can result from using different expression systems. The draft guidance states that "in general, proteins can differ [with respect to] . . . modification to amino acids, such as sugar moieties

158. Center for Drug Evaluation and Research, Application Number: 21-905 Approval Letter (2007).

159. European Medicines Agency, European public assessment report (EPAR) for Valtropin (2012).

160. European Medicines Agency, European public assessment report (EPAR) for Silapo (2012)

161. *Scientific Considerations*, *supra* note 62, at 8.

162. *Quality Considerations*, *supra* note 92, at 9.

(glycosylation) or other side chains . . ."[163] and that applicants should conduct tests to compare the post-translational modifications (such as glycosylation and phosphorylation) of the follow-on biologic and reference products.[164]

While the FDA is likely to allow follow-on biologics with transformation differences, its draft guidance documents caution that "[m]inimizing differences between the proposed and reference expression systems to the extent possible can enhance the likelihood of producing a highly similar protein product."[165] Thus, follow-on biologic companies must still take care to ensure that the changes they make do not have a substantial effect on the final product.

Because the FDA will allow follow-on biologic companies to make a range of transformation changes, follow-on biologic manufacturers will often be able to avoid infringing under the doctrine of equivalents. However, courts will probably find more infringement for transformation changes than for pre-transformation changes. This is because while the changes are still far back on the manufacturing chain and thus will likely be less scrutinized by the FDA, giving the follow-on biologic company greater latitude to make changes that take it outside the area the reference drug's patent claims, the FDA will still carefully scrutinize transformation changes. In certain cases, it is possible that a transformation change could cause the FDA to categorize a drug designed as a follow-on biologic to be a completely new drug and not allow it to use the abbreviated pathway outlined in the BPCIA.

Transformation changes have already been litigated in the context of biotechnology—though not quite to the same extent as pre-transformation changes—lending some predictability to how courts will treat them. Courts have found that using a different vector to transform a cell or using a different cell line does not infringe by equivalents.

In *Enzo Biochem, Inc. v. Calgene, Inc.*,[166] the court found that using a different transformation vector did not infringe under the doctrine of equivalents. Both parties in the case make genetically modified tomatoes. Calgene genetically modified the tomato using cDNA, whereas Enzo used an inverted gene.[167] Enzo argued that Calgene infringed under the doctrine of equivalents. The court found that cDNA and an inverted gene have the same effect (they shut off the function of the target gene) but that the method they use to do so is different.[168] The court held that Calgene did not infringe by equivalents.[169]

Courts have also found that using a different cell line does not infringe

163. *Scientific Considerations*, *supra* note 62, at 5.
164. *Id.* at 9.
165. *Quality Considerations*, *supra* note 92, at 10.
166. Enzo Biochem, Inc. v. Calgene, Inc., 14 F. Supp. 2d 536 (D. Del. 1998).
167. *Id.* at 560.
168. *Id.*
169. *Id.*

under the doctrine of equivalents. *Enzo v. Calgene* deals briefly with the question, pointing out that Enzo uses a prokaryote cell whereas Calgene uses a eukaryotic cell and concluding that this difference in cell types was not insubstantial.[170] *Carnegie-Mellon Univ. v. Hoffman-La Roche Inc.*[171] deals with the question in greater detail. In *Carnegie-Mellon*, the plaintiff held a patent on "the recombinant plasmid containing a DNA coding sequence for the expression of DNA polymerase activity . . . wherein the bacterial host system and the bacterial source are each E. coli."[172] Hoffman-La Roche makes a recombinant plasmid that causes cells to express *Taq* DNA polymerase, which is derived not from *E. coli* but from *Thermus aquaticus*, a different type of cell.[173] Because *Taq* polymerase is not from *E. coli*, it does not literally infringe on the patent, but plaintiff argued that it infringed by equivalents. The court held that there was no infringement by equivalents.[174]

As these cases show, the trend in transformation change cases is to find that the defendant's product does not infringe by equivalents. In addition, these cases all involved relatively well understood technologies, meaning that the plaintiff will be more likely to be able to prove its case. This suggests that changing the transformation vector or cell line is a sufficient change that such a product would not infringe by equivalents on a patent that included a cell type as a limitation.[175] However, all infringement cases are fact-dependent, and it is easy to imagine a scenario where a closely related cell line, or a vector with an insubstantial difference, was used in the defendant's product, in which case that product might infringe by equivalents. Nevertheless, transformation differences are overall a good target for follow-on biologic companies seeking to make a change that is does not infringe.

3. *Cell culture changes*

Cell culture is the process of growing the cells that produce the biologic drug. Depending on the drug, it may not have a cell culture step (if it is made synthetically or is harvested from tissue) but most biologics on the market are produced through cell culture. Changes to cell culture may include growing cells at a different temperature, in different media, or increasing the reactor turnover. Changes in cell culture can be closely related to changes in purity, because the method of cell culture can often affect the purity of the resulting

---

170. *Id.*

171. Carnegie-Mellon Univ. v. Hoffman-La Roche Inc., 541 F.3d 1115 (Fed. Cir. 2008).

172. *Id.* at 1128.

173. *Id.* at 1119.

174. *Id* at 1129.

175. Though of course the outcome is fact-dependent. It is easy to imagine a scenario where a closely related cell line was used for the defendant's product, in which case that product might infringe by equivalents.

product. It is likely that it will be very difficult for follow-on biologic companies to make changes in cell culture large enough to avoid infringing without running afoul of the FDA's "highly similar" regulations.

Unlike the unwillingness to find infringement by equivalents seen in pre-transformation and transformation cases, the court in the leading (and only) cell culture change case upheld a jury verdict of infringement by equivalents. In *Boehringer Ingelheim Vetmedica, Inc. v. Schering-Plough Corp.*,[176] the plaintiff held a patent on growing and isolating a pig virus using a process of incubating monkey cells containing the virus until a cytopathic effect (a visible change in the monkey cells due to viral infection) was observed.[177] The defendant also grew the pig virus in monkey cells, but instead of incubating until a cytopathic effect was observed, defendant incubated for a specified time period.[178] The jury was presented with evidence that the defendant was aware of the time required to achieve a cytopathic effect, and in fact may have based their time measurements off that period, and that the incubation period was similar to plaintiff's incubation period. The jury found that Schering-Plough's process did infringe.[179] On appeal for judgment as a matter of law, the Federal Circuit held that the jury was presented with evidence that "Schering's practice of incubating the viral culture for a defined period of time performs the same function, in the same way, with the same result as incubating the viral culture until a defined degree of [cytopathic effect] is observed," and did not overturn the verdict.[180]

One lesson from this case is the danger of encouraging follow-on biologic companies to change a drug just enough to avoid infringement. Although the court in this case came to the right conclusion, it is easy to imagine a follow-on biologic company using Schering-Plough's strategy. While a cleverer change might have avoided infringement, it could also place patients in danger. Boehringer's process for determining the incubation period relied on examining the cells to determine that enough viruses had grown. Schering-Plough's process for determining the incubation period involved using a proxy. Proxies are often less accurate than the measures they are based on. Thus there might be a small difference in the amount of virus produced in each batch. While most pigs would probably be effectively vaccinated, perhaps a few pigs would not be. Schering-Plough's maneuvering to get around the patent introduced unnecessary risk to the patient. This is not an outcome courts should encourage.[181]

---

176. Boehringer Ingelheim Vetmedica, Inc. v. Schering-Plough Corp., 320 F. 3d 1339 (Fed. Cir. 2003).

177. *Id.* at 1343.

178. *Id.* at 1344.

179. *Id.* at 1350.

180. *Id.* at 1353.

181. However, courts should also not be responsible for determining issues of safety, which are the responsibility of the FDA. Courts are notoriously poor at resolving questions

Besides the court's unfavorable treatment of the above cell culture case, a further reason to believe that cell culture will be a difficult place for follow-on biologic companies to make changes is because of FDA scrutiny. Process is a crucial part of ensuring that follow-on biologics are similar to the reference drug. Unlike small molecule drugs, where identical drugs can be made by very different processes, small changes in the biologic manufacturing process can produce disproportionately sized changes in the final product.[182] Because of the

---

of science, making it difficult for them to adequately determine whether a product is safe or not. For discussion of courts' difficulty with scientific questions *see, e.g.*, Peter Lee, *Patent Law and the Two Cultures*, 120 YALE L.J. 2, 7 (2010); *see also* Arti K. Rai, *Engaging Facts and Policy: A Multi-Institutional Approach to Patent System Reform*, 103 COLUM. L. REV. 1035, 1040 (2003) ("Generalist trial judges, and the juries empanelled by trial judges, may be overwhelmed by the technology involved in patent cases."); Dan L. Burk & Mark A. Lemley, *Is Patent Law Technology-Specific?* 17 BERKELEY TECH. L.J. 1155, 1196 (2002) ("[J]udges are at a rather serious disadvantage in trying to put themselves in the shoes of an ordinarily skilled scientist."); Kimberly A. Moore, *Are District Court Judges Equipped to Resolve Patent Cases*, 15 HARV. J.L. & TECH. 1, 38 (2002) (concluding that "judges are not, at present, capable of resolving these [scientific patent] issues with sufficient accuracy"). Furthermore, judges themselves do not like scientifically complex cases. In the wake of *Daubert*, Judge Kozinski wrote that judges now face "a far more complex and daunting task in a post-*Daubert* world than before" (because judges are now responsible for claim construction). Daubert v. Merrell Dow Pharms., Inc., 43 F.3d 1311, 1315 (9th Cir. 1995). In a judicial panel discussion on science and law, the Honorable Alfred V. Covello stated, "I don't see how you could try a patent matter to a jury. Goodness, I've gotten involved in a few of those things. It's like somebody hit you between your eyes with a four-by-four." Judicial Panel Discussion on Science and the Law, 25 CONN. L. REV. 1127, 1145 (1993). Judge William Schwarzer wrote that science and technology issues "share one characteristic: They challenge the ability of judges and juries to comprehend the issues—and the evidence—and to deal with them in informed and effective ways." William W. Schwarzer, INTRODUCTION TO FED. JUD. CTR., REFERENCE MANUAL ON SCI. EVIDENCE 1, 1 (1st ed. 1994). The Supreme Court agreed, writing that "patent litigation can present issues so complex that legal minds, without appropriate grounding in science and technology, may have difficulty in reaching decision." Blonder-Tongue Labs., Inc. v. Univ. of Ill. Found., 402 U.S. 313, 331 (1971). Justice Scalia famously dismissed his understanding of scientific issues by quipping, "I told you before I'm not a scientist. (Laughter.) That's why I don't want to deal with global warming, to tell you the truth." Transcript of Oral Argument at 8, Massachusetts v. EPA, 127 S. Ct. 1438 (2007) (No. 05-1120).

182. Grabowski, *supra* note 78, at 515. Note that making changes production process is a very effective way for small molecule drugs to work-around a patent. *See, e.g.*, Sanofi-Aventis U.S. LLC v. Sandoz, Inc., 2009 WL 1741571 (D.N.J. 2009), *vacated*, 345 Fed. App'x. 594 (Fed. Cir. 2009) (plaintiff's patent covered a method of using High Performance Liquid Chromatography (HPLC) to purify their product. Both defendants found ways to purify the drug using other methods and the court found that they did not infringe.); SmithKline Beecham v. Apotex, 2005 WL 941671, *2 (E.D. Pa. 2005) (The plaintiff's product is Paxil, a blockbuster anti-depression drug. The plaintiff's conducted experiments with paroxetine, the active ingredient in Paxil, to "identify processes suitable for industrial scale production of paroxetine." They settled on a process that involved reacting an arecoline compound with a grignard reagent. This process could only be conducted in a non-ether solvent. The defendants created a synthesis process that works in an either solvent, which plaintiff's does not. Based on this, the court found that the defendants' process did not infringe.).

importance of process in the creation of a follow-on biologic product and because there are safety concerns attendant on a change in process, it is unlikely that the FDA will allow follow-on biologic manufacturers to make large changes in the cell culture process. Thus, it is unlikely that the follow-on biologic manufacturers will be able to make changes significant enough to avoid infringing.

4. *Purification changes*

It will be difficult for follow-on biologic companies to make purification changes in order to avoid infringing. Purification is part of the process of making the biologic, and, as with cell culture, small changes in the purification process could result in major changes to the safety of the drug.[183] Thus, it will be an area watched closely by the FDA, and the scope of changes follow-on biologic companies will be allowed to make will likely be narrow. The statute itself references purity, requiring that there be "no clinically meaningful differences between the biological product and the reference product in terms of the safety, purity, and potency of the product."[184]

There are two types of purity changes that follow-on biologic companies can make. They can make changes in the *level* of purity or they can make changes in the purification *process*. Because of the statutory requirement that there be "no clinically meaningful differences . . . in terms of the . . . purity"[185] of the product, follow-on biologic companies are unlikely to be able to make changes in the level purity that escape infringement by equivalents.

There are no cases involving the doctrine of equivalents and a level of purity change in a biologic, however there are several such cases involving generic small molecule drugs. The courts in the small molecule drug cases always found that the drug infringed.[186] This trend is likely to extend to follow-

183. Grabowski, *supra* note 78, at 516.

184. 42 U.S.C. § 262(i)(2)(B) (2006).

185. Id.

186. Pozen Inc. v. Par, Pharma., Inc., 800 F. Supp. 2d 789, 809 (E.D. Tex. 2011); Aventis Pharma Deutschland GmbH v. Lupin, Ltd., 2006 WL 1582412, *5 (E.D. Va. 2006). In *Pozen*, the plaintiff produced Treximet, a painkiller used to treat migraines. Treximet combines two active ingredients, sumatriptan and naproxen, in two layers. It is a multilayer tablet protected by a patent which claims "substantially all of said triptan is in a first layer . . . and substantially all of said naproxen is in a second, separate layer." During claim construction, the court defined "substantially all" to mean "at least 90%." Defendants Par and Dr. Reddy's Laboratories each created an ANDA product that did not literally infringe. Par's product contained 85% of the naproxen in the first layer and 100% of the sumatriptan and 15% of the naproxen in the second layer. Dr. Reddy's product contained 85% of the sumatriptan in the first layer and 100% of the naproxen and 15% of the sumatriptan in the second layer. The court found that the products infringed by equivalents. In *Aventis*, the plaintiff's product is Altace, a medication made of the compound ramipril and used to treat high blood pressure. Plaintiff's patent covers ramipril "substantially free of other isomers." The defendant's generic product was made of ramipril containing between 0.06% and 0.5%

on biologics. A purity level change that receives FDA approval will not change the *function* of the drug because the active ingredient and bioavailability are the same. It will not change the *way* a drug functions because the active ingredient and bioavailability are again the same. It will not change the *result* because the active ingredient and bioavailability are the same. Therefore a follow-on biologic with a purity level change small enough that it is still approved as a follow-on biologic should always infringe by equivalents.

The second type of purity change a follow-on biologic company can make is in the purification *process*. These changes will be more likely to succeed in getting FDA approval and avoiding infringement. The FDA will still closely monitor the range of changes allowed relative to the brand name product, but it may be possible for a generic company to design a process that purifies a different *way*. There is only one case involving the doctrine of equivalents and a purification process change in biotechnology. In *Genentech, Inc. v. Boehringer*[187] the plaintiff patented a process for purifying proteins which included the step of "removing high molecular weight impurities using a molecular sieve or high speed centrifugation techniques."[188] The court construed 'molecule sieve' to mean gel permeation chromatography or gel filtration.[189] The defendant used a depth filter, not a gel, to remove high molecule weight impurities, so it did not literally infringe. Genentech argued that the depth filter is equivalent to high speed centrifugation since both have the same function—removing high molecular weight impurities from a solution.[190] The court disagreed. It found that centrifugation and filtration operate in different ways, the former by spinning a solution, and the latter by pouring a solution through a membrane. In addition, centrifugation separates particles by weight and solubility, whereas filtration separates particles by size.[191] Therefore Boehringer did not infringe by equivalents.[192]

Genentech is an example of a change in purification *process*. As the case demonstrates, it should be possible for follow-on biologic companies to make a change in process that does not infringe by equivalents. However, the results of small molecule drug cases suggest that it will not be possible to make a change in purity level that does not infringe by equivalents. Thus, brand name companies should patent their products by degree of purity, rather than by purification process (or ideally, by both). Follow-on biologic companies should seek to create work-arounds by inventing different methods of purification, rather than by changing the degree of purity.

---

by weight of isomer-1. The court held that the defendant's product infringed.

187. Genentech, Inc. v. Boehringer Mannheim GmbH, 47 F. Supp. 2d 91 (D. Mass. 1999).

188. *Id.* at 116.

189. *Id.*

190. *Id.* at 117.

191. *Id.* at 116.

192. *Id.* at 120.

5. *Formulation changes*

Formulation changes involve changing inactive ingredients. Formulation changes will likely be some of the easiest changes for follow-on biologic companies to make. In general, changes in the formulations of small molecule drugs did not infringe by equivalents. In over 80% of cases, generic manufacturers that created work-arounds involving formulation changes did not infringe by equivalents.[193] The BPCIA gives follow-on biologic companies latitude to make formulation changes by explicitly allowing changes in inactive ingredients: "the biological product [must be] highly similar to the reference product *notwithstanding minor differences in clinically inactive components*."[194] It remains to be seen whether "minor differences" will allow manufacturers to make changes wide enough to avoid the patent infringement, but based on the experience of generic manufacturers it seems likely.

The experience of generic drug manufacturers is not completely analogous to biologics. Most formulation changes made in generic drugs involved incorporating the active ingredient into pills that can be taken orally while most biologics are given by injection, and cannot be given orally (yet). However,

193. In 80% of cases, generic drugs involving a formulation work-around did not infringe. *Acorda* Therapeutics Inc. v. Apotex, Inc., 2011 WL 4074116 (D.N.J. 2011) (brand name drug: tizanidine on beads; generic drug: tizanidine granulation; court found no infringement); Cephalon, Inc. v. Watson, Pharm., Inc., 769 F. Supp. 2d 729 (D. Del. 2011) (brand name drug: sodium bicarbonate; generic drug: potassium bicarbonate; court found no infringement); Elan Corp. v. Andrx, Pharm., Inc., 2008 WL 4709251 (S.D. Fla 2008) (brand name drug: multi-particulate pellet form surrounded by multi-layer membrane; generic drug: pellet that does not dissolve completely during use, is not completely spherical and is not completely enclosed by membrane; court found infringement by equivalents); In re Omeprazole Patent Litig., 490 F. Supp. 2d 381 (S.D.N.Y. 2007) (brand name drug: talc and hydroxypropylmethylcellulose to stabilize core; generic drugs: other chemicals used to stabilize core; court found no infringement); Ranbaxy Lab. Ltd. v. Abbott, Lab., 2005 WL 3050608 (N.D. Ill. 2005) (brand name drug: drug mixed with "a pharmaceutically acceptable polymer;" generic drug: drug mixed with glycerin monostearate; court found infringement by equivalents); Janssen Pharm. N.V. v. Eon Laboratories, Lab. Mfg., 374 F. Supp. 2d 263 (E.D.N.Y. 2004) (brand name drug: beads with diameter of 600-700um; generic drug: beads with diameter of 700-800um; court found no infringement); Bristol-Myers Squibb Co. v. Andrx, Pharm., 343 F. Supp. 2d 1124 (S.D. Fla. 2004) (brand name drug: pregelatinized starch; generic drug: microcrystalline cellulose; court found no infringement); Bristol-Myers Squibb Co. v. Teva, Pharm. USA, Inc., 288 F. Supp.2d 562 (S.D.N.Y. 2003) (brand name drug: lubricant selected from stearyl fumarate or hydrogenated vegetable oil; generic drug: lubricants sodium lauryl sulfate and glyceryl behenate; court found no infringement); Biovail Corp. Int'l v. Andrx, Pharm., Inc., 158 F. Supp. 2d 1318 (S.D. Fla. 2000) (brand name drug: drug in admixture with wetting agent; generic drug: drug over core of sucrose and starch; court found no infringement); Upjohn Co. v. Mova Pharm. Corp., 31 F. Supp. 2d 211 (D.P.R. 1998) (brand name drug: spray-dried lactose making up 70% of composition; brand name drug: spray-dried lactose making up 49% of composition; court found no infringement); A.H. Robins Co. v. Erbamont, Inc., 1991 WL 229150 (S.D. Ohio 1991), *vacated* (brand name drug: hydrophilic surfactant external to microcapsule; generic drug: myristic acid in shell wall of microcapsule; court found no infringement).

194. 42 U.S.C. § 262(k)(2)(A)(i)(I)(aa) (2012) (emphasis added).

most biologics contain excipients dissolved in solution with the active ingredient; therefore, excipient changes in generic small molecule drugs are still relevant.

Biologics manufacturers in Europe have made changes to excipients. Filgrastim Hexal uses glutamate as its buffer, whereas the reference product, Neupogen, uses acetate.[195] The EMEA determined that the two buffer components were equally effective in maintaining the stability of the active ingredient.[196] The manufacturers of Eprex, an epoetin compound, originally used human serum albumin (HSA) as a stabilizer. They later switched to glycine and polysorbate 80, a formulation change (though Eprex is a pioneer drug, not a follow-on biologic).[197]

Only one U.S. court case has dealt with the doctrine of equivalents in the context of formulation changes in biologics. While the case does not give a definitive holding, it gives a hint at how courts will treat formulation changes. In *Amgen v. Hoffman-La Roche*,[198] the Federal Circuit addressed whether a change in Pegylation infringed by equivalents. Pegylation is the process of adding a polyethylene glycol (PEG) chain to a drug, which improves its water solubility and circulation time. Amgen held a patent claiming "a pharmaceutical composition comprising an effective amount of a glycoprotein product effective for erythropoietin therapy. . ."[199] Roche produces Pegylated erythropoietin. At trial, a jury found that Roche's product infringed by equivalents and Roche appealed for a judgment as a matter of law.[200] Amgen argued that it had presented evidence that Pegylation was equivalent to a composition "effective for erythropoietin therapy" because both compounds had the same function (to stimulate the maturation of bone marrow cells into red blood cells), way (by binding to an erythropoietin receptor) and result (making more blood cells).[201] The District Court found that Amgen's testimony was given as part of its literal infringement case, not as particularized testimony concerning the doctrine of equivalents, which is required, therefore the jury heard no evidence on the doctrine of equivalents, and so the jury verdict should be overturned. The Federal Circuit agreed.[202]

With little guidance from biologics cases, my predictions on how patent law will interact with formulation changes in follow-on biologics are derived

---

195. European Medicine Agency Evaluation of Medicines for Human Use. Assessment Report for Filgrastim Hexal, 7, available at http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_Public_assessment_report/human/000918/WC500022471.pdf (last visited Dec. 12, 2012).

196. *Id.*

197. Schellekens, *supra* note 107, at i.30.

198. Amgen Inc. v. F. Hoffmann-La Roche Ltd., 580 F.3d 1340 (Fed. Cir. 2009).

199. *Id.* at 1380.

200. *Id.* at 1382.

201. *Id.* at 1381.

202. *Id.* at 1382.

from cases on small molecule drugs. As mentioned, many formulation changes made in small molecule drugs cannot be applied to biologics because the changes were to oral formulations, whereas most biologics are injected. This means that it will be more difficult for follow-on biologic manufacturers to make formulation changes that do not infringe by equivalents simply because there is a lower number of acceptable changes open to them, because biologics are offered in fewer types of dosage forms.

Furthermore, follow-on biologic manufacturers will have to focus on making changes that do not have a substantially similar function, way, or result on an element-by-element basis.[203] This means that the specific excipient that is switched will have to perform a different function, do so a different way, or achieve a different result. Courts also appear to be looking for true innovation, rather than mere copying. Courts in small molecule drug cases have shown that they have little patience for copies that generic companies tried to disguise as substantial changes.[204]

In addition, courts look for detailed evidence of how the excipient functions. In small molecule drug cases, courts have refused to find infringement because the plaintiff did not provide sufficient evidence of how the inactive ingredient performed the particular function.[205] It is possible that follow-on biologic companies will have an advantage when it comes to this information requirement because of the BPCIA's heightened data requirements for biologics. Data submitted to the FDA is not automatically evidence of legal equivalence, but the individual studies can be used to support an argument of equivalence (as long as the studies directly compare the drugs).[206]

In the process of completing the abbreviated BLA, follow-on biologic companies will have to provide evidence that "the biological product and

---

203. *See* Acorda v. Apotex, 2011 WL 4074116 (D.N.J. 2011).

204. *See, e.g.*, Elan v. Andrx, 2008 WL 4709251 (S.D. Fla. 2008). In *Elan*, the plaintiff's product is Naprelan, a controlled release formulation of naproxen sodium, which is a pain reliever. Elan's patent claims naproxen in a multi-particulate pellet form, which creates the controlled release layer. Each pellet has a core of naproxen surrounded by a multi-layer membrane. The defendant claimed its product is different from the plaintiff's because (1) it is not a multi-particulate form because it disintegrates partially; (2) it is not a pellet because it is not completely spherical; and (3) it does not have a multi-layer membrane surrounding a core because its multi-layer membrane coating does not "completely enclose" the core. The court begins its analysis by noting that once scientists at Andrx decided to create generic naproxen sodium, they obtained a copy of Elan's patent to study. They tried various methods of mixing together ingredients until they settled on a formulation, which they felt was sufficiently different from Naprelan. The court was very skeptical of Andrx's changes, finding that all aspects of the generic drug performed substantially the same function in substantially the same way to obtain substantially the same result, and thus infringed under the doctrine of equivalents.

205. *See, e.g.*, In re Omeprazole Patent Litigation, 490 F. Supp. 2d 381 (S.D.N.Y. 2007); Cephalon v. Watson, 769 F. Supp. 2d 729 (D. Del. 2011).

206. *See* Adams Respiratory Therapeutics, Inc. v. Perrigo Co., 616 F.3d 1283, 1287-89 (Fed. Cir. 2010).

reference product utilize the same mechanism or mechanisms of action for the condition or conditions of use prescribed. . .but only to the extent the mechanism or mechanisms of action are known for the reference product."[207] This requirement is addressed to the active ingredient, not the excipient, but it seems likely the company will also need to understand how the excipients interact with the drug as compared to the reference drug's excipients. The FDA's draft guidelines also recommend that applicants assess "excipients and any formulation effect on purity, product- and process-related impurities, and stability."[208] This evidence will be helpful for proving their case in court.

Overall, formulation changes are a good target for follow-on biologic companies. It may be a less fruitful path for follow-on biologics relative to generics, because there are fewer dosage forms and thus fewer possible changes to make. However, the BPCIA's explicit allowance of excipient changes suggests the FDA will allow follow-on biologic companies to make a reasonably wide range of excipient substitutions, giving them more ground to avoid infringing by equivalents. Moreover, if the follow-on biologic companies are able to make a substantial change, the heightened data requirements means that they will likely have better data to use to prove that their change is substantial.

6. *Packaging changes*

Packaging changes involve changing the external container that holds the biological product. Packaging can be a very important part of a biologic because it can affect the stability of a product (for example, packaging may be required to keep out heat, light, or moisture) or improve ease of delivery. There are no packaging cases involving biologics. Therefore, my predictions in this section are based on packaging cases involving small molecule drugs. Although packaging biologics presents different concerns from packaging small molecule drugs, the way that the courts address the problem is likely to be similar.

The FDA has indicated a willingness to approve follow-on biologics that have different packaging from the reference product. In its draft guidance, the FDA stated that "some design differences in the delivery device or container closure system used with the proposed follow-on biologic product may be acceptable."[209] The FDA gave the example of auto-injectors, a syringe that already contains a pre-determined amount of a drug, which cuts out the step of filling the needle before injection. The draft guidance states that it "may be possible, for example, for an applicant to obtain licensure of a proposed biosimilar product in a pre-filled syringe or in an auto-injector device. . .even if

207. 42 U.S.C. § 262 (k)(2)(A)(i)(II).
208. *Scientific Considerations*, *supra* note 62, at 9.
209. *Questions and Answers*, *supra* note 90, at 5.

the reference product is licensed in a vial presentation."[210] However, the FDA emphasizes that it is still necessary for a follow-on biologic with a packaging change to meet "the statutory standard for biosimilarity."[211]

Several cases involving the packaging of small molecule drugs and the doctrine of equivalents have been litigated.[212] The results were mixed, with half the courts finding infringement. The difference lies in the 'way' prong of the doctrine of equivalents test (although the courts do not all use this language). In *Abbott Laboratories v. Baxter Healthcare Corp*, the plaintiff solved the problem of degradation by mixing its drug with water.[213] The defendant solved the problem of degradation by creating a package lining that physically blocked the drug from coming into contact with the walls of the container, the source of the degrading chemical.[214] The court found that there was no infringement because the products prevented degradation in a different *way*.[215] Conversely, in *Mead Johnson & Co. v. Barr Laboratories, Inc.*, the plaintiff patented a method of scoring a pill in order to help a patient divide it into sections.[216] The plaintiff's product had opposing score notches, whereas the defendant's product had transverse score notches. The court found that there was infringement, noting that both products facilitated tablet breakage the same *way*—by directing pressure applied by the patient to achieve a more uniform fracturing.[217]

Packaging is a promising area for follow-on biologic companies because it appears both that the doctrine of equivalents will not bar all packaging changes and that the FDA will allow a relatively broad range of packaging changes. The comparison of *Abbott Laboratories v. Baxter Healthcare Corp.* and *Mead Johnson & Co. v. Barr Laboratories, Inc.* indicates that courts are willing to find that a packaging change does not infringe by equivalents as long as they are convinced that the containers function in different ways. Follow-on biologic companies could overcome the doctrine of equivalents by changing the packaging of their products if that packaging improved the stability of the product or ease of delivery in a different way. The FDA's draft guidance indicates that the FDA will allow packaging changes as long as the products

---

210. *Id.*

211. *Id.*

212. Abbott Laboratories v. Baxter Healthcare Corp, 660 F. Supp. 2d 882 (N.D. Ill. 2009); Abbott Laboratories v. Baxter, Pharmaceutical Products, Inc., 2005 WL 2347221 (N.D. Ill. 2005); EKR Therapeutics, Inc. v. Sun Pharma., Pharmaceuticals, Ltd., 633 F. Supp. 2d 187 (D.N.J. 2009); Bio-Technology General Corp. v. Duramed, Pharmaceuticals, Inc., 174 F. Supp. 2d 229 (D.N.J. 2001); Bio Technology General Corp. v. Duramed, Pharmaceuticals, Inc., 325 F.3d 1356 (Fed. Cir. 2003); Mead Johnson & Co. v. Barr, Laboratories, Inc., 38 F. Supp. 2d 289 (S.D.N.Y. 1999).

213. 660 F. Supp. 2d 882, 884 (N.D. Ill. 2009).

214. *Id.*

215. *Id.* at 888.

216. 38 F. Supp. 2d 289, 295 (S.D.N.Y. 1999).

217. *Id.* at 296.

remain "highly similar". It is likely to be easier for a follow-on biologic company to change the packaging while remaining "highly similar" than for the follow-on biologic company to change the formulation or manufacturing process while remaining "highly similar."

CONCLUSION

Questions of patent infringement, particularly under the doctrine of equivalents, are difficult even with longstanding, well characterized technologies. The complexities of biotechnology present unique challenges that courts struggle to resolve. The advent of follow-on biologics is yet another hurdle that the legal system is, at the moment, not well positioned to face. Because the development of follow-on biologics is so expensive, it is tremendously important that the products be treated consistently and predictably when they arrive in court.

This Article has presented a guide for industry, courts and scholars on how questions of patent infringement—principally the under doctrine of equivalents—will develop and how they should be resolved. Both the BPCIA and patent law guide the shape of infringement suits. Follow-on biologics companies will be most successful when they make a change in the pre-transformation process, the transformation process, the formulation, or the packaging. They will be least successful when they make a change in the cell culture conditions or the purification process. This is because the FDA will more closely regulate the latter category of changes, giving follow-on biologics companies less scope to make changes that will bring them outside the range of equivalents for the brand-name product. It remains to be seen how courts will address issues of infringement for follow-on biologics, but all parties should be aware of the complexity of the scientific and legal issues and the importance of addressing them properly.

STANFORD TECHNOLOGY LAW REVIEW
VOLUME 16, NUMBER 1 FALL 2012

MOORE IS LESS: WHY THE DEVELOPMENT OF INDUCED PLURIPOTENT STEM CELLS MIGHT LEAD US TO RETHINK DIFFERENTIAL PROPERTY INTERESTS IN EXCISED HUMAN CELLS

Osagie K. Obasogie* & Helen Theung**

CITE AS: 16 STAN. TECH. L. REV. 51 (2012)
http://stlr.stanford.edu/pdf/mooreisless.pdf

ABSTRACT

Since Moore v. Regents of the University of California*, there has been a wide-ranging debate regarding the holding of the case and its implications for property law.* Moore *stands for the notion that individuals do not have a property interest in ordinary cells taken from their bodies during medical procedures nor the commercial products that researchers might develop from them. At the same time, cases such as* Davis v. Davis *and* Hecht v. Superior Court *have asserted that individuals maintain a property interest in other types of cells—namely embryos and gametes (eggs and sperm)—once they are removed from the body. This, among other developments, has led to a fragmented regime in property law pertaining to excised biological materials that turns, in large part, on the type of*

* Associate Professor of Law, University of California Hastings College of the Law with a joint appointment at the University of California, San Francisco (UCSF) Department of Social and Behavioral Sciences; Senior Fellow, Center for Genetics and Society. B.A. Yale University; J.D. Columbia Law School; Ph.D. University of California, Berkeley. Robin Feldman, Lisa Ikemoto, Stuart Newman, Krishanu Saha, and David Winickoff kindly and thoughtfully reviewed earlier drafts of this Article, in which we are most grateful.
** Attorney, Sidley Austin LLP in Palo Alto, CA. B.A. University of California, Santa Barbara; J.D. University of California, Hastings College of the Law. The opinions expressed herein are those of the authors and do not necessarily reflect the views of the firm, clients, or any affiliates of any of the foregoing. This Article has been prepared for informational purposes only and does not constitute legal advice.

cell in question: individuals have a diminished interest in regular somatic cells (skin, muscle, etc.) while courts have recognized that people retain a heightened property interest in reproductive cells such as sperm, eggs, and embryos. The articulated reason for the differential property interests in these two cell types is that embryos and gametes have the "potential for human life" while individuals are thought to have little use for ordinary body cells once they are excised.

This default rule has framed property law regarding excised human cells for over two decades. It exists to balance the need for scientists to have access to research materials with individuals' reproductive autonomy. To the extent that the dividing line determining the property interest in excised cells turns largely upon their "potential for human life," the recent development of induced pluripotent stem cells (iPSCs) suggests that this default rule is becoming increasingly untenable. Research has shown that iPSCs can create the ability to genetically reprogram somatic cells into a pluripotent state that may allow them to differentiate into other types of cells—including eggs and sperm—that can be used to create new organisms. While these developments have not yet been fully applied to human iPSCs, they nonetheless suggest that iPSCs may soon be able to give ordinary somatic body cells the same potential for human life as naturally produced embryos and gametes but without the corresponding property interest.

This Article argues that given this new technology, its relative success in animal models, and its impending application to human cells, the current default rules precluding individuals' property interest in excised somatic cells needs substantial reconsideration. We propose a three-part approach to manage the challenges that iPSCs create for this aspect of property law. This includes (1) a self-imposed moratorium on human applications of iPSC research that can lead to human reproduction (2) Congressional action that vests property interests in the donors of somatic cells once their cells have been reprogrammed to a pluripotent state and differentiated into reproductive cells and (3) Judicial action that distinguishes Moore and related cases by acknowledging the reversion of property interest to donors once somatic materials are reprogrammed to a state of pluripotency and differentiated into reproductive cells. This proposal offers the best way to deal with the profound legal issues created by this new technology with the least disruption to existing rules and policy preferences.

INTRODUCTION

Like *Marbury v. Madison* and *Erie Railroad Co. v. Tompkins*, *Moore v. Regents of the University of California* is one of a handful of cases that virtually every first-year law student reads as part of her introduction to the American legal system. *Moore* stands for what has become a proverbial default rule in property law: individuals do not have property interests in their own cells once they are removed from their bodies. (This default rule pertains to cells removed from living human beings during medical procedures, not (a) organs removed for transplant, (b) tissues subject to routine-removal statutes such as corneas from cadavers,[1] or (c) anatomical gifts. These transactions are governed by the National Organ Transplant Act, state level implementations of the Uniform Anatomical Gift Act, and other laws.) Thus, in *Moore*, researchers at UCLA Medical Center that used John Moore's spleen cells without his knowledge or consent to develop a profitable cell line were not liable for conversion[2] since the Court found that Moore no longer had a property interest in these excised cells. Yet, *Moore's* holding regarding individuals' diminished property interests in excised cells does not apply to *all* human cells. There are two exceptions: gametes (eggs and sperm) and embryos. Courts have found that since these cells have the potential to create or become independent human beings, individuals retain a property interest in them after being removed from the body that does not exist for somatic cells, or the ordinary non-reproductive cells that make up various parts of the body such as hair, skin, or Moore's spleen. Thus, the critical dividing line in property law with regards to individuals' interest in their excised cells is whether or not they have the "potential for human life."[3]

To the extent that the default rule regarding individuals' diminished property interest in excised human cells largely exists for policy reasons such as promoting efficient research and respecting individuals' reproductive decision-making,[4] this dividing line between gametes and embryos on one hand and ordinary non-reproductive somatic cells on the other appears coherent. Excised somatic cells that do not have any reproductive capacity are important

---

1. *See* COMMITTEE ON INCREASING RATES OF ORGAN DONATION, ORGAN DONATION: OPPORTUNITIES FOR ACTION 206-07 (JAMES F. CHILDRESS & CATHARYN T. LIVERMAN eds., 2006).

2. Conversion is a tort that reflects the "unauthorized and wrongful exercise of dominion and control over another's personal property, to the exclusion of or inconsistent with rights of owner." BLACK'S LAW DICTIONARY 332 (6th ed. 1990).

3. Davis v. Davis, 842 S.W.2d 588, 597 (Tenn. 1992).

4. Individuals' intent and institutional informed consent play a significant role in the disposition of these cells. See discussion in Part II.

for scientific advancement and often have little value to most patients. Therefore, a default rule that gives sole property interests to scientists and research entities to the exclusion of individual patients encourages efficiency that promotes and rewards innovation. In contrast, gametes and embryos that have the potential to create independent human life may be extraordinarily valuable to individuals, which law recognizes by acknowledging a continued property interest in these types of cells once outside of the body.

However, new developments in human biotechnology are making this dividing line increasingly untenable—to the point where current default rules espousing individuals' diminished property interest in somatic cells may need substantial reconsideration. In 2007, research groups headed by Shinya Yamanaka[5] and Jamie Thompson[6] demonstrated the ability to reprogram human somatic cells into a pluripotent state. This means that regular somatic cells like those from Moore's spleen—the very types of cells that individuals have a diminished property interest in once excised—can be reverted back into a condition (pluripotency) whereby they can develop into several different types of cells—*including eggs or sperm*.[7] Known as induced pluripotent stem cells (iPSCs), these cells are hailed as offering an end-run around the ethical quagmire surrounding embryonic stem cells since they offer the promise of regenerative medicine (where such pluripotency might allow researchers to "grow" patient-specific cells to cure diseases) without the ethically fraught issue of destroying embryos.

Yet at the same time that iPSC research has been heralded as resolving a particularly thorny ethical issue, it has created a profound challenge for property law that has gone almost wholly unnoticed. If the legal justification for diminishing individuals' property interests in their own somatic cells and acknowledging scientists' claims to own such material is that these cells do not have the "potential for human life," then the impending ability to reprogram such cells into a pluripotent state where they can then differentiate into reproductive cells with the potential to become autonomous human beings radically upends this logic. In short, human iPSCs, if they achieve the same potential as has already been demonstrated in animal experiments, can become a profound game changer in that *every somatic cell would have the "potential for human life*"; the proverbial spleen cells from John Moore and any other ordinary cells removed during medical procedures would potentially be just a few steps away from being turned into gametes that could then be used for reproductive purposes. This suggests that this new technology might blur the

---

5. Kazutoshi Takahashi et al., *Induction of Pluripotent Stem Cells from Adult Human Fibroblasts by Defined Factors*, 131 CELL 861, 861 (2007).

6. Junying Yu et al., *Induced Pluripotent Stem Cell Lines Derived from Human Somatic Cells*, 318 SCIENCE 1917, 1917 (2007).

7. While iPSC research with mice has shown the ability to induce somatic cells into a pluripotent state that can then differentiate into reproductive cells, this research has not yet been done with humans though it is considered to be feasible. See Part III.

dividing lines and default rules in property law and that serious reconsideration may be needed.

Despite the wide-ranging post-*Moore* discussion of property rights in excised human cells, this Article is the first to identify and articulate the profound challenges raised by iPSCs for property law. These challenges are likely to have important implications. Given an estimated 270 million human tissue samples held in domestic biobanks, the development of iPSCs adds a new and qualitatively different dimension to an endearing question in property law: who can own your body (in terms of excised cells and tissues) and under what circumstances?

This Article proceeds in three parts to address these issues. Part I assesses existing rules pertaining to individuals' property interests in excised somatic cells as well as gametes and embryos. Part I also examines key cases in the development of this jurisprudence to identify and substantiate a basic underlying premise in property law pertaining to excised human cells: in the absence of a contract or preexisting agreement, default property interest in excised human cells is given to researchers and scientists *except* in the case of eggs, sperm, and embryos, whereby individuals retain a property interest since these cells have the "potential for human life." Thus, Part I highlights how this potentiality is a key dividing line in establishing differential property interests in excised human cells. Part II discusses the development of iPSCs and how this technology might complicate the logic of this dividing line by giving human somatic cells the potential to become life through cellular reprogramming that reverts them to a pluripotent state where they can differentiate into many types of cells, potentially including eggs and sperm. Part III situates this issue in the landscape of the current scholarly debate on property interest in human cells to highlight the transformative nature of this technology; existing conversations have entirely missed the significance of iPSC research for property law. We then offer a three-part proposal for how law and science should respond to the challenges raised by iPSCs. We argue that (1) the scientific community should engage in a self-imposed moratorium on human applications of iPSC research that may lead to human reproduction; (2) that Congress should enact legislation that vests property interests in excised somatic cells in donors once these cells have been reprogrammed to a pluripotent state and differentiated into reproductive cells; and (3) that courts should acknowledge the heightened property interest that vests in excised somatic cells once they are reprogrammed to a pluripotent state and differentiated into reproductive cells. After discussing various objections some may have with this proposal and offering rebuttals, we conclude with a brief discussion of this Article's significance for the future of property law.

I. PROPERTY LAW REGARDING SOMATIC CELLS, GAMETES, AND EMBRYOS: A BRIEF OVERVIEW

Since *Moore v. Regents of the University of California*, the law regarding individuals' property interests in excised cells has been fragmented and unevenly developed. Starting with *Moore*, this section discusses key cases that have laid the broad foundations for the current default rules that confer differential property interests to human cells in a manner that depends heavily on their potential to create or become human life. Although the case law on this topic is scarce and dispersed, courts have consistently concluded that individuals have little to no property interests in excised somatic cells while also acknowledging individuals' significant property interests in gametes and embryos that exist outside the body. The potential to create or become full human beings plays a large role in the court's justification for these rules.

A. *Somatic Cells*

Discussions about property rights and the human body are not new.[8] However, recent developments in new reproductive and genetic technologies have given rise to novel questions about the rights individuals have in excised tissues and cells. *Moore v. Regents of the University of California* represents one of the earliest judicial considerations of this question, and its disposition has had cascading effects on how the law understands individuals' property interests in their own cells once they are disconnected from their bodies.

John Moore was a patient at UCLA Medical Center in the mid 1970's, where he received treatment for hairy-cell leukemia. In the course of this treatment, Dr. David Golde, Moore's physician, took "extensive amounts of blood, bone marrow aspirate, and other bodily substances" all while knowing—

8. *See generally* ALAN HYDE, BODIES OF LAW (1997). Radhika Rao offers an interesting discussion of an early theory of the body as property as articulated by John Locke: "The image of the body as a form of property possessed by its 'owner' dates back at least to John Locke, whose influential theory of property derived all ownership from the property possessed by individuals in their own persons. In his treatise "Of Property," written around 1690, Locke asserted: "Though the Earth and all inferior Creatures be common to all Men, yet every Man has a Property in his own Person. This no Body has any Right to but himself." According to Locke, individual ownership of the physical body entailed ownership of those external things that are the product of the body's labor. Yet Locke apparently envisioned the body as property of a special sort, held in trust rather than as an individual owner. As a result, he believed that a person's rights to life and liberty were inalienable because they were not his own, but belonged to another. These limits upon bodily property followed from the fact that ultimate ownership rested with the deity. Thus Locke apparently viewed individuals as stewards over their bodies, possessing themselves in trust rather than as outright owners. Therefore, despite his reliance upon property rhetoric, his image of the rights individuals possess in their bodies clearly does not rise to the level of complete ownership." Radhika Rao, *Property, Privacy, and the Human Body*, 80 B.U. L. REV. 359, 367-68 (2000).

and without disclosing to Moore—that these biological materials "were of great value in a number of commercial and scientific efforts," that could provide "competitive, commercial, and scientific advantages."[9] Moore also had his spleen removed at Golde's recommendation. While Golde received consent for the splenectomy, he did not disclose to Moore that his spleen and other biological materials would be used for research. Between 1976 and 1983, Moore travelled to UCLA Medical Center from his home in Seattle to give more samples of blood and other tissues because "he had been told that the procedures were to be performed only there and only under Golde's direction."[10]

Little did Moore know that Golde was working with others to develop a cell line from T-lymphocytes derived from Moore's tissues. Golde, Shirley Quan, and the Regents of the University of California patented the cell line and shared in the royalties while excluding Moore from any compensation.[11] Moore brought suit, claiming a breach of fiduciary duty and lack of informed consent. Moore also brought a claim for conversion, or that the defendants interfered with his interest in his personal property—his blood, cells, etc.—and that he subsequently had an interest in the products derived from his bodily materials.[12] While the Court found that Dr. Golde did not fulfill his fiduciary duty to Moore and impermissibly failed to obtain informed consent by not disclosing his financial interests, it ruled against Moore's conversion claims. As a descriptive matter, the Court stated that current law simply did not support Moore's claim that he *owned* these excised biological materials, which is a predicate to making any successful claim that they were illegally subject to conversion.[13] The Court refused to extend the principle of conversion to this

9. Moore v. Regents of the Univ. of Cal., 793 P.2d 479, 481 (Cal. 1990).

10. *Id.* at 482.

11. "With the Regents' assistance, Golde negotiated agreements for commercial development of the cell line and products to be derived from it. Under an agreement with Genetics Institute, Golde 'became a paid consultant' and 'acquired the rights to 75,000 shares of common stock.' Genetics Institute also agreed to pay Golde and the Regents 'at least $ 330,000 over three years, including a pro-rata share of [Golde's] salary and fringe benefits, in exchange for . . . exclusive access to the materials and research performed' on the cell line and products derived from it. On June 4, 1982, Sandoz 'was added to the agreement,' and compensation payable to Golde and the Regents was increased by $ 110,000. '[T]hroughout this period, . . . Quan spent as much as 70 [percent] of her time working for [the] Regents on research' related to the cell line." *Id.* at 482.

12. "Moore also attempts to characterize the invasion of his rights as a conversion – a tort that protects against interference with possessory and ownership interests in personal property. He theorizes that he continued to own his cells following their removal from his body, at least for the purpose of directing their use, and that he never consented to their use in potentially lucrative medical research. Thus, to complete Moore's argument, defendants' unauthorized use of his cells constitutes a conversion. As a result of the alleged conversion, Moore claims a proprietary interest in each of the products that any of the defendants might ever create from his cells or the patented cell line." *Id.* at 487.

13. Quoting *Del E. Webb Corp. v. Structural Materials Co*, the *Moore* Court noted that "to establish a conversion, plaintiff must establish an actual interference with his

novel area of excised cells for three reasons. First, the Court noted that this analogy—treating excised biological materials like personal property—had not been supported by any other court, suggesting a general consensus that individuals did not retain an ownership interest in cells akin to personal property once removed from the body. Second, state statutes limited what individuals could do with excised biological materials, which suggested that any ownership interest in them has been significantly curtailed.[14] Lastly, with regard to Moore's claim that he had a property interest in the cell line derived from him, the Court reasoned that "[it] cannot be Moore's property . . . . because the patented cell line is both factually and legally distinct from the cells taken from Moore's body."[15]

But the Court also made a series of *normative* claims regarding whether a property interest in excised cells *ought* to be recognized to support Moore's conversion claims, which allowed the Court to discuss various policy issues implicated by this situation in a manner that has substantially affected this area of law. The Court cited three policy reasons to explain why the use of excised cells in medical research did not constitute conversion. First, the Court explained that these issues are better handled by legislatures. Second, the Court acknowledged that a tort of conversion is not necessary to affirm or protect the rights of patients. But the Court seemed most disturbed by the third reason: that potentially adverse impacts might follow from extending the tort of conversion into the area of biomedical specimens. The Court noted:

> Research on human cells plays a critical role in medical research. This is so because researchers are increasingly able to isolate naturally occurring, medically useful biological substances and to produce useful quantities of such substances through genetic engineering. These efforts are beginning to bear fruit. Products developed through biotechnology that have already been approved for marketing in this country include treatments and tests for leukemia, cancer, diabetes, dwarfism, hepatitis-B, kidney transplant rejection, emphysema, osteoporosis, ulcers, anemia, infertility, and gynecological

---

ownership or *right of possession* . . . Where plaintiff neither has title to the property alleged to have been converted, nor possession thereof, he cannot maintain an action for conversion." *Id.* at 488.

14. "Pursuant to Health and Safety Code section 7054.4, '[n]otwithstanding any other provision of law, recognizable anatomical parts, human tissues, anatomical human remains, or infectious waste following conclusion of scientific use shall be disposed of by interment, incineration, or any other method determined by the state department [of health services] to protect the public health and safety.' Clearly the Legislature did not specifically intend this statute to resolve the question of whether a patient is entitled to compensation for the nonconsensual use of excised cells. A primary object of the statute is to ensure the safe handling of potentially hazardous biological waste materials. Yet one cannot escape the conclusion that the statute's practical effect is to limit, drastically, a patient's control over excised cells. By restricting how excised cells may be used and requiring their eventual destruction, the statute eliminates so many of the rights ordinarily attached to property that one cannot simply assume that what is left amounts to "property" or "ownership" for purposes of conversion law." *Id.* at 491-92.

15. *Id.* at 492.

> tumors, to name but a few . . . . [T]he extension of conversion law into this area will hinder research by restricting access to the necessary raw materials. . . . [The] exchange of scientific materials, which still is relatively free and efficient, will surely be compromised if each cell sample becomes the potential subject matter of a lawsuit.[16]

Thus, the burden that might befall the scientific community drives much of the Court's concern. For the *Moore* court, the legislature should lead such far-reaching changes rather than the judiciary since they involve policy issues far beyond individual property rights. The end result is that after *Moore*, individuals are thought to have radically diminished property interests in their excised somatic cells, although physicians and researchers still have a duty to inform patients that their biological materials may be used in research and for commercial purposes.

Subsequent cases have further discussed and reaffirmed the *Moore* court's conclusion that individuals do not have a property interest in excised cells. *Greenberg v. Miami Children's Hospital Research Institute*, a 2003 decision by the United States District Court for the Southern District of Florida, involved the disposition of human tissues given to researchers for the purpose of identifying the genes responsible for Canavan diseases and developing carrier tests that would permit prenatal screening. This collaboration, where affected patients gave tissues and other biological materials to researchers, led to a breakthrough that identified the Canavan-associated gene. While the plaintiffs expected that any developments stemming from research using their blood and tissues would be offered in an affordable and accessible manner that stayed in the public domain, the researchers patented the gene. This gave the researchers and the hospital "the ability to restrict any activity related to the Canavan disease gene, including without limitation: carrier and prenatal testing, gene therapy, and other treatments . . . ."[17] Soon after, the hospital allegedly threatened other hospitals that infringed their patent through unauthorized testing and began negotiating licenses and royalty fees that restricted the tests' availability. The tissue donors sued, saying that they were not aware of the researchers' intent to patent the research or commercialize it.

The plaintiffs made several claims in their suit, including lack of informed consent, unjust enrichment, and fraudulent concealment. The plaintiffs' conversion allegation claimed a property interest in research stemming from their donated tissues and blood. This was based on an underlying claim that they continued to possess a property interest in these biological materials once excised from their bodies. The court declined to extend conversion theory to excised tissues, stating that the tissues were "donations to research without any contemporaneous expectations of return of the body tissue and genetic

16. *Id.* at 494-99.

17. Greenberg v. Miami Children's Hosp. Research Inst., 264 F. Supp. 2d 1064, 1067 (S.D. Fla. 2003).

samples . . . ."[18] Citing *Moore*, the court noted "[t]he California Supreme Court . . . held that the use of the results of medical research inconsistent with the wishes of the donor was not conversion, because the donor had no property interest at stake after the donation was made . . . . Similarly [in *Greenberg*], the property right in blood and tissue samples also evaporates once the sample is voluntarily given to a third party."[19] It is also important to note that the *Greenberg* court adopted the same consequentialist reasoning articulated in *Moore* for denying the plaintiffs' conversion claim. The Court plainly stated, "if adopted, the expansive theory championed by plaintiffs would cripple medical research as it would bestow a continuing right for donors to possess the results of any research conducted by the hospital."[20]

Washington University v. Catalona presented a similar issue before the Eighth Circuit Court of Appeals in 2007. Dr. William Catalona was a urologist at Washington University where his primary research area was prostate cancer. During his nearly three-decade tenure at Washington University, he amassed a large biorepository of blood and tissue for prostate cancer research—both from his patients and through larger scale recruiting.[21] Catalona accepted a new position at Northwestern in 2003 and sent letters to the research participants asking that they sign a form allowing their samples to be transferred from Washington University to his new employer. Washington University filed a declaratory action in 2003 that sought to establish their ownership of the repository amassed by Catalona and all of the biological samples that it contained.

The court held in favor of Washington University, finding that the patients' stored samples were *inter vivos* gifts donated to the university and remained its property whereby the individual donors could not re-assign the gifts to Catalona. This rationale is similar to that used by the *Greenberg* court; by framing the research participants as "donors" who made a "gift," the court

---

18. *Id.* at 1074.

19. *Id.*

20. *Id.* at 1076.The *Greenberg* court uses a similar argument in rejecting the plaintiff's claim that the researchers failure to disclose their economic interests amounted to a lack of informed consent: "[D]isclosing economic interests has no support in established law, and more ominously, this requirement would have pernicious effects over medical research, as it would give each donor complete control over how medical research is used and who benefits from that research." *Id.* at 1070.

21. "At the time of the district court's permanent injunction hearing in this case, more than 30,000 [research participants] were enrolled in WU prostate cancer research studies. About 2,500 to 3,000 [research participants] had been patients of Dr. Catalona. The Biorepository contains: (1) approximately 3,500 prostate tissue samples taken from patients of Dr. Catalona and other WU physicians within the Division; (2) about 100,000 blood or serum samples donated by over 28,000 men, 75% of whom were not patients of any WU physician, but rather were volunteers recruited through the media; and (3) DNA samples provided by approximately 4,400 men, which included patients of different WU physicians and relatives of those patients." Wash. Univ. v. Catalona, 490 F.3d 667, 671-72 (8th Cir. 2007).

implicitly acknowledged that any residual property interest that the individuals might have in their excised tissue was relinquished by agreeing to participate in medical research.[22] This leads to what has now become a jurisprudentially familiar result: individuals have diminished property interests in excised somatic cells and tissues used for research purposes. Taken together, these three cases indicate the parameters of a general default rule that excised human cells and tissues used in medical research are the property of the researcher and/or research institution rather than the individual donor.

B. *Embryos and Gametes*

Courts have taken a different approach to understanding individuals' property interests in other forms of human cells, namely embryos and gametes. This section briefly describes two influential cases that highlight the way courts approach individuals' property interests in this area.

Davis v. Davis, a 1992 opinion from the Supreme Court of Tennessee, was one of the earliest judicial opinions to consider the proper disposition of frozen embryos held in a fertility clinic where there was not any preexisting agreement or contract to determine how unused embryos should be handled. The genetic parents, Mary Sue and Junior Davis, divorced; Mary Sue initially wanted to gestate the embryos and then wanted to donate them to an infertile couple while Junior wanted the embryos discarded. Not only did the couple not stipulate what should happen to any unused embryos prior to their divorce, but there was also no relevant state statute to determine the embryos' fate. Thus, the court examined a number of scientific, ethical, and legal perspectives to determine how to proceed. *Davis* remains a leading case regarding the disposition of gametes and embryos because it provides a broad framework for courts to use when assessing this issue:

> Disputes involving the disposition of pre-embryos produced by in vitro fertilization should be resolved, first, by looking to the preferences of the progenitors. If their wishes cannot be ascertained, or if there is dispute, then their prior agreement concerning disposition should be carried out. If no prior agreement exists, then the relative interests of the parties in using or not using the pre-embryos must be weighed.[23]

But what is particularly important for this Article is the court's reasoning on whether embryos are "persons" or "property" in determining progenitors' rights with regards to their disposition in the absence of any contract or other agreement. The court concluded that embryos:

> [A]re not, strictly speaking, either "persons" or "property," but occupy an interim category that entitles them to special respect because of their *potential*

---

22. The court referenced the consent form and brochure distributed to research participants in coming to this decision. *See id.* at 674-75.

23. Davis v. Davis, 842 S.W.2d 588, 604 (Tenn. 1992).

> *for human life* . . . . [The Davis' interest] is not a true property interest. However, they do have an interest in the nature of ownership, to the extent that they have decision-making authority concerning disposition of the pre-embryos.[24]

Thus, the *Davis* court distinguished the property interest that individuals might have in embryos from other types of excised somatic cells and tissues, largely based upon their "potential for human life" or the capacity to develop into an autonomous human being. This potential is not enough to give individuals a standard property interest over embryos. But, it is enough, according to the court, to give the genetic parents authority to determine how the embryos will ultimately be used—which is the *precise* property interest that was sought by and denied to the cell and tissue progenitors in *Moore*, *Greenberg*, and *Catalona*.

Decided four years after *Davis*, *Hecht v. Superior Court* is significant in that it applied the same "potential for human life" rationale for giving individuals a heightened property interest in gametes and not only embryos. *Hecht* concerned the disposition of fifteen sperm vials left to Deborah Hecht by her partner, William Kane, after his death. Hecht intended to use Kane's sperm to attempt to conceive and give birth to a child. Kane's two adult children objected, leading to the suit. Central to this case is whether or not Kane "owned" the sperm vials in a manner that allowed him to give this type of property to Hecht. Relying on *Moore*, Kane's children argued that Kane could not have a property interest in his sperm or control its disposition once outside of his body much like the California Supreme Court held that Moore did not have a property interest in his excised cells. The California Court of Appeals rejected this application of *Moore* to gametes in the first of three *Hecht* decisions, noting that the "decedent had an interest, in the nature of ownership, to the extent that he had decision making authority [ ] to the sperm. . . . [that] falls within the broad definition of property."[25] Even though Kane was dead at the time of this dispute, the court's inquiry focused largely on what type of interest Kane had in his sperm while alive and whether that allowed him to give it to Hecht.[26] The court used *Davis* to distinguish *Hecht* from *Moore* to find a residual property interest in gametes that is akin to that found in embryos because of their shared potential to create life. The *Hecht* court wrote that sperm "is unlike other human tissue because it is 'gametic material' that can be used for reproduction . . . . [T]he value of sperm [as in embryos] lies in its *potential to create a child* . . . . [Therefore] decedent had an interest, in the nature of ownership . . . . [that] is sufficient to constitute 'property.'"[27] As in the embryo cases, donors' intent remains primary in determining the

---

24. *Id.* at 597 (emphasis added).

25. Hecht v. Superior Court of L.A., 16 Cal. App. 4th 836, 846 (1993).

26. Kane's will also stated that his sperm should be given to Hecht after his death. *See* Hecht v. Superior Court of L.A., 50 Cal. App. 4th 1289, 1292 (1996).

27. *Hecht*, 16 Cal. App. 4th at 850 (emphasis added).

disposition of gametes and frames the nature of the property interest involved. For example, in *Estate of Kievernagel v. Kievernagel*, the court cited *Hecht* in acknowledging that "gametic material, with its potential to produce life, is a unique type of property and thus not governed by the general laws relating to gifts or personal property or transfer of personal property upon death"[28] so as to deny a widow's claim to her late husband's sperm precisely because he wanted the sperm destroyed when he died. To be sure, most cases regarding the disposition of excised embryos and gametes focus on ascertaining and giving effect (where discernable) to the intent of the parties as articulated in informed consent forms or other written agreements at the time of the procedure.[29] However, for the purposes of this Article, *Davis* and *Hecht* are important in that they exist outside of the realm of contractual or statutory interpretation to understand the residual property interests that parties retain in excised gametes and embryos as distinct from ordinary somatic cells because of their potential to become human life.

C. *Distinctions Between Cases Involving Somatic Cells and Gametes/Embryos*

There are at least two distinctions between these two sets of cases that are worth noting. First, the cases pertaining to somatic cells and tissues (*Moore*, *Greenberg*, *Catalona*) deal with disputes between patients/research participants and biomedical research entities while the embryo and gamete cases (*Davis*, *Hecht*) deal with disputes between individual litigants outside of the context of medical research. Thus, the courts' discussion of property rights occurs in two different contexts where there are different sets of competing interests. For example, the cases dealing with somatic cells and tissues each stress the adverse impact that might occur to scientific research if patients and research participants retained a controlling property interest in their excised materials and how this may work against the public interest. This broader notion of efficient use of cells and tissues to yield a wider social benefit is absent from disputes concerning gametes and embryos. Instead, the litigants are dealing with issues pertaining to reproduction that have a different set of social and legal concerns such as the appropriate dispensation of property in probate matters, privacy, the right to procreate, and the right to not be a parent. This, however, may be a distinction without much of a difference. What is of interest

---

28. In re Estate of Kievernagel, 166 Cal. App. 4th 1024, 1030 (2008).

29. *See, e.g.*, Kass v. Kass, 91 N.Y.2d 554 (1998), citing *Davis* in stating that "the relevant inquiry thus becomes who has dispositional authority over [the embryos]. Because that question is answered in this case by the parties' agreement . . . we have no cause to decide whether the pre-zygotes are entitled to 'special respect.'" *Id.* at 564-65; Litowitz v. Litowitz, 48 P.3d 261 (Wash. 2002), noting that the court "based[d] [its] decision . . . solely upon the contractual rights of the parties under the pre-embryo cryopreservation contract with the Loma Linda Center for Fertility and In Vitro Fertilization." *Id.* at 271.

to the many scholarly and judicial disquisitions on *Moore* and its progeny is the specific question of what property interests individuals have in excised cells and tissue, regardless of whether the person exerting a competing claim is a scientist or an ex-spouse. The differential contexts that underlie the juxtaposition of property interests in somatic cells and tissues versus those in gametes and embryos should not disqualify such analyses as an apples and oranges comparison; indeed, it is not uncommon for courts to look across contexts to clarify the precise property interest involved.[30] Rather, these cases should be looked at broadly and comparatively to draw greater scrutiny to default rules conferring differential property interests in somatic cells and gametes/embryos.

Second, the cases pertaining to gametes and embryos place a higher premium on fulfilling the progenitors' intent at the time of removal. For example, the *Hecht* court found that Kane had sufficient interest in his sperm for it to constitute property in terms of conferring jurisdiction of the dispute to the Probate Court.[31] But much of this finding is driven by the court's acknowledgement that gametes are a special type of property that should be disposed of according to the donor's *intent*. In rejecting the trial court's order to destroy the sperm, the appellate court noted that Kane's "will evidences the decedent's intent that Hecht, should she so desire, is to receive his sperm stored in the sperm bank to bear his child posthumously."[32] The *Hecht* court relied heavily upon the test set forth by *Davis* to determine the dispensation of gametes when there is a dispute, where primacy is given to the "preferences of the progenitors."[33] Given the clarity of Kane's intent as expressed in his will, the *Hecht* court simply applied Kane's intent as articulated.[34]

---

30. *See, e.g.*, Janicki v. Hosp. of St. Raphael, 744 A.2d 963, 970-71 (Conn. 1999).

31. *Hecht*, 16 Cal. App. 4th at 850.

32. *Id.* at 850-51.

33. The full test elaborated by the *Davis* court states: "[W]e hold that disputes involving the disposition of pre-embryos produced by *in vitro* fertilization should be resolved, first, by looking to the preferences of the progenitors. If their wishes cannot be ascertained, or if there is dispute, then their prior agreement concerning disposition should be carried out. If no prior agreement exists, then the relative interests of the parties in using or not using the pre-embryos must be weighed. Ordinarily, the party wishing to avoid procreation should prevail, assuming that the other party has a reasonable possibility of achieving parenthood by means other than use of the pre-embryos in question. If no other reasonable alternatives exist, then the argument in favor of using the pre-embryos to achieve pregnancy should be considered. However, if the party seeking control of the pre-embryos intends merely to donate them to another couple, the objecting party obviously has the greater interest and should prevail." Davis v. Davis, 842 S.W.2d 588, 604 (Tenn. 1992).

34. "Given the procedural posture of this case, and the fact that, for purposes of addressing real parties' arguments, we are assuming that decedent intended to allow Hecht to use his sperm for posthumous artificial insemination, it is premature for us to apply the *Davis* test. At this point, the only issue which we address is whether artificial insemination with the sperm of a *decedent* violates public policy. There is nothing in *Davis* which indicates that such artificial insemination violates public policy." *Hecht*, 16 Cal. App. 4th at 859.

Intent arguably cuts the other way in cases dealing with somatic cells and tissues. In each of these cases, litigation arose out of plaintiffs' contention that their cells or tissues were being used in a manner inconsistent with their intent or expectations at the time they donated the biological materials. Yet, courts have consistently disregarded these claims in favor of conferring a controlling property interest to researchers once the cells or tissues have been excised—even when, as in *Moore*, the court simultaneously finds that the researcher breached his disclosure obligations in a manner that may very well have altered the patient's medical decision-making.[35] The *Moore* court specifically denied the conversion claim because they concluded that Moore "clearly did not expect to retain possession of his cells following their removal"[36] and that it would be inappropriate to acknowledge any lingering property interest that legitimizes the conversion claim since such recognition had not been granted by previous courts. Such a ruling would be contrary to California statutes that restrict patient interest in excised cells. Moreover, the Court also found that the patented cell line was sufficiently different to preclude Moore's claim to ownership. Thus, the Court disregarded Moore's intent and expectations as expressed in the suit. Viewed another way, they only looked at what his intent/expectations *could legally be* at the time the cells were excised rather than taking seriously his *actual* intent/expectations as expressed during the litigation. The *Greenberg* court also articulated this comparatively thin understanding of intent in the context of somatic cells and tissues, saying that the tissues in question were donated[37] and that, following *Moore*, researchers' use of the materials in a manner inconsistent with the donors' intent or expectations was irrelevant since they no longer had a legal interest in them.[38] Where *Greenberg* used the language of "donation" to characterize the transfer of property interest in somatic tissues and to curtail any serious consideration of the plaintiff's intent, *Catalona* embraced the language of "gift" to achieve similar ends:

> Our conclusion that the [research participants] intended to make gifts of their biological samples at the time of their donation is bolstered further by the language of the brochure, which characterized the [research participants'] donations as "a free and generous gift of [biological materials] to research that may benefit society." The brochure's acknowledgment that donated materials may be shared with non-WU researchers, without any further authorization

---

35. Though the *Moore* court did not allow Moore's conversion claim, they nonetheless held that "a physician who is seeking a patient's consent for a medical procedure must, in order to satisfy his fiduciary duty and to obtain the patient's informed consent, disclose personal interests unrelated to the patient's health, whether research or economic, that may affect his medical judgment." Moore v. Regents of the Univ. of Cal., 793 P.2d 479, 485 (Cal. 1990).

36. *Id.* at 488-89.

37. Greenberg v. Miami Children's Hosp. Research Inst., 264 F. Supp. 2d 1064, 1074 (S.D. Fla. 2003).

38. *Id.* at 1074 (citing *Moore*, 51 Cal.3d 120).

> needed from the [research participants], informed the [research participants] they would relinquish or abandon the right to designate the particular destination of their biological materials upon agreeing to participate in a medical research study. Such language, considered together with the consent form, cannot reasonably be characterized as reflecting the [research participants'] intention either to entrust their samples solely to Dr. Catalona or to transfer the samples in some legal form other than a gift.[39]

Thus, the potential for gametes and embryos to develop into autonomous human beings constitutes a legally significant default rule that amplifies donors' intent so as to acknowledge a property interest that exists after removal. The inability of somatic cells and tissues to generate human beings diminishes the seriousness in which courts consider donors' intent—even when the use of these materials is so egregiously inconsistent with their desires that it generates litigation—leading courts to effectively decline any property interests that individuals have to them or any derivative product. Therefore, the type of cell or tissue in question, in terms of its ability to create or become a human being, appears to be the most legally significant variable in how the court determines whether individuals maintain a residual property interest in extracorporeal materials. Individuals' diminished property interest in somatic cells and tissues and heightened property interest in gametes and embryos are justified by what is perceived to be a clear dividing line: the potential to become human life. In the next section, we discuss how the development of iPSCs poses serious challenges to this default rule.

II. INDUCED PLURIPOTENT STEM CELLS: WHAT THEY ARE, HOW THEY WORK, AND THEIR IMPLICATIONS FOR PROPERTY LAW

A. *A Brief Description of the Technology*

The biomedical promise and excitement surrounding stem cells lies in their ability to differentiate into different types of cells. Stem cells are unlike other cells in three regards:[40] (1) they are unspecialized (i.e. they do not have tissue-specific structures that limit them to only performing particular functions[41]);

---

39. Wash. Univ. v. Catalona, 490 F.3d 667, 674-75 (8th Cir. 2007). The court noted that the research participants' (RP's) options with regards to directing the use of their tissues were quite limited: "The RPs' subsequent rights to their biological materials were expressly limited to the option to discontinue participation in the study to avoid answering additional questions, donating more biological materials, or allowing their biological materials to be used for further research. Beyond these particular and limited rights, the RPs retained no greater interest with regard to their biological materials. Such rights cannot be equated with or interpreted to include the broad privileges or proprietary interests advocated by the defendants." *Id*. at 675.

40. *See generally* NAT'L INST. OF HEALTH, STEM CELL BASICS (2009), *available at* http://stemcells.nih.gov/info/basics/basics2.asp.

41. *Id.*

(2) they can divide and replicate for a long time; and (3) they can differentiate into specialized cells. Typically, stem cells come in two types: adult and embryonic. Adult stem cells are mature stem cells found in particular tissues (bone marrow, heart, etc.) of mature organisms and "are responsible for renewing and repairing the body's specialized cells."[42] Adult stem cells are *multipotent*, meaning that they ordinarily can only differentiate (become more specialized) into a limited number of cell types. Embryonic stem cells, which are coaxed from early stage human embryos, are unique in that they are *pluripotent*, meaning that they can differentiate into derivatives of any of the three main categories of human tissues—"ectoderm (skin, nerves, brain), mesoderm (bone, muscle), and endoderm (lungs, digestive system)"[43]—as well as germ cells that are the precursors for gametes.

Pluripotency is a key trait; it lays the foundation for the promise of regenerative medicine where patient-specific tissues can be developed to treat many illnesses. For example, one approach might involve treating heart disease by employing embryonic stem cells to develop patient-specific replacement cells that might be able to repair damaged heart tissue.[44] Similarly, embryonic stem cells might be used to develop patient-specific nerve tissue that might regenerate spinal cord tissues to help give movement back to paralyzed individuals.[45] Thus, many researchers believe that embryonic stem cells' pluripotency creates an avenue of research that is more promising than the limitations associated with adult stem cells, which do not exhibit this trait.

However, the promise behind embryonic stem cells is not without ethical controversy. Obtaining access to these pluripotent stem cells requires destroying embryos. Many people consider embryos to be no less a form of human life than an actual human being, making it ethically problematic to destroy one form of life to save or heal another. This has led to a form of stem cell politics that mirrors abortion politics, where competing definitions of when life begins often determines how individuals understand the legitimacy of this medical technique.

It is in this context that the 2007 discovery of iPSCs was heralded as a new technology that might resolve this ethical and political problem. By adding a handful of genes to somatic cells and providing the right laboratory environment, these cells can be "induced" or "reprogrammed" to revert back to a pluripotent state that exhibits the same characteristics as embryonic stem

---

42. Insoo Hyun, *Stem Cells*, *in* FROM BIRTH TO DEATH AND BENCH TO CLINIC: THE HASTINGS CENTER BIOETHICS BRIEFING BOOK FOR JOURNALISTS, POLICYMAKERS, AND CAMPAIGNS, 159, 159 (Mary Crowley ed., 2008).

43. *Id.*

44. *See generally* Siamak Davani et al., *Can Stem Cells Mend a Broken Heart*, 65 CARDIOVASCULAR RESEARCH 305, 310 (2005).

45. *See generally* M.A. Woodbury, *Hans Keirstead Can Make Mice Walk Again (and Humans Too?)*, ESQUIRE MAGAZINE, Nov. 17, 2009, *available at* http://www.esquire.com/features/best-and-brightest-2009/human-embryonic-stem-cell-research-1209.

cells—all without having to destroy an embryo.[46] These reprogrammed cells can, in effect, give rise to an entirely new organism from somatic cells—which has not yet been done with humans but was demonstrated by two separate teams in 2009 with mice.[47] Using somatic cells taken from the skin of adult mice, these researchers used a virus to inject four genes into mice cells, which reprogrammed the cells to a state of pluripotency, causing them to exhibit the same plasticity as embryonic stem cells. They were then implanted in the acellular surrounding material of a nonviable "tetraploid" embryo that had its own cells modified; the new embryo with the reprogrammed somatic cells then developed into new baby mice. iPSCs have also been used to reprogram mouse somatic cells into pluripotent cells that were then differentiated into precursor germ cells that were used in fertilization.[48] While these mouse experiments pertaining to reproduction via iPSCs have not yet been demonstrated with human somatic cells, there is growing evidence that human applications are feasible and a logical extension of these animal experiments.[49]

B. *Potential Applications of iPSC Research With Human Cells*

There are several potential applications of iPSC research with human cells that might directly implicate the property interests of individual cell and tissue donors. It has been estimated that there are over 270 million tissue samples stored in U.S. biobanks alone, with an additional 20 million samples being added every year.[50] While some of these biobanks contain eggs, sperm, and embryos, millions of these samples are ordinary somatic cells and tissues used by scientists to conduct research that may lead to new therapies and treatments. While there is a robust debate regarding biobank governance to manage their disposition and to protect patients from privacy intrusions or from other harms

---

46. *See generally* Takahashi et al., *supra* note 5; Yu et al., *supra* note 6.

47. *See generally* Lan Kang et al., *iPS Cells Can Support Full-Term Development of Tetraploid Blastocyst-Complemented Embryos*, 5 CELL STEM CELL 1, 1-4 (2009); Xiao-yang Zhao et al., *iPS Cells Produce Viable Mice Through Tetraploid Complementation*, 461 NATURE 86, 86-89 (2009).

48. Masanori Imamura et al., *Induction of Primordial Germ Cells From Mouse Induced Pluripotent Stem Cells Derived from Adult Hepatocytes*, 77 MOLECULAR REPROD. & DEV. 802, 808 (2010); Natalie de Souza, *Gametes from Stem Cells*, 8 NATURE METHODS 789, 789 (2011).

49. *See* Charles A. Easley IV et. al., *Direct Differentiation of Human Pluripotent Stem Cells into Haploid Spermatogenic Cells*, 2 CELL REPORTS 440, 443-44 (2012) (evidencing advances in using iPSC technology to turn human somatic cells into gametes); Rosa Silverman, *Scientists Create Sperm from Skin Sample*, UK TELEGRAPH, Aug. 29, 2012, http://www.telegraph.co.uk/science/science-news/9505267/Scientists-create-sperm-from-skin-sample.html. The study's lead author told the UK Telegraph "[n]o one has been able to make human sperm from pluripotent stem cells . . . in the lab, but this research indicates it might be possible."

50. Susanne B. Haga & Laura M. Beskow, *Ethical, Legal, and Social Implications of Biobanks for Genetics Research*, 60 ADVANCES IN GENETICS 505, 506 (2008).

that may lead to having their tissues identified,[51] the possibility that somatic cells in biobanks might be used for iPSC research raises important questions for property law that have not even been articulated yet alone addressed.

Although iPSC research has been touted as resolving debates regarding the ethics of destroying embryos to harvest pluripotent stem cells, some commentators have pointed to new ethical problems created by this emerging technology. One looming ethical issue is the notion of consent. It is common practice that individuals provide written consent to physicians and researchers to use their excised cell and tissue samples for research. However, patients and research participants often do this with an implicit understanding that there are certain biological limitations regarding what can be done with their somatic cells and tissues. The advent of iPSCs, however, suggests radically different possibilities—including the reprogramming of somatic cells into precursor germ cells that can mature into gametes that can be fertilized and grow into a living person.

Standard informed consent forms typically do not contemplate this possibility. To use somatic cells obtained under existing informed consent processes for iPSC research without raising the new possibilities to patients and research participants is ethically problematic for the obvious reasons that many people would object to their cells being used in a manner that may indiscriminately create new life that would be genetically related to them.[52]

Another concern is privacy. iPSC lines derived from a living individual contains genetic information about the donor and his/her relatives that may be sensitive. De-identification may not always be desirable in the context of iPSC research for both clinical and technical reasons.[53]

The third major ethical issue with iPSC research is that it is a relatively straightforward process that is not difficult to replicate.[54] In short,

---

51. *See generally* David E. Winickoff & Richard N. Winickoff, *The Charitable Trust as a Model for Genomic Biobanks*, 349 NEW. ENG. J. MED. 1180, 1182-83 (2003).

52. *See generally* Katriina Aalto-Setälä et al., *Obtaining Consent for Future Research with Induced Pluripotent Cells: Opportunities and Challenges*, 7 PLOS BIOLOGY 204, 207 (2009).

53. Amy Zarzeczny et al., *iPS Cells: Mapping the Policy Issues*, 139 CELL 1032, 1033 (2009). "One way for researchers to address these [privacy] concerns is to de-identify or anonymize the data at the time of donation. However, there are various problems with this approach. First, there are clinical, research, and policy reasons why anonymization (that is, de-linkage from identifiable information) may not be an ideal approach. For instance, future clinical applications (e.g. transplantation) may necessitate obtaining follow up information about the donor's health status. Second, with the advent of large-scale genome-wide association studies, technology now exists to detect a specific individual's single nucleotide polymorphism, even when de-identified and in a pooled data set."

54. David Cyranoski, *5 Things to Know Before Jumping on the iPS Bandwagon*, 452 NATURE 406, 406 (2008). Although the process is rather straightforward, iPS cells can be rather difficult to develop. Nature's David Cyranoski reports that "as simple as this procedure might seem, iPS cells are not easy to make. Kathrin Plath at the University of California, Los Angeles estimates that each of the reprogramming genes (she used six) has

reprogramming somatic body cells into a pluripotent state only requires inserting four genes in culture to trigger the process of de-differentiation that turns a regular somatic cell (e.g. skin) into a pluripotent state[55] that can differentiate into reproductive cells. Shinya Yamanaka, who led one of the first teams to discover iPS cells stated "we are presenting new ethical issues, maybe worse ones, because many people can do this—and without telling anybody."[56] As an editorial from *Nature* noted, "the facility with which iPS cells can be derived could make it easier to derive gametes from any person, living or dead."[57] This combination of (a) a transformative technology that is (b) straightforward to conduct with (c) widely accessible raw materials (somatic cells and tissues) stored by the millions in biobanks and (d) a legal regime whereby the progenitors of these raw materials have no legal interest in their own cells and tissues (while also remaining vulnerable to third party usage that could be quite damaging) might produce disputes and unwanted outcomes of an order of magnitude that was previously unimaginable. Put bluntly, any person who has had a tissue biopsy stored in a biobank may now, at least theoretically, become a genetic parent without their consent.

One particularly sensitive area of research is the use of iPSCs for reproductive medicine. The specter of this type of research draws attention not only to the fraught nature of iPSC research in the current bioethical and biomedical environments but also to its legally problematic dynamics. The advent of iPSCs implicate property interests in ways that obliterate the rationale for the existing default rules that apportion differential property interests according to cells' potential to create new human beings. For example, to the extent that iPSC research might be able to reprogram and differentiate somatic cells into reproductive cells that can become mature gametes, researchers may find iPSC-derived gametes "useful both for understanding gametogenesis and as a potential infertility treatment [whereby] gametes derived from iPS cells would have virtually the same DNA as the somatic cell donor."[58] A less scrupulous researcher could use iPSC methods and readily available somatic cells and tissues to pursue human cloning—even in jurisdictions that ban reproductive cloning by other means.[59] Beyond notions of privacy and consent,

---

only a 15% chance of making it into a given cell. Even if they all make it, the cell has only a 5% chance of being fully reprogrammed. The low efficiency presents a riddle for scientists, but with millions of cells available in a biopsy sample, it is not a roadblock."

55. Justin Lowenthal et al., *Specimen Collection for Induced Pluripotent Stem Cell Research: Harmonizing the Approach to Informed Consent*, 1 STEM CELLS TRANSLATIONAL MED. 409, 409 (2012). "Although recent evidence has tempered the hope that translating these technologies toward new therapies will be easy, there is great interest in using iPSC lines to advance translational goals. A broad range of human tissue types are currently being procured to facilitate the generation of iPSC lines."

56. Cyranoski, *supra* note 54, at 408.

57. *New Sources of Sex Cells,* 452 NATURE 913, 913 (2008)

58. Aalto-Setälä et al., *supra* note 52, at 206.

59. "In theory, injecting human iPSCs into a human tetraploid blastocyst could create

it is likely that scenarios such as these would be quite troubling to the somatic cell donor if they came to pass for the precise reasons that law has recognized a residual property interest in gametes and embryos: the potential to create new individuals intuitively gives donors a heightened interest in how their cells are used, lest a genetically related version of themselves be created without their knowledge or against their wishes. Yet, in the existing legal regime, donors of somatic cells and tissues would have no recourse to even prevent a scientist from pursuing the most questionable types of research with their biological materials.

III. NORMATIVE PROPOSALS

iPSCs pose a dramatic challenge to this default rule in property law, which recognizes an individual's property interests in excised gametes and embryos based on their "potential for human life" while refusing any corollary property interest in excised somatic cells. While research with iPSCs in humans is in its early stages, the developments with mice cells suggest that it may soon be possible for excised human somatic cells to be reprogrammed and differentiated into reproductive cells, which would have the very potential for life that courts have identified as a prerequisite for having a residual property interest in excised cells. Therefore, iPSCs can make it technically feasible for any regular somatic cell—from an individual's hair, skin, or Moore's storied spleen—to be reprogrammed in a manner that can create life. Somatic cells now arguably have the same potential to become new human beings that embryos and gametes do, but without a corresponding property interest to prevent misuse or to respect individuals' wishes concerning the disposition of such cells. While the development of iPSCs have led to important reconsiderations regarding informed consent[60] and intellectual property,[61]

---

a child who is a clone of the somatic cell donor and whose placenta comes from the donor(s) of the blastomeres. . . . [M]any current policies ban only human reproductive cloning by somatic cell nuclear transfer (SCNT)." Bernard Lo et al., *Cloning Mice and Men: Prohibiting the Use of iPS Cells for Human Reproductive Cloning*, 6 CELL STEM CELL 16, 16 (2010).

60. *See* Lowenthal et al., *supra* note 55. The development of iPSCs creates important questions regarding informed consent, i.e. how can a patient consent to donating tissues that may be subject to iPSC research when it is not yet entirely clear what that research might entail. It is interesting to note that while they do not speak to the exact issues discussed in this Article, existing regulations for human embryonic stem cells do not give donors downstream control over their usage. For example, the NIH Guidelines on Human Stem Cell Research only states that donors "should have been informed that they retained the right to withdraw consent for the donation of the embryo" and that they should also be informed that "the results of research using [human embryonic stem cells] may have commercial potential, and that the donor(s) *would not* receive financial or any other benefits from any such commercial development." NATIONAL INSTITUTES OF HEALTH GUIDELINES ON HUMAN STEM CELL RESEARCH, II(A)(3)(d)(iii), II(A)(3)(e)(vi) (2009) (emphasis added), *available at* http://stemcells.nih.gov/policy/2009guidelines.htm. Guidelines from the California Institute

scholars have not identified nor discussed the tensions created for property law. This section fills this void by first discussing the existing relevant literature. We then offer a three-part proposal that addresses the challenges posed by iPSCs for property law and then discuss concerns that may be raised by our recommendations.

A. *Existing Literature*

Legal scholars have not discussed the transformative challenges that iPSCs pose for property law. The scholarly literature regarding property interest in human cells and tissues has largely focused on exploring different theoretical or doctrinal bases to rethink the default rules established by *Moore*—i.e., that individuals do not have a property interest in their excised somatic cells. Scholars have been largely critical of *Moore's* holding and have offered alternative paradigms as solutions. For example, Robin Feldman has recently argued that "[i]t defies common sense to say that an individual lab can hold property rights in the tangible cells removed from a person's body while the person whose body supplied the cells cannot."[62] This intuitive understanding that *Moore* was incorrectly decided leads Feldman to urge courts to revisit the issue in light of the knowledge gleaned through decades of experience since the initial decision.[63] Pilar Ossorio shares a similar normative sensibility[64] in arguing that a formal transfer (gift, sale, etc.) should be needed before a researcher can use someone else's tissues for their own purposes.[65] Radhika

---

of Regenerative Medicine (CIRM)—the largest state level program for stem cell funding – note that "donor[s] will have no legal or financial interest in any commercial development resulting from the research" and that "a donor must be given the opportunity to impose restrictions on future uses of donated materials [but that] researchers may choose to use materials only from donors who agree to all future uses without restriction." CAL. CODE REGS. tit. 17, § 100100(b)(1)(I), § 100100(b)(2)(2012).

61. *See* Robin Feldman & Deborah Furth, *The Intellectual Property Landscape for iPS Cells*, 3 STAN. J.L. SCI. & POL'Y 16 (2010).

62. Robin Feldman, *Whose Body Is It Anyway? Human Cells and the Strange Effects of Property and Intellectual Property Law*, 63 STAN. L. REV. 1377, 1384 (2011).

63. "Our enthusiasm and appreciation for the miraculous advances of science should not blind us to the necessity of thinking through the interests of the people whose cells provide the raw materials, nor should it obviate the necessity of ensuring that those raw materials are properly obtained. Perhaps courts in the appropriate jurisdictions will feel moved to revisit these issues, now that we have decades of experience with this type of scientific research." *Id.* at 1385.

64. "If one accepts the proposition 'my body belongs to me,' then I think there is a strong argument that extracorporeal bodily materials should be considered, initially, the property of the person from whom it was derived. . . . Does changing the location of bodily material from within my body to outside my body change my property rights in that material? I do not think so." Pilar N. Ossorio, *Property Rights and Human Bodies*, *in* WHO OWNS LIFE?, 223, 234-35 (David Magnus, Arthur Caplan, & Glenn McGee eds., 2002).

65. Ossorio states that "the individual from whom bodily materials are derived [should be] the initial owner of those materials, and that legitimate transfers from them to scientists

Rao examines the overlapping legal regimes of property, contract, and privacy to tease out the incoherence that law applies to individuals' rights and relationships to excised cells and tissues. She suggests a model of property in the human body as stewardship, where individuals collectively possess themselves in a public trust, rather than being outright owners.[66]

Not all commentators find the default rules established by *Moore* and subsequent cases troubling. For example, Hakimian and Korn argue that the *Moore* regime of diminished individual property interest in excised cells is justifiable:

> Because the benefits of medical knowledge derived from tissue research potentially accrue to all individuals and future generations, society may justify the expansive use of these valuable resources based on the principle of justice. Human tissue specimens are a unique and irreplaceable research resource, and society's strong interest in the advancement of medical knowledge deserves a coherent and internally consistent legal, regulatory, and ethical framework to govern specimen use.[67]

Similarly, Russell Korobkin argues that *Moore* correctly established an important default rule—one based on contract rather than property—of no compensation for transactions pertaining to human tissues.[68] Nevertheless, much of the literature surrounding *Moore* has expressed dissatisfaction with both the outcome and default rules established by the case, which has led to proposals for alternative models. For example, Charlotte Harrison has suggested a hybrid approach that falls between notions of altruism and private property, such that there would be "a general rule of donation for research tissue at the time it is acquired" which would "provide an objective, non-market mechanism for compensation after research use for unusual cases in which samples prove to have significant commercial utility."[69] Donna Gitter

---

must take place before scientists can rightfully possess, use, or sell those materials, or exclude others from doing so." *Id.* at 241.

66. Radhika Rao, *Genes and Spleens: Property, Contract, or Privacy Rights in the Human Body*, 35 J.L. MED. & ETHICS 371, 379-80 (2007). *See also* Rao, *supra* note 8.

67. Rina Hakimian & David Korn, *Ownership and Use of Tissue Specimens for Research*, 292 JAMA 2500, 2504 (2004).

68. Korobkin notes "[i]n the twenty-first century, biotechnology is becoming increasingly important in medical research. If biomedicine is able to fulfill the hopes of the scientific community by creating a new paradigm for the treatment of disease – one in which biological agents regenerate diseased or dead tissues – disembodied tissues could become the cures for a variety of ailments. It is likely that scientists around the world will need a tremendous amount of human tissues of all types just to mount the research effort, regardless of whether the promise is ever actually fulfilled. In the new era of biomedical technology, it is critically important for the law to facilitate tissue transactions efficiently. This, in turn requires understanding and embracing the underlying wisdom of *Moore*." Russell Korobkin, *"No Compensation" or "Pro Compensation":* Moore v. Regents *and Default Rules for Human Tissue Donations*, 40 J. HEALTH L. 1, 27 (2007).

69. Charlotte H. Harrison, *Neither* Moore *nor the Market: Alternative Models for Compensating Contributors of Human Tissue*, 28 AM. J.L. & MED. 77, 78 (2002).

takes a different approach in proposing that Congress recognize individuals' enduring property interest in excised human tissues by enacting legislation that not only allows research participants to sell human tissues for research purposes but also expressly provides for a tort of conversion if researchers wrongly use such materials.[70]

Despite these wide ranging perspectives, there has been no conversation in the literature—legal, biomedical, or otherwise—on how the development of iPSCs significantly complicates the existing conversation on property interests in excised human cells and tissues.[71] What is needed is a discussion of how iPSCs might shift the empirical footing underneath these theoretical, policy, and doctrinal conversations by making the existing logic behind the differential treatment of somatic cells and gametes/embryos largely incoherent. Prior to the development of iPSCs, the default rule regarding individuals' property interests in excised cells worked reasonably well by efficiently giving scientists access to research materials without exposing donors to the risk that their cells could be used by third parties to create new humans with a genetic tie to them. The default rule acts as a firewall in acknowledging a residual property interest in gametes and embryos outside of the body so that progenitors can control their disposition. But iPSC research disrupts the protections afforded by this default rule by giving somatic cells the potential to create human life without extending a residual property interest to donors so that they may limit the use of somatic cells for reproductive purposes. A model is needed that allows the law to co-evolve with technology in a manner that is least disruptive to the existing default rules. Yet it is also important to maintain the sensibilities of promoting efficient research with somatic cells and tissues while also giving donors a heightened property interest in excised cells with reproductive abilities. Here, we propose a three-part approach.

B. *A Three-part Proposal*

1. *Self-imposed moratorium on using iPSCs for human reproduction*

To address the jurisprudential instability in property law created by iPSCs, we first argue that the research community should implement a self-imposed moratorium on applications of iPSC research that pursue human reproduction

---

70. Donna M. Gitter, *Ownership of Human Tissue: A Proposal for Federal Recognition of Human Research Participants' Property Rights in Their Biological Material*, 61 WASH. & LEE L. REV. 257, 339-41 (2004).

71. Zarzeczny et al. come the closest to raising these issues (albeit in a non-legal journal article) by discussing the possibility of "reach through rights" stemming from iPSC research, whereby donors may have a controlling interest in the use of cell lines resulting from their tissues. However, this does not speak to the issue identified in this Article, i.e. that iPSCs complicate existing default rules pertaining to individuals' property interest in excised biological materials. *See* Zarzeczny et al., *supra* note 53, at 1034.

with somatic cells from individuals who have not specifically given informed consent for their biological materials to be used in this manner. We are not the first to suggest that certain types of iPSC research should not proceed. Bernard Lo et al. have argued for a moratorium in relation to the potential use of iPSCs for reproductive cloning purposes, citing both safety and ethical concerns.[72] However, the manner in which iPSC research potentially destabilizes the existing default rules pertaining to property law suggests that this provides another important reason to stop human reproductive applications with somatic cells from individuals who have not specifically consented to this use until further legal and policy deliberations can take place.

We do not suggest that iPSC research pertaining to animal models should stop. Nor do we necessarily suggest that all iPSC research with human cells should cease. We do, however, find the safety and ethical concerns pertaining to human reproductive cloning raised by Lo and his colleagues to be quite persuasive. Accordingly, we specifically tailor this first part of our proposal to the indeterminate property interest raised by using somatic cells from non-consenting individuals for iPSC research,[73] to the extent that this technology may give somatic cells the same potential to become human life as gametes and embryos, where progenitors retain a property interest after excision. We believe that the scientific community should take a leadership role in assessing and resolving the tensions regarding the ownership of excised biological cells created by iPSC research. This should start with a self-imposed moratorium on the most questionable aspect of this technology—research that results in the creation of reproductive cells that have the potential to create life and whose donors did not specifically consent to this purpose. While some may argue that a self-imposed moratorium would not be effective in stopping this research, we believe that research institutions and funding organizations can provide the appropriate combination of incentives and disincentives to require clear consent from donors that unequivocally demonstrates their understanding that their cells will be used for iPSC research and that they appreciate the full range of potential outcomes. For the scientific community to not take the lead with this

---

72. "There continue to be compelling safety reasons to ban human reproductive cloning. Existing laws and professional guidelines should be carefully revised to cover tetraploid complementation with iPSCs and other technologies in addition to [somatic cell nuclear transfer], thereby broadening the ban on attempts at reproductive cloning to existing and future technologies." Lo et al., *supra* note 59, at 20.

73. Aalto-Setälä et al. discuss the importance of consent in iPSC research, noting that "iPS cells are an exciting new approach to developing pluripotent stem cell lines that are genetically identical to people with known phenotypes. While they avoid the ethical issues inherent in embryonic stem cells, they do raise some ethical concerns regarding consent for future research. Obtaining consent for fundamental downstream research with iPS cells, together with offering the options of allowing recontact by researchers and giving permission for additional sensitive types of future research, will show respect for somatic cell donors, promote public trust in stem cell research, and allow optimal use of scientific discoveries." Aalto-Setälä et al., *supra* note 52, at 207.

rather nominal proposal risks creating the perception that they are more interested in preserving the unadulterated commercial potential of iPSC research rather than respecting the law's concern with preserving individuals' property interests in excised cells with the potential to create human life. This perception, if fostered by the research community's reluctance, may lead to distrust among potential research participants. This may stifle development of human applications of this technology as well as others. We believe that this moratorium should only be in place for a limited period, at least until the proposal's second part is in place.

2. *Legislative action*

Second, we argue that Congress should consider legislation that acknowledges a property interest in excised somatic cells that vests back to the progenitor once these cells are reprogrammed to a state of pluripotency and differentiated into reproductive cells. Since the characteristic that destabilizes the existing default rules is induced pluripotency that can lead to the creation of reproductive cells—the outermost point at which a somatic cell has the potential to create an independent human organism—it makes clear legal and scientific sense to protect individuals' property interests at this point like any other excised reproductive cell with this potential. To the extent that this has been the key rationale for the differential property interests given to these cells, Congress can quickly and logically bring uniformity to this area of property law by extending existing property interests in gametes and embryos to somatic cells at the moment they cease to function as ordinary body cells and obtain reproductive capacity through the processes of induced pluripotency and differentiation. Such Congressional action would provide the most protection to potential research participants and encourages iPSC researchers to specifically seek informed consent from the progenitors of somatic cells and tissues.

3. *Judicial action*

Lastly, in light of these scientific developments and proposed Congressional actions, we argue that courts should distinguish *Moore* to clearly identify a property interest in excised somatic cells that are reprogrammed and differentiated into reproductive cells through the processes of induced pluripotency. This will add further consistency to the existing jurisprudence by demonstrating that the acknowledgement of a property interest in somatic cells that have been reprogrammed and differentiated into reproductive cells is not a departure from the existing default rules. Rather, it is consistent with the court's longstanding emphasis on protecting individuals' interests in excised biological materials that have the potential to create autonomous human beings. To the extent that iPSCs may not be the last technique to confer pluripotency to somatic cells, such judicial pronouncements can create a stable paradigm that protects individual property interests. It can also provide a predictable research

environment for researchers seeking consistency in how to handle excised cells and tissues and how to assess the commercial viability of their research.

C. *Three Possible Objections*

There are at least three different concerns to this proposal that merit discussion. First, there is the argument that the proposed approach is overly broad. Many in the research community may argue that iPSC research can go forward without any type of self-imposed moratorium, Congressional action, or judicial affirmation as scientists can be trusted to not abuse donors' samples, the public's trust, or existing ethical guidelines. However, we argue that this issue is too important to allow human applications of iPSC research to exist within the status quo default rules as this may lead to the existing paradigm's perversion. Scientists are given a default property interest in excised somatic cells because they are ostensibly of little use to most individuals. Yet, this rule can now "lock out" individuals from any property interest in their own cells once they are reprogrammed and differentiated into reproductive cells with the potential to become life. Our proposal is a logical extension of the current default rules. We simply argue that law should extend the same property interest to somatic cells that are conferred to gametes and embryos once these somatic cells are reprogrammed and differentiated to exhibit traits that are substantially similar to those of ordinary reproductive cells—and therefore have the potential to become new human beings.

A second concern is that the vesting feature of our proposal—where Congressional action would confer a property interest to somatic cell donors once their cells have been reprogrammed and differentiated into reproductive cells—is too burdensome to be functional; it would require each and every somatic sample to be tracked in light of the possibility that it might be used for iPSC research. However, this concern overstates the issue. Our proposal is designed to incentivize the creation of separate biobanks where somatic cell and tissue donors have specifically consented to the use of their samples for iPSC research. Thus, the point of Congressional action of this nature is for the market and scientific community to develop a workable model outside of requiring the tracking of all banked biological samples on the off chance that they may be useful to iPSC researchers. The creation of separate biobanks with donors fully consented about the prospects of their samples being used for iPSC research can create new efficiencies in that researchers can have confidence that they are using samples in an ethically appropriate manner and that donors are fully aware of the potential use of their samples.

A third objection that may be raised by this proposal is that it introduces unnecessary instability in this area of law and threatens the established property interests conferred to researchers through the default rules established by *Moore*, *Davis*, *Hecht*, and other cases. Put differently: leave well enough alone. This possible objection both understates and overstates the significance of the proposal being made. It understates it in that what might be considered "well

enough" for the research community may not be the case for individual cell and tissue donors. Yet, it also overstates things in that the proposal does not take anything away from researchers, nor does it change the existing default rules. Indeed, the default differential property interests are preserved. Our proposal simply addresses the new middle ground created by somatic cells that are reprogrammed to a pluripotent state, and conservatively errs on the side of individual donors to the extent that reprogrammed pluripotency and subsequent cell differentiation can give somatic cells the very potential to become life that the existing default rules acknowledge as a key reason for extending property interest to donors. Rather than destabilizing the existing regime, this proposal respectfully adheres to the preferences and concerns embedded in the existing default rules.

CONCLUSION

Discussions concerning property rights in the body have been ongoing for hundreds of years. *Moore v. Regents of the University of California* was a watershed moment in establishing the foundation for the current default rules concerning the type of property interests individuals have in excised somatic cells, with subsequent cases adding the bricks and mortar. This Article has shown the evolving nature of these property interests, especially in the context of new genetic technologies such as iPSCs. By giving somatic cells the potential to become life, iPSCs challenge the entire modern regime concerning property interests in excised human cells and tissues and create new opportunities for rethinking this area of law. Our proposal provides a coherent framework to deal with the implications iPSCs might have for property law without upending this entire jurisprudence. By granting the same property interest in gametes and embryos to somatic cells that are reprogrammed and differentiated into reproductive cells, the current legal framework, policy preferences, and expectations among scientists and research participants are preserved. In addition, this recognition of property interests creates an incentive for the iPSC research community to develop separate biobanks composed of samples from donors who fully consent to their cells and tissues being subjected to this new technology. This provides a viable means by which the integrity of both the iPSC research agenda and cell and tissue donors can be preserved.

STANFORD TECHNOLOGY LAW REVIEW
VOLUME 16, NUMBER 1 FALL 2012

NEGOTIATING CLOUD CONTRACTS: LOOKING AT CLOUDS FROM BOTH SIDES NOW

W. Kuan Hon, Christopher Millard & Ian Walden*

CITE AS: 16 STAN. TECH. L. REV. 79 (2012)
http://stlr.stanford.edu/pdf/cloudcontracts.pdf

ABSTRACT

Contract terms for cloud computing services are evolving, driven by users' attempts to negotiate providers' standard terms to make them more suitable for their requirements, as well as market developments, particularly among cloud integrators. This Article, drawing on sources including interviews with cloud computing providers, users and other market actors, is the first in-depth research into how cloud contracts are negotiated. In particular, we have focused on instances where users have requested changes to providers' standard terms, and the extent to which providers agreed to those changes. We have found that the terms that generated the most negotiation were provider liability, service level agreements, data protection and security, termination rights, unilateral amendments to service features, and intellectual property rights. Changes to providers' standard terms are likely to filter down from large deals where users have negotiated amendments, and filter up from regulatory action affecting the consumer market. This Article suggests a multiplicity of approaches are emerging, rather than a de facto 'cloud' model, with market participants developing a range of cloud services with different contractual terms, priced at different levels, and embracing standards and certifications that aid legal certainty and compliance, particularly for small

* The authors are, respectively, a consultant to the QMUL Cloud Legal Project, Professor of Privacy and Information Law at the Centre for Commercial Legal Studies, Queen Mary, University of London, and Professor of Information and Communications Law at the Centre for Commercial Legal Studies, Queen Mary, University of London. This Article was written as part of the QMUL Cloud Legal Project at the Centre for Commercial Law Studies, Queen Mary, University of London. The authors are very grateful to their sources—most of which wished not to be named—for their participation and assistance in this research. The authors are grateful to Microsoft for generous financial support making this project possible. Views herein, however, are solely the authors'.

and medium-sized enterprise users.

INTRODUCTION

The top ten strategic technology trends for 2013 include, are based on, or incorporate cloud computing; those trends include personal cloud, hybrid in-

formation technology and cloud computing, cloud-based analytics, in memory computing and integrated technology ecosystems.[1] However, the cloud market is still relatively immature. The state of providers' standard contract terms seems to reflect this. In a 2010 survey of some thirty standard terms of cloud providers,[2] most terms surveyed were found, unsurprisingly, to be weighted in favor of the provider, and many were potentially non-compliant, invalid, or unenforceable in some countries.

This Article summarizes our subsequent qualitative research into negotiated cloud contracts, where users have requested changes to providers' standard cloud contract terms and the extent to which providers agreed those changes. Based on our research, users consider that providers' standard contract terms or offerings do not sufficiently accommodate customer needs in various respects. The top six types of terms most negotiated, according to our sources, were as follows, with the third and fourth issues ranking roughly equally in importance (depending on type of user and service):

1. exclusion or limitation of liability and remedies, particularly regarding data integrity and disaster recovery;
2. service levels, including availability;
3. security and privacy,[3] particularly regulatory issues under the European Union Data Protection Directive;[4]
4. lock-in and exit, including term, termination rights, and return of data on exit;
5. providers' ability to change service features unilaterally; and
6. intellectual property rights.

This Article proceeds as follows: in Part I we describe our research methodology and the scope of our analysis. Then, in Part II, we outline providers' perspective on cloud contract terms. Next, in Part III, we discuss users' perspective on cloud contracts, including factors that influence why users request certain terms and why providers may agree to them, including the role of systems integrators. In the course of that analysis we discuss in detail the key types

---

1. *Gartner Identifies the Top 10 Strategic Technology Trends for 2013*, GARTNER NEWSROOM (Oct. 23, 2012), http://www.gartner.com/it/page.jsp?id=2209615

2. Simon Bradshaw, Christopher Millard & Ian Walden, *Contracts for Clouds: Comparison and Analysis of the Terms and Conditions of Cloud Computing Services*, 19(3) INT'L J.L. & INFO. TECH. 187 (2011).

3. *See* Cloud Industry Forum, *Cloud UK: Paper One Adoption and Trends 2011* (2011) [hereinafter *CIF1*], *available at* http://www.cloudindustryforum.org/downloads/whitepapers/cif-white-paper-1-2011-cloud-uk-adoption-and-trends.pdf (62% of enterprise decision-makers polled cited data security, and 55% data privacy, as their biggest concerns with cloud adoption); *cf.* Cloud Industry Forum, *Cloud UK: Paper Four Cloud Adoption and Trends for 2012* (2011) [hereinafter *CIF4*], *available at* http://www.cloudindustryforum.org/downloads/whitepapers/cif-white-paper-4-cloud-adoption-and-outlook-for-2012.pdf (these figures were 64% and 62% respectively, in a Cloud Industry Forum poll conducted one year prior).

4. Council Directive 95/46/EC, on the Protection of Individuals with Regard to the Processing of Personal Data and on the Free Movement of Such Data, 1995 O.J. (L 281) 31.

of terms negotiated. Finally, we conclude with how cloud contracts are likely to develop as the cloud market matures.

I. METHODOLOGY AND SCOPE

This Article offers a qualitative analysis of negotiations of cloud computing contract terms[5] over a particular time period. Our sources discussed specific deals they had been involved with, and their experiences and personal views. Some providers and law firms provided generalized experiences regarding users and clients. From our research, this Article identifies themes regarding cloud contracts that we believe are reasonably representative of the key issues of concern in the cloud market at this relatively immature stage of cloud adoption.

Our sources comprised reports of experiences at various public conferences and seminars, or in informal discussions with cloud actors, plus detailed confidential interviews[6] with over twenty organizations[7] of at least an hour each, conducted between December 2010 and early 2012. Our interviewees included: cloud providers, such as United Kingdom-based or global software-as-a-service (SaaS), platform-as-a-service (PaaS) and infrastructure-as-a-service (IaaS) providers, including integrators, and European Union and non-European Union telecommunications providers; cloud users, such as businesses serving consumers, financial services businesses, and United Kingdom public sector and educational organizations; and other cloud market actors, such as law firms and insurance industry firms.[8]

This Article analyzes only contracts between cloud users and providers of SaaS, PaaS or IaaS *services* to users, where providers could include integrators.[9] We do not analyze end-user software licenses relating to cloud infrastruc-

---

5. In discussing negotiated contracts with our sources, we relied solely on the information these sources provided regarding their contracts. Generally, we were not able to verify terms as stated, since contracts were confidential to the parties. Not all types of terms were discussed with every source, as some users had not yet reached the stage of needing to scrutinize certain terms (e.g., ongoing audit rights).

6. Most interviewees were United Kingdom-based, and from legal rather than information technology departments. In some organizations, we spoke to both legal and technical experts.

7. For qualitative research purposes, fifteen to twenty interviews generally is considered adequate to obtain a good range of views before reaching "saturation," the point where no new information or themes are observed in the data and the same information starts recurring. *See, e.g.*, Greg Guest, Arwen Bunce & Laura Johnson, *How Many Interviews Are Enough? An Experiment with Data Saturation and Variability*, 18 FIELD METHODS 59 (2006). We did not attempt any quantitative surveys, although we reference certain surveys carried out by others, such as the Cloud Industry Forum.

8. We do not provide more detailed breakdowns, as that could identify those who agreed only to be interviewed anonymously. Mention of an organization's name in this Article does not imply its participation in a CLP interview.

9. Several entities may be involved in providing a single cloud service, including

ture or applications running on cloud infrastructure. In cloud computing, users obtain services, which may include rights to run software owned by the provider (or licensed by it from the rights owner for the purpose of providing services to third parties who, in using the provider's services, run such software in the cloud). However, cloud users procure primarily services, not licenses.

The possibility of chains of services and providers (not always known to users)[10] means users may rely on multiple parties, with multiple possible points of failure. Cloud users generally also depend on internet connectivity—usually involving contracts with telecommunications providers.[11] We do not discuss such contracts, contracts for consultancy or advisory services for users' adoption of cloud computing (sometimes forming part of a larger package), or contracts for supporting services for, or working with, the primary cloud service.

Also, while cloud users may have their own individual end users (e.g., employees or customers using the service procured) we focus only on the cloud user's relationship with its cloud provider, not its own end users. The sole exception involves integrators, where we discuss wholesale contracts with their providers, where they are users, or their end user customers, where they are providers.

II. CLOUD PROVIDERS' PERSPECTIVE

The starting point for cloud contracts is usually the providers' standard terms and conditions, which are provider-favorable and, generally, are designed for high-volume, low-cost, standard, commoditized services on shared multitenant infrastructure. However, as many providers' standard terms are not suitable to accommodate enterprise users' requirements, cloud users have sought changes to make the terms more balanced and appropriate to their own circumstances. It appears that there has been some movement in this direction, particularly for large users. Nevertheless, our research indicates that some providers'

---

suppliers of hardware, software, and data center services. There may be chains of contracts between them. Contractual requirements may vary with the cloud "stack" component under consideration. *See* Bradshaw et al., *supra* note 2, at 11; W. Kuan Hon, Julia Hörnle & Christopher Millard, *Data Protection Jurisdiction and Cloud Computing—When are Cloud Users and Providers Subject to EU Data Protection Law? The Cloud of Unknowing Part 3*, 26 INT'L REV. L. COMPUTER & TECH. (2012), *available at* http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1924240.

10. *See* W. Kuan Hon, Christopher Millard & Ian Walden, *Who is Responsible for 'Personal Data' in Cloud Computing?—The Cloud of Computing Part 2*, 2 INT'L DATA PRIVACY L. 3 (2012), *available at* http://idpl.oxfordjournals.org/content/2/1/3; W. Kuan Hon & Christopher Millard, *Data Export in Cloud Computing—How can Personal Data be Transferred Outside the EEA? The Cloud of Unknowing Part 4*, 9 SCRIPT-ED 25 (2012), *available at* http://script-ed.org/wp-content/uploads/2012/04/hon.pdf.

11. Indeed, some sources felt that the biggest issue for cloud, although not discussed much, was responsibility for the underlying network, and who takes the risk of its failure (sometimes, integrators).

negotiations are very process-driven, particularly at the lower price end of the market, where providers seemed unable or unwilling to accommodate differences such as corporate structures entailing (for users) separate localized contracts for non-United States affiliates.

III. CLOUD USERS' PERSPECTIVE

Before discussing contract terms changes that users have requested, and areas where they have succeeded, we first discuss why they may want changes, and factors that may influence whether providers accede. Drivers for users to seek changes to providers' standard terms may be internal or external. Internal reasons obviously include commercial issues such as required high service levels for mission-critical services, and allocation of risk between user and provider (particularly provider liability). External drivers are chiefly regulatory, namely, the need to comply with laws and regulations, including regulatory action. However, other external drivers are possible, such as insurers, whose role in the evolution of the cloud market is likely to increase.[12] Insurers may influence terms by, for example, insisting on certain certifications before agreeing to insure services—although according to one specialist United Kingdom cloud insurer, more providers than users were insuring, mainly for liability and errors and omissions.

Factors affecting development of cloud contract terms will be led by both demand and supply. More customer-friendly terms are being demanded by large users such as government, and being offered or agreed by integrators, smaller or specialist providers, and market entrants, thus making contract terms a source of competitive advantage for providers. As such, we believe that standard terms will improve from users' perspective as the market develops further. Our findings tie in with a survey for the Cloud Industry Forum of 450 senior information technology and business decision makers in enterprises, small and medium-sized businesses, and public sector organizations, published in early 2011. This found that, while forty-eight percent of organizations polled were already using cloud, only fifty-two percent of those (particularly larger organizations) had negotiated their contracts, with forty-five percent stating that they had no opportunity to negotiate (click-through terms, discussed below). In other words, according to that survey, about half of such users' cloud contracts were negotiated.[13]

---

12. Currently, many insurers view cloud as a form of outsourcing, which may already be within the scope of policies covering outsourcing.

13. Cloud Industry Forum, *Cloud UK: Paper Three – Contracting Cloud Services: A Guide to Best Practices* (2011) [hereinafter *CIF3*], *available at* http://www.cloudindustryforum.org/downloads/whitepapers/cif-white-paper-1-2011-cloud-uk-adoption-and-trends.pdf.

A. *Risk Management, Internal Controls, Governance, and Awareness: The Click-Through "Trap"?*

Many providers' roots lie in click-through web services offered to consumers or small and medium-sized enterprises, where users are presented with providers' standard contract terms, and "click through" to accept the terms, without any opportunity for negotiation. With many services, the only additional step is for users to enter credit card details, whereupon they may start using the service immediately. This history seems reflected in contractual terms and sign-up procedures for cloud services generally. Cloud providers' contracts are often "click-through," as the nature of cloud services enables use of a click-through consumer-based distribution model, and some providers deliberately choose that model for cloud. Other providers maintain generic click-through online terms for "self-service" customer-managed services, aimed at smaller or trial users, but offer (possibly regionalized) framework or master agreements for larger users enabling specific services to be purchased online.

For providers, click-through can eliminate negotiation costs and may reduce legal liabilities and other risks.[14] Some users noted that some providers, even large ones, did not have sufficient in-house legal resources to deal with users' requests to change terms, which might be another reason why they refuse to negotiate. One IaaS provider noted that they had never had a customer try to negotiate its standard click-through terms. Another provider was considering moving to 'click-through only' specifically to avoid the cost and time of negotiating terms.

However, the "click-through" model poses risks for users, such as fostering a bypass, sometimes deliberate, of institutional procurement processes.[15] A survey in 2010 found that fifty percent of information technology and information technology security specialists (forty-four percent in Europe) were not aware of all cloud computing resources deployed in their organizations.[16] Of

---

14. However, non-negotiated standard terms may be unenforceable in some circumstances, even against businesses; *see* Unfair Contract Terms Act, 1977, c. 50 (U.K.). A future CLP article will discuss consumer protection law issues.

15. Possibly encouraged by some traditional in-house information technology functions' perceived lack of responsiveness, coupled with the desire of those concerned to deploy quickly without the perceived delays of dealing with internal information technology, security or legal functions. *See, e.g.*, David Linthicum, *3 Reasons Cloud App Development is Taking Off*, INFOWORLD (Oct. 20, 2012), http://www.infoworld.com/d/cloud-computing/3-reasons-cloud-app-development-taking-176439 ("Employees frustrated with the wait for corporate IT [this is a direct quote, so reinstate the abbreviation here?] to solve a business problem simply hire their own developers and use PaaS as a cheap way to get their applications built, tested, and deployed."). Perhaps also, users are habituated to "clicking through" to agree terms automatically, from websites' standard processes for online sales of consumer products or services. In this sense, click-through for cloud contracts reflects the influence of consumer distribution models.

16. Larry Ponemon, PONEMON INST., SECURITY OF CLOUD COMPUTING USERS: A STUDY OF PRACTITIONERS IN THE U.S. AND EUROPE 6 fig.7 (2010), *available at*

decision makers accountable for cloud services, a "surprisingly high" number responded "don't know" to several key questions in a 2011 survey.[17]

While click-through may enable more efficient and flexible provisioning of information technology services, users' risk exposures may be affected. As certain German data protection regulators noted,[18] click-through's speed and ease means some customers may be tempted to accept providers' standard terms online, in order to start using the desired service quickly, without considering fully the nature or effect of those terms, or going through their organization's standard procurement procedures (such as legal or data protection review or security or other risk assessments).

It is not unknown for users to discover their employees had subscribed to cloud services on providers' standard terms, and then attempt to negotiate more acceptable terms. One SaaS provider noted that sometimes users were unaware of other internal departments using its service; only when the user's information technology department discovered the position was the user able to consolidate services. Some cloud services were even being used without any written contract, whether because they were free or being trialed on a pilot or test basis.[19]

Perhaps some employees bypass internal procurement procedures in situations where services are free of charge, at least for basic services. However, "free of charge" or "low cost" does not necessarily mean "free of risk" or "low risk." Legal, regulatory or reputational risks may exist. This is particularly true if data involved are not "fake" test data but constitute "real" data, perhaps even confidential or personal data. Furthermore, organizations may still be charged for essential supporting services or "extras" beyond the "free" component, for example Postini spam filtering for Google Apps SaaS.

Even where contracts were negotiated, some users' lawyers commented that their involvement had not been sought at an early enough stage to influence negotiations: "Business or procurement people negotiate without lawyers in the room; eventually the contract gets to legal staff, but not soon enough!" It is difficult for a user's lawyers to change contract terms when they are told that the contract has already been agreed upon, the service is going live shortly, and

---

http://www.ca.com/~/media/Files/IndustryResearch/security-cloud-computing-users_235659.pdf

17. *CIF3*, *supra* note 13, at 3-4 (polling some 450 decision makers).

18. DER ARBEITSKREISE TECHNIK UND MEDIEN DER KONFERENZ DER DATENSCHUTZBEAUFTRAGTEN DES BUNDES UND DER LÄNDER, ORIENTIERUNGSHILFE – CLOUD COMPUTING (2011) § 4.1.2 (Ger.), *available at* http://www.datenschutz-bayern.de/ technik/orient/oh_cloud.pdf [hereinafter DATENSCHUTZBEAUFTRAGTEN].

19. *See, e.g.*, Letter from Chelsea and Westminster Hosp., Aug. 3, 2011 (on file with author). In response to a freedom of information request regarding Chelsea and Westminster Hospital's tests of cloud computing for healthcare involving Flexiant, the Data Capture and Auto Identification Reference Project (DACAR) based at Chelsea and Westminster Hospital, and Edinburgh Napier University, the hospital replied, "There is no contractual agreement between Flexiant and Chelsea and Westminster Hospital Foundation Trust at this time."

to "just do a quick review."

Not involving lawyers in transactions from the outset may, in some organizations, be prevalent in relation to other types of contracts too. But while attempts to circumvent "the lawyers say 'no'" may speed up deals, they may also expose organizations to liability and other legal risks. Sometimes, users' lawyers were even asked to draft or review cloud contracts without being told the service's nature or purpose. This also poses risks: IaaS is quite different from SaaS, while a SaaS data storage service for managing an individual customer's personal data should be approached differently from a service intended only to record meteorological readings.

Providers have greater control in the cloud, particularly with SaaS, with correspondingly reduced user control, as illustrated by publicity regarding limitations in Microsoft Office 365's "P" Plan, where users were restricted in how many e-mails they could receive in twenty-four hours.[20] This provider constraint, which would not be possible with traditional e-mail applications installed on internal servers, illustrates the need for users to scrutinize contractual terms before signing up. However, procedurally, sometimes providers make terms available only late in the sign-up process, which may make it harder for users to compare different offerings at an early stage.

The above suggests that there may be internal governance and cultural issues which some organizations need to consider and address as part of their overall risk management strategy, both internally and in relations with providers. Scrutiny of procedures and practices, including training, may be in order. Otherwise, some organizations may find themselves committed[21] to cloud contracts on terms unfavorable to them, exposing them to possible legal risk, including breach of legal or regulatory obligations which may subject them to civil or even criminal liability.

Some users, such as News International, flag credit card charges by employees signing up for Amazon Web Services IaaS and the like, partly to secure better block pricing.[22] However such flagging may not be possible with sign-ups for free services. Another user, which is also a provider, has a policy banning employees from clicking through on standard term cloud contracts.

There also seems to be an educational issue for users' lawyers. Some providers' lawyers pointed out that users' lawyers sometimes raised points on providers' standard contract terms for the "wrong reasons," such as "going to town" to justify their fees! The most important reason was lack of understand-

---

20. Ed Bott, *Small Businesses, Beware the Office 365 Fine Print*, ZDNET (Oct. 21, 2011, 3:00 AM GMT), http://www.zdnet.com/blog/bott/small-businesses-beware-the-office-365-fine-print/4151.

21. Assuming the employees concerned at least have ostensible authority to bind the organization, which seems likely.

22. Paul Cheesbrough, Presentation at Amazon Summit: AWS Use at News International (June 14, 2011), *available at* http://plugyourbrand.com/AWS_Summit_2011/#trail_head/presentations.

ing amongst such users' lawyers about cloud computing and what the user was buying. Many users' lawyers' approached cloud contracts as software licenses or technology acquisitions, rather than as contracts for services. Others treated them as classic information technology outsourcing, without taking proper account of cloud computing's unique features. One SaaS provider found the largest users negotiated the least, perhaps because the deal was relatively small for them, or perhaps because they better understood the nature of SaaS, so they did not seek terms inappropriate to cloud computing. However, it must be said that users' lawyers have also suggested that some providers' terms are too software license-orientated, and there is room for them to be more cloud-appropriate. Indeed, our sources considered that, even in the information technology industry, many intermediaries have been treating cloud computing as a supply of products or licenses rather than services. All this seems to indicate the relative immaturity of the market.

While not the key focus of this Article, pre-contractual due diligence should not be overlooked. This may include security,[23] disaster recovery, data retention and return of data on exit (including data formats), and exit strategies.[24] We discuss contractual provisions on those issues below.

B. *To Negotiate, or Not to Negotiate?*

Users may decide that, for small initial pilots or tests of moving specific workloads or processes to the cloud, the time and costs of negotiating providers' standard terms are not worthwhile. For example, the Alpha.gov.uk prototype, for a single site for United Kingdom government online services, was based on Amazon's IaaS, using Amazon's standard terms on a click-through basis.[25] Similarly, Warwickshire County Council's pilot of Google Apps[26] commenced in late 2011, on Google's standard terms. Only when considering full migration of the relevant function, or processing of "real" data, particularly personal data, might such users then scrutinize contract terms.

However, users may not be able to negotiate providers' terms. As with any commercial agreement, much depends on relative bargaining power. Large providers generally decline any changes to their standard terms, insisting their

---

23. For pre-contractual security checks (not limited to data protection), European Union data protection authorities have suggested European Network and Info. Sec. Agency, *Cloud Computing: Information Assurance Framework* (Nov. 2009).

24. However, it seems only 45% of United Kingdom cloud users had a plan to migrate to another provider on any service interruption or termination (with 12% responding "don't know"). Similarly, only 45% of integrators had such a plan. CIF3, *supra* note 13, at 9.

25. Government Digital Service response to freedom of information request (Dec. 22, 2011) (on file with author).

26. Kathleen Hall, *Warwickshire County Council Signs Google to Pilot G-Cloud E-mail Service*, COMPUTERWEEKLY.COM (Sept. 19, 2011, 5:00 AM), http://www. computerweekly.com/news/2240105636/Warwickshire-County-Council-signs-Google-to-pilot-G-Cloud-e-mail-service.

services are only on a "take it or leave it" basis—even when requested by large users, such as integrators. Some users have had to "take it," negotiating only commercial terms (price, payment frequency, perhaps availability levels), because "we needed them more than they needed us." Other users have walked away.

Some users appear to accept the general inability to negotiate providers' standard terms and agree to those terms notwithstanding.[27] Nevertheless, even large providers have departed from their standard terms to secure deals they perceive to be sufficiently worthwhile in terms of financial, strategic or reputational "trophy" value. For example, there was publicity about Google's deals to provide Google Apps SaaS to the City of Los Angeles,[28] and more recently to Cambridge University.[29] Some smaller SaaS providers consider requests to change standard terms from larger customers, such as those expected to pay more than a certain amount annually.

Generally, bigger users, particularly from regulated industries, try to negotiate more. Indeed, some go further, insisting on cloud contracts being on *their* own standard information technology services or outsourcing terms, on a "take it or leave it" basis. Such users mainly comprise government bodies and financial institutions. Partly they may have more purchasing power, but also their internal procedures may make it difficult and time-consuming to contract on terms other than their own (for example, banks may require director sign-off for changes to their standard terms). To secure such users' business quickly, some providers may accept the user's terms, although some terms may not suit cloud services.[30]

Thus, both providers and large users want to contract on their own standard terms, although generally neither set of terms seems optimal for cloud contracting arrangements.

C. *Role of Integrators*

Of growing significance in cloud is the information technology channel: termed by the Cloud Industry Forum "Reseller, Service Provider and Out-

---

27. *See CIF3*, *supra* note 13, at 4-5.

28. Reportedly, however, there have been problems. *See, e.g.*, Jon Brodkin, *Google Apps Hasn't Met LAPD's Security Requirements, City Demands Refund*, ARS TECHNICA (Oct. 20, 2011, 8:09 AM), http://arstechnica.com/business/news/2011/10/google-apps-hasnt-met-lapds-security-requirements-city-demands-refund.ars.

29. Google Apps Education Edition Agreement, UNIVERSITY OF CAMBRIDGE, http://www.ucs.cam.ac.uk/googleapps/google-apps-cambridge-contract.pdf (last visited May 9, 2012).

30. W. Kuan Hon, et al., *UK G-Cloud V1 and the Impact on Cloud Contracts – Part I*, 17 COMM. L. 78 (2012); W. Kuan Hon, et al., *UK G-Cloud V1 and the Impact on Cloud Contracts – Part II*, 17 COMM. L. __ (forthcoming 2013).

sourcer," comprising information technology consultancies,[31] managed services providers, systems integrators, specialist resellers, technical value-added resellers, information technology outsourcers, distributors and volume resellers, and information technology retailers. We use "integrator" to mean "reseller, service provider or outsourcer."

Integrators contract with both end users and providers, unless their end users contract directly with providers.[32] Therefore, in many ways, integrators are like providers who use other providers to offer services to their own end users. Integrators are potentially very large users of IaaS or PaaS, based on which they provide services (particularly SaaS) to their end users. Integrators may also provide cloud services on infrastructure they control, such as IBM's Smart Enterprise Cloud, or HP's Enterprise Cloud which includes SaaS and IaaS. They also offer customized cloud services, particularly private or hybrid cloud.

According to our research, integrators are better able than end users to negotiate improved terms with providers.[33] This may be because integrators may have ongoing relationships with providers, and perhaps better bargaining position with larger business volumes, as an integrator may use the same provider to service multiple customers. However, with large providers, we found even global integrators had difficulty obtaining the changes they or their customers needed for data protection or security purposes.

We found some end users contracted with integrators, rather than with providers, because some integrators were prepared to give more contractual assurances than providers. For instance, some interviewees mentioned users contracting with integrators for Office 365, rather than using Office 365 on Microsoft's infrastructure, enabling users to obtain from integrators contractual assurances they could not have obtained from Microsoft, in terms of liability or liability caps, service levels and credits, support, and backups.

However, if in such cases the integrator's contract with the provider is not truly "back to back," integrators bear the risks of any mismatch in obligations

---

31. Which may include advice on migration to cloud, building cloud hardware or software (on clients' or third party infrastructure), support or training.

32. Forty-three percent of United Kingdom cloud services resellers contracted directly with end user customers, 12% required customers to contract with providers (typically on commission), 28% offered a mix dependent on solutions required. 48% offered back-to-back terms between supplier and customer. *CIF3*, *supra* note 13, at 5.

33. Cloud Industry Forum findings were similar: a (provider-favorable) right for providers to change contracts unilaterally by posting a new version online was imposed in contracts with 32% of users, but only 19% of integrators who were resellers. *See CIF3*, *supra* note **Error! Bookmark not defined.**, at 5. Despite their seemingly better bargaining position, however, only 40% of integrators (compared to 52% of end users) were reported to have "consciously negotiated" contracts with providers (since presumably those surveyed were aware whether their organization had negotiated). *Id.* at 4; Cloud Industry Forum, *Cloud UK: Paper Two – The Impact Upon the IT Supply Chain* 12 (2011) [hereinafter *CIF2*], *available at* http://www.cloudindustryforum.org /downloads/whitepapers/cif-white-paper-2-2011-cloud-uk-impact-upon-the-it-supply-chain.pdf.

and liabilities. "If we use a cloud service, as a systems integrator we have to be very careful about what the customer requires, because we might not be able to get that from the cloud service provider," one stated. "We're taking a big slice of the risk pie," another noted. In many cases, integrators make a calculated decision to take that risk, to gain or retain users' business by better meeting their needs. They consider risks in some respects may be spread over end users, and take a view on aggregated risk. Such a position may also receive some encouragement from the insurance market.[34]

However, integrators' greater willingness to assume risk is not unlimited. Integrators noted that where users insisted on a particular provider, if the integrator could not persuade the provider to amend standard terms to meet customer requirements, it had to make that position clear to the customer, and sometimes leave the customer to contract directly with the provider. Hence, while some integrators do not wish to invest in infrastructure, others already are, or are considering, providing cloud services on infrastructure they control, in order to offer users such as banks the assurances they need while being able to manage their own exposure better.

Integrators may also offer mixed cloud and non-cloud services, particularly to large users, where cloud computing is just one service forming part of a larger package of services or "whole business" deal, including for example internet connectivity or consultancy services. Similarly, customers may seek SaaS management or monitoring tools within more traditional large outsourcing deals. Contracts for such deals will obviously reflect the cloud component's relatively minor position.

D. *Other Relevant Factors*

The extent to which users may need to negotiate contracts will obviously depend on how much relative control the particular system's design affords users and providers over users' applications or data, and how "customer friendly" a provider's standard terms are. With paid-for services, providers are generally more willing to accept liability (or greater liability), and agree other user-requested commitments or measures, than with free services.[35] The more providers are paid, the more they are willing to concede. Market factors also play their part, with a global user noting that one large provider was more flexible than others because it was trying to "catch up" in the cloud space, and indeed was more flexible on cloud than on other types of contracts.

---

34. *See, e.g., The Cloud Risk Framework: Informing Decisions About Moving to the Cloud*, MARSH & MCLENNAN COMPANIES (2012), http://f.datasrvr.com/fr1/812/29871/3424_MA12-11623_Cloud_Computing_Frmwk_UK_04-2012_final_nocrps.pdf.

35. Bradshaw et al., *supra* note 2.

IV. CLOUD CONTRACT TERMS: DETAILED ANALYSIS

A. *Liability: Exclusion, Limits and Remedies for Breach of Warranties and Indemnities*

Providers' exclusion of liability, particularly for outages and data loss, was generally the biggest issue for users.[36] Providers try to exclude liability altogether,[37] or restrict liability as much as possible,[38] because they provide commoditized services: understandably, providers may not wish to be exposed to say $100 million of liability for a deal worth $1 million; and unlimited liability could put smaller providers out of business.

According to our research, providers state liability is non-negotiable, and "everyone else accepts it." Even large users had difficulty getting providers to accept any monetary liability, with one global user stating that generally it "had to lump it," and another saying, "they won't move." Refusal to accept any liability was cited as a "deal breaker" by several users. Although liability exclusion is more widespread and more acceptable under United States jurisdictions, some United States users still refused to deal with some providers who excluded liability.

However, some global users negotiated successfully for provider liability. This might be more common where cloud services formed part of a larger deal, such as telecommunications. If liability was agreed, it was almost invariably accepted only in limited circumstances, restricted to narrowly-defined types of damage: typically, "direct" losses only, with no liability for indirect or consequential losses. Where liability was accepted for "direct losses," that term's definition might be much discussed.

Some users, particularly those who could insist on their own standard terms, such as financial institutions or government, secured unlimited liability for defined types of breach or loss, notably breach of confidentiality, privacy or data protection laws, or breach of regulatory or security requirements such as breaches giving rise to regulatory fines. One integrator commented, "For privacy breaches there is no cap that people will agree [to] that would be sufficient."

However, more commonly, liability was capped, sometimes with different caps for different types of losses, and often limited by reference to amounts paid by the user in total or over a period like a year, such as 100% or 125% of

---

36. In the Cloud Industry Forum survey, providers excluded liability for data loss for 34% of United Kingdom cloud users, accepting capped liability for data loss and breach of contract for 54%. With integrators, providers accepted more liability: 20% of providers excluded liability, 31% agreed capped liability; and 27% accepted full liability. *CIF3*, *supra* note 13.

37. Bradshaw et al., *supra* note 2.

38. For example, in the United Kingdom, liability cannot be excluded for negligence causing death or personal injury.

six months' fees.[39] Some providers would agree a higher percentage or longer period of fees for certain deals, such as where fees were paid up front.[40] Users who could impose their own terms might cap liability at a higher proportion, say 150% of charges paid over the last year.

Some telecommunications providers appeared more willing to accept some liability. This may be because they control internet connections to users and therefore have greater control regarding connectivity and service availability. Smaller providers also seemed more flexible, with one enterprise-oriented SaaS provider accepting liability for outages as standard if caused by internet connectivity failures at its data centers, whereas many large providers would not accept any monetary liability for outages however caused. As previously mentioned, some integrators also seemed more willing to accept liability, so that, for example, a global user contracted with an integrator who accepted liability for data loss, when the provider would not accept liability.

Conversely, providers have argued that, in trying to remove or reduce liability exclusions and limitations or increase service levels for commoditized services, customers want to have their cake and eat it too—seeking the cheapest services while requesting the highest levels of assurances. More technologically-sophisticated users stated that they arranged their own backups, for example to their own servers, while a provider noted that users were beginning to understand that they cannot expect much provider liability for low cost services.

Some SaaS providers emphasized that they provide services, rather than licensing software, and preferred not to include any contractual software licenses, to avoid associated risks. With open source software, providers would exclude liability for intellectual property rights infringement. Many users wanted providers to warrant that intellectual property rights relating to application software used in the cloud were the provider's and that the service did not infringe third party rights, with appropriate indemnities. This might not be possible where providers licensed applications from third parties to offer SaaS services. To address third-party intellectual property rights, for example intellectual property rights belonging to providers' own suppliers, some users requested copies of third party indemnities to providers and sought back-to-back indemnities from providers. The indemnity's scope might require scrutiny for suitability. For example, one global provider, whose SaaS service included the supply of certain content, limited its liability only to copyright infringement, excluding patent-related infringements. Generally, as with other kinds of liability, intellectual property rights indemnities were limited to direct losses

---

39. Interestingly, an integrator commented that some providers who indicated willingness to offer capped liability subsequently reverted to total liability exclusion. Also, one SaaS provider observed that its liability ceiling was not negotiated as often as one might think, although it was questioned more in higher value deals.

40. Cloud Industry Forum recommends providers consider offering higher caps for higher fees. *See CIF3*, *supra* note 13, at 13.

only, and capped. One global user declined to deal with a SaaS provider who restricted intellectual property rights indemnities to certain countries' intellectual property rights only, and limited the amount and losses covered.

When considering provider liability, contractual provisions are not the only factor. Several sources pointed out that providers offering unlimited liability may not be creditworthy should losses arise, but users may find it easier to put forward to directors contracts where providers assume rather than exclude liability. Too much focus on contractual terms rather than, for example, financial standing may thus be too narrow for risk management purposes.

B. *Resilience, Availability, Performance and Service Levels*

1. *Data integrity, resilience and business continuity*

A common theme was business continuity and disaster recovery, i.e., how to ensure integrity and availability of cloud data and applications. One user noted, "Providers tend to say cloud is very redundant, fault-tolerant, there's no need for disaster recovery—but there is."

Many providers take two or three backups of data in practice, perhaps even onto backup tapes, although they may not commit contractually to doing so. While some will undertake to make the necessary number of backups, most will not warrant data integrity, or accept liability for data loss. Where a SaaS service included the supply of certain data, liability was limited to replacing any lost data.

Confidentiality provisions may result in liability upon any data security breach, but not for data loss or corruption, as unauthorized access to user data may result in confidentiality breaches, but data loss or corruption may not. Therefore, specific warranties (with liability) in relation to data loss or corruption may be important in addition. Some global users have secured such warranties, for example unlimited liability for data loss or corruption, for one financial institution, and even monetary compensation for data loss including recovery costs, for another global user.

One non-European telecommunications provider undertakes to backup data and guarantees their integrity, commenting that Amazon did not because its charges are low. Many providers, including Amazon, offer backup as a separate service: if the user pays extra, the provider undertakes to make backups, and assumes liability for the integrity of backups and data loss.[41]

Providers such as Amazon stress that cloud involves shared responsibility: both users and providers have responsibility for data integrity, backup and se-

41. One enterprise-oriented SaaS provider noted that force majeure was often raised in relation to provider liability, depending on whether it was required to make backups, and whether the force majeure incident also affected the backup.

curity, and allocation of responsibilities and risks needs careful consideration.[42] Users generally have more control with IaaS or PaaS than with SaaS, because IaaS users instantiate or terminate virtual servers and choose what to install on those servers, such as firewalls, anti-malware and other security measures; and users decide what applications they wish to install and host on IaaS or PaaS, such applications often being user-developed and therefore user-controlled. In contrast, SaaS users use standardized applications, provided by SaaS providers in environments which users cannot control, relying on providers to secure applications as well as environments. The users we interviewed who were more technically aware, such as technology businesses or integrators, tended to recognize the need to implement their own backup strategy, rather than expecting providers to backup as part of their basic service. An integrator using cloud to provide SaaS services to its own users implemented its own extensive disaster recovery procedures, backing up or failing over to the same or separate data centers or another provider, depending on end users' risk tolerance.

2. *Service levels, service credits*

There were different approaches to service level agreements (SLAs), i.e. commitments on availability levels and performance. This was probably because availability levels are often quite high, and capacities, performance, and service levels are normally negotiable commercial issues varying with user requirements, rather than legal issues—for instance, guaranteeing different service levels at different prices.[43]

Standards are still lacking, making it difficult for users to compare different services. In large deals, methods for measuring service levels were often debated, with users wanting numerous key performance indicators, although at the low end providers stipulate the metrics, generally not exceeding five or ten key performance indicators.

Providers justify refusing to negotiate service levels on the basis that they provide commoditized services. One enterprise-oriented SaaS provider stated it did not even offer SLAs. Very few of its users, about 0.2%, had requested SLAs; fewer still wanted as high as 99.7% uptime. Possibly this was because it had designed its service for very high uptime levels, and published information on historic service levels.

If availability, reliability and performance are vital, such as with mission-critical applications or real time services, users may consider issues such as

---

42. *See, e.g.,* AMAZON *Web Services: Overview of Security Processes,* AMAZON.COM (May 2011), http://awsmedia.s3.amazonaws.com/pdf/AWS_Security_Whitepaper.pdf.

43. Cloud Industry Forum recommends that providers document management systems, processes and resources for consistent service delivery, specify whether users may audit business continuity or disaster recovery processes, and publish average availability times. *See* CIF3, *supra* note 13, at 12.

how the load on providers' infrastructure from other users can affect the user's application performance; how well the service handles peak spikes; how sufficient robustness is ensured, and so on.[44] One issue which seemingly has not seen much discussion is the risk of additional users adversely affecting the service, because capacity in the cloud is not in fact unlimited. This may result in SLA breaches, yet reportedly users generally did not try to restrict how quickly providers added new users or how much capacity they could offer new users.

Users with mission-critical applications may accordingly seek higher availability levels, warranties regarding response times, undertakings not to terminate services without notification and consent, longer prior notification of proposed maintenance downtime, perhaps even notification of usage exceeding agreed limits (to allow the user to investigate and manage the situation), rather than immediately throttling usage and slowing performance without consulting users. However, providers are unlikely to warrant latency (response time), unless perhaps they control the network, such as some telecommunications providers or integrators, who for a price will warrant low latencies. One provider considered that what users truly want, but providers did not offer, was guaranteed application performance, for example X simultaneous users with maximum Y response time; it believed there would be increasing focus on application performance management and monitoring.

Users may also consider how and how long it takes to restore data from backup if systems go down or data are lost. One integrator commented that providers' response time requirements were "nothing like what clients insist upon."[45] The time lag varies with providers, from seconds for some business-oriented services (which accordingly charge more), to days.

SLAs are often referenced by linking to providers' published website details, which providers may amend, thus putting the onus on users to monitor providers' sites for changes.[46] As this is burdensome, some users have required prior notification of impending SLA changes.

For SLA breaches, remedies are normally excluded except as specifically provided. Providers generally exclude any remedies other than service credits, even for total service failures, although some allow optional termination if SLAs fall below a certain percentage, and may even accept liability for some monetary compensation if (but only if) terminated. Standard terms generally circumscribe the circumstances when service credits are given; for example, they may give credits only for failures arising from matters under the provider's control, or only if credits are claimed within a certain period. One global user, who uses cloud computing to provide real-time services to end users, did not insist on performance warranties or even service credits, as any outage would

44. Bradshaw et al., *supra* note 2, at § 4.19.

45. Cloud Industry Forum recommends providers should specify protocols and service level agreements for restoring data from backups. *CIF3*, *supra* note 13, at 11.

46. Bradshaw et al., *supra* note 2.

be highly detrimental to its reputation so even monetary compensation would probably be inadequate, and service credits were difficult to quantify. Therefore it decided simply to accept standard SLAs.

Even where service levels are non-negotiable, service credits for SLA breaches may be. However, while preferring service credits, some providers do offer benchmarked "money back" rebates or monetary compensation.

3. *Transparency*

Our research indicated that there are two main ways in which availability data is provided to users: users may monitor availability themselves, or providers may provide the information to users, such as specific webpages which are kept updated, and which users may check.

A global user stated that if the contract included SLAs, it would further require reporting of statistics. Similarly, a non-European telecom provider mentioned that it offered SLAs according to users' requests, and would commit to keeping users informed proactively within certain boundaries ("tell them before they tell us").

Where availability was critical, users might monitor the service or application themselves, for example for failure of virtual machine ("VM") instances, using the provider's tools or (as many providers do not allow user access to their tools) public tools or the user's own tools. However, too much user monitoring may itself affect application performance (and increase bandwidth usage charges). After assessing the impact of monitoring, one large user was persuaded not to monitor. One global user also pointed out that providers' usage monitoring for billing purposes may affect service performance. Therefore, for services where near real-time response times are critical, some users sought rights to require providers to pause or stop monitoring if materially detrimental to performance. One integrator stated that it usually monitored charges itself, depending on the setup of the service.

It seems providers generally could do more to improve transparency regarding availability and service levels by providing such data to users proactively. Illustrating that greater transparency is possible technically, trust.salesforce.com and Microsoft Dynamics CRM were cited as services enabling users to check, even in real time, data on performance and availability of services. Such transparency may involve technical changes and costs for providers. Generally it is providers with large enterprise customer bases who are providing such data for users.

4. *Users' liability*

Another issue is users' liability to providers, particularly where users employ cloud computing to deliver services to their end users. Unsurprisingly, users who could negotiate their contracts would not accept liability to providers. One global user declined to contract with a provider who required such liabil-

ity.

Cloud users, not having control over their own end users' actions, also declined to indemnify providers for such actions, even if constituting breaches of providers' terms, even if causing loss to the provider. A compromise was for providers to terminate or suspend the service, with sufficient prior notice for users to investigate and terminate the culprit's account if necessary. Such users also ensured that their contracts with end users allowed termination for misconduct, with indemnities from end users.

C. *Regulatory Issues*

Generally, providers' role regarding their users' compliance obligations still seems not well defined, understood or accommodated by providers. A common theme was that many providers, in standard terms or even in negotiations, would not take into account that users have regulatory or other legal obligations and may need to demonstrate compliance to regulators. Some users expressed frustration at providers' lack of empathy with their compliance obligations, especially in Europe. For some users, the solution was to use cloud computing only in less highly regulated jurisdictions, such as some Asian countries. Integrators are in an interesting position, being "caught in the middle," as the compliance responsibilities are generally not theirs but their end users'.

It seems some providers simply have not considered regulatory requirements' impact on their terms.[47] Several users noted this. One discovered, after protracted negotiations, that a global provider had never conducted a regulatory review of its own services or terms, let alone of its contract with its own subprovider. Reportedly some users had, for cost-saving reasons, decided to use cloud despite inadequate contract terms, "taking a view" on regulators discovering and enforcing any non-compliance. However, one integrator commented that, although providers refused to negotiate privacy and security issues, when users refused to sign, some would agree changes.

Regulatory issues varied with jurisdiction; more issues reportedly arose with, say Germany or France, than the United Kingdom. But with European users generally, and users outside government and financial sectors, data protection laws were the most commonly-cited regulatory issue. This seems unsurprising, as data protection law is "horizontal" rather than "vertical," meaning that it regulates all sectors, and controllers of personal data remain responsible

47. To some extent, this may reflect some users' lack of focus on regulatory or legal issues when procuring cloud services. *See supra* Part III.B; *see also* Edward F. Moltzen, *Analysis: Dropbox Carries Risks For SMBs*, *CRN*, (Nov. 4, 2011), http://www.crn.com/news/cloud/231902380/analysis-dropbox-carries-risks-for-smbs.htm (noting that the then new Dropbox for Teams cloud storage service did not meet Payment Card Industry requirements or the United States Health Insurance Portability and Accountability Act or Sarbanes-Oxley Act; Dropbox responded that customers who beta-tested the service prior to launch were concerned with collaboration and ease-of-sharing, not the cited requirements or laws).

if processing data in the cloud. Financial sector regulation was another key regulatory issue.

Data location and data confidentiality were the top data protection law concerns, followed by data processor and transfer agreements and the role of sub-providers, while for financial institutions the biggest issues were security requirements and audit rights.

1. *Data location and data export*

Users were not concerned about colocation[48] within a third party's data center, so much as geographical location of data centers. Many users were concerned about the location of data center(s) employed by their providers, particularly if using data centers outside the European Economic Area. Conversely, other users specifically wanted data kept offshore, such as in the Channel Islands, but never in the United States.

The Data Protection Directive prohibits transfers of personal data outside the European Economic Area except in specified circumstances, such as when recipients are certified under the United States Safe Harbor scheme, or where using standard model clauses or binding corporate rules.[49] This prohibition is significant. Global users have refused to contract with providers who declined to include terms to comply with the Data Protection Directive's international transfers requirements.

The Cloud Industry Forum recommends providers offering European Economic Area-only locations should inform users accordingly. Indeed, European Economic Area users might consider this a selling point. It also recommends providers should disclose all data center locations, including those used for backups, and whether data may be transferred outside the European Economic Area.[50] However, some providers will not disclose data center locations.

Users have sought, and sometimes obtained, warranties or undertakings that all data centers used for their data were in the European Union and European Economic Area or that data were located only within the European Union. One United Kingdom-based global user stated that, although it did not process personal data in the cloud, it still required its global provider to confine processing to European Union data centers, and, should it nevertheless transfer data to the United States, it must be certified under Safe Harbor and comply with its principles.

Some services allow users to choose locations of data centers used to process users' data, e.g., European Union-only,[51] while providers are increasingly offering, albeit with exceptions, to restrict data to users' chosen locations as

48. Namely, sharing hardware or software with other customers in the same location.
49. Hon et al., *Data Export in Cloud Computing*, *supra* note 10, at 30-31, 41-47.
50. *CIF3*, *supra* note 13, at 11.
51. Hon et al., *Data Export in Cloud Computing*, *supra* note 10, at 25.

standard.[52]

Verifying that data are actually processed in the data centers claimed by providers is difficult, technically.[53] One provider noted that some providers were misleadingly labeling servers as "EU" when they could process data elsewhere. One public sector user felt warranties of United Kingdom-only data location could be untrue, citing a hosted services provider who stated that data storage was limited to the United Kingdom, but whose IP address indicated a United States location.[54]

Users may need location information for reasons other than the restriction on transferring personal data outside the European Economic Area. The Data Protection Directive requires controllers to choose processors providing "sufficient guarantees" regarding security measures for processing, and to ensure compliance with those measures. This may be difficult without more transparency regarding providers' systems, data center locations and transmissions.

Although the Data Protection Directive allows personal data transfers *within* the European Economic Area, many users, particularly public sector users, further require data centers to be within their own country. One provider stated that, because it did not currently use any data centers in the United Kingdom, it could not offer its solution to United Kingdom central government departments such as the Cabinet Office or Department for Works and Pensions, who required use of United Kingdom data centers only.[55] Integrators have also noted this position.

Our research ties in with the Cloud Industry Forum surveys, which found that seventy-five percent of United Kingdom users considered it important for data to be stored in the European Economic Area, with forty percent wanting data confined to the United Kingdom. Interestingly, while forty-one percent of users surveyed by Cloud Industry Forum wanted corporate data confined to the

---

52. *See, e.g.*, *AWS Customer Agreement*, AMAZON WEB SERVICES (March 15, 2012), http://aws.amazon.com/agreement (section 3.2) ("We will not move Your Content from your selected AWS regions without notifying you, unless required to comply with the law or requests of governmental entities."). *See also* Hon et al., *Data Export in Cloud Computing*, *supra* note 10, at 25-26.

53. Hon et al., *Data Export in Cloud Computing*, *supra* note 10, at 33 n.36.

54. However, an IP address may not always reflect geographical location. For example, a multinational corporate may route all communications through its United States headquarters, thus identifying them externally as having United States IP addresses, irrespective of local offices' locations.

55. This may seem little different from United States regulation, such as the International Traffic in Arms Regulatory framework, requiring that certain cloud services for federal government be housed in information technology infrastructure located in the United States accessible only to vetted United States citizens. For example, in 2011 Amazon launched a cloud service to comply with these requirements. *See* Werner Vogels, *Expanding the Cloud – The AWS GovCloud (US) Region,* ALL THINGS DISTRIBUTED (Aug. 16, 2011), http://www.allthingsdistributed.com/2011/08/aws_govcloud_region.html; *AWS Security and Compliance Center*, AMAZON WEB SERVICES, http://aws.amazon.com/security/ (last visited Dec. 12, 2012).

United Kingdom, the proportion increased for small and medium-sized enterprises as well as public sector users.[56]

Providers may attempt to address these issues by using partners with data centers in the required countries, or by using private clouds. While data location can be circumscribed by such commercial or technical means, this involves greater costs, because providers may not be able to use resources as efficiently.

Salesforce was cited as allowing users to check their data's location in near real-time with its trust.salesforce.com webpage. Possibly this is due to Salesforce's history of servicing business users, whereas many other cloud services' initial customer base is mainly comprised of individual consumers or small and medium-sized enterprises. Technically, providers may be able to engineer their systems to offer similar services. However, this will involve expense which may be passed on to customers, and, as mentioned above, verifying claimed data locations is still difficult technically.

Data location is a problematic, in some ways emotionally charged, issue. One global user said that if it could not ensure its data were confined to the European Union, it would nevertheless avoid the United States because of the United States PATRIOT Act and litigation issues. Financial institutions particularly raised this legislation. European Union providers and others have suggested that data would be safer in, for example, a German-only cloud.[57] However, the United States is not the only nation that may access data for anti-terrorism or anti-crime purposes, and the current high profile of this Act may perhaps reflect some marketing opportunism, and certain political concerns regarding the United States exercising its powers extra-territorially, more than legal differences.[58] We discuss law enforcement access provisions further below.

Given the prevalence of remote access to data, data center location is not the only factor affecting data location. Many sources pointed out that a European Economic Area provider, with data centers confined to the European Economic Area, may, in order to provide round-the-clock "follow the sun" services, use support staff or sub-contractors outside the European Economic Area who have (or are given) access to customer data or metadata. Remote access to user data by affiliates or support staff or sub-contractors may involve "transfer." Even where such personnel cannot login to user accounts, they may be able to view metadata, such as when an e-mail was delivered. To troubleshoot account issues, sometimes users may give login details, including passwords, to support staff, in which case user consent may be assumed. However, where

---

56. *CIF3*, *supra* note 13, at 8.

57. Cornelius Rahn, *Deutsche Telekom Wants "German Cloud" to Shield Data From US,* BLOOMBERG (Sept. 13, 2011), http://www.bloomberg.com/news/2011-09-13/deutsche-telekom-wants-german-cloud-to-shield-data-from-u-s-.html.

58. Ian Walden, *Accessing Data in the Cloud: The Long Arm of the Law Enforcement Agent,* QUEEN MARY SCH. L. LEGAL STUD. RES. PAPER 74, 2 (2011), *available at* http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1781067.

non-European Economic Area support staff have remote access to personal data (even when not given login details by individuals seeking assistance), and users have not been informed, it is not certain that users must be assumed to have consented to the transfer. One user called potential access by support staff "a huge hole," citing a provider who insisted its service involved no international data transfers because its data center was located in the European Union, seemingly overlooking the relevance of the fact that its support, maintenance and even code debugging might occur in India and that therefore personal data could be transferred to staff located in India. Some providers deal with this by using a recognized method of transfer to non-European Economic Area support staff or support sub-contractors, such as establishing contracts based on model clause terms.

Irrespective of data protection, existing contracts may restrict data location. For example, one global user's service, provided to end users over the cloud, includes provision to end users of licensed content subject to intellectual property rights. Certain content licenses required this global user to store the content on its own secured servers, and to know its location at all times. Obviously, such content cannot be stored in the cloud in circumstances where it could be distributed or moved across different data centers. Even within one data center, a user's stored data may migrate between locations, depending on systems used. Thus, some intellectual property rights licensing schemes may hold back cloud adoption, and licensors may need to be made more aware of how cloud storage works.

Other regulatory issues may be relevant. For example, export control laws were mentioned as restricting transfer of certain information or software to particular countries indicating that locations of providers and data centers may arise in that context as well, with remote access being, again, a thorny issue.

2. *Data processor agreements, and sub-processors*

Users who are "controllers" of "personal data" under the Data Protection Directive must comply with certain requirements when engaging "processors" to process personal data for them. The Data Protection Directive's minimum requirements have been increased by some states, so rules are not uniform across the European Economic Area.[59]

Many providers' standard terms are silent on the point, or at most state that the provider acts only as "data processor." This seems mainly for providers' protection, to try to ensure they are not regarded as "controllers" (with greater obligations and liabilities) even though contractual labels are not determinative.[60]

---

59. *See* Hon et al., *Who is Responsible for 'Personal Data' in Cloud Computing*, *supra* note 10, at 4.

60. *Id*. at 7-9.

The Data Protection Directive requires the controller or user's contract with the processor or provider to contain an agreement, often referred to as a "data processor agreement", which states that the processor must process personal data only according to the controller's instructions and must take certain security measures. Some enterprise-oriented providers' standard terms may, to assist users to comply with the Data Protection Directive, specifically include a data processor agreement. Otherwise, users who are controllers of personal data may want it included. This is common, for example, in financial institutions' standard contracts.

However, some providers would not go further than stating their processor status. The underlying reason may be that the Data Protection Directive does not cater well to cloud processing. In particular, even treating infrastructure providers as processors may be inappropriate, as they provide users with information technology resources and suitable (but standardized) environments within which users may use those resources rather than actively processing data for users.[61] Providers have therefore considered it inappropriate to agree contractually to process data only on users' instructions as it is for users to control their own processing. With shared common multi-tenant infrastructure involving one standardized environment for all users, it would be difficult, if not impossible, to comply should different customers issue different instructions regarding the environment or resources.[62]

That said, some providers have accepted a data processor agreement. Conversely, where providers would not agree anything more specific than general obligations to comply with data protection laws, one global user stated that it had accepted that with large providers who, from previous dealings, the user knew had implemented proper data protection processes.

Providers' terms generally entitle them to use sub-contractors, for example, to provide support services. One fundamental issue is whether cloud sub-providers are sub-processors of personal data. This Article assumes they are sub-contractors and sub-processors, although the matter is arguable.[63] Use of cloud "sub-processors" has arisen in several European Union data protection authorities' decisions or guidance; they consider that compliance requires controllers to know all possible sub-processors, perhaps even down to the data-center operator level.

Whether or not personal data were involved, some users wished to restrict sub-contracting (including of support services), and to provide contractually that providers' contractual obligations were unaffected by any sub-contracting.

---

61. *Id.* at 9-14 (analyzing these issues in detail).

62. A possible exception involves private clouds outsourced to integrators, where integrators may be able to comply with instructions regarding infrastructure dedicated to and customized for the user concerned.

63. *See* Hon et al., *Who is Responsible for 'Personal Data' in Cloud Computing*, *supra* note 10, at 6.

Users in large deals normally prohibited sub-contracting without their consent, except to providers on a pre-approved list and on certain mandatory terms. They sometimes asked to see the sub-contract, or even contracted directly with sub-contractors for obligations including confidentiality. One global user stated that, as with managed services, if any sub-contractors had access to data stored or otherwise processed by the user in the cloud (whether personal data or otherwise), or if any sub-contractors provided services worth more than a certain percentage of the contract value, it required rights to vet and veto them. For smaller deals the user wanted notification of any such sub-contractors' identities, but not necessarily veto rights. With SaaS, the user still inquired specifically as to sub-providers and their geographical locations. Similarly, other global users indicated they would not allow sub-contracting or assignments without express prior consent. One SaaS provider confirmed that some users insisted on consent as a pre-condition for assignments. Therefore, it seems that at least some users who have been able to negotiate contracts have required similar safeguards and pre-contract information, as European Union data protection authorities have required.

In summary, current data protection law puts both users and providers in a difficult position, with one or the other being required to take a view on the contract terms, accepting that a data processor agreement is necessary by law where personal data are involved, but noting that it may be meaningless or impossible to comply with a data processor agreement in the cloud. Investigation of any sub-providers, and possibly even data centers ultimately, may also be necessary for compliance.

3. *Data subject rights*

If data subjects request access to their personal data from users, one SaaS provider pointed out that, because users have direct access to and control over data (including any personal data) they choose to process using cloud computing, it should be unnecessary to involve providers. Users can retrieve the requested personal data directly themselves. Nevertheless, one global user secured from its global provider an obligation to co-operate with the user as necessary to respond to subject access requests, and similar wording is standard in United Kingdom government contracts, including the G-Cloud v1 framework agreement.[64]

Again, these sorts of contractual provisions, while they made sense for processing agents who have sole control of outsourced data, do not recognize the "self-service," tools-based nature of cloud computing.

D. *Confidentiality, and Rights to Monitor, Access, Disclose or Use*

64. Hon et al., *supra* note 30, at 22 n.159.

Customer Data

1. *Confidentiality*

Providers may receive or have access to two main kinds of confidential data: data disclosed by users during contract negotiations (such as information on intended purpose and usage of the service), and data processed via the service. In both cases, these may include personal or commercially sensitive data, in the latter case, possibly even customer or other third party personal data. Users may accordingly want confidentiality or non-disclosure agreements from providers. In some cases, users have secured unlimited liability for the former, but liability for the latter has proved more difficult to allocate, and can be a "show stopper."

Many users have persuaded providers to accept liability, usually capped, for breach of confidentiality, and generally users pushed for higher caps here. Users preferred capped liability from integrators to no liability or restricted liability under providers' standard terms. It was rare, but not unknown, for providers to accept uncapped liability for breach of confidentiality, data disclosure, or data protection breaches, at least where limited to direct losses and excluding contingent liabilities or consequential damage. Enterprise-oriented providers seemed more flexible with highly regulated users, for example even giving indemnities for breach of data protection obligations. In contrast, one non-European Union provider stated it would never agree to any liability for breach of confidentiality.

Definitions of breach of confidentiality and the amount of the cap obviously involved careful consideration, for both users and providers, as did definitions of direct, indirect, and consequential losses. Some providers also considered that data loss should not be treated as breach of confidentiality per se, particularly with SaaS, capping such liability accordingly.[65]

Users negotiated for confidentiality obligations to survive termination of the agreement, typically for 5 to 7 years, depending on the data's nature. Some users even wanted such obligations to continue indefinitely.

2. *Access to user data; disclosure*

Users fear possible unauthorized access to their data. Many providers have "back doors" to access users' data, and contractually reserve access rights,[66]

65. On "personal data" in cloud computing, *see* W. Kuan Hon, Christopher Millard & Ian Walden, *The Problem of 'Personal Data' in Cloud Computing: What Information is Regulated?—The Cloud of Unknowing*, 1(4) INT'L DATA PRIVACY LAW 211 (2011). On responsibility for personal data in cloud computing, *see* Hon et al., *Who is Responsible for 'Personal Data' in Cloud Computing*, *supra* note 10.

66. Bradshaw et al., *supra* note 2.

which they consider are needed for maintenance, servicing, support, or even security purposes. Their users may have to accept that support and perhaps other staff may access their data. Other providers stated they had no such access, although they could terminate users' access.

Even some providers who stated that they cannot access user data without customer login details reserved rights to access data for maintenance and the like, and acknowledged that certain employees could do so in "emergency" situations. In addition, users may volunteer login details to support staff when seeking assistance, although the extent of resulting access depends on that individual user's own access rights.

Regarding service usage monitoring, some users had not negotiated the issue, but others stipulated expressly that information obtained from such monitoring, or from support or maintenance activities, must be treated as confidential information subject to confidentiality provisions. Some users restricted contractually the purposes for which the provider could monitor use: for example, only for security purposes (such as filtering e-mail for spam or malware), or to establish and substantiate charges payable, such as how many end users had used a particular service, or to verify compliance with terms, for example size, capacity and bandwidth limits. Some users wanted to prohibit use of resulting data for any other purpose, and indeed prohibit monitoring for any other purposes. However, some providers would not agree to any restriction on their monitoring rights. Others stated they could not technically monitor processing performed by users, e.g., within users' VMs.

Standard terms usually authorize providers to disclose users' data on court order—or even if simply *requested* by law enforcement authorities.[67] Some providers would contractually agree to notify a user immediately upon law enforcement or other official authorities' requests for that user's data. However, certain laws may forbid such notification. Thus, the term would be qualified accordingly. Absent such standard terms, users have required providers to notify them promptly on receipt of any third party requests (unless prohibited by law), passing requests to the user (which integrators would send on to end users or customers), and allowing the user to make representations to resist disclosure. Indeed, financial institutions' standard terms may even prohibit disclosure altogether except with their consent, although they would accept that disclosures may be required under laws or court orders preventing providers from notifying users. In some cases, providers considered that users subject to specific lawful intercept requirements, notably telecommunications users, endeavored to pass off those obligations to providers.

One SaaS provider had users with specific requirements as to data location

67. Bradshaw et al., *supra* note 2; Walden, *supra* note 58. Nevertheless, 54% of United Kingdom users believed providers would give data to third parties only if required by court order, while 93% of United Kingdom users expected providers to contact them before releasing data. *CIF3*, *supra* note 13.

and ability to access data (e.g., never outside country X, must be duplicated in country Y as well as elsewhere, never in country Z, and so forth). It addressed those requirements by using affiliates and partners operating data centers in different countries. If an entity in one country was required by that country's laws to access data held in a data center in another country, it could not do so, because it had no control over that data—the affiliate or partner did.

E. *Security Requirements, Audit Rights, Security Breaches or Incidents and Incident Response*

Security is often cited as an issue in cloud computing—partly because of general concerns arising from loss of user control, and partly because data protection laws require controllers to take appropriate security measures to protect personal data. It may also be easier for users to obtain board approval for contracts specifying detailed security requirements and audit rights, demonstrating that a considered, structured decision had been taken on the security risks.

Accordingly, security often arose in pre-contract discussions, particularly as many providers were not forthcoming regarding their security arrangements. One global user noted, "It is a challenge to find out what protection providers are providing." Worryingly, a report in early 2011 on the security of cloud providers found that most providers, including large ones, did not prioritize security.[68]

Several sources cited problems with audit rights, particularly for financial services users, who require extensive audit rights for themselves, their financial auditors, and regulators.[69] Deals have fallen through because providers would not compromise on audit rights. One integrator commented that its users did not appreciate the level of resistance from providers on audit rights. Audit rights and certifications have been in issue with shared services generally, not just cloud computing.

Education is important, as it appears some users still lack sufficient knowledge about cloud components and services. For example, an integrator was asked to agree to audit rights where it only provided application software running on the user's own infrastructure.

---

68. PONEMON INST, SECURITY OF CLOUD COMPUTING PROVIDERS STUDY (2011) (reporting a survey of 103 cloud service providers in the United States and twenty-four in six European countries). In contrast, a more recent survey of 300 senior European technology leaders, including executives, investors, policy makers and officials aligned with the technology market, found that, of respondents who ranked cloud computing as a top three area of potential growth, security (protection and centralization of data) was one of the features stated to most drive their company's adoption of cloud computing (at 58%). DLA PIPER INT'L LLP, THE EUROPEAN TECHNOLOGY INDEX (2012).

69. *See, e.g.,* FINANCIAL SERVICES AUTHORITY, SENIOR MANAGEMENT ARRANGEMENTS, SYSTEMS AND CONTROLS, 2009, 8.1.8 (U.K.).

1. *Providers' security measures: pre-contractual audits*

What security measures should be taken, and who is best placed to take them, will vary with the nature and type of service. With IaaS and PaaS, providers consider users have much more control over security measures, while SaaS providers generally have more control.

Regarding pre-contractual audits, for data protection and other reasons users wanted to know what physical and digital security measures providers took, to ensure providers had adequate security policies and underpinning systems, and that any issues were followed up in practice, with appropriate approval processes for configuration and change management. Users' security questionnaires might include hundreds of detailed questions, such as fire extinguisher locations.

However, providers generally considered that, particularly with shared infrastructure and multi-tenancy, it would be detrimental to security and against their own security policies to provide full details of their security policies and practices to all prospective customers, or allow data center visits. In other words, too much transparency about security can itself compromise security.

One global user asks all potential providers for their security standards, which its security team reviews, and said, "Where we know we will be struggling with the provider because of their security standards, and they're not prepared to negotiate or change them, we'll go elsewhere." Some providers provided documentation or other information showing they take security measures. Generally, providers would at most allow users (or users' security teams) to see a summary or high-level overview of security policies, measures, and standards. Only rarely would a provider provide more detailed information than it made generally available. However, for the right deal or customer, typically government or financial services, one enterprise-oriented provider had, subject to non-disclosure agreements, allowed prospective users' security-vetted personnel to make escorted data center visits, view specific documentation such as its ISO27001[70] policies and procedures and other detailed information given to its certifier to support its certification, and discuss issues with teams providing or supporting services and application security, security monitoring, and so forth. However, they were restricted to viewing hard copies in closed rooms, with no ability to take copies.

One global user commented that Salesforce emphasizes its willingness to allow users to visit its data centers, whereas some others are less receptive to audits of their physical or logical storage and systems. For small deals, understandably providers may not be willing to allow numerous prospective customers to visit data centers and the like, although one SaaS provider, who disallows

70. ISO27001 is an industry standard security framework for implementing information security management systems within an organization.

"pen testing,"[71] would allow "physical walk rounds" of sites, and another global provider also permitted site visits. This may depend on how much control the provider has. If it controls all relevant data centers, escorted "tours" are obviously less problematic, although they still involve resources and costs for providers (and, accordingly, increased costs for users). Some queried the value of physical visits, and indeed the Swedish data protection authority questioned their appropriateness with shared infrastructure, "since data from several players are stored in the same premises and access opportunities for all who store the data would result in security risks in itself."[72]

Sometimes, having information about the provider's security measures is sufficient for users. One global user stated it did not necessarily negotiate detailed terms regarding security if, having seen the provider's documentation and been permitted access to its systems, it was satisfied with its security.

2. *Whose security policy?*

To what extent could users dictate security policies or practices? In a pre-cloud, single user scenario, such as traditional outsourcing to a managed services provider, a provider might well agree to follow the user's security policy.

However, where multiple users share standardized infrastructure, it would be difficult, if not impossible, for providers to comply with all users' separate security policies, with possibly different, even conflicting, requirements. Nevertheless, users often wanted to specify their own security policies (which they considered appropriate to the data's sensitivity), typically scheduled or annexed to the contract, with the provider undertaking to comply with those minimum measures.

Providers generally refused. At most, they undertook to comply with their own policy, perhaps specifically stated to be based on industry best practices such as ISO27001, usually reserving rights to change their own policy unilaterally—essentially on a "take it or leave it" basis. One global user noted that, having reviewed the provider's policy, for reasons of pragmatism it was willing to accept that position, at least for relatively low risk data. Even where providers' security policies already covered, for example, encrypting stored data, they might not necessarily make that a contractual commitment, although they might agree that, for example, only qualified and vetted employees who needed access to user data would have access.

Depending on the data's sensitivity, users may want more assurance on minimum-security levels than simply compliance with providers' own policies. For instance, some users wanted all data at rest and all connections to be en-

---

71. *See infra* Part IV.E.4 (for definition).

72. Datainspektionen, *Tillsyn enligt personuppgiftslagen (1998:204) – Salems kommunstyrelse* (2011), *available at* http://www.datainspektionen.se/Documents/beslut/ 2011-09-30-salems-kommun.pdf.

crypted.[73] Most users could not compel providers to undertake additional security measures, although some large ones, such as banks, could. One global user, after much negotiation, persuaded some providers to agree to follow the user's own security policies, although others refused. For another user, some providers agreed minor changes, but about half refused to meet higher standards. This may partly be due to reluctance to incur costs, sometimes significant, in obtaining and maintaining certifications such as ISO27001. A non-European telecommunications provider stated that, in one large deal, the user considered even ISO27001 inadequate, although there the provider was willing to receive the user's specific requests and consider implementation costs. Deal value seems highly relevant. For the right price and users, such as financial institutions, some smaller providers were willing to accept higher security requirements—one SaaS provider even proved willing to agree to a financial institution's full standard security schedule.

According to our sources, many deals have fallen away, or "not even gotten to first base," because providers would not agree to follow users' security policies. One global provider felt that in such situations, the user was simply not ready to move to the cloud.

3. *Certifications*

Independent certifications to objective industry standards may be a possible compromise to address security issues. Industry standards and certifications specific to cloud security have not been fully developed, although organizations like the Cloud Security Alliance, Open Data Centre Alliance and Cloud Industry forum are progressing matters. Whatever standards are adopted need to be cloud-appropriate. One provider commented that some customers, especially governments (United States, European Union including United Kingdom) or government contractors, required Evaluation Assurance Level certification under Common Criteria security evaluation standards.[74] However, those specify particular hardware or software rigidly, with certifications being invalidated on changing the relevant hardware or software, —so they do not suit the cloud, where hardware used may be unknown, and products evolve quickly. That provider said, "It is complex and expensive to get official certification for so little certainty."

One enterprise-oriented SaaS provider stated that all of its customers raised security as an issue, but after conducting their due diligence and noting the service's certifications such as for credit cards (PCI/DSS),[75] ISO27001, and

---

73. Whether providers have access to decryption keys is a different, important issue. *See* Hon et al., *Who is Responsible for 'Personal Data' in Cloud Computing*, *supra* note 10.

74. COMMON CRITERIA, COMMON CRITERIA FOR INFORMATION TECHNOLOGY SECURITY EVALUATION (2012).

75. PCI SEC. STANDARDS COUNCIL, PAYMENT CARD INDUSTRY (PCI) DATA SECURITY

SAS70[76] certifications, they were more reassured; indeed, some were satisfied that the security level far exceeded their internal security protections. Some providers undertook to obtain and maintain industry standard certifications such as ISO27001, providing users with copy certifications and so on. One global provider undergoes regular SAS70 type II audits by an independent auditor, sharing reports with users.

Other certifications or security assurances may be considered regarding software infrastructure, such as on the effectiveness of virtualization platforms for segregating users. In September 2011 the CESG (United Kingdom National Technical Authority for Information Assurance) assured VMWare's VSphere 4.0 hypervisor for hosting, on the same platform, VMs for United Kingdom public sector information, protectively marked Restricted (Business Impact Level 3 or IL3[77]) and below.[78]

4. *Pre-contractual penetration testing*

Pre-contractual due diligence measures for users may include security penetration testing ("pen testing"), to check security issues such as integrity and robustness of providers' security policy and information technology systems, and how (and how well) users' data or instances are separated from other users' data.

Many users, particularly from regulated sectors, wished to conduct pre-contractual pen testing. However, most providers would not agree, because of potential adverse impact on other users' services or data. An enterprise-oriented public SaaS provider confirmed it often received such requests from large organizations, particularly financial institutions. That provider had occasionally agreed, subject to the user's agreement to accept unlimited liability for any damage caused and to constrain testing as regards timing, from which IP address, and so on. Any agreed pen testing would usually be confined to a sand-

---

STANDARD (2010). Note that a provider's PCI/DSS compliance does not automatically mean its users are compliant, as that requires further action "above the hypervisor," within the purview and control of users rather than the provider. Marcia Savage, *PCI DSS Compliant Cloud Providers: No PCI Panacea,* SEARCH CLOUD SECURITY (Mar. 22, 2011), http://searchcloudsecurity.techtarget.com/news/2240033583/PCI-DSS-compliant-cloud-providers-No-PCI-panacea.

76. AM. INST. OF CERTIFIED PUB. ACCOUNTANTS, STATEMENT ON AUDITING STANDARDS NO. 70 (1992). Replaced in June 2011 by AM. INST. OF CERTIFIED PUB. ACCOUNTANTS, STATEMENT ON STANDARDS FOR ATTESTATION ENGAGEMENTS NO. 16 (2011).

77. *See* CESG & CABINET OFFICE, EXTRACT FROM HMG IA STANDARD NO.1 BUSINESS IMPACT LEVEL TABLES (2009).

78. CESG, *CAPS Product Results*, http://www.cesg.gov.uk/finda/Pages/CAPSProduct.aspx?PID=176 (last visited Nov. 15, 2012). *See also* CESG, CESG AND VMWARE DELIVER TRUSTED PLATFORM FOR HOSTING MULTI-LEVEL ENVIRONMENTS (Sept. 14, 2011), *available at* http://www.cesg.gov.uk/Publications/Documents/cesg-vmware_joint-statement14-09-11.pdf.

box, a special segregated area, to avoid possible damage to systems that could affect other users. Numerous pen tests were conducted on this basis annually. Some automated scanning of the application was permitted, at weekends only. However, that provider still did not permit unlimited pen testing.

Providers that disallow user pen testing may conduct their own tests (or use a third party), sharing results with current or prospective users. One SaaS provider organizes its own regular third party pen tests, including after application upgrades, and shares summarized results with users. Some users required sight of such results before contracting. Others still wanted their own or third party pen tests. Another SaaS provider stated that it was obtaining SAS70 certification[79] to address users' desire to conduct their own pen testing.

Ongoing user pen tests were unusual. One global user, after conducting pre-contractual pen testing on a global provider, without a physical site audit, was satisfied by that test together with undertakings to comply with security measures, without requiring rights to conduct future pen tests.

As a non-European telecommunications provider pointed out, it would be impossible to stop users from pen testing if they wished to conduct them. Interestingly, although that provider had several large users including banks, it commented that none sought prior security testing, although all requested security certifications.

5. *Ongoing audit rights*[80]

Regarding post-contract audits, many financial services users felt they needed providers to commit to allowing audits at least when the user's regulator (or its end user's regulator) required it. However, most providers refused for reasons of security and costs, although providers were more willing to allow audits if users met all costs. Even enterprise-oriented SaaS providers that allowed audits would not offer the unfettered audit rights required by financial institutions. They generally restricted them narrowly, for example once a year per user, only if the user's regulator required the audit, or only with their prior consent. Or, they might agree to provide only "commercially reasonable" cooperation rather than full audit rights.

Several European Union data protection authorities consider that users need technical and practical means to investigate suspected unauthorized accesses to personal data, whether within the user or provider, meaning contractual rights to logs or audits. Some providers offer users tools to monitor access-

79. Replaced by SSAE 16 from June 2011. *See* AM. INST. OF CERTIFIED PUB. ACCOUNTANTS, *supra* note 76.

80. For a detailed survey and guide to ongoing security service level agreements, *see* European Network and Info. Sec. Agency*, Survey and Analysis of Security Parameters in Cloud SLAs Across the European Public Sector* (2011); European Network and Info. Sec. Agency, *Procure Secure: A Guide to Monitoring of Security Service Levels in Cloud Contracts* (2012).

es to user data, not just logs but also real-time 24-hour monitoring, so users may check who has accessed which accounts, what they viewed, and what they did. One SaaS provider stated that it undertakes to log all accesses as standard. It commented that, once this undertaking was pointed out, most users who requested audit rights were content to rely on access logs instead. Provision of such tools by more providers may help increase user trust by increasing transparency, as well as assisting with legal compliance. Some authorities also consider that, apart from access logs, for users to check providers' compliance with required security measures, ongoing post-contract audit rights against all sub-providers would also be needed—IaaS, PaaS, perhaps even down to the data center level.[81] Clarity is needed regarding how far down the "stack" users must go.

A global integrator commented that providers' reluctance to allow audits might partly arise from their not necessarily being able to pinpoint exact locations of users' data.[82] Also, providers may not have sufficient control or rights to allow audits, especially if using sub-providers. In concrete terms, consider a SaaS provider using a PaaS provider's service, which may itself use a third party's IaaS infrastructure or data centers. The SaaS provider may, if it chooses, permit inspection of its application code to verify application security – but it may not be entitled to permit users to audit the PaaS or IaaS provider or visit data centers, or even conduct its own audits of sub-providers. Therefore, rights to audit sub-providers may be problematic. That said, one financial institution was able to require comprehensive audit rights from a United States SaaS provider for itself, affiliates and regulators, including sub-providers and sub-contractors.

One compromise in which integrators were involved was for providers to agree to share their own audit reports with the integrator, or at least allow the integrator to view the reports, and ideally (though many providers might not agree) allow it to share the results with its own users and their regulators. Where audits are disallowed and no independent third party audits are conducted or shared with users, users must rely on providers' undertakings regarding security (if any). Compliance verification there would be impossible. A global user noted, "The only way to find out if they have actually complied is if they have a major breach or loss of confidential information!"

The general lack of audit rights causes difficulties, especially for integrators whose customers require audit rights, possibly even including rights to select auditors. This issue may loom larger in future as audit rights increasingly

---

81. Datainspektionen, *supra* note 72. This view was also taken in mid-2012 by European Union data protection regulators collectively. *See* Article 29 Working Party, *Opinion 05/2012 on Cloud Computing*, Working Paper 196 (Jul. 1, 2012), *available at* http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/2012/wp196_en.pdf.

82. *See* Hon et al., *Data Export in Cloud Computing*, *supra* note 10, at n.36.

come to the fore—described by one user as a "headache," with users being dependent on the (not necessarily consistent) views and actions of individual national regulators. Ongoing third party audits to industry standards could be considered as a possible solution, as with pre-contractual audits, but legislators and regulators need to consider these issues, particularly the extent to which users may rely on such audits for liability purposes.[83]

6. *Security breach notification*

Many providers' standard terms did not require security incidents to be reported to users. This may have been for operational reasons as many providers' systems and processes were not originally set up to enable them to notify users of incidents easily and quickly.

Users have requested notification of data losses or security breaches at, the provider level in, say, twenty-four hours, sometimes even when only other users were affected. Some providers agreed to notify users promptly of breaches or losses, at least where affecting that user (but possibly not if only affecting other users). Others, while not committing contractually to notify users, would in practice attempt to notify users as soon as possible, although in some cases faster notification was possible if the user paid for higher support levels. Still other providers would not commit to notifying users, except perhaps for specialized users such as telecommunications providers who by law need that capability, but are putting into place systems and organizational measures to enable them to do so in future.

For large users, some providers have been willing, if the matter goes to a senior enough level, to accept additional obligations to use commercially reasonable efforts to monitor for and detect breaches, and to notify of breaches (at least where affecting that user), perhaps also with rights for the user to terminate for the breach. The notification period may be short, for example within one business day after the provider becomes aware of an actual or potential threat, at least where it may adversely impact on the user's service, such as unauthorized access. As with other important notices, users may require notification to be in writing to a particular address, not just by e-mail.

As regards actions providers must take after security breaches, the agreed security standards usually addressed the position, for example by requiring isolation or quarantine of affected areas, and otherwise remedying the situation.

On receiving notification, some users wanted their security team to investigate and consider whether the provider had met the required security standards, and if they had not, the user wanted rights to serve notice to remedy the breach,

83. For example, some data protection authorities (for instance, Sweden's, *see* Datainspektionen, *supra* note 72, at 16) consider that independent third party audits would not absolve controllers from responsibility to ensure appropriate security measures were in fact taken.

if not already remedied, or even to terminate the agreement. Unlike with traditional outsourcing, however, many providers would not agree to any joint analysis of the breach with users, but insisted on dealing with breaches themselves.

F. *Lock-In and Exist*

Many users cited lock-in as one of their top concerns with the cloud. There are several aspects to lock-in. Exit strategy and end-of-contract transition were major concerns amongst users, including data portability, and the importance of retaining metadata as well as data. Users may not wish to be "locked in" or tied down for too long an initial contract term—this seems essentially a commercial issue, often linked to pricing, sometimes negotiable. We cover this later.

A major lock-in concern is risk of dependence (or over-dependence) on one provider's, often proprietary, service. If the service is terminated for whatever reason, users wanted to recover all their data and metadata in formats that are easily accessible, readable, and importable into other applications, whether running internally or in another provider's cloud. This is commonly called "data portability." Application portability is one aspect of dependence risk which is not discussed as much as data portability, whether by our sources or in the literature. However, it is equally if not more important, particularly for IaaS and PaaS. As cloud use becomes more widespread and sophisticated, we believe future contract terms may extend to cover application portability, virtual machine portability, and perhaps even interoperability.

1. *Data retention, deletion and data portability*

Two issues arise with data retention and deletion. The first is, will providers retain users' data when needed by users, so users can retrieve data in usable format? The second is, will providers delete user data when required?

There are two main circumstances when users may wish providers to retain data: where data needs to be retained for purposes of regulation, litigation or other legal reasons affecting users; and after contract termination, where users want providers to retain users' data for a long enough period after termination to allow users to recover their data.

Data retention for legally required purposes, such as litigation e-discovery or preservation as evidence upon law enforcement request, has not been negotiated much yet, according to our research. A large enterprise-oriented SaaS provider stated no prospective user had raised the issue until early 2011. We think it will become more important in future. One aspect is how much, if any, assistance providers give users. For example, an enterprise-oriented provider stated, if users needed to retain data for longer periods, such as for tax reasons, they had to arrange their own storage. A global user was concerned about using a particular large provider because it did not know what, if any, processes that provider had implemented for e-discovery, whereas another platform it used did have such processes. Another global user stated it had secured a global pro-

vider's agreement to retain, segregate and secure data if specifically requested, while a SaaS provider specializing in e-mail continuity provides, as standard, tools to assist with e-discovery, offering users a choice of retention periods, longer periods costing more.

These examples illustrate that data retention for compliance is technically possible. However, e-discovery tools for the cloud still seem to be lacking. One user commented that providers often had difficulty understanding that the user would need their help should it receive a relevant request. Users have defined contractually exactly what assistance they may need from providers here.

Users' ability to have data returned upon contract termination has seen more negotiation. Process simplicity may be as important as data format. A United Kingdom public sector educational institution chose one provider's SaaS service over another's partly because it believed it was easier to retrieve data from the former.

There are several aspects here: data format, what assistance (if any) providers will give users, what if anything providers charge for such assistance, and data retention period.

Some providers, especially enterprise-oriented ones, commit routinely to return users' data in standard format (typically CSV) on termination, at least upon the user so requesting. Several users mentioned they use Salesforce SaaS, one benefit being the return of data in common CSV format, as well as, with certain services, the possibility of weekly downloads or e-mails of their data. Some global users requested contractual commitments regarding return of data in their required format after termination for whatever reason, which some providers agreed (at least for reasonable formats), although sometimes this would be at additional cost if substantial quantities of data were involved.

Most providers did not offer any assistance, even contracted paid assistance packages. However, a SaaS provider stated that it assists in providing data in a different format for a set fee if requested by the user, while a global user noted that some providers or integrators would agree to provide assisted migration at contract end. A non-European telecommunications provider stated that it assists with migration in ("onboarding"), and, while willing to assist with migration out, found even its smaller users did not need it, as they took their own direct copies.[84]

Another lock-in issue is how long after termination users have to recover data before deletion. Many providers delete all data immediately or after a short period (often thirty days), but some users obtained longer grace periods, for example two months, perhaps requiring notice to users before deletion. Some providers offered longer periods, such as ninety days. For large deals, parties could agree on eighteen months to two years, with assistance for migration,

84. Cloud Industry Forum recommends that providers should assist with migration or at least allow users sufficient time to self-migrate. *CIF3*, *supra* note 13, at 15.

more like classic outsourcing deals—although six months might suffice if providers agreed to provide "reasonable assistance." The period needed to migrate user applications and data would of course depend on the circumstances.[85]

Regarding deletion, standard terms may require providers to delete data after termination only if the user so requests, or are silent on deletion. However, depending on the type of service and intended usage,[86] users have sought contractual commitments to ensure deletion of data from the provider's systems (including any duplicates or backups). This is relevant both when users delete data while using a cloud service, and after contract termination. Deletion after termination may be particularly important with personal data, including data held by any sub-processors.

As with local computers, when "deleting" cloud data, often data are not actually deleted. Instead, "pointers" recording locations of different data fragments, are deleted, and data are overwritten by fresh data over time.[87] Deleting data, rather than pointers, is more complex, and for better security requires overwriting a minimum number of times, the ultimate measure being secure destruction of physical storage media. Deleting data from backup tapes is even more difficult. Also, if cloud data move between different equipment automatically, data may remain in previous devices until overwritten.[88]

One SaaS provider splits data into smaller portions randomly after termination of the user's account, so it would be virtually impossible for users (although not the provider) to reconstruct data. It had received requests to commit to stronger deletion, but would not agree, although it would certify it no longer held data after their return. A global provider would agree to delete all duplicates (although it was unclear to what standard[89]), and also to certify deletions.

There are different locations and media from which data may require deletion, and different levels of deletion, up to destroying physical hardware securely. The degree of deletion required depends on the level of security required for the data concerned.

More secure data deletion is more expensive for providers (particularly destroying rather than redeploying equipment). Accordingly, they did not employ

---

85. This paragraph was derived from our interviews. Additionally, Cloud Industry Forum recommends that providers should give at least thirty days' notice before deletion. *Id.*

86. Data deletion is often an issue in the cloud, but may be considered unimportant for some services. For example, for a SaaS service involving temporary processing but no permanent storage of data, a global user was not concerned about data deletion provisions even though personal data might be involved.

87. Google notes this in its standard terms for Google Apps. *Google Apps for Business (Online) Agreement*, http://www.google.com/apps/intl/en/terms/premier_terms_ie.html (accessed Nov. 15, 2012), at section 10.4.

88. This issue was not generally covered in contracts, although it is noted by, for example, German data protection authorities' cloud guidance, *see* DATENSCHUTZBEAUFTRAGTEN, *supra* note 18.

89. For an example deletion standard, *see* CODE OF PRACTICE, Secure destruction of confidential material, BS EN 15713 (British Standards Inst. 2009).

more secure methods unless necessary, and would want to pass on costs to users. Deletion seems as much a financial as technical issue. Costs may be partly why, as a global user noted, data deletion has not hitherto been a particular focus of cloud. A global integrator commented that providers would offer to delete data, but were reluctant to do so to ISO standards as they wished to reuse hardware. Where secure deletion was vital, as with financial institutions or telecommunications providers, some providers were willing to guarantee it—but usually only at greater cost.

Some global users had successfully required providers to delete data if requested, providing evidence of permanent deletion of all copies. Some users would not contract with providers who did not so agree. A non-European telecommunications provider agreed to contractual terms requiring deletion, including specific provisions for overwriting data to recognized data deletion standards to make data unrecoverable, if larger users so requested. One global user secured additional obligations to deliver all backups of stored data and applications within a short period after termination. However, the most another global user secured was acknowledgement of data confidentiality, and warranties that data would be deleted following termination.

During the term, some users wanted rights to ensure data they deleted were deleted permanently, including all duplicates, such as after receiving complaints about end user intellectual property rights breach, or law enforcement requests for deletion. To assure the third party regarding deletion of all offending content, one compromise was to agree that providers must use reasonable endeavors to delete data and erase relevant storage media, when specifically requested by the user. Providers might also need capabilities to quarantine rather than delete data, for example with intellectual property rights disputes.

There may be an educational issue: users and third parties may need more explanation regarding degrees of deletion, depending on nature and sensitivity of content. If certain data are no longer accessible for most purposes, that might suffice in some situations, whereas more sensitive data such as customer payment records might require absolute destruction, or at least more secure deletion (and more stringent or frequent audits). The same situation arises where third party contracts, like certain content licenses, restrict where users may store data or applications. One provider noted another educational issue: users may need to make employees aware that, for example with SaaS services, 'deletion' often merely transfers data to a "recycle bin," stored for say 90 or 180 days, before deletion.

G. *Term and Termination*

1. *Minimum term, renewals and notice periods*

Our sources felt that, before contracting, users needed to consider carefully the minimum and maximum term acceptable to them and their exit strategy, and ensure their practical requirements were addressed contractually if neces-

sary. A long initial term may be one aspect of lock-in.

One global user felt there was some "cloud washing," in that certain services described as "cloud" did not in fact utilize typical cloud pricing models or cloud technologies, for example because those services required minimum revenue and term commitments (unlike the "pay as you go" model meant to typify cloud), new servers had to be procured if more capacity was needed (instead of instantly increasing capacity on demand). However, even with acknowledged cloud services like Salesforce, a minimum term was often said to be required. A global user pointed out that larger users were better able to "beat cloud providers down on price," and therefore deals might be less commercially viable for providers unless longer-term. Bigger deals, or deals involving more sophisticated or customized services, were more likely to involve an initial minimum term, and there are indications that initial fixed terms are increasingly common. The question of initial minimum term is more a commercial than a legal issue, with some providers being willing to reduce unit or per-user prices for longer initial terms. Some providers wanted early termination fees (which may be "huge") if users terminated a fixed-term contract earlier for convenience, as recovery of fixed set-up costs were designed to be spread over the term.

Providers preferred fixed-term, at least initially, to rolling monthly contracts. Standard terms typically stipulate a one to three-year initial term, sometimes renewing automatically unless terminated. Cloud Industry Forum research[90] indicated forty-six percent of United Kingdom users' contracts renew automatically, particularly with smaller organizations, but only thirty-eight percent of integrators' contracts. The critical difference was that sixty-four percent of integrators' internal business practices incorporated early warning systems to manage renewals proactively.

As contracts often require notice of non-renewal within a set period before expiry, users could miss the window. Many users successfully requested deletion of automatic renewal provisions, or increases in the period before term expiry within which users could give notice, say from thirty days to sixty days. Because some users had apparently failed to understand the rollover mechanism, one SaaS provider even sends automated reminders at intervals in advance of that period. It seems users may need to improve contract terms or internal processes regarding renewals.

Length of initial term, and therefore period of lock-in, varied with type of service and deals. Some basic click-through SaaS services may be on a rolling basis, say monthly or ninety days. However, other SaaS services, particularly large deals, may have an initial fixed term and may involve master agreements. Terms sought by integrators depend on their customers' requirements. Some users, like financial institutions, might require even longer initial terms with guaranteed renewals, because they needed price ceilings over a longer period

90. *CIF3*, *supra* note 13, at 6; *see also CIF2*, *supra* note 33, at 12.

and required significant transition periods. However, providers generally did not offer initial periods as long as five years, so such users had to agree to shorter terms than they wished, and must devise advance plans for dealing with end-of-term issues.

Where terms permit termination for convenience by either party, notice period was another issue. To give enough time to migrate providers, users often needed longer notice from providers regarding service termination, than for notice merely of service changes, such as several months instead of one. This was particularly so with mission-critical applications, or with IaaS/PaaS where users might need to modify their hosted application's code to run on another provider's cloud. Accordingly, many users stated they always requested longer than the typical thirty days. Thus, portability and migration of applications, not just data, can be important for users. Ideally those aspects should be checked during due diligence (such as testing ability to export data to desired formats), rather than relying on the contract, as data and applications may require migration not only on contract termination but also should providers become insolvent or close down.

Conversely, some users wanted rights to terminate for convenience on shorter notice. For example, one global user negotiated longer notice from its provider than it had to give itself, although identical notice periods for both is not uncommon.

Users also risk lock-in in practice if, for example, their developers mainly use one provider's IaaS/PaaS service. While they can develop and leverage expertise in that service, they may also prefer to continue using it, given inevitable learning curves with other providers' services. Some users therefore encouraged employees to use several providers, to avoid over-reliance on one provider's service and its (possibly proprietary) application programming interfaces. If cloud services become standardized, the "internal expertise" issue may become less important. The use of proprietary versus open source cloud infrastructure will probably become a more significant issue in future.

2. *Termination events*

Insolvency and material breach are common events allowing termination. One global user did not want providers to terminate for anything except non-payment, as even an insolvent user may need continued use of cloud services while winding down business. A financial institution using its own standard terms stipulated non-payment as the only event entitling providers to terminate. However, generally providers seemed unwilling to remove other termination events, such as material breach and insolvency. One provider considered its termination events "set in stone" and would never agree to changes, unless perhaps the deal was large enough and the user would walk away without that change.

Regarding termination for non-payment, some users increased the notice period given by providers before such termination. For other types of breaches,

integrators and other users who use cloud to service their own multiple end users had problems with providers' rights to terminate immediately on issues such as material breach, breach of acceptable use policies (AUPs) (covered below), or upon receiving third-party complaints regarding breach of their intellectual property rights. Such users did not wish actions of one end user customer to trigger rights to terminate the whole service.[91]

However, many services lack granularity. For instance, an IaaS provider may not be able to locate and terminate the offending VM instance, and therefore need to terminate the entire service. Providers, while acknowledging this deficiency, still refused to change terms, but stated they would take a commercial approach to discussions should issues arise. Nevertheless, some users managed, though not without difficulty, to negotiate for notification from providers of any third party complaints regarding intellectual property rights infringement, no termination or suspension of service without further notification, and cooperation with the user as it tried to resolve the matter with the relevant end user or third party, perhaps by terminating just the offending end user's account with it. Users with multiple applications also sought to limit termination to a particular application, rather than all applications hosted with that provider.

Breach of providers' AUP may be a specific termination event, or material breach entitling providers to terminate.[92] AUPs tend to be "take it or leave it" and were not often negotiated, possibly because such terms were generally accepted as reasonably standardized, and providers considered them "hardcore." For instance, continually exceeding agreed usage limits would be considered unacceptable. Nevertheless, one global user negotiated successfully with a global provider for a less restrictive AUP, with fewer usages being stipulated as unacceptable.

For many users, AUPs were unimportant, although rights for providers to change AUPs unilaterally were not, as AUP breaches usually entitle providers to terminate. As mentioned, AUPs were problematic for integrators and others using cloud to serve their own end users, particularly consumers. One end user or customer breaching a provider's AUP may put the integrator or user in breach, enabling the provider to terminate services for the user and all other end users. Therefore, such users ensured that their end users were contractually obliged to comply with providers' AUPs, with end user accounts being terminable for breach. Instead of standard providers' terms enabling instant termination, they required providers to give notice before termination for breach of AUPs (except perhaps for material breaches), of say thirty days to enable users

---

91. Forty-six percent of United Kingdom integrators had actively engaged with providers to define and determine termination events. CIF3, *supra* note 13, at 9. Perhaps smaller integrators had not yet focused on this issue, and/or few integrators polled were providers themselves.

92. Cloud Industry Forum research showed 70% of United Kingdom users had checked AUPs to ensure they were comfortable with them. *Id*.

to remedy the breach, and obliging providers to consult before termination. One global integrator considered this a "huge" issue that did not get attention because of emphasis on liability and service levels. In its experience, most providers were willing to give some notice before termination, but not necessarily to afford users opportunities to remedy breaches, which it considered vital.

Regarding users' termination rights, some users ("only the better lawyers," one SaaS provider noted) wanted termination rights for change in control of providers, for example if taken over by the user's competitor, with the usual issues regarding definitions and scope of "change of control," while users such as financial institutions wanted rights to terminate if required by regulators, law or regulation. Providers generally did not agree, except perhaps for large deals in specific circumstances with clear definitions of "competitor." While material breach is a common basis for termination, some users secured specific termination rights for defined breaches, such as breach of confidentiality, security policies or intellectual property rights provisions.

3. *Suspension*

Standard terms usually reserve rights to suspend services, such as for non-payment. This obviously affects users' services, and some requested changes, which were sometimes agreed. However some providers had no suspension rights, preferring simply to terminate for breach.

Like termination, suspension for breach of AUPs etc, was a particular issue for integrators or users with multiple end users, as services for all end users could be suspended for one end user's actions or omissions. Accordingly, one global user would not permit suspension except with prior notice and its agreement. An integrator would not allow suspension for any reason other than non-payment, unless prior notice was given, including reasons for suspension, so it could notify its end users and discuss it with them as appropriate. Another global user similarly required prior written notice of non-payment, with a final notice, before suspension was allowed, and a commitment to restore services within a certain number of days after payment. Other global users, including an integrator, agreed to allow suspension for breach, but again only after reasonable prior notice (which might be quite long in some cases), good faith consultation with the user, or other requirements. One financial institution did not permit suspension on any grounds, which the provider accepted.

Of course, suspension for reasons unrelated to users may be necessary, such as following a security incident, or to deal with technical service problems, so users generally have agreed to that kind of suspension.

H. *Changing Service Description or Features*

Many standard terms allow providers to change certain or all contract terms unilaterally.[93] This was seen by interviewed users as unacceptable. Again, enterprise-oriented providers were more likely to agree, or stipulate as standard, that amendment was only permitted if agreed in writing by users, or at least that users would receive prior notification whereupon they could terminate.[94] The latter was more common.

In particular, providers' right to change unilaterally service features, functions or even service description, was much negotiated. One non-European telecommunications provider stated that it notified its (enterprise) users of service feature changes, but would not necessarily contractually undertake to do so. A SaaS provider stated it never discontinues features, only introduces improvements, which customers could choose to enable. Its standard terms already stipulated that new features must not materially decrease users' functionality. The importance of this issue, and how much users negotiated it, varied with type and usage of service, extent of termination rights. For example, where users could terminate for convenience on giving notice, if a user disliked a feature change it might simply terminate.

With commoditized SaaS services, users might have to accept providers' rights to change features, although many users still wanted a qualification that changes must not adversely affect their service. For IaaS and PaaS, however, changes might be more significant, as they could result in users having to rewrite application code created to integrate with proprietary provider application programming interfaces, which users obviously wished to avoid. With large deals for mission-critical services, for example a global user's main platform for servicing its own end users, a provider's refusal to change this term could be a deal breaker. Therefore, users have insisted providers cannot change core services without consent, although minor changes to service features or support aspects might be permitted without notification. Service improvements were permitted. If changes were materially detrimental to their service, however, some users negotiated rights to terminate (or even to reject changes), at least upon the current contract period expiring, without liability but possibly without any rebates either.

Users generally also wanted longer prior notification of key changes and their impact, of at least thirty days or more. Certain providers who would not relinquish rights to change service features unilaterally still agreed to notify such changes, although not necessarily for longer periods or even in advance. Users wanted notice periods long enough to allow them to assess changes, dis-

93. Bradshaw et al., *supra* note 2, at § 5.2.

94. Cloud Industry Forum recommends providers should not be entitled to change terms without consent, or at least should give users notice and allow them to terminate. CIF3, *supra* note 13, at 14.

cuss them with their own customers where relevant and, for changes considered detrimental, perhaps commence negotiations with new providers and give notice to terminate the contract. If providers guaranteed a longer lead time, say a year, before introducing notified changes, users were more comfortable with being able to adapt in time.

I. *Intellectural Property Rights*

Intellectual property rights issues frequently arise regarding cloud-processed data and, or, applications, including the cloud service itself. Indeed, for one global SaaS provider, they were negotiated the most.

Providers' terms may specify they own deliverables, for example documentation. Some users wanted clarification that users retained ownership of cloud-processed data, confidential information and so on. Some providers' standard terms do so provide.[95]

Standard terms may not address who owns rights to applications users develop or deploy on IaaS/PaaS. Some users wanted clarification that users own such intellectual property rights. However, the line is sometimes unclear between a user's application and the provider's platform and integration tools. In one deal, involving customization, a global user secured only an exclusive use period. Where integrators develop applications for their own customers, customers might require intellectual property rights ownership, or at least rights to use the software free after contract termination or transfer.

Another issue of contention concerned ownership rights to service improvements arising from user suggestions or bug fixes. Providers may require users to assign such rights. Yet users may not want their suggested improvements to be made available to competitors. Such users sought to prohibit provision of those improvements to providers' other customers, without the user's consent. This issue has arisen even in a deal involving open source software, where the user could not claim rights, let alone forbid using bug fixes for competitors.

Some cloud services include application licenses, and some users wanted clarification that pricing covered such licenses. However, other services do not include application licenses. To install user-sourced third party applications on IaaS/PaaS, or even some SaaS, users must "bring their own" licenses. Providers wanted express clarification of users' entitlement to load and use such applications on providers' infrastructure, even if the provider could not run them itself.

---

95. Cloud Industry Forum research found nearly 75% of United Kingdom users (and 68% of integrators) were content that providers' contracts did not allow providers to take ownership of data/IPR. *CIF3*, *supra* note 13. *See* Chris Reed, *Information "Ownership" in the Cloud,* QUEEN MARY UNIVERSITY OF LONDON, SCHOOL OF LAW, Legal Studies Research Paper No. 45/2010 (2010), *available at* http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1562461.

Only some application licensors allow users to "port" on-premise licenses to a cloud environment. Although not directly affecting contract terms, such licensing may be problematic for users. For example, logging usage in a VM may be unworkable, as they are continually instantiated or terminated, which may make it impossible to identify VM locations and which run licensed software. Rights to bring a set number of licenses to the cloud, irrespective of VM location, would assist. One SaaS provider noted that, even when existing licenses could be "converted" to cloud, including the payment model, licensors' sales or marketing teams did not necessarily publicize that benefit, possibly because the commission structure was not so remunerative.

Licenses are charged on different bases, such as annually in advance for on-premise, and monthly rolling per user for "in-cloud." Some licensing schemes were preferred by SaaS providers for better matching provider-user payment models and being more suitable for public multi-tenanted cloud environments, such as Citrix's, which allows monthly per-user payments, as compared with, for example, Oracle's, which charges based on number of processor cores in the system used.[96]

CONCLUDING REMARKS

With initial adopters of cloud computing mostly being individual consumers or small and medium-sized enterprises, cloud computing epitomizes information technology's increasing consumerization as well as commoditization. The common use of providers' standard contract terms in cloud computing reflects the consumer distribution model. However, many factors are combining to force providers to become more flexible in their terms.

From the supply side, integrators, traditional information technology services vendors and telecommunications providers that are willing to accept more risk are increasingly entering the market, seeing the opportunity to sell more robust, enterprise-grade services, with contract terms to match—rather than the "as is, where is" services such as Amazon EC2 and Google Apps offer. Increasing availability of open source components such as OpenStack may also facilitate market entry by new providers, which may increase competition.

The market, while becoming more sophisticated and transparent than three years ago, seems to be fragmenting. There will still be bigger providers offering generalized "one size fits all" commodity services. However, niche providers and integrators are emerging, who are more willing to tailor services to user

96. *See, e.g.*, Bill Claybrook, *Warning: Not All Cloud Licensing Models Are User-Friendly*, SEARCHCLOUDCOMPUTING (Aug. 2011), http://searchcloudcomputing.techtarget.com/feature/Warning-Not-all-cloud-licensing-models-are-user-friendly. Oracle's charges equate each virtual core with a physical core (although one physical core could support several virtual cores) or vary with the number of virtual cores that Amazon EC2 instances used possess.

needs, whether contract terms or service features. Therefore, in order to remain competitive, providers may have to be more aware of user concerns, more flexible in negotiations, and more willing to demonstrate the security and robustness of their services.

Even large providers are realizing that, to gain or keep customers from certain sectors, they must adapt accordingly. Several global providers are offering different services with different pricing and sets of terms, from consumer-orientated to enterprise-oriented, with specific terms for certain market sectors or functionality, such as third party-certified services.

Large users who were able to require that contracts be on their own standard terms are moving to make their terms more cloud-friendly, with some financial institutions beginning to consider producing standard SaaS and possibly even IaaS terms. Indeed, there may be scope for coordinated collective efforts on the part of financial institutions to collaborate on producing suitably balanced standard terms for cloud contracts in consultation with providers, obviously bearing in mind competition law restrictions.

In the middle or low value markets, choice and information are still limited, and many contract terms are still inadequate or inappropriate for users' needs; yet they may lack the bargaining power to force contract changes. Negotiations by large users with large providers, which have helped educate providers about users' issues such as privacy and security so that providers take due account of users' compliance concerns, will probably filter down to the middle market at least. This is because that market is potentially very large, and therefore attractive to providers. That may result in standard terms offered for mid-sized deals becoming more user-appropriate over time.

As for low value deals, action regarding inadequate contract terms may be needed by legislators or regulators, such as consumer protection or data protection authorities. However, legislators and regulators also need education regarding cloud computing technologies and business structures, so that laws and regulation enable cloud computing to be used in a balanced way so as to realize the potential economic benefits. Some of the current difficulties arise, not necessarily because contract terms are poor, but because data protection and financial services laws assume certain things that are not true in the cloud. In particular, current laws assume controllers' absolute control over processors, one-to-one relationships rather than one to many, dedicated instead of shared infrastructure, and processors who actively process data for controllers instead of renting out self-service resources.

With customized managed private cloud services on dedicated infrastructure, providers may be more flexible on contract terms. However, commoditized public cloud services on shared infrastructure are a very different proposition. They are cheap because they are standardized. One provider felt the biggest challenge was that users wanted the lowest price, but the highest specifications and features, such as location monitoring or audit rights. Forcing providers to accept more liability and incur the expense of upgrading their infrastructure, while asking them to maintain low commodity prices, does not seem

an appealing proposition for providers, which may itself undermine market development. If one provider's infrastructure is not secure enough for personal data, users should choose another provider that does provide sufficient security.

In other words, rather than stipulating mandatory terms for commodity cloud, a better course for policymakers may be to encourage a greater range of available cloud services (with different sets of terms) which users can assess, choosing the service best suiting their needs—whether cheap public clouds for data that are neither personal nor commercially confidential data, more expensive "personal data" clouds, or even more expensive, high security, auditable private or community clouds, such as sector-specific clouds for financial services or healthcare institutions.

Even mid-sized organizations, which may wish to process confidential or personal data in the cloud, are unlikely to have expertise to assess providers' security measures. Therefore, in order to enable consumers and small and medium-sized enterprises to consider and compare cloud services properly, work is also needed on industry standards and certifications, for sub-providers as well as direct providers, including standards on data portability and interoperability as well as security. Suitable standards and certifications, with provision for both self-certification and independent third party certifications, might then form the basis for laws that recognize appropriate certificates, trustmarks or seals as adequate for various compliance purposes.

In order for appropriate standards and certifications to be accepted by legislators and regulators, however, more openness and transparency are needed from providers. In particular, they need to explain more clearly and in more detail how data location relates to data security, including whether data may be secured against remote access by unauthorized persons even if located outside the European Economic Area or moved between different equipment, including data fragmentation and data structures. They need to spell out clearly the manner by which they are able to access users' data or monitor users' processing, if at all. They should perhaps also develop tools to, for example, enable users to verify accurately the locations of data and VM instances.

The increasing involvement of insurers in cloud is also relevant, as insurers may be better able to assess risks than smaller users. However, users may also need to consider actively insuring against providers' breaches, outages, data loss, and so on, checking that the coverage is appropriate.[97] One provider pointed out that, almost irrespective of how much liability providers accept, it was critical for users to understand cloud "layers" and data location.[98] Some

97. Forty-three percent of users had insurance for business interruption due to the provider's disaster or data leak, but 37% did not know if they would be covered, although 65% of users expected their providers to cover these risks! *CIF3*, *supra* note 13, at 9.

98. As a consumer example illustrating the risks of ignorance regarding layers, users of the online backup service Backify, which is based on the Livedrive service, lost their data after a dispute between Backify and Livedrive. *See* Brid-Aine Parnell, *Punters Lose Backups in Cloud Storage Biz Spat,* THE REGISTER (Nov. 17 2011), http://www.theregister.

smaller users have been contracting without understanding virtualization, the vertical supply chain or layering of providers; some even had no idea who provided the service. Ignorance of cloud structures may result in risks not being properly addressed. There is clearly also a need to help users climb the cloud learning curve. Users may for example need to consider which if any functions should be migrated to cloud and on what basis, such as starting with pilots only, conducting risk assessments, and implementing internal controls. They need to recognize that, quite apart from contract terms, they may need to take other practical measures, in particular pre-contractual due diligence and testing, encryption of data as appropriate, backing up internally or to another provider when using low-cost services without guaranteed backup, and post-contract monitoring.

In summary, while contract terms have been negotiated for larger deals, the small and medium-sized enterprise user market is unlikely to benefit in the short to medium term. While changes to providers' standard contract terms should filter down from large deals, and up from regulatory action regarding consumer and other deals, a multi-pronged approach may be the best solution, where all types of players are involved to encourage the development of a full variety of cloud services and contract terms priced at different levels, with standards and certifications to assist with legal certainty regarding compliance.

---

co.uk/2011/11/17/livedrive_backify_dispute/.

STANFORD TECHNOLOGY LAW REVIEW
VOLUME 16, NUMBER 1 FALL 2012

UNCHAINING RICHELIEU'S MONSTER: A TIERED REVENUE-BASED COPYRIGHT REGIME

Martin Skladany*

CITE AS: 16 STAN. TECH. L. REV. 131 (2012)
http://stlr.stanford.edu/pdf/richelieusmonster.pdf

ABSTRACT

This Article proposes a tiered revenue-based copyright regime, which would require copyright holders to select one of two different copyright terms. The first tier would provide a fixed, nonrenewable copyright term of 10-14 years, while the second tier would offer a one-year copyright term that could be indefinitely renewed as long as the work is successful enough to meet or exceed a revenue threshold.

A tiered revenue-based copyright regime will break the gridlock between copyright proponents lobbying for longer copyright terms and public domain advocates insisting that terms are already remarkably excessive. It will entice Hollywood to set orphans free and to accept dramatically reduced copyright terms for most artwork in exchange for gaining longer protection for the most successful commercial works. It will require all artists seeking copyright protection to register each work and will immediately transfer all works that are not registered—most newly created noncommercial art and all existing orphans—to the public domain. It will increase the speed at which the overwhelming majority of commercial art moves into the public domain, because artists selecting the first tier would have only 10-14 years of copyright protection—a much shorter term than current law provides. Moreover, it will free much of the commercial art in the second tier within one or a few years of copyright registration, for the revenue-based annual renewal system will be a final filter to ensure that only the most profitable works continue to stand aloof from the commons.

---

* J.D. Yale Law School, M.P.A. Princeton University, M.Phil. Cambridge University. I would like to thank Ian Ayres, Yochai Benkler, Mary Boniece, Jeffrey Cunard, John J. Donohue III, and Caroline R. Sherman for their helpful comments and suggestions.

INTRODUCTION

> Because a knowledge of letters is entirely indispensable to a country, it is certain that they should not be indiscriminately taught to everyone. A body which had eyes all over it would be monstrous, and in like fashion so would be a state if all its subjects were learned; one would find little obedience and an excess of pride and presumption. The commerce of letters would drive out that of goods, from which the wealth of the state is derived.
>
> —Cardinal Richelieu[1]

Unlike Cardinal Richelieu, we do not actively suppress education in the hopes of preserving the public's obedience and the state's wealth. Instead, we stifle individual creativity and our society's cultural vibrancy by maintaining an excessively restrictive copyright regime alleged to spur economic growth. But the notion that indiscriminate increases in copyright protection will further stimulate our economy is a myth. Numerous respected scholars have shown that almost all of the economic benefits from copyrights accrue early on in

---

1. CARDINAL RICHELIEU, THE POLITICAL TESTAMENT OF CARDINAL RICHELIEU 14-15 (Henry Bertram Hill trans., Univ. of Wisc. Press 1961) (1687). Richelieu continues,

> It would ruin agriculture, the true nourishment of the people, and in time would dry up the source of the soldiery, whose ranks flow more from the crudities of ignorance than from the refinements of knowledge. It would, indeed, fill France with quibblers more suited to the ruination of good families and the upsetting of public order than to doing any good for the country. If learning were profaned by extending it to all kinds of people one would see far more men capable of raising doubts than of resolving them, and many would be better able to oppose the truth than to defend it. It is for this reason that statesmen in a well-run country would wish to have as teachers more masters of mechanic arts than of liberal arts.

Id.

almost all copyright terms.[2] Professor James Boyle observes:

> For most works, the owners expect to make all the money they are going to recoup from the work with five or ten years of exclusive rights. The rest of the copyright term is of little use to them except as a kind of lottery ticket in case the work proves to be a one-in-a-million perennial favorite.[3]

Nobel laureates George A. Akerlof, Kenneth J. Arrow, James M. Buchanan, Ronald H. Coase, and Milton Friedman and 12 other economists submitted an Amici Curiae in Support of Petitioners in *Eldred v. Ashcroft*[4] showing that the copyright term extensions enshrined in the 1998 Sonny Bono Copyright Term Extension Act[5] would only increase an author's expected income from a book, assuming the author lives for 30 years after the book's release, by no more than roughly 0.33% or one-third of one percent.[6] At this point in the history of United States copyright, the returns from extending copyright protection have essentially vanished. In his dissent in *Eldred*, Justice Breyer states, "The present extension will produce a copyright period of protection that, even under conservative assumptions, is worth more than 99.8% of protection *in perpetuity*."[7]

This push to extend copyright does not make sense given the historical data. When copyright owners were allowed to renew their holdings after 28 years, roughly only 15% did so in fiscal 1959.[8] Similarly, fewer than 11% of the "copyrights registered between 1883 and 1964 were renewed at the end of their twenty-eight-year term, even though the cost of renewal was small."[9] At least two reasons explain these low figures. First, as mentioned above, demand for most copyrighted artwork exists primarily at its release—not decades later. Second, a dollar earned 40 or 50 years from now is worth only a few pennies today; the longer the copyright terms, the smaller the financial incentive to create with each additional year of protection.

Moreover, copyright protection appears to be little more than an

---

2. HM TREASURY, GOWERS REVIEW OF INTELLECTUAL PROPERTY, 2006, at 52 (U.K.) *available at* http://www.official-documents.gov.uk/document/other/0118404830/0118404830.pdf.

3. JAMES BOYLE, THE PUBLIC DOMAIN: ENCLOSING THE COMMONS OF THE MIND 11 (2008).

4. 537 U.S. 186 (2003) (affirming the legitimacy of the 20-year extension to copyright's duration).

5. Sonny Bono Copyright Term Extension Act, Pub. L. No. 105-298, 112 Stat. 2827 (1998).

6. Brief of George A. Akerlof et al. as Amici Curiae in Support of Petitioners at 10-12, Eldred v. Ashcroft, 537 U.S. 186 (2003) (No. 01-618).

7. Eldred v. Ashcroft, 537 U.S. 186, 255-56 (2003) (Breyer, J., dissenting).

8. SUBCOMM. ON PATENTS, TRADEMARKS, AND COPYRIGHTS OF S. COMM. ON THE JUDICIARY, 86TH CONG., 2D SESS., COPYRIGHT LAW REVISION, STUDY NO. 31, at 187 (Comm. Print 1961) (prepared by Barbara Ringer), *available at* http://www.copyright.gov/history/studies/study31.pdf.

9. WILLIAM M. LANDES & RICHARD A. POSNER, THE ECONOMIC STRUCTURE OF INTELLECTUAL PROPERTY LAW 212 (2003).

afterthought for many artists, not a critical factor in deciding whether to create. When registration was a prerequisite for protection prior to 1909, scores of artists never even sought to obtain copyright for their works. For example, in the United States between 1790 and 1800, of the over 21,000 works published, only 648 works were registered in order to obtain copyright protection[10]—authors only bothered to seek copyright on less than four percent of works created. Examining registration and renewal statistics together, it seems unlikely that potential artists will be motivated to create by the prospect of receiving pennies in copyright royalties decades into the future.

The insignificant added benefit to copyright holders of extending copyright terms does little to increase the incentives that artists face in deciding whether and how much to create.[11] William Patry states, "[I]n my 27 years of practicing copyright law, I have never seen a study presented to Congress that even makes a stab at demonstrating that if the proposed legislation is passed, X number of works that would not have been created will be."[12] In fact, Kai-Lung Hui and I.P.L. Png provide empirical evidence that the Sonny Bono Act's extension of copyright did not increase the creation of United States movies.[13]

Instead, such copyright term extensions leave society with millions of copyright orphans—artwork that is commercially unavailable and without a known copyright owner. A majority of film and book holdings are estimated to be orphan works.[14] By definition, no benefit arises from the protection of these orphans because there are no known copyright holders to receive royalties. Yet millions of orphan works are caught in the equivalent of legal purgatory, with society bearing the harm of being unable to read, listen to, or watch them for free. The effect of the dramatic increase in the length of copyright over the last few decades

> is simply to toll, or delay, the passing of works into the public domain. This latest extension means that the public domain will have been tolled for thirty-nine out of fifty-five years, or 70 percent of the time since 1962. Thus, in the twenty years after the Sonny Bono Act, while one million patents will pass into the public domain, zero copyrights will pass into the public domain by virtue of the expiration of a copyright term.[15]

Such lack of access is not a well functioning market, nor is it in line with the spirit of the Constitution's Progress Clause.

---

10. William J. Maher, *Copyright Term, Retrospective Extension, and the Copyright Law of 1790 in Historical Context*, 49 J. COPYRIGHT SOC'Y U.S.A. 1021, 1023-24 (2002).

11. *Id*.

12. WILLIAM PATRY, MORAL PANICS AND THE COPYRIGHT WARS 62 (2009).

13. *See generally* Kai-Lung Hui & I.P.L. Png, *On the Supply of Creative Work: Evidence from the Movies*, 92 AM. ECON. REV. PAPERS & PROC. 217 (2002).

14. BOYLE, *supra* note 3, at 9 (citing reports by the Center for the Study of the Public Domain at Duke University School of Law for film holdings).

15. LAWRENCE LESSIG, FREE CULTURE: HOW BIG MEDIA USES TECHNOLOGY AND THE LAW TO LOCK DOWN CULTURE AND CONTROL CREATIVITY 134-35 (2004).

Not only does society suffer from not having legal access to a vast collection of existing artwork, but it also loses potential artistic creations; for borrowing from the past can be essential to the creative process, and uniformly long copyright terms make it harder for artists to borrow. Copyright attempts to mediate between two supposed goods: facilitating the "free flow of ideas, information and commerce" and motivating authors to produce by using economic incentives.[16] Extending copyright terms further tightens these constraints and means that "no one can do to the Disney Corporation what Disney did to the Brothers Grimm."[17] This is anything but a minor limitation. Shakespeare and Milton are but two authors of many who created great works unencumbered by such constraints.[18] Individual creativity and our society's culture are significantly impeded not only because copyright terms on works are too long and too broad but also because the continued protection of copyright orphans prevents artists from uncovering such artwork and presenting it anew to the culture. The rediscovery of abandoned works is much more powerful now than even 15 years ago because we have a medium that makes widespread availability almost costless—the Internet.

Numerous other harms result from the current copyright regime and the relentless push to extend terms. Excessively long copyright terms impair free speech principles.[19] By restricting the commons and hence reducing the diversity in sources of information, they enable media to more easily selectively disclose information or skew the nature of what is communicated.[20] Further, a copyright system with exceedingly long terms limits the ability of citizens to inform themselves, because such terms increase the "power over the price of information . . . in the hands of intellectual property owners."[21]

Commercial entities whose revenue derives largely from copyright—what I am calling here "Big Copyright"[22]—have a long history of lobbying for copyright term extensions. Thomas Babington Macaulay and others successfully fought off an attempt to extend copyright over a century and a half ago in the United Kingdom. In an 1841 speech to the House of Commons,

---

16. Sony Corp. v. Universal City Studios, Inc., 464 U.S. 417, 429 (1984).

17. Lawrence Lessig, *The Creative Commons*, 55 FLA. L. REV. 763, 764 (2003). Disney has borrowed many stories and characters from the Grimm brothers, including, for example, Pinocchio and Cinderella.

18. MARK ROSE, AUTHORS AND OWNERS: THE INVENTION OF COPYRIGHT 2 (1993).

19. *See generally* Neil W. Netanel, *Copyright and the First Amendment; What* Eldred *Misses—and Portends*, *in* COPYRIGHT AND FREE SPEECH: COMPARATIVE AND INTERNATIONAL ANALYSES (Jonathan Griffiths & Uma Suthersanen eds., 2005).

20. Yochai Benkler, *Freedom in the Commons: Towards a Political Economy of Information*, 52 DUKE L.J. 1245, 1267 (2003).

21. PETER DRAHOS & JOHN BRAITHWAITE, INFORMATION FEUDALISM: WHO OWNS THE KNOWLEDGE ECONOMY? 4 (2002).

22. Big Copyright does not include all of corporate America. Numerous multinationals, like consumer electronics firms, in fact have traditionally come into conflict with Big Copyright in regard to copyright policy.

Macaulay vividly and skillfully described the costs of extending copyright well beyond the death of artists:

> Dr. Johnson died fifty-six years ago. If the law were what my honourable and learned friend wishes to make it, somebody would now have the monopoly of Dr. Johnson's works. Who that somebody would be it is impossible to say; but we may venture to guess. I guess, then, that it would have been some bookseller, who was the assign of another bookseller, who was the grandson of a third bookseller, who had bought the copyright from Black Frank, the Doctor's servant and residuary legatee, in 1785 or 1786. Now, would the knowledge that this copyright would exist in 1841 have been a source of gratification to Johnson? Would it have stimulated his exertions? Would it have once drawn him out of his bed before noon? Would it have once cheered him under a fit of the spleen? Would it have induced him to give us one more allegory, one more life of a poet, one more imitation of Juvenal? I firmly believe not. I firmly believe that a hundred years ago, when he was writing our debates for the Gentlemen's Magazine, he would very much rather have had twopence to buy a plate of shin of beef at a cook's shop underground.[23]

It has been recognized that the recent successful campaign for an additional 20 years of copyright protection was lobbied for,[24] depending on one's perspective, by the likes of Disney to propagate their own economic welfare, not the cultural, political, and economic interests of society at large.[25] In fact, critics deride the legislation as the "Mickey Mouse Protection Act."[26] Even pro-market publications like *The Economist* talk of "absurdly long copyright periods."[27] The publication states, "Starting from scratch today, no rational, disinterested lawmaker would agree to copyrights that extend to 70 years after an author's death, now the norm in the developed world."[28]

Some have argued that "corporate capture can only be part of the explanation" for the push for longer copyright terms; other factors like the "honest delusion" of "maximalism," which equates more copyright with more innovation, and "authorial romance," which equates invention with absolute originality, are partially to blame.[29] While this is inevitably true, citizens' mistaken precepts do not mean that anyone but Big Copyright hires an army of lobbyists to push for extended copyright terms. Further, through public

---

23. Thomas Babington Macaulay, speech delivered in the House of Commons (Feb. 5, 1841), *in* 8 THE LIFE AND WORKS OF LORD MACAULAY: COMPLETE IN TEN VOLUMES 200-01 (Edinburgh ed., Longmans 1897).

24. Joyce Slaton, *A Mickey Mouse Copyright Law?*, WIRED, Jan. 13, 1999, *available at* http://www.wired.com/politics/law/news/1999/01/17327.

25. For other reasons to possibly dislike Disney, see generally CARL HIAASEN, TEAM RODENT: HOW DISNEY DEVOURS THE WORLD (1998).

26. *Copyrights: A Radical Rethink*, ECONOMIST, Jan. 23, 2003, at 15, *available at* http://www.economist.com/node/1547223.

27. *Digital Publishing: Google's Big Book Case*, ECONOMIST, Sept. 5, 2009, at 18.

28. ECONOMIST, *supra* note 26.

29. James Boyle, *Deconstructing Stupidity*, FIN. TIMES (Apr. 21, 2005), *available at* http://www.ft.com/cms/s/2/39b697dc-b25e-11d9-bcc6-00000e2511c8.html.

relations campaigns Big Copyright deliberately perpetuates misleading justifications of the theory underlying copyright law. Lewis Hyde describes how the Motion Picture Association of America (MPAA) persuaded Californian legislators to mandate "that all public schools must develop an 'education technology' plan" that instructs kids on copyright law through the distorted lens of Big Copyright's willful misunderstanding.[30] Even the Boy Scouts in Los Angeles offers a "Respect Copyright" merit badge; "the MPAA wrote the curriculum for that, too."[31] These campaigns "teach a series of simplifications, even falsehoods, when it comes to the ownership of art and ideas," for example, one lesson falsely states that "intellectual property is no different than physical property."[32]

This extension of copyright terms has been but a part of the larger expansion of intellectual privilege in the last few decades,[33] which some have deemed to be another "enclosure movement."[34] Not only has the length of copyright protection increased—"tripled in the past thirty years"[35]—but also penalties for violating copyright have become more draconian.[36] In addition, what is protected under copyright was expanded in the fields of software,[37] architectural works,[38] and choreographed works.[39] Copyright holders' power to stop derivative uses of their art has increased, and "copyright's reach has changed, as every action becomes a copy and hence presumptively regulated."[40] Further, it has also become generally illegal to circumvent digital rights management (DRM) technologies.[41]

The Constitution states that Congress shall have power "to promote the Progress of Science and useful Arts, by securing for limited Times to Authors and Inventors the exclusive Right to their respective Writings and Discoveries."[42] It goes without saying that the phrase "for limited times" is an explicit Constitutional limitation to the duration of copyright. While the initial copyright term for new works was rather modest—under the first copyright law in the United States in 1790, 14 years plus an option to renew for an additional

---

30. LEWIS HYDE, COMMON AS AIR: REVOLUTION, ART, AND OWNERSHIP 6-7 (2010).

31. *Id.* at 7.

32. *Id*. at 8.

33. Mark A. Lemley, *Property, Intellectual Property, and Free Riding*, 83 TEX. L. REV. 1031, 1042 (2005).

34. *See generally* James Boyle, *The Second Enclosure Movement and the Construction of the Public Domain*, 66 LAW & CONTEMP. PROBS. 33 (2003).

35. LESSIG, *supra* note 15, at 161.

36. No Electronic Theft (NET) Act, Pub. L. No. 105-147, 111 Stat. 2678 (1997).

37. 17 U.S.C. § 117 (2000).

38. 17 U.S.C. § 102(a)(8) (2000).

39. 17 U.S.C. § 102(a)(4) (2000).

40. LESSIG, *supra* note 15, at 161.

41. Digital Millennium Copyright Act, 17 U.S.C. §§ 1201-1205 (2000).

42. U.S. CONST. art. I, § 8, cl. 8.

14 years if one survived—the long-term upward tick in copyright terms is staggering. The next ratchet in term length was to 42 years in 1831, to 56 years (28 years with a renewable option of another 28 years) under the Copyright Act of 1909, to the life of the author plus 50 years (corporate authors—works-for-hire—receiving 75 years from publication) under the Copyright Act of 1976,[43] and then to the life of the author plus 70 years (corporate authors receiving 95 years)[44] in the Copyright Term Extension Act of 1998.

This last act granted copyright extensions not simply to future artwork but to existing artwork, even when the artists who created the pieces are already dead. This change demonstrates that incentives are often irrelevant to the drive to expand terms, since incentives will clearly neither motivate people to do something they have already done nor inspire the dead.[45] More often than not, legislators and judges have not used historical understanding to analyze this upward surge. For example, many countries' borders or forms of government do not survive 70 years, yet copyright terms can easily last for 100-130 years depending on an artist's age when she created a work. In fact, when the United Nations was established less than 70 years ago in 1945, there were only 51 U.N. member nations, compared to 193 in 2011.[46]

Any social justice movement has to decide whether it is strategically worth directly opposing corporate America on a particular point, or whether it should instead formulate a proposal that improves the status quo without upsetting corporate America's prerogatives. The decision whether to fight the giant or search for a policy solution that does not significantly unsettle it depends on numerous factors, including whether the social movement has a practicable plan to improve the status quo and also how much harm corporate America is actually inflicting on society.

---

43. Numerous Congressional enactments extended the second term for renewed copyrights that were to expire between September 19, 1962 and December 3, 1976, to the end of 1976. *See* Pub. L No. 87-668, 76 Stat. 555 (1962); Pub. L. No. 89-142, 79 Stat. 581 (1965); Pub. L. No. 90-141, 81 Stat. 464 (1967); Pub. L. No. 90-416, 82 Stat. 397 (1968); Pub. L. No. 91-147, 83 Stat. 360 (1969); Pub. L. No. 91-555, 84 Stat. 1441 (1970); Pub. L. No. 92-170, 86 Stat. 490 (1971); Pub. L. No. 92-566, 86 Stat. 1181 (1972); and Pub. L. No. 93-573, §104, 88 Stat. 1873 (1974).

44. For works for hire, the 1976 Act provides either 75 years from publication or 100 years from creation, whichever ends first. The Copyright Term Extension Act lengthened these terms to 95 and 120 years, respectively. Pub. L. No. 105-298, §102, 112 Stat. 2827 (1998).

45. Thomas Jefferson designed the initial United States copyright term to last no longer than the life of the copyright holder. While his 19-year term recommendation, which he calculated using actuarial tables, was never passed, the 14-year provision under the 1790 law, mentioned above, with an option for a second 14-year renewal if the artist was still alive, essentially freed the protected works at the grave of the author. PAUL K. SAINT-AMOUR, THE COPYWRIGHTS: INTELLECTUAL PROPERTY AND THE LITERARY IMAGINATION 125 (2003).

46. UNITED NATIONS, *Member States: Growth in United Nations Membership, 1945-Present*, http://www.un.org/en/members/growth.shtml (last visited Dec. 20, 2012).

Jon Pareles has stated: "Any song that is well enough known to make a takeoff worthwhile has probably already raked in plenty of profits from sales, licensing agreements, sheet music, etc. Sometimes I'm tempted to suggest that any song that has sold more than a million (or maybe two million or five million) copies ought to go directly into the public domain, as if its fans have ransomed it from the copyright holders."[47] I find Pareles's suggestion not only innovative but also appealing—its ransom imagery touches on copyright as a monopoly that was supposedly bargained for with society and hence should be theoretically subject to renegotiation. Yet Pareles's idea obviously provides no enticement to Big Copyright to agree to such a bargain. If the public fought a protracted, rough battle, it might prevail with Pareles's suggestion as law, yet even so such a prize would not get society what is more important—free and quick access to the over 99% of less successful commercial and noncommercial artwork that is currently restricted by copyright. Further, if as a society we want rapid and marked improvement on the issue of copyright's length, we need to attract, not repel, Big Copyright.

In the words of Alexander Bickel, "No society, certainly not a large and heterogeneous one, can fail in time to explode if it is deprived of the arts of compromise, if it knows no ways of muddling through. No good society can be unprincipled; and no viable society can be principle-ridden."[48] Edmund Burke previously expressed a similar sentiment, stating, "All government, indeed every human benefit and enjoyment, every virtue and every prudent act, is founded on compromise and barter. We balance inconveniences; we give and take; we remit some rights, that we may enjoy others; and we choose rather to be happy citizens than subtle disputants."[49] Compromise can at least be a necessary evil.

When some of the more prominent copyrights held by Disney once again approach their expiration dates, corporate America will lobby anew for increased copyright terms.[50] The danger of this is, as Professor Peter Jaszi has notably suggested before the Senate Judiciary Committee, that Congress can legislate a perpetual term "on the installment plan."[51] Crusaders against such

---

47. Jon Pareles, *Parody, Not Smut, Has Rappers in Court*, N.Y. TIMES, Nov. 13, 1993, *available at* http://www.nytimes.com/1993/11/13/arts/critic-s-notebook-parody-not-smut-has-rappers-in-court.html.

48. ALEXANDER M. BICKEL, THE LEAST DANGEROUS BRANCH: THE SUPREME COURT AT THE BAR OF POLITICS 64 (Yale Univ. Press, 2d ed. 1986).

49. Edmund Burke, On Moving His Resolutions for Conciliation with the Colonies (Mar. 22, 1775), *in* 1 SELECT WORKS OF EDMUND BURKE 221, 278 (E.J. Payne ed., Liberty Fund 1999).

50. Professor Bell has argued that copyright "exhibits means and ends remarkably similar to those of social welfare programs" and that such an analogy can instruct us on "understanding copyright as a statutory mechanism for redistributing rights." Tom W. Bell, *Authors' Welfare: Copyright as a Statutory Mechanism for Redistributing Rights*, 69 BROOK. L. REV. 229, 229 (2003).

51. *The Copyright Term Extension Act of 1995: Hearings on S. 483 Before the S.*

encroachment have an important strategic decision to make: to outright reject further increases in copyright protections or to propose a new copyright system that accommodates both the segment of corporate America that reveres copyright and those of us who desire a culture less restricted by it.

Taking this second approach, I propose a tiered revenue-based copyright regime.[52] It would give the "one-in-a-million" copyright holder the ability to cash out her lottery ticket, without having to derail our culture. It would do this by presenting all artists with two different copyright terms, which they would have to choose between. The first tier would provide a fixed, nonrenewable copyright term of 10 to 14 years, while the second tier would offer a one-year copyright term that could be indefinitely renewed as long as the work is successful enough to meet or exceed a revenue threshold.

A two-tiered revenue-based copyright regime will break the gridlock between Big Copyright lobbying for longer copyright terms and public domain advocates insisting that terms are already remarkably excessive. It will solve the problem of exceedingly long copyright terms for most artwork in exchange for giving Big Copyright the opportunity to have much longer copyright protection on its most successful commercial works. It will immediately free millions of orphaned works, and its structure will institutionally preclude orphans from reemerging. It will require all artists seeking copyright protection to register each work, and thus keep most noncommercial art in the public domain. It will increase the speed at which the overwhelming majority of commercial art moves into the public domain, because artists selecting the first tier would have only 10 to 14 years of copyright protection—a much shorter term than current law provides. Moreover, it will free much of the commercial art in the second tier within one or a few years of copyright registration, for the revenue-based annual renewal system will be a final filter to ensure that only the most profitable works continue to be excluded from the commons.

This proposal addresses only copyright length, not other ills plaguing copyright like its excessive breadth and depth. Further, the proposed copyright regime is limited to artwork—e.g., I do not consider copyright on software. Also, it is assumed that if a tiered revenue-based copyright regime is implemented, the United States would have to withdraw from at least the Berne Convention, which compels all signatories to subscribe to a minimum copyright term of the life of the artist plus 50 years and requires that no formalities be placed on artists.[53]

Comm. on the Judiciary, 104th Cong. 72 (1995) (statement of Professor Peter Jaszi, American University, Washington College of Law).

52. There is a different and unrelated proposal for using tiers within copyright law based on the originality of copyrighted material. *See generally* Gideon Parchomovsky & Alex Stein, *Originality*, 95 VA. L. REV. 1505 (2009).

53. Landes, Posner, and I do not see withdrawal from Berne as an insurmountable impediment. Landes and Posner make explicit the need to withdraw from Berne for their proposed indefinitely renewable copyright scheme: "This would require the United States to

Part I expounds on the details and advantages of a tiered revenue-based copyright system. It also discusses numerous ways to modify such a regime. Part II demonstrates the advantages of such a proposal over other proposed revisions to the length of copyright's term. Part III addresses six potential objections to a tiered revenue-based copyright regime. Finally, Part IV concludes by arguing that the proposed regime represents a significant step in our progress toward a more expansive and vibrant public domain.

I. THE PROPOSAL

In the tiered revenue-based copyright regime,[54] copyright holders would have to select one of two tiers or tracks of copyright protection for their artwork.[55] Both tiers would require registration of artwork, preferably online given the lower transaction costs and ease with which the public can check on the copyright status of registered works. Online registration might also encourage the public to track the revenue claims of copyright holders.[56]

Tier One would grant a work automatic copyright protection for a set period of time—e.g., 10 or 14 years—without any option to renew the copyright. Protection for 10 years is reasonable because, as already stated by Boyle, this is the upper range of protection from which almost all copyrighted artwork will bring in revenue. The suggested alternative of 14 years is simply a historical nod to the length of the initial term in United States copyright law, minus the possibility of a 14-year extension. While such relatively short terms might seem radical to some, *The Economist* has proposed going back to this original term from 1790.[57] A third possible term length would be 20 years to align copyright terms with patent terms. The point of Tier One is to introduce

---

withdraw from the Berne Convention . . ." LANDES & POSNER, *supra* note 9, at 215 n.15. They cannot see such withdrawal as insurmountable or they would not have suggested such a scheme. *But see* Christopher Sprigman, *Reform(alizing) Copyright*, 57 STAN. L. REV. 485, 552 (2004).

54. In researching another article, I stumbled across the work of Hala Essalmawi from Egypt. She tangentially mentions using revenue as a determinant of copyright length. Hala Essalmawi, *Options and Alternatives to Current Copyright Regimes and Practices*, *in* ACCESS TO KNOWLEDGE IN THE AGE OF INTELLECTUAL PROPERTY 627, 630 (Gaëlle Krikorian & Amy Kapczynski eds., 2010).

55. This regime would also mandate altering when copyright protection generally starts to some formulation of publication or registration, again harking back to previous copyright requirements.

56. This last possible function is similar to Beth Simone Noveck's idea of peer-to-patent, allowing the public to comment on the appropriateness of patent applications. BETH SIMONE NOVECK, WIKI GOVERNMENT: HOW TECHNOLOGY CAN MAKE GOVERNMENT BETTER, DEMOCRACY STRONGER, AND CITIZENS MORE POWERFUL 3-15 (2009).

57. ECONOMIST, *supra* note 26, at 15. Yet, "to provide any incentive at all, more limited copyrights would have to be enforceable, and in the digital age this would mean giving content industries much of the legal backing which they are seeking for copy-protection technologies." *Id*.

copyrighted works into the public domain as soon as possible.

Tier Two would grant only one year of automatic copyright protection but would allow the protection to be renewed indefinitely for a fee as long as the copyrighted work meets or exceeds a revenue threshold. Every year the copyright holder would have to submit verification that the work meets or exceeds the revenue benchmark in order to obtain another year of copyright protection. Different revenue benchmarks could exist depending on the type of work copyrighted.

The structure of Tier Two would include five additional key features. First, the revenue threshold would be set high—i.e., the revenue that a copyrighted work would have to produce each year would be substantial. Second, the revenue threshold for each copyright would increase from year to year at a rate higher than the inflation rate. For example, the adjusting revenue threshold could be set to increase at the rate of inflation plus two percent each year.[58] Third, copyright holders would have to pay a substantial fee to renew their copyright if it meets the vigorous revenue requirements. Fourth, the renewal fee would be ever increasing at a rate higher than inflation, e.g., it could be benchmarked to the rate of inflation plus two percent per year. Fifth, a limit would be placed on the percentage of works copyrighted through Tier Two that could be renewed each year. For example, a maximum of one-tenth of one percent or one percent of all copyrighted material could be renewed each year. If more than one-tenth of one percent or one percent of copyrights meet the revenue threshold for renewal in a given year, the threshold would automatically increase. If this occurs, only the highest revenue-generating copyrighted works would continue to receive copyright protection. The number of works in Tier Two that are allowed to retain their copyright each year could be either a percentage of all current copyrighted works or a percentage of only the works in Tier Two. The latter restriction would make it more difficult for a work to be renewed.

Tier Two would cater to large business enterprises that are confident that they have just created the next Mickey Mouse. Tier One would most likely cater to the vast majority of creators who are risk-averse, doubtful that they could satisfy the revenue benchmark of Tier Two, or not willing to pay the substantial annual renewal fees.

The value of having two tiers in a copyright regime is similar to the value of the legal formalities that artists formerly fulfilled in order to obtain copyright protection. Both ideas attempt to filter art into different categories in order to get more works into the public domain more quickly. Since the proposed tiered revenue-based copyright system includes a registration requirement for all works, most noncommercial works would immediately flow into the commons,

---

58. This proposal is similar to the idea "that the older a copyrighted work is, the greater the scope of fair use should be." Joseph P. Liu, *Copyright and Time: A Proposal*, 101 MICH. L. REV. 409, 410 (2002).

and copyright orphans would be kept to a minimum. Plus, a tiered regime not only would distinguish between commercial and noncommercial art, but also it would divide commercial art into two tiers to accelerate the speed at which most of it enters the public domain. The two tiers would allow commercial artists to judge how risk-averse they are in their calculations of the likely revenue from their artwork. For example, if an artist has doubts about the earning potential of a piece, she would likely opt for the nonrenewable protection provided through Tier One. This calculus would partly determine how quickly commercial art enters the public domain, depending on which tier the artist chooses.

The revenue-based structure of Tier Two is attractive not only because it would move more commercial artwork more quickly into the public domain, but also because it would make the consuming public the final arbiter of copyright protection.[59] The public's implicit consent to renewal would come in the form of a good number of people having enough interest in a copyrighted work to pay for access to it. For this reason, a revenue-based renewal system is more likely to be deemed constitutional than a copyright system that allows for unlimited renewals based simply on the actions of the copyright holder, as with Landes and Posner's proposal for the automatic unlimited renewal for a fee, discussed below.

A tiered revenue-based copyright regime could be modified in numerous ways. First, the number of years of automatic copyright protection under Tier One could be reduced or extended to ensure that copyright reform is significant yet feasible within the current climate.

Second, there could be more than two tiers. For example, a Tier Three could grant five to seven years of automatic copyright protection, plus annual renewal into perpetuity as long as the copyrighted work meets a revenue benchmark. Of course, Tier Three's revenue benchmarks would be significantly higher than those of Tier Two. Tier Three would provide more upfront security to copyright holders but at the cost of greater difficulty in renewing because of higher revenue requirements.

Third, the revenue requirements under Tier Two could be timed differently. Instead of having annual revenue benchmarks, Tier Two could have two- or three-year benchmarks.[60] Or the length of time of each successive benchmark

---

59. A copyright holder could advertise heavily to boost revenue targets, but she could not buy her way out of the revenue requirements by selling use rights to herself so that she would meet the revenue requirements. This ban would have to include careful restrictions on internal transfer pricing between subsidiaries of conglomerates or possibly not counting such sales. While subsidiary X could generate revenue from its copyright by selling royalties to subsidiary Y, the transaction would not be included in revenue totals determining whether or not copyrights would be renewed.

60. Under such a revision—e.g., a three-year renewal period—the revenue threshold could be formulated in several ways: (1) an average revenue threshold for all three years, (2) a peak threshold whereby the copyright would be renewed if in one of the three years

could steadily increase or decrease over time. For example, the first benchmark period could be one year, the second benchmark period two years, and so on.

Fourth, the annual revenue benchmarks could be set to increase by only the inflation rate instead of a rate higher than inflation, or by a rate that would actually decrease over time—i.e., make benchmarks easier to meet.

Fifth, the registration requirement could be limited to Tier Two, with Tier One being the default option that would automatically apply if an author does not take the affirmative steps to select Tier Two.

Sixth, the renewal fee under Tier Two could be increased to deter strategic copyrighting. Landes and Posner suggest a "stiff renewal fee" because their proposal for "indefinite renewals" is potentially vulnerable to "[a] more serious concern" that "copyright holders might renew their copyrights for strategic purposes, hoping one day to 'hold up' an author who wanted to copy their work. This practice would resemble strategic patenting."[61] Regardless of whether a "stiff renewal fee" ameliorates the problem,[62] a tiered revenue-based copyright system is not likely to be susceptible to such risk because the holder's decision is not the final factor determining renewal.

Seventh, the renewal fee for artwork in Tier Two could be lowered to offset the difficulty of meeting revenue benchmarks. The fee could be nominally constant, without taking inflation into account, over the life of a copyright. Alternatively, the real value of the fee could be held constant—i.e., adjusted for inflation regularly. Or this substantial renewal fee could nominally decrease over time. It could even go in the opposite direction of the trend established for the amount at which the revenue threshold is set.

This proposal to reform copyright is feasible because corporate America can be persuaded to accept a system that offers indefinite copyright protection on its blockbuster creations for as long as such an arrangement increases its overall profits. The motivation behind a tiered revenue-based copyright regime is to give Big Copyright what we know it wants—the promise of possibly infinitely extendable copyrights—in exchange for increasing the scope and vitality of the public domain. Should the open commons movement care whether Mickey Mouse is perpetually copyrighted as long as the copyright term is significantly shortened for most works, the problem of orphan works is solved, the fair use doctrine is broadly construed, and substantial similarity provisions are narrowly tailored? Lessig thinks not.[63] Even if we should, the current regime is suboptimal because it ensures that Big Copyright will always demand longer copyright terms, paralyzing society under a one-size-fits-all copyright regime. A tiered revenue-based copyright system would dissolve corporate America's insistence on a monolithic copyright system.

---

revenue exceeds the threshold, or (3) an annual revenue threshold for each of the three years.

61. LANDES & POSNER, *supra* note 9, at 221.

62. *Id*.

63. LESSIG, *supra* note 15, at 221.

II. ADVANTAGES OVER OTHER PROPOSED REVISIONS TO THE COPYRIGHT TERM

The benefit of a tiered revenue-based copyright system is that it dramatically expands the public domain without trampling on the toes of Big Copyright. Numerous other schemes attempt to do the same: The Public Domain Enhancement Act; the Shawn Bentley Orphan Works Act of 2008 and the Orphan Works Act of 2008; Sprigman's reformalization of copyright proposal; and the indefinitely renewable copyright regime suggested by William Landes and Richard Posner.[64]

A. *Public Domain Enhancement Act*

The proposed Public Domain Enhancement Act,[65] now dead, was a practical attempt to lessen the harm brought on by exceedingly long copyright terms. The bill would have introduced into the public domain abandoned copyrighted works after 50 years by requiring copyright holders to pay a registration tax of $1 "due 50 years after the date of first publication or on December 31, 2006, whichever occurs later, and every ten years thereafter until the end of the copyright term."[66]

The MPAA ultimately opposed this act on what Professor Lessig states were "embarrassingly thin" grounds.[67] He goes on to argue that the underlying reason for such opposition was an

> effort to assure that nothing more passes into the public domain. It is another step to assure that the public domain will never compete, that there will be no use of content that is not commercially controlled, and that there will be no commercial use of content that doesn't require *their* permission first. . . . Their aim is not simply to protect what is theirs. *Their aim is to assure that all there is is what is theirs*. . . . [T]hey fear the competition of a public domain connected to a public that now has the means to create with it and to share its own creation.[68]

The Public Domain Enhancement Act was much better than the status quo, but it was neither as progressive nor as palatable as a tiered revenue-based copyright regime.

My proposal is more progressive because it would more quickly move

---

64. This list is not meant to be exhaustive.

65. H.R. 2408, 109th Cong. (2005).

66. Cong. Research Service, *H.R. 2408 (109th): Public Domain Enhancement Act Official Summary*, http://www.govtrack.us/congress/bill.xpd?bill=h109-2408&tab=summary (last visited Oct. 31, 2012). To avoid violating the Berne Convention, this requirement would have applied only to copyright holders of art created by American artists. Article 5(2) of the Berne Convention only prohibits a signatory state from imposing formalities on foreign authors. Berne Convention for the Protection of Literary and Artistic Works, Paris Act art. 5(2), July 24, 1971, 25 U.S.T. 1341, 828 U.N.T.S. 221.

67. LESSIG, *supra* note 15, at 253.

68. *Id*. at 255-56.

most artwork into the public domain. The registration requirement would immediately free all orphans and most noncommercial creations, while the shorter terms and the revenue requirements would free a substantial portion of commercial artwork much more rapidly.

Further, my proposal would have a better chance than the Public Domain Enhancement Act of getting Big Copyright to cooperate with those desiring a more open commons, because it offers the incentive of much longer terms on the most successful commercial artwork. It is impossible to know whether Big Copyright would view such an enticement as attractive enough to overcome its desires to keep the public domain debilitated and orphan works locked in legal limbo, yet there is cause to be optimistic. First, most copyright orphans have been abandoned for a reason—they either were never commercially successful or have already outlived any commercial usefulness. Second, Big Copyright's successful commercial art almost by definition has been more successful than copyright orphans, from a market perspective. Third, if orphans are freed, roughly two groups could make use of the material: Big Copyright and amateurs or individual creators. Big Copyright competes with itself routinely and there is no reason why any one firm in Hollywood would benefit substantially more than any other in being able to potentially exploit orphans that have been released into the public domain. At the same time, Big Copyright cannot truly fear that amateur or individual artists would use such newly freed orphans more effectively than Hollywood itself.

B. *Shawn Bentley Orphan Works Act of 2008 & Orphan Works Act of 2008*

The Shawn Bentley Orphan Works Act of 2008[69] proposed, among other things, to significantly reduce remedies, under certain circumstances, for infringement of orphan works. While it passed the Senate, a similar House bill, the Orphan Works Act of 2008, died.[70]

The Shawn Bentley Orphan Works Act of 2008 tried to limit "the remedies in a civil action brought for infringement of copyright in an orphan work, notwithstanding specified provisions and subject to exceptions, if the infringer meets certain requirements."[71] These conditions included "perform[ing] and document[ing] a reasonably diligent search in good faith to locate and identify the copyright owner before using the work" and, if the copyright holder was known, providing attribution to her.[72] Compensation would be restricted to

69. S. 2913, 110th Cong. (as passed by Senate, Sept. 27, 2008).

70. H.R. 5889, 110th Cong. (2008).

71. Cong. Research Service, *S. 2913 (110th): Shawn Bentley Orphan Works Act of 2008 Official Summary*, http://www.govtrack.us/congress/bill.xpd?bill=s110-2913&tab=summary (last visited Oct. 31, 2012).

72. *Id*.

reasonable compensation for the copyrighted artwork. No compensation would be necessary if the use was by a nonprofit institution and was "performed without any purpose of commercial advantage and is primarily educational, religious, or charitable in nature."[73]

Like the Public Domain Enhancement Act, both orphan works acts had real promise in attempting to increase access to copyright orphans, yet not as much as a tiered revenue-based copyright system. Neither orphan works act aspired to reduce the length of copyright: each simply aimed to reduce the potential cost of using orphan works, if one follows the procedures within the proposed acts. A tiered revenue-based copyright regime is superior to both orphan acts on the same three grounds discussed above with regard to the public domain act: (a) it gets the vast majority of noncommercial artwork into the public domain much more quickly, (b) it also moves more commercial artwork into the commons more rapidly, and (c) it has a greater chance of enticing, not antagonizing, Big Copyright because it offers Hollywood a substantial incentive in the form of much longer copyright terms on the most successful works.

C. *Sprigman's Reformalization of Copyright*

Christopher Sprigman has argued that the reformalization of copyright by creating new-style formalities would allow for substantial reform to "take place without damaging the interests of copyright owners who would otherwise have strong incentives to oppose the creation of a less restrictive copyright regime."[74] He writes:

> The simplest solution would be to preserve formally voluntary registration, notice, and recordation of transfers (and reestablish a formally voluntary renewal formality) for all works, including works of foreign authors, but then incent compliance by exposing the works of noncompliant rightsholders to a "default" license that allows use for a predetermined fee. The royalty payable under the default license would be low. Ideally, the royalty to license a work that a rightsholder has failed to register . . . should be set to approximate the cost of complying with these formalities (i.e., the total cost of informing oneself about the details of compliance and then satisfying them).[75]

Sprigman argues that such a reform would "ease[] access to commercially valueless works for which protection (or the continuation of protection) serves no purpose and [would] focus[] the system on those works for which protection is needed to ensure that the rightsholder is able to appropriate the commercial value of the expression."[76]

Sprigman's new-style formalities reform is a reasoned policy option that should be seriously considered. It has at least one benefit over a tiered revenue-

---

73. *Id*.
74. Sprigman, *supra* note 53, at 568.
75. *Id*. at 555.
76. *Id*.

based copyright regime: Sprigman's view that the Berne Convention would permit such new formalities, though he admits "there are arguments both ways."[77] While his reform would improve access to commercially unsuccessful work, it would not immediately place it into the public domain like a tiered revenue-based copyright system would. My proposal opens the door to bringing much more commercial artwork into the public domain much more quickly for two reasons. First, the nonrenewable term of Tier One would free the vast majority of registered artwork within 10 to 14 years. Second, artwork registered under the annual renewal system of Tier Two would also quickly enter the commons if it fails to meet the revenue thresholds necessary to maintain copyright protection. This would especially be the case if there is a yearly percentage cutoff as to how much commercial artwork could continue to be protected.

D. *Landes & Posner's Indefinitely Renewable Copyright Regime*

Testifying in 1906 before Congress against the need for copyright term limits, Samuel Clemens, a.k.a. Mark Twain, said, "There is only about one book in a thousand that can outlive forty-two years of copyright. Therefore why put a limit at all? You might just as well limit a family to 22. It will take care of itself."[78] Following in the footsteps of Twain, Landes and Posner propose a copyright regime of indefinitely renewable copyrights in which copyright holders could pay a fee to have their copyrights renewed after short fixed terms.[79] Under their proposal, all new and existing copyrighted artwork would need to be registered, and copyright holders could extend their copyrights as many times as they desire.

Landes and Posner's proposal has numerous attractive characteristics, yet a tiered revenue-based copyright regime has more advantages.

First, my proposal is more effective in moving commercial artwork into the commons. Whether we consider it a good thing or a tragedy,[80] a substantial portion of our culture comprises commercially successful artwork (films, music, TV, etc.). Landes and Posner's scheme would lead (unless stiff renewal fees are contemplated) to most commercially successful artwork being absent from the public domain for an extremely long time. My proposal is more capable of moderating the amount of successful commercial art that stays

---

77. *Id*. at 556.

78. Clemens's testimony was reprinted in Samuel L. Clemens, *Copyright in Perpetuity*, 6 GREEN BAG 2d 109, 111 (2002).

79. William M. Landes & Richard A. Posner, *Indefinitely Renewable Copyright*, 70 U. CHI. L. REV. 471 (2003). Their idea is also articulated in LANDES & POSNER, *supra* note 9, at ch. 8.

80. *See generally* MAX HORKHEIMER & THEODOR ADORNO, DIALECTIC OF ENLIGHTENMENT (Gunzelin Schmid Noerr ed., Edmund Jephcott trans., 2002) (criticizing the value of commercial artwork).

locked up; it pushes all but the most profitable copyrighted works into the public domain and does so within a reasonable timeframe.

Second, both proposals create some transparency by requiring registration, but my proposal is more transparent and less susceptible to abuse because it prevents copyright holders from having full control over the terms of their copyright protection. Under Landes and Posner's proposal, the decision to renew lies solely with the copyright holders, who can continue pay for copyright protection indefinitely. They can refuse to ever allow anything into the commons, either to prevent their opponents from potentially benefiting from their creations or out of a pack rat mentality. Under my proposal, copyright holders may choose to pursue renewal, but whether a renewal is granted depends on whether the work in question meets the revenue threshold. Ultimately, it would fall to the public to decide, through their pocketbooks, whether a copyright should be extended.

This benefit of the public's implicit consent as the determining factor for renewal ties into the third advantage: a tiered revenue-based copyright regime would have a better chance of meeting constitutional objections than the system proposed by Landes and Posner. This is because a tiered revenue-based copyright regime would not guarantee copyright holders direct control or indefinite protection. Landes and Posner simply state that their "concern is with the economics rather than the constitutionality of indefinite renewal."[81] In a footnote they go on to say, without explanation, that "[i]n light of" *Eldred v. Ashcroft*, "it is unlikely that a system of indefinite renewals, which has more to commend it than the Sonny Bono Act, would be held unconstitutional."[82]

Fourth, a tiered revenue-based copyright system can be modified. For example, while my proposal requires registration of all artwork, it could easily be altered to eliminate that requirement for the fixed term tier. Such flexibility is not possible with Landes and Posner's proposal.

While my proposal has more advantages than Landes and Posner's, it does have at least two comparative drawbacks.

First, the transaction costs of my proposed reform are slightly higher, yet even Landes and Posner suggest collecting some revenue figures because "a single fee for all types of copyrighted work is unlikely to be optimal. An alternative that would minimize legislative and regulatory discretion, and hence rent seeking, would be to make the fee equal to a fixed percentage of the first year's inflation-adjusted revenues from the sale or rental of the copyrighted work."[83] More important, the transaction costs in my proposal are borne by those expecting to benefit from society's largesse, not by the public in general. Plus, the higher transaction costs to those seeking further copyright protection are outweighed by the benefit of significantly expanding the public domain.

---

81. LANDES & POSNER, *supra* note 9, at 211.

82. *Id*.

83. *Id*. at 219-20.

Second, while a tiered revenue-based copyright regime would reduce rent seeking by Big Copyright relative to our existing system, Landes and Posner's suggested revisions would even further diminish rent-seeking activity—though not entirely eliminate it given that Big Copyright could still lobby for "lower renewal fees and longer renewal terms."[84]

Finally, Landes and Posner's reform could import my proposal of multiple tiers into their formulation. For example, one tier could provide a nonrenewable term at little or no cost, while a second tier could allow for renewals but have expensive renewal fees.

III. POTENTIAL PROBLEMS

A tiered revenue-based copyright system is not a perfect solution, but it is better than the current copyright system and the proposals described in Part II above. Below are responses to some of the most common criticisms of this proposition that have not already been discussed.

A. *Too Difficult to Track Revenue*

Very few, if any, variables can be perfectly and costlessly measured. While measuring copyright revenue will not be immune from some abuse and some complications in calculation,[85] the proposed regime's features will not be easy to abuse or impossible to assess.[86] Also, Landes and Posner's suggestion for their proposal of indefinite copyright renewals is applicable to a tiered revenue-based copyright system:

> The aggregate transaction costs [of the proposal] . . . would depend on the number and possibly the value of licenses (holding tracing costs constant), the transaction costs per license, and the administrative cost of operating the renewal system. Since the number of licenses would depend in part on the total number of works renewed, aggregate transaction costs could actually fall compared either to a system of automatic renewals or to a single term of life plus seventy years.[87]

Past and present day examples of copyright payment and/or registration systems that are arguably more complex than the measurements required by my proposal also suggest that a tiered revenue-based copyright regime is practically feasible. Historically, the expense of obtaining copyright was more costly and time-consuming than it is under current law. Copyright was intelligently structured as a quasi-test of an artist's intent to seek copyright

---

84. *Id*. at 221.

85. The same is true for other related possible scenarios such as gauging the value of individual copyrights that are combined to create a larger work, for example, a film.

86. Simpler variables, if any decent candidates exist, would cost less to measure but would be less useful for the present purposes.

87. LANDES & POSNER, *supra* note 9, at 217.

protection. These historical requirements, which began to be eroded from 1909 onwards, entailed registration of the artwork, the deposit of copies of the artwork with the Copyright Office, and placing a notice of copyright protection on every published copy of the work. What did society gain from the easing of such requirements? Arguably little more than a copyright regime that automatically defaults to extending copyright protection to almost everything, including our shopping lists. Such historical hoops served a similar purpose to the aims of my proposal—keeping the public domain robust—while not being too burdensome. Simply returning to the previous requirements would be a major step forward, yet Big Copyright would have no incentive to do so.

One contemporary real world example of a complex copyright arrangement is the 1992 Audio Home Recording Act (AHRA). In addition to requiring Serial Copy Management System (SCMS) controls and stating that no action claiming copyright infringement can be brought against individuals making musical copies for private, noncommercial use, the AHRA enables manufacturers of digital audio equipment to sell digital tapes and recorders if they pay royalties on all such sales.[88] The royalties are divided among background musicians, vocalists, featured recording artists, record companies, composers, and music publishers.[89] While the percentage each of the groups receives is fixed by statute, the law does not mandate how individuals within these groups must be compensated. This example demonstrates that law can be functional even if many variables cannot be perfectly measured or observed.

A second example is the American Society of Composers, Authors and Publishers (ASCAP). It is a performing rights organization of over 450,000 composers, lyricists, songwriters, and publishers[90] that licenses billions of nondramatic public performances of their copyrighted artwork each year and then distributes the royalties to its members.[91] ASCAP is

> guided by a "follow the dollar" principle in the design of [its] payment system. In other words, the money collected from television stations is paid out to members for performances of their works on television, the money collected from radio stations is paid out for radio performances, and so on . . . . The value of each performance is determined by several factors, including the amount of license fees collected in a medium (television, cable, radio, etc.), how much we receive in fees from the licensee that hosted the performance, and the type of performance (feature performance, background music, theme song, etc.).[92]

---

88. Audio Home Recording Act of 1992, Pub. L. No. 102-563, 106 Stat. 4237 (1992) (codified as 17 U.S.C. §§ 1001-1010).

89. 17 U.S.C. §§ 1004-1008.

90. *What Is ASCAP?*, ASCAP, http://www.ascap.com/about/ (last visited Dec. 20, 2012).

91. *ASCAP Payment System: How You Get Paid at ASCAP*, ASCAP, http://www.ascap.com/members/payment/ (last visited Dec. 20, 2012).

92. *Id*.

In fact, royalty calculations for an individual musical work are more complicated than the above summary suggests. ASCAP multiplies five variables (use weight, licensee weight, "follow the dollar" factor, time of day weight, and general licensing allocation) together and then adds radio feature premium credits and TV premium credits to arrive at a final tally.[93] The general licensing allocation drives home the point that intricate systems for approximating values that cannot practically be precisely measured can successfully work. The general licensing allocation is calculated by the following method: "Fees collected from non-broadcast, non-surveyed licensees (bars, hotels, restaurants and the like) are applied to broadcast feature performances on radio and all performances on television, which serve as a proxy for distribution purposes."[94]

While these two examples are not perfect precedent for proving that revenues can be measured accurately enough without bankrupting artists and regulators, they serve as positive indicative guides. Measuring the revenue of copyrighted artwork will not be flawless or even elegant, but it is practically achievable on a large scale.

B. *Negative Effects of Locking Up the Most Successful Commercial Art*

This proposal would make access to the most successful commercial artwork more expensive and hence more restricted.[95] In this regard, it could be viewed as harmfully revising the definition of a free society. Yet, as will be argued below, such concerns are minor impediments relative to the benefits the proposed reform will bring.[96] Professor Lawrence Lessig forcefully states that

---

93. *ASCAP Payment System: Royalty Calculation*, ASCAP, http://www.ascap.com/members/payment/royalties.aspx (last visited Dec. 20, 2012).

94. *Id*.

95. A related potential concern is the possibility expressed by Felix Cohen many years back: "The vicious circle inherent in this reasoning is plain. It purports to base legal protection upon economic value, when, as a matter of actual fact, the economic value of a sales device depends upon the extent to which it will be legally protected." Felix Cohen, *Transcendental Nonsense and the Functional Approach*, 35 COLUM. L. REV. 809, 815 (1935). While Cohen's observation can be piercing in other contexts, it does not fit the facts or the nature of this proposal. As is clearly evidenced by the existence of millions of copyright orphans, legal protection does not always automatically create economic value.

96. It is feasible, though not necessarily likely, that with the enactment of this proposal, the most successful commercial works may actually lose cultural significance. Because such works will always cost more, they will be used less by other artists and—at least after their initial splash—will be consumed less and less by individuals. While this effect would decrease the extent to which the most successful works are engrained in our cultural DNA, Big Copyright could ward off such a fate by advertising its star earners year in and year out to increase their impact or at least counteract their decline. Yet constant advertising would carry the risk of overexposing the public to the advertised works, thus causing consumers to revolt against them. For the effects of marketing on copyright's ability to spur new creation, see generally Mark S. Nadel, *How Current Copyright Law Discourages Creative Output: The Overlooked Impact of Marketing*, 19 BERKELEY TECH.

at some point artwork should be free for others to take and criticize in whatever way they want.

> It should be free, that is, not only for the academic, who would certainly be allowed to quote the book in a critical essay; it should be free as well for authors . . . as well as film directors or playwrights to adapt or attack as they wish. That's the meaning of a free society, and whatever compromise on that freedom copyright law creates, at some point that compromise should end.[97]

I have previously argued, and still maintain, that copyright needs to be abandoned in rich countries because the overabundance of successful commercial art harms citizens.[98] In the United States, for example, copyright has done such a good job of supporting the production of polished commercial art that it has turned the average citizen into a passive overconsumer. Americans on average consume 8.54 hours a day of entertainment and news.[99] This statistic does not even include hours spent surfing the Internet. Rich countries need to eliminate the source of this overconsumption—i.e., copyright—so that more individuals have the inclination as well as the time to create for themselves. The elimination of copyright would decrease the amount of commercial art produced. It would encourage individual productivity by putting a large dent in the amount of time citizens spend passively consuming others' artwork. This newly available time could be used for many different ventures, including, for some, spending a few hours a week creating on their own.

I have also claimed, and still contend, that copyright should be abandoned in poor countries because access to rich country artwork facilitates the embrace of liberal values.[100] One of the most effective ways to promote democracy and reduce intolerance in developing countries is by exposing citizens to developed country artwork, which even in its most commercial forms communicates liberal values subtly, or not so subtly, in the background. While some poor country art may do a better job of communicating liberal values than some rich country art, and while some rich country art may be terrible at conveying such values, on average rich country art, warts and all, does a better job of demonstrating the vitality and necessity of democracy, liberty, freedom of expression, equality, and human rights. Eliminating copyright in developing countries would allow rich country art to be freely distributed and, over the long run, to be a factor in convincing large numbers of individuals in poor countries of the ethical necessity of adopting democratic values.

L.J. 785 (2004).

97. LAWRENCE LESSIG, THE FUTURE OF IDEAS: THE FATE OF THE COMMONS IN A CONNECTED WORLD 199 (2001).

98. *See generally* Martin Skladany, *Alienation by Copyright: Abolishing Copyright to Spur Individual Creativity*, 55 J. COPYRIGHT SOC'Y U.S.A. 361 (2008).

99. *Id*. at 366.

100. Martin Skladany, Culture and Copyright in Developing Countries (Dec. 20, 2012) (unpublished manuscript) (on file with author).

I mention the above to assure the reader that I am a romantic when it comes to believing copyright should and can radically change to enlarge our lives through promoting freedom of thought and action. As will be discussed below, anything is possible in the long term—e.g., the civil rights movement, the fall of communism. Professor Lessig's claim about what a free society necessitates is alluring and convincing to me, yet this Article is about being brutally honest about how much cultural freedom we can realistically expect to win in the short term. We should be willing to give Disney more of what it wants so that it stops deforming and shackling most of our culture. Such a calculation is by necessity utilitarian. Lessig is no stranger to such compromises, as he explicitly mentions,[101] yet he asserts that "at some point" the compromise that copyright engenders "should end." My assertion is simply to redo the calculus—lock up a much smaller amount of content for a longer time in order to allow a vast amount of content to become free much sooner. Unlike Landes and Posner's proposal, which would likely be in violation of the Constitution, mine asserts that copyright should not be unilaterally perpetual for holders who simply pay renewal fees. Yet I admit that practically speaking for the most successful commercial artwork, I am pushing back the date significantly.

Lessig, many copyright scholars, and I desire to significantly shorten copyright's length. If we do nothing now because we do not have the necessary mobilization for radical reform, then the most successful commercial artwork will be locked up for a long time regardless, given Big Copyright's ability to simply lobby for a Cher Copyright Term Extension Act as a follow-up act to the Sonny Bono Copyright Term Extension Act.

C. *Overexposure Risk for Big Copyright*

Trademark and right-of-publicity laws recognize the possibility that the underlying property can be devalued by overexposure. Landes and Posner ask whether this is a factor that needs to be considered for copyrightable expression. They state, "There is some evidence that it is a concern of the Walt Disney Company with regard to its copyrighted characters, such as Mickey Mouse."[102] They continue:

> We must not press the congestion argument . . . too far. While examples can be given of works even of elite culture that may have been damaged by unlimited reproduction (the *Mona Lisa*, the opening of Beethoven's *Fifth Symphony*, and several of Van Gogh's most popular paintings come immediately to mind), there are counterexamples: the works of Shakespeare seem unimpaired by the uncontrolled proliferation of performances and derivative works, some of them kitsch, such as Shakespeare T-shirts and the

101. LESSIG, *supra* note 97, at 199.

102. LANDES & POSNER, *supra* note 9, at 224.

> movie *Shakespeare in Love*. And in the field of popular culture, think only of Santa Claus as an example of the power of an iconic character to survive incessant use, apparently undamaged.[103]

Landes and Posner are correct that the *Mona Lisa* and the works of Shakespeare are different, but they take this difference, which rests on the *Mona Lisa* being "damaged" too far. Sure, the *Mona Lisa* may be kitsch because of its immense popularity, unlike Shakespeare's work, but at the same time it continues to be a highly respected masterpiece that caps a visit to the Louvre for millions. More to the point, from the perspective of Big Copyright, this dual personality of the *Mona Lisa*, if it were still under copyright, would not harm its revenue stream; rather, it would very likely increase it. Kitsch can sell brilliantly on its own, but when combined with genius it is an asset Big Copyright would love to own.

I do not doubt that Disney manages the proliferation of its copyrighted characters in order to maximize profits without risking overexposure. However, given the existence of Disney World, Disney Land, its foreign theme parks outside of Paris and Tokyo, Disney stores, etc., this overexposure concern does not appear to pose a real threat to the viability of a tiered revenue-based copyright regime. Essentially, what would overexposing a work mean for Disney when its existing promotional efforts are so extensive?

Under my proposal, all copyright holders who choose to protect their works under Tier Two would understand that they potentially face this overexposure concern. Even if only a small percentage of works registered under Tier Two achieve renewal, no copyright holders would have an incentive to overexpose their holdings if they are reasonably confident that such holdings could relatively easily meet the revenue thresholds. For example, would anyone doubt that Mickey Mouse will be one of the consistently highest revenue-producing works? And if Disney does not think that Mickey Mouse is currently being overexposed, how could anyone think that an icon less famous and less assured of meeting the revenue threshold is in danger of overexposure?[104]

D. *Destroying the Market for Copyright Artwork*

Some might claim that a tiered revenue-based copyright regime with one tier having an annual revenue requirement would annihilate the commercial art market because the public would simply wait a year before paying to see or listen to any artwork with a renewable copyright. Such an argument overlooks current marketing practices and consumer behavior.

For example, the current Hollywood practice is to roll out a film gradually in different forms. Many movies are first available only in theaters. When

---

103. *Id*. at 226-27.

104. Granted such a copyrighted artwork on the revenue threshold boundary could fill a small niche market and hence risk overexposure within that small segment.

released on DVD or online, some movies can initially only be bought, with the option of renting coming a month or two later. Most of these steps occur within a year, and a marketing push often precedes each step to create and maintain a movie's "must-see" status.

Even if a large enough group of individuals is willing to wait for a copyrighted work to fail to meet a revenue benchmark so that it would be released into the public domain, there would be no certainty *ex ante* that the work would not meet the revenue cutoff. Hence, such a group could wait for decades or longer to see a Mickey Mouse movie for free. Such uncertainty could even create a situation similar to the prisoner dilemma: while it would be in the group's best interest to wait a year and deny the copyright holder enough revenue to meet the benchmark, individual members might prefer to purchase the product the day of its release instead of having to deal with the uncertainty of trusting others not to buy it immediately. This is not to claim that some individuals might happily resist all the marketing, live with the uncertainty, and wait for a work to go off copyright; but this group is likely to be small and hence would not significantly chip away at the commercial art market.

E. *Dangers of Striking a Deal with Big Copyright*

Another concern is that Big Copyright might agree to this proposal but then, over the long term, fight to change the provisions of the bargain. While such a possibility is unfortunate, whenever reaching across the aisle, one has to consider such behavior. *Si vis pacem, para bellum*: if you wish for peace, prepare for war.[105] In fact, Big Copyright should plan for the same contingency—public domain advocates continuing to push for shorter copyright terms—though neither side should necessarily expect any success if it ventures away from a compromise built on a tiered revenue-based system.

Two main concerns exist.[106] First, Big Copyright might attempt to compromise the stringency of revenue thresholds or increase the length of renewal periods. Yet Big Copyright has little incentive to prolong copyright on all artwork if it does not own most of it and if much of what it owns is essentially worthless after a decade. Second, Big Copyright might also strategize to bring back copyright orphans in one form or another to reduce the size of the public domain so that there is less competition for its holdings. Such potential competition from works in the public domain is uncertain, given that orphan works are often orphaned because they were unsuccessful

---

105. FLAVIUS VEGETIUS RENATUS, VEGETIUS: EPITOME OF MILITARY SCIENCE 63 (N.P. Milner trans., 2d rev. ed., Liverpool Univ. Press 1996) (ca. 430-435).

106. Another concern is that Big Copyright could always reissue a lapsed copyright in a form different enough to get a new copyright. This already occurs and is difficult to eliminate outside of abandoning copyright entirely. The saving grace is that the related works are not identical.

commercially. Also, Big Copyright could benefit from a clearing of copyright orphans into the public domain, because all artists—commercial as well as noncommercial—would gain an enormous amount of newly available free material to borrow from. Furthermore, Big Copyright would have to seriously consider whether any added revenue that might result from breaking the compromise is worth (a) the potential financial cost of lobbying to make an extra buck on lackluster holdings and (b) the risk of breaking faith with society, given the danger of being painted evil like Big Pharma. Finally, it must be remembered that one of the biggest strengths of a tiered revenue-based copyright regime is that relative to the current copyright system it would reduce rent seeking on copyright's term length.

If Big Copyright accepts a tiered revenue-based copyright regime, incentives could be built into the new legislation to discourage powerful commercial interests from later lobbying to loosen the requirements. First, Big Copyright could be required to contribute money to a nonprofit that would lobby to ensure that the new copyright system's term provisions are not altered in the future to favor Big Copyright. Second, a poison pill could be attached to the tiered revenue-based copyright regime bill—i.e., if the regime is altered for the benefit of Big Copyright, the poison pill dilutes its copyright ownership but not everyone else's. Alternatively, the bill could require a transition period during which all copyright holders (or just those who pushed for the bill that alters the tiered copyright regime) have their copyright diluted (i.e., scope of protection weakened).

F. *Practical Impossibility of Legislative Reform*

While the above objection points to the dangers of concluding an agreement with Big Copyright, a further objection is that a deal cannot be struck given how copyright law is made in practice. Essentially, Congress will listen to and broker compromises only with those showering them with campaign contributions and bringing along celebrities.

The process of copyright legislation has been characterized by the need for consensus among opposing stakeholders, where all parties must benefit and where no party and its interests are deprived of a seat at the table. A series of conferences of different parties with a stake in copyright guided legislation forward that ultimately revised copyright law into the Copyright Act of 1976.[107] During this period, Register Abraham Kaminstein of the Copyright Office stated "that the key to general revision would be to draft a copyright bill that benefited each of the competing interests. In that, the conferences succeeded. The bill that emerged from the conferences enlarged the copyright pie and divided its pieces among conference participants so that no leftovers

107. JESSICA LITMAN, DIGITAL COPYRIGHT 53 (2001).

remained."[108] Such legislation by negotiated settlement where copyright stakeholders significantly influence copyright law has been the norm for almost a century.[109] This norm has created a troubling situation where, as Professor Jessica Litman writes, "current stakeholders are unwilling to part with short-term statutory benefits in the service of long-term legal stability" and interested parties "disfranchised by current law lack the bargaining chips to trade for concessions."[110]

The picture painted above is not at first encouraging. Yet Litman later lists libraries, schools, consumer groups, and civil liberties nonprofits as some of the interests "who employ paid Washington lobbyists to speak up for the needs of unrepresented citizens."[111] So there are at least some public-spirited groups that can make it to the negotiating table. Further, commercial interests at the table can fight vigorously against each other when their interests conflict. Moreover, as pointed out by Boyle, the situation has looked at least as bleak in the past for other issues like the environment, yet broad and robust actors eventually coalesced into a powerful movement.[112]

Whether a deal can be struck also depends on the nature of the deal. A tiered revenue-based copyright system is designed to make the proposed reform enticing to Big Copyright, unlike other proposals that do not offer Big Copyright any incentives.

CONCLUSION

Big Copyright will forever engage in rent-seeking activity. As Mancur Olson has demonstrated, this is the nature of political systems.[113] In the short run, activists dedicated to reducing the length of copyright protection have few options, if any, besides negotiating a deal that will entice Big Copyright to set orphans free and to accept dramatically reduced copyright terms for the vast majority of artwork in exchange for gaining longer protection for its most successful commercial works.

Because most of this Article has been focused on demonstrating how Big Copyright will deem such a proposal attractive, it seems appropriate to briefly touch on a less immediate concern. Dangle a large enough carrot in front of business interests and they will not resist because their *raison d'être* is profit. Yet the same does not work for activists, given the nature of their beliefs.

---

108. *Id*.

109. *Id*. at 62.

110. *Id*. at 63.

111. *Id*. at 193.

112. *See generally* James Boyle, *A Politics of Intellectual Property: Environmentalism for the Net?*, 47 DUKE L.J. 87 (1997).

113. *See generally* MANCUR OLSON, THE LOGIC OF COLLECTIVE ACTION: PUBLIC GOODS AND THE THEORY OF GROUPS (rev. ed., Harvard Univ. Press 1971) (1965).

Activists do not have to love this proposal's trade-off, nor should they give up their efforts to build, over the long term, society-wide support for their positions; in the short term, however, they must compromise to stay true to their stated goals. Some communist intellectuals undermined socialism because they felt that temporary measures slowed progress and sullied their purity of purpose. Oscar Wilde wrote of denying the poor charity so that society would more quickly open its eyes to the horrors of capitalism and hence more readily embrace communism.[114] His twisted logic infantilized the poor—as if they did not already know how hard their lives were or forgot such hardship when it was temporarily relieved by a private charity, and as if they could not understand the effects of capitalism or charity's relationship to different economic systems. Copyright activists have the high moral ground. A tiered revenue-based copyright system will quickly get them significantly closer to their desired goal. They should not be co-opted by copyright abolitionists or by their own ideal vision of copyright into keeping the chains around Richelieu's Monster, especially given that partial freedom now will strengthen the movement for the arduous long-term fight ahead.

114. *See generally* OSCAR WILDE, THE SOUL OF MAN UNDER SOCIALISM AND SELECTED CRITICAL PROSE (Linda Dowling ed., Penguin Classics 2001) (1885-1891). Marx advocated a similar argument in his critique of petty-bourgeois socialism. KARL MARX, THE COMMUNIST MANIFESTO 81-82 (Frederic L. Bender ed., Samuel Moore trans., Norton 1988) (1848).

STANFORD TECHNOLOGY LAW REVIEW
VOLUME 16, NUMBER 1 FALL 2012

NOTE

SECTION 101 AND COMPUTER-IMPLEMENTED INVENTIONS

Robert D. Swanson*

Cite as: 16 Stan. Tech. L. Rev. 161 (2012)
http://stlr.stanford.edu/pdf/computerimplementedinventions.pdf

ABSTRACT

The law surrounding the patentability of computer-implemented inventions is extraordinarily unclear. Thankfully, the Federal Circuit has granted rehearing en banc to CLS Bank v. Alice *to determine the test for computer-implemented inventions under § 101. This Note identifies three current approaches in Federal Circuit doctrine, and finds each lacking. In their place, this Note proposes and defends the data manipulation test, which would hold that a claim is not directed to an abstract idea if it contains a computer element such that: (1) the computer manipulates data, rather than merely being present, (2) the data being manipulated is inherent to the computer, and (3) the data manipulation is directed to one or more particular applications. As this Note argues, the data manipulation test is an easily administrable compromise test that the Federal Circuit, or the Supreme Court upon review, can adopt in* CLS Bank.

---

* J.D. Stanford Law School, 2013. The author is deeply grateful to Professor Mark Lemley for his invaluable commentary on earlier drafts, as well as to countless others for stimulating conversations and criticism.

INTRODUCTION

The current case law on whether many computer-related inventions constitute patentable subject matter under § 101 of the Patent Act[1] is extraordinarily unclear. In the past two years, the Supreme Court has twice taken § 101 cases,[2] producing opinions that have generated serious controversy among the patent bar.[3] As a result, the Federal Circuit has struggled to determine when computer-related claims pass § 101, with its post-*Bilski* case law being defined by a major circuit split over computer-implemented inventions.

With the Federal Circuit's *en banc* review of *CLS Bank* on the horizon, this Note attempts to end this confusion. This Note does not engage in the debate over whether broad subject matter eligibility is preferable to the Supreme Court's recent restrictions. Rather, this Note accepts the Court's rulings and proposes an administrable test drawn from current precedent that the Federal Circuit can adopt in its upcoming *CLS Bank en banc* decision. I follow a

1. 35 U.S.C. § 101 (2012).

2. Bilski v. Kappos, 130 S. Ct. 3218 (2010); Mayo Collaborative Services v. Prometheus Laboratories, 132 S. Ct. 1289 (2012).

3. *See, e.g.*, Shuba Ghosh, *Guest Post on Bilski: Throwing Back the Gauntlet*, PATENTLYO (June 29, 2010, 1:33 PM), http://www.patentlyo.com/patent/2010/06/guest-post-on-bilski-throwing-back-the-gauntlet.html; Donald Chisum, *Notes on Bilski*, CHISUM.COM (June 29, 2010), http://www.chisum.com/current-developments/bilski-watch/notes-on-bilski#more-298; Robert Sachs, *Punishing Prometheus*, PATENTLYO (Mar. 26, 2012, 9:10 AM), http://www.patentlyo.com/patent/2012/03/punishing-prometheus-the-supreme-courts-blunders-in-mayo-v-prometheus.html.

pragmatic approach, demonstrating how courts can apply the Supreme Court's § 101 jurisprudence and existing Federal Circuit doctrine to focus on the real issue: the effect a computer element has on a claim's patentability. As one Federal Circuit panel has already done,[4] courts should examine the computer's role in the invention.

In Part III, this Note proposes a test to guide that inquiry. To build to that point, I explore in Part I the recent case law on § 101. There, I identify three tests the Federal Circuit has employed, pinpointing useful methodology in each that the court can extend in *CLS Bank*. Part II discusses common approaches to abstract ideas in the literature, demonstrating why they would fail in practice. After introducing the data manipulation test for computer-implemented inventions in Part III, I explore the test's boundaries through several real and hypothetical examples. This Note concludes by surveying a number of the test's advantages and responding to conventional criticisms.

I. RECENT CASE LAW ON COMPUTER-IMPLEMENTED INVENTIONS

Section 101 of the Patent Act states that patentable subject matter encompasses "any new and useful process, machine, manufacture, or composition of matter."[5] Often, courts mention that this "include[s] anything under the sun that is made by man."[6] However, there are three common law exceptions to § 101:[7] "laws of nature, physical phenomena, and abstract ideas."[8]

Courts typically invalidate computer-implemented inventions under the abstract idea exception.[9] The *Bilski* decision is the most recent Supreme Court

---

4. Bancorp Services, L.L.C., v. Sun Life Assur. Co. of Canada (U.S.), 687 F.3d 1266 (Fed. Cir. 2012).

5. 35 U.S.C. § 101 (2012).

6. Diamond v. Chakrabarty, 447 U.S. 303, 309 (1980).

7. It is worth noting that § 101 was itself a codification of common law. Peter S. Menell, *Forty Years of Wandering in the Wilderness and No Closer to the Promised Land: Bilski's Superficial Textualism and the Missed Opportunity to Return Patent Law to Its Technology Mooring*, 63 STAN. L. REV. 1289, 1302-03 (2011). Moreover, the Supreme Court's *Bilski* opinion erroneously treated § 101 as an area of statutory law rather than common law. *Id*. at 1301-02.

8. *Chakrabarty*, 447 U.S. at 309.

9. *See, e.g.*, CyberSource Corp. v. Retail Decisions, Inc., 654 F.3d 1366 (Fed. Cir. 2011); Dealertrack, Inc. v. Huber, 674 F.3d 1315 (Fed. Cir. 2012); Fort Properties, Inc. v. Am. Master Lease LLC, 671 F.3d 1317 (Fed. Cir. 2012). Note, however, that the boundaries between the different exceptions have never been very well defined. No court or commentator has authoritatively distinguished between laws of nature and abstract ideas when it comes to computer-related claims. Earlier cases characterized algorithms—like mathematical relationships—as laws of nature. Courts now generally analyze computer-related inventions as abstract ideas; computing technology has advanced beyond merely implementing mathematical relationships that exist in nature, and has developed its own concepts. While the relationship between decimal and binary numbers may exist in nature,

case directly dealing with abstract ideas, so it marks a good place to start.

A. Bilski v. Kappos

Bilski v. Kappos came to the Supreme Court from an *en banc* decision in the Federal Circuit.[10] The *In re Bilski* court held that the machine or transformation test, drawn from a trilogy of prior Supreme Court cases,[11] is the sole patent eligibility test for process claims under § 101.[12] Under the machine or transformation test, if a process (1) "is tied to a particular machine," or (2) "transforms a particular article into a different state or thing," the process is not abstract (or, for laws of nature, the process does not preempt the law).

Many in the patent bar viewed this decision in terms of the Federal Circuit's relationship with the Supreme Court. The Supreme Court had long criticized the Federal Circuit for being too pro-patent,[13] but the *In re Bilski* opinion drew heavily from specific language in Supreme Court precedent to invalidate the patent.

The Supreme Court, however, disagreed with the Federal Circuit's reasoning. To the surprise of many, the Supreme Court granted certiorari in *Bilski* and affirmed under different analysis. Justice Kennedy's majority opinion disapproved of the Federal Circuit's use of the machine or transformation test, holding that while it "is a useful and important clue," it "is not the sole test for deciding whether an invention is a patent-eligible 'process.'"[14] The Court refused to provide any further guidance, stating that it "need not define further what constitutes a patentable 'process,' beyond pointing to the definition of that term in § 100(b) and looking to the guideposts in *Benson*, *Flook*, and *Diehr*."[15] In concluding, the Court emphasized that it "by no means foreclose[s] the Federal Circuit's development of other limiting criteria that further the purposes of the Patent Act and are not inconsistent with its text."[16]

The Supreme Court thus did little more in *Bilski* than disapprove of the machine or transformation test as the sole test for process claims and punt the

---

concepts in computer science such as linked lists or hash tables are artificial, human-made constructs, and should accordingly be analyzed under abstract ideas doctrine.

10. In re Bilski, 545 F.3d 943 (Fed. Cir. 2008) (en banc) *aff'd but criticized sub nom.* Bilski v. Kappos, 130 S. Ct. 3218 (2010).

11. *Benson*, 409 U.S. 63; Diamond v. Diehr, 450 U.S. 175 (1981); Parker v. Flook, 437 U.S. 584 (1978).

12. *In re Bilski*, 545 F.3d at 954 ("A claimed process is surely patent-eligible under § 101 if: (1) it is tied to a particular machine or apparatus, or (2) it transforms a particular article into a different state or thing.")

13. *See, e.g.*, KSR Int'l Co. v. Teleflex Inc., 550 U.S. 398, 415-419 (2007); eBay Inc. v. MercExchange, L.L.C., 547 U.S. 388, 392-94 (2006).

14. Bilski v. Kappos, 130 S. Ct. at 3221.

15. *Id.* at 3231.

16. *Id*.

issue back to the Federal Circuit. The result was to only deepen the divisions already present in the Federal Circuit. Any judge that approved of the machine or transformation test could continue to apply it, so long as she considered "other limiting criteria." Likewise, any judge that disliked the machine or transformation test could reject it as mostly inapplicable to "Information Age" inventions and consider any other criteria he preferred instead.[17]

Post-*Bilski*, this split is most manifest in cases involving computer-implemented inventions. Since *Bilski*, the Federal Circuit has arguably employed three principal approaches to computer-implemented inventions: (1) invalidating claims only when they are manifestly abstract, (2) upholding claims where the computer serves as a meaningful limit on the claim's scope, and (3) focusing on whether the computer is integral to the claimed process.

B. *The "Manifestly Abstract" Test*

Under the manifestly abstract test, a claim survives a § 101 challenge "[u]nless the single most reasonable understanding is that a claim is directed to nothing more than a fundamental truth or disembodied concept, with no limitations in the claim attaching that idea to a specific application."[18] This version of the test was articulated in the now-vacated *CLS Bank International v. Alice Corp. Pty. Ltd.*, but the test has its origins in *Research Corp. Technologies v. Microsoft* ("*RCT*").[19] In *RCT*, the court held that it "will not presume to define 'abstract' beyond the recognition that this disqualifying characteristic should exhibit itself so manifestly as to override the broad statutory categories of eligible subject matter."[20]

Clearly, this test is extraordinarily broad—nearly any claim will have at least one limitation attaching the abstract idea to a specific application.[21] If a

---

17. *See id.* at 3227 (questioning the machine or transformation test's applicability to Information Age inventions).

18. CLS Bank Int'l v. Alice Corp. Pty. Ltd., 685 F.3d 1341, 1352 (Fed. Cir. 2012).

19. Research Corp. Tech., Inc. v. Microsoft Corp., 627 F.3d 859 (Fed. Cir. 2010).

20. *Id.* at 868.

21. Under previous, less restrictive law, patent drafters would often attempt to impart patent eligibility on computer-related claims by re-writing the claim so that it principally claims a computer. The Federal Circuit at one time essentially blessed this strategy *en banc* in *In re Alappat*, 33 F.3d 1526 (Fed. Cir. 1994), *abrogated by In re Bilski*, 545 F.3d 943. The court held that installing new "programming creates a new machine, because a general-purpose computer in effect becomes a special purpose computer once it is programmed to perform particular functions pursuant to instructions from program software." *Alappat*, 33 F.3d at 1545. This decision gave patent drafters free rein to write software process claims as system claims to the computer with the software installed. The Federal Circuit abrogated *Alappat* in *In re Bilski*, although the Supreme Court's affirmation of *Bilski* on alternate reasoning has allowed one post-*Bilski* Federal Circuit panel to rely on it. *See Ultramercial LLC v. Hulu LLC*, 657 F.3d 1323, 1328-29 (Fed. Cir. 2011); *see also CyberSource,* 654 F.3d at 1375 (distinguishing *Alappat*). Especially after the *Prometheus* decision, which requires "unconventional" steps beyond the natural law (or algorithm, in this case), and which also

method or system has any use whatsoever, it will not be manifestly abstract. Only claims directed to the idea alone, such as Samuel Morse's famous claim to electromagnetism,[22] would fail this test.

Although a number of scholars have argued for a broad § 101 filter,[23] such a test is no longer appropriate, as it is inconsistent with Supreme Court precedent. In both *Bilski* and *Mayo Collaborative Services v. Prometheus Laboratories, Inc.*,[24] the Supreme Court tightened § 101's requirements, rejecting the idea that § 101 is merely a coarse eligibility filter.

Prometheus—a law of nature case and thus not directly applicable to abstract ideas doctrine—emphasized the Court's *Bilski* decision, signaling its intention that § 101 restrict patentable subject matter. Two concepts from the *Prometheus* case are most likely to find their way into abstract ideas cases: (1) its step-by-step method of analyzing a claim, and (2) its focus on whether the invention adds any unconventional activity to the background principle.

Without going into unnecessary detail, the *Prometheus* decision somewhat revived a method of analyzing a claim that the Supreme Court introduced in the *Parker v. Flook* case and subsequently buried in *Diamond v. Diehr*. The *Prometheus* opinion, like in *Flook*, followed the process step by step, asking whether each step adds anything to the law of nature, which the court assumes is known.[25] Instead of stopping at the end of the process, the *Prometheus* Court, like in *Diehr*, considers the claim as a whole, asking whether the ordered combination adds to the law of nature.[26]

At first glance, the manifestly abstract test as articulated in *CLS Bank* arguably follows this approach. The test requires that at least one claim limitation direct the idea to a specific application, which is similar to going step by step through the claim and asking whether the step adds unconventional activity to the background principle. However, upon closer inspection, the manifestly abstract test is not nearly as restrictive as the methodology in *Prometheus*. In particular, having a specific application is vastly broader than adding unconventional activity to the background principle.

This distinction is plainly apparent in *CLS Bank*. The invention in *CLS Bank* is "a computerized trading platform for exchanging obligations in which a trusted third party settles obligations between a first and second party so as to eliminate 'settlement risk.'"[27] In other words, the patents claim a business method that uses an intermediary to reduce risk in certain financial

---

emphasizes the insignificant post-solution activity doctrine, *Alappat* is (or at least should be) a relic of the past. *Prometheus*, 132 S. Ct. at 1298.

22. *See* O'Reilly v. Morse, 56 U.S. 62 (1854).

23. *See, e.g.*, Michael Risch, *Everything is Patentable*, 75 TENN. L. REV. 591 (2008).

24. 132 S. Ct. 1289 (2012).

25. *Id.* at 1297-98.

26. *Id.* at 1298 ("to consider the three steps as an ordered combination adds nothing to the laws of nature that is not already present when the steps are considered separately.").

27. *CLS Bank*, 685 F.3d at 1343.

transactions—in essence, a two-sided escrow arrangement. Besides the two-sided escrow idea, the invention encompasses a general-purpose computer, which is used to exchange data and perform certain complex calculations.

Using the *Prometheus* methodology, the claims surely fail. The escrow idea is a background principle, and the computer activity is merely conventional data processing. The *CLS Bank* court, using the manifestly abstract test, concluded otherwise, holding that the claims "cover the practical application of a business concept in a specific way."[28] Looking at each of the limitations, the majority noted a number of specific steps in the business method that "do not appear to preempt much in the way of innovation."[29] Therefore, under the manifestly abstract test, the *CLS Bank* claims were not manifestly abstract because they were directed to a specific application. In the end, despite using seemingly similar language to *Prometheus*, the manifestly abstract test circumvents *Prometheus*' intent.

The Federal Circuit used a slightly modified version of the manifestly abstract test in the vacated *Ultramercial, LLC v. Hulu, LLC*[30] decision. The claims in *Ultramercial v. Hulu* recited a process for showing a consumer on the Internet an advertisement in return for giving her access to a "media product." Clearly, this process requires the Internet, as streaming video is not possible without it, but the real invention was not the ability to stream video—rather, the claims described a strategy of generating revenue from showing advertisements before giving access to the media product.

The *Ultramercial* court held that the claims constitute patentable subject matter because they "disclose a practical application" of "the mere idea that advertising can be used as a form of currency."[31] This "practical application test"—which appears in the *CLS Bank* version of the manifestly abstract test—is itself slightly controversial.[32] However, the court stated later in the opinion

---

28. *Id*. at 1355.

29. *Id*. at 1356.

30. 657 F.3d 1323 (Fed. Cir. 2011), *cert. granted, judgment vacated sub nom.* WildTangent, Inc. v. Ultramercial, LLC, 132 S. Ct. 2431 (2012).

31. *Id*. at 1328.

32. This statement is quite similar to the *State Street Bank* test that was rejected in *In re Bilski* and *Bilski v. Kappos*. State St. Bank & Trust Co. v. Signature Fin. Group, Inc., 149 F.3d 1368, 1373 (Fed. Cir. 1998), *abrogated by In re Bilski*, 545 F.3d 943, 959 (Fed. Cir. 2008) (holding that an invention is patentable if it is directed to a "useful, concrete, and tangible result"). Moreover, there is a very fine distinction in the Supreme Court's jurisprudence between disclosing a practical application and limiting an abstract idea to a particular technological field. *Cf. Bilski v. Kappos*, 130 S. Ct. at 3218 ("[T]he prohibition against patenting abstract ideas 'cannot be circumvented by attempting to limit the use of the formula to a particular technological environment' or adding 'insignificant postsolution activity'" (quoting *Diehr*, 450 U.S. at 191-92)); *Bilski v. Kappos*, 130 S. Ct. at 3230 ("an *application* of a law of nature or mathematical formula to a known structure or process may well be deserving of patent protection." (emphasis added) (quoting *Diehr*, 450 U.S. at 187)). The *Ultramercial* invention could easily fall on the unpatentable side of these statements (if they can in fact be distinguished), as it merely limits the use of the idea that advertising can

that the claim's "breadth and lack of specificity does not render the claimed subject matter impermissibly abstract."[33] This second statement refers to the claim's failure to specify a particular mechanism for delivering media content to the consumer, but conflicts with the initial reasoning that limited the claim's scope to a specific application of the abstract idea. Thus, the *Ultramercial* opinion simultaneously upheld the claims for being limited in scope and posited that scope is irrelevant to the inquiry.

The court also noted that many of the claimed steps "are likely to require intricate and complex computer programming,"[34] and that "the invention involves an extensive computer interface."[35] Although the opinion refused to "define the level of programming complexity required before a computer-implemented method can be patent-eligible," it incorporated this consideration in the manifestly abstract test.[36] Yet this cannot be a workable factor—it would be impossible for a group of computer scientists, much less a group of judges, to agree on a definition of "complex computer programming." Further, even if judges could define the term, its definition may become obsolete rapidly as science progresses. Most significantly, the claims at issue in *Ultramercial* do not even claim the complex computer programming—they claim the business method by which the complex computer programming can be made profitable. The *Ultramercial* claims, therefore, do not even pass the *Ultramercial* test. It should not be surprising, then, that subsequent cases did not follow this approach.

C. *The Scope Test*

Three post-*Bilski* cases, *Cybersource v. Retail Decisions*,[37] *Dealertrack v. Huber*,[38] and *Fort Properties v. American Master Lease*,[39] roughly follow an approach that examines whether the computer serves as a meaningful limit on

---

be used as currency in the technological environment of the Internet. This contrasts with the invention in the Supreme Court's famous *O'Reilly v. Morse* case, 52 U.S. 62 (1853), which applied the laws of electromagnetism to create the telegraph. If the "application of a law of nature" and "limit to a particular technological environment" language can be distinguished, it is likely in this manner—knowledge of the law of nature can be applied to create new inventions that rely on that law of nature, but the law of nature itself cannot be claimed, even if the claim is limited to a particular industry. Applying this thinking to *Ultramercial*, then, the abstract idea—that advertisement can be used as currency—is not applied to create a new invention, as video streaming and similar technology were already widely known. Rather, the claims merely limit the abstract idea to the particular technological field of accessing media products online.

33. *Ultramercial*, 657 F.3d at 1329 (Rader, J., concurring).
34. *Id.* at 1328 (majority opinion).
35. *Id*.
36. *Id*.
37. *CyberSource*, 654 F.3d 1366 (Fed. Cir. 2011).
38. *Dealertrack*, 674 F.3d 1315 (Fed. Cir. 2012).
39. *Fort Properties*, 671 F.3d 1317 (Fed. Cir. 2012).

the claim's scope. This language originates from the machine or transformation test: "the use of a specific machine or transformation of an article must impose meaningful limits on the claim's scope to impart patent-eligibility."[40]

Unlike the manifestly abstract test, which *CLS Bank* explicitly stated, the Federal Circuit has elaborated this test's bounds principally through the three cases that have applied it. This Note will thus explore these cases in chronological order.

Cybersource v. Retail Decisions, decided after *RCT*, involved a method for detecting credit card fraud on the Internet by comparing the Internet address on the purchase to the credit cards used at that address. The claim was not limited to any particular formula for detecting fraud, but was directed toward "constructing a map of credit card numbers," and "utilizing the map of credit card numbers to determine if the credit card transaction is valid."[41] The patent also included system claims to "a computer-readable medium" storing instructions for executing the same process as in the process claim.[42]

Much of the *Cybersource* opinion is devoted to characterizing the method claims as mental processes, which previous Supreme Court cases had held were unpatentable abstract ideas.[43] Addressing the system claims, the court found that "the incidental use of a computer to perform the mental process of claim 3 does not impose a sufficiently meaningful limit on the claim's scope."[44] The panel also rejected the notion that "simply reciting the use of a computer to execute an algorithm that can be performed entirely in the human mind" is patentable.[45]

Since *Cybersource*, however, the Federal Circuit has been less apt to apply the mental processes doctrine to computer-implemented inventions. Most recent computer-implemented inventions require computer calculations so complex or voluminous that the human mind could not practically perform them. Instead, combining *Cybersource*'s rulings on the method and system claims, the Federal Circuit has held that such inventions simply apply a general-purpose computer to these calculations, indicating that the computer does not serve as a meaningful limit on the claim's scope. Although the Federal Circuit has not characterized it as such, a general-purpose computer can, in this context, be logically considered an extension of the human mind. The ability of a computer to perform complex calculations is not the invention, and calculations are abstract mental processes, so routine applications of a computer to perform complex calculations are unpatentable abstract ideas. While proponents of the manifestly abstract test initially fought the idea that

40. *In re Bilski*, 545 F.3d at 961.
41. U.S. Patent No. 6,029,154 claim 3 (filed July 28, 1997).
42. *Id.*
43. *See Benson*, 409 U.S. at 67; *Flook*, 437 U.S. at 586.
44. *CyberSource*, 654 F.3d at 1375.
45. *Id.*

inserting a general-purpose computer into an abstract process does not necessarily make the process patent eligible, this contingent has since acknowledged that computers cannot necessarily confer patent-eligibility.[46]

The *Dealertrack* and *Fort Properties* cases apply this framework to invalidate claims. The patent in *Dealertrack* claimed a process involving a credit application clearinghouse for car dealerships. Essentially, the method comprised of "receiving credit application data," "obtaining credit report data," forwarding the data to various lenders, and receiving decisions back from the lenders.[47] Unlike in *Cybersource*, a computer system and network are entirely necessary to the process, both because the process is automated and because of the sheer amount of data that must be manipulated. This fact made it difficult for the court to invalidate the claims purely under the mental processes doctrine.

The *Dealertrack* majority held that the claim's scope was so broad as to preempt the abstract idea of processing information through a clearinghouse.[48] While the patentee—surely mindful of the manifestly abstract test[49]—argued that the phrase "computer aided" in the claims limited the abstract idea to a practical application, the court rejected the computer as failing to place any real limit on the claim's scope.[50]

The *Fort Properties* claims forced the court to directly address the effect of a computer element on a claim's patentability. The *Fort Properties* case involved two sets of claims, one without any computer limitation, and the second with a computer element. Apart from the computer limitation, the claims were identical, as both were directed to a method for manipulating a real estate portfolio to take advantage of a favorable tax provision.[51]

After rejecting the first set of claims based on *Bilski*, the court considered

---

46. *See CLS Bank*, 685 F.3d at 1353, 1355 ("mere computer implementation cannot render an otherwise abstract idea patent eligible," "the mere fact of computer implementation alone does not resolve the patent eligibility question"). However, the manifestly abstract test arguably sidesteps this doctrine by allowing any use of a general-purpose computer in a specific application to pass the test.

47. U.S. Patent No. 7,181,427 (filed Sept. 3, 1997).

48. *Dealertrack*, 674 F.3d at 1333 ("In this case, however, we are compelled to conclude that the claims are invalid as being directed to an abstract idea preemptive of a fundamental concept or idea that would foreclose innovation in this area.").

49. Addressing *Ultramercial*, the court noted that the *Ultramercial* claims require "an extensive computer interface," whereas this process is merely computer-aided. *Id.* at 1334. Ultimately, however, this distinction is unsatisfying. Although the *Ultramercial* invention, according to the *Ultramercial* court, necessitates "complex computer programming," the *Dealertrack* invention surely requires more, as the claims encompass a large interactive computer network.*Ultramercial*, 657 F.3d at 1328. The *Ultramercial* claim describes its programming steps in more detail, but the code only functions to register user mouse clicks, count the number of times an advertisement has been shown, and give the consumer access to a media product. *Ultramercial*, 657 F.3d at 1324-25.

50. *Dealertrack*, 674 F.3d at 1333-34.

51. *Fort Properties*, 671 F.3d at 1318-19.

whether the computer element transformed the otherwise unpatentable claims into patentable subject matter. Quoting *Cybersource* and *Ultramercial*, the opinion asked whether the computer "impose[d] meaningful limits on the claim's scope," or whether the invention involved "advances in computer technology."[52] As the computer element did not substantively alter the claimed process, the court concluded that "the computer limitation [was] insignificant post-solution activity," and did "not impose meaningful limits on the claim's scope."[53] Therefore, the claims constituted unpatentable subject matter.

The Federal Circuit has yet to uphold claims under the scope test, so its exact boundaries are difficult to discern. Presumably the *RCT* claims, which were upheld under the manifestly abstract test, would also pass this test. The *RCT* patent claimed a digital halftoning process useful for generating images in printing and computer displays. Halftoning allows computer displays and printers to simulate the full spectrum of colors using only a limited number of primary colors. Because the method places dots or pixels of the primary colors in particular patterns, the image can appear to the viewer as a continuous tone image. The particular method claimed a pixel-by-pixel comparison process using a "blue noise mask."[54]

Clearly, the computer is heavily involved in every step of the process. This invention certainly does not seem to be an abstract idea, but because the test is so vague, it is difficult to generate a principled reason why the computer imposes a meaningful limit on the claim's scope. The claim has no scope beyond the computer because the computer is necessary to implement the invention, but the computer is also necessary in *Dealertrack*.

The mental processes doctrine, as extended by *Cybersource*, is also little help. The *RCT* invention uses a general-purpose computer. Moreover, a human mind could presumably compare pixels using a blue noise mask even though there is no need for halftoning absent computer technology. At their most basic level, computers simply manipulate data by performing various calculations, something that human minds can also do, albeit much slower.

More fundamentally, purely scope-based tests fail because different inventions deserve different scope—some are major breakthroughs and others are incremental improvements. It is impossible to methodically determine the correct scope for each invention. Examining whether a computer sufficiently limits a claim's scope is similarly doomed—there is no principled way to discern how much limitation is enough. If courts allow any limitation to be enough, they end up right where they started—the mere presence of a computer as necessary to implement an invention confers patentability. The Federal

---

52. *Id.* at 1323 (quoting *CyberSource,* 654 F.3d at 1375 and *Ultramercial*, 657 F.3d at 1328).

53. *Id.* at 1324.

54. *See, e.g.*, U.S. Patent No. 5,111,310 (filed Dec. 4, 1990). For a fuller description of the claims and an explanation of a "blue noise mask," see *RCT*, 627 F.3d at 862-66.

Circuit has definitively rejected the notion that this is sufficient.

Therefore, the scope test provides little assistance in separating patent-eligible inventions from unpatentable claims. Although developing this doctrine using test cases over time could be possible, the Federal Circuit is so split on computer-implemented inventions that this strategy would be unlikely to succeed. Computing technology is so complex that courts could likely find some reason to distinguish nearly identical cases. The scope test simply cannot provide the needed clarity to be a workable solution.

D. *The "Integral" Test*

What this Note will call the "integral test" was articulated in *Bancorp Services, L.L.C. v. Sun Life Assurance Co. of Canada*,[55] the last case decided before the Federal Circuit granted the *en banc* petition in *CLS Bank*. The integral test resembles the scope test explained in *Fort Properties*, but differs sufficiently in its focus to be considered a separate test. The test, as elaborated by *Bancorp*, states that "[t]o salvage an otherwise patent-ineligible process, a computer must be integral to the claimed invention, facilitating the process in a way that a person making calculations or computations could not."[56] The *Bancorp* court apparently considers this to be a test for when the computer serves as a meaningful limit on the claim's scope. However, because the integral test focuses on the computer's role in the process and not the claim's breadth, it is distinct.

Bancorp invalidated claims to a method of managing a stable value protected life insurance policy. The process involved numerous precise calculations that required substantial computing power.[57] The *Bancorp* claims thus were substantively similar to those in *Fort Properties* and *CLS Bank*. The court disposed of these claims quickly under the integral test, albeit using "meaningful limits" language: "The computer required by some of Bancorp's claims is employed only for its most basic function, the performance of repetitive calculations, and as such does not impose meaningful limits on the scope of those claims."[58] Distinguishing this case from *CLS Bank*, the panel observed that "the computer limitations do not play a 'significant part' in the performance of the claimed invention."[59] These statements demonstrate that the court is less concerned with the claim's scope than the computer's role in the invention.

Although this Note builds upon the integral test's approach, the test is

---

55. Bancorp Services, L.L.C. v. Sun Life Assurance Co. of Canada, 687 F.3d 1266 (Fed. Cir. 2012).

56. *Id*. at 1278.

57. *Id*. at 1269-72.

58. *Id*. at 1278.

59. *Id*. at 1280 (quoting *CLS Bank*, 685 F.3d at 1355).

currently too undefined to be consistently administered. It is entirely unclear what the term "integral" means, despite the specification that a computer must "facilitat[e] the process in a way that a person making calculations or computations could not."[60] A person cannot feasibly perform all of the calculations required by the inventions in *Bancorp*, *Fort Properties*, *Dealertrack*, or *CLS Bank*. The court likely intends the foregoing language to exclude inventions that apply a general-purpose computer to make routine calculations. Even if the language were meant as a per se exclusion, which is unlikely, it does not adequately establish the boundaries of patentable subject matter because it defines "integral" in the negative—stating that something is not integral is useful, but it does not define what *is* integral. More elaboration, as provided by the test proposed in this Note, is needed.

E. *Summary*

The current state of abstract idea doctrine is thus characterized by total discord. Through seven cases, the Federal Circuit is split four to three. The Supreme Court has weighed in twice, but *Bilski* has only caused more confusion and *Prometheus* has so far been criticized and ignored.

The Federal Circuit has proposed three principle tests in its post-*Bilski* case law. One, contrary to the Supreme Court's intentions in *Bilski* and *Prometheus*, operates as a broad eligibility filter, rejecting claims that are so manifestly abstract as to not even be directed to a specific application. Such claims would likely be directed explicitly to fundamental ideas or algorithms in the abstract. A second test examines whether the computer serves as a meaningful limit on the claim's scope. This test is highly indeterminate, as the "meaningful limit" language is overly subjective. This scope test has developed into a third test, focusing on whether the computer is integral to the claimed invention. The integral test is also too indeterminate in its current form; however, concentrating on the computer's role in the process is a useful approach to be developed further in this Note.

Given the extent of current disagreement, courts are in dire need of a new compromise test that is easy to administer and derivative of current precedent.[61] The test proposed in this Note attempts to draw upon concepts from all three current approaches to arrive at a workable definition of an abstract idea for computer-implemented inventions. The Federal Circuit, in *en banc* consideration of *CLS Bank*, or the Supreme Court, in a later case, could adopt this test to resolve current disputes in doctrine and to provide owners of patents on computer-implemented inventions with much-needed predictability.

---

60. *Id*. at 1278.

61. No test can completely harmonize the case law, as the cases clearly contradict each other. However, the most effective proposal will use current precedent to the greatest extent possible, making the test simpler to adopt.

II. CURRENT APPROACHES IN THE LITERATURE TO ABSTRACT IDEAS

As Lemley, Risch, Sichelman, and Wagner summarize nicely in "Life After Bilski," academics traditionally conceptualize § 101 as a gatekeeper.[62] Under this view, § 101 excludes entire categories of inventions as contrary to the policies of patent law. However, there are several drawbacks to this approach. First, innovation is unpredictable; excluding certain subject areas per se may unintentionally stunt future innovation that a good patent policy would protect.[63] Second, delineating the lines between subject areas is exceedingly difficult.[64] What constitutes a business method? What about a medical diagnostic test? As an example, do the claims in *Ultramercial* recite a business method or a software process? Just as the courts have struggled to define an abstract idea, the same would happen with per se subject matter exclusions. The strategy of categorically excluding inventions has the effect of kicking the can down the road without providing additional clarity. Finally, such an approach is impractical. Courts have repeatedly declined to impose per se limitations on patentable subject matter, most recently in *Bilski*, and there is no indication the Court will go that route in the future.[65]

Another approach is to use § 112 instead of § 101 to regulate overly vague or broad claims.[66] The theory is that applicants cannot successfully enable or describe inventions when they claim broad property rights. While this idea has some merit, as it can prevent claims that go beyond the scope of the actual invention, it ultimately fails. Its principal defect is that § 112 pertains to the disclosure only at the time of filing. Although an invention may be the only known application of a law of nature or abstract idea at the time of filing, future innovation could render that understanding obsolete. Sometimes, as was the case with Morse's claim 8, the claim is clearly not enabled by the specification. Other times, a patent may claim the only conceivable application of a natural law at the time of filing, when in fact downstream innovation may discover other applications. This often happens with computer-aided inventions because the related technology develops rapidly. Therefore, § 112 is not designed to effectively screen unpatentable subject matter from patentable subject matter.

A third prominent category defines abstract ideas as inventions that do not

---

62. Mark Lemley et al., *Life After Bilski*, 63 STAN. L. REV. 1315, 1326 (2011).

63. *Id*. at 1326-27.

64. *Id*. at 1327.

65. Justice Stevens' concurring opinion in *Bilski* rejected business methods per se under § 101. Some commentators believe that Justice Stevens initially wrote a majority opinion and lost the majority before the opinion issued. *See, e.g.*, Chisum, *supra* note 3. Now that Justice Stevens is no longer on the Court, and in light of the unanimous decision in *Prometheus*, it looks unlikely that courts will impose any per se subject matter exclusions in the near future.

66. Donald S. Chisum, *Weeds and Seeds in the Supreme Court's Business Methods Decision: New Directions for Regulating Patent Scope*, 15 LEWIS & CLARK L. REV. 11, 13-14 (2011).

relate to a practical end. This category encompasses two distinct proposals. The first proposal is from Michael Risch, who argues that § 101 should be resituated as an inquiry into an invention's usefulness.[67] This approach has a number of disadvantages. First, Risch's test requires that § 101 be decided partially as a factual issue. Certain factors, such as whether there is a market demand for the invention, necessitate factual testimony. Incorporating factual questions into § 101 would compel a substantial shift in patentable subject matter doctrine, which currently regards § 101 as a wholly legal issue. Moreover, as the *Prometheus* decision reinforces, § 101 is a threshold inquiry that courts must decide at the outset.[68] Thus, a factual approach that requires expert testimony is impractical.

Second, Risch's test is either difficult to administer or woefully underinclusive—nearly any invention can be defended as useful.[69] As long as the invention is not anticipated by some prior art—the requirement of § 102—and is enabled under § 112, it likely represents improvement in the art, and thus satisfies any threshold usefulness test. Furthermore, the Supreme Court conclusively rejected this inclusive approach in *Bilski* and *Prometheus*, rendering the test unhelpful for this Note's practical purposes.[70] Alternatively, if usefulness is defined more restrictively, the test becomes as confusing as the current abstract ideas doctrine. "Usefulness" is entirely indeterminate, and, even if given more teeth, is still too vague to provide a workable test.

Finally, Risch's proposal fails because it applies a textual approach to a common law area. As mentioned earlier in this Note, the 1952 Patent Act codified existing common law on patentable subject matter.[71] Section 101 is intentionally broad, so as to not foreclose future common law development in the area.[72] Thus, extrapolating any real limitation from the text of § 101 is contrary to the provision's intent and structure.

The second proposal in the "practical end" category comes from "Life After Bilski." The authors suggest that "[c]laims are proper when the scope of the patentee's claims is commensurate with a practical, real world contribution the patentee has made."[73] In addition, they provide an incomplete list of five factors relevant to the determination, ranging from whether the technological

---

67. Michael Risch, *Reinventing Usefulness*, 2010 BYU L. REV. 1195 (2010).

68. *Prometheus*, 132 S. Ct. at 1303-04.

69. While Risch may advocate underinclusiveness as a policy matter, this paper refuses to engage in those debates. Given the Supreme Court's clear directive that the Federal Circuit must apply § 101 with some force, an underinclusive test is not a viable solution.

70. Although the patents at issue in *Bilski* and *Prometheus* were generally bad (and likely would not have survived other statutory requirements), the Supreme Court has clearly moved away from an inclusive attitude toward § 101.

71. Menell, *supra* note 7, at 1302-03.

72. *Id.* at 1296-97, 1302-03.

73. Lemley, *supra* note 62, at 1339.

field is fast-moving to whether the patentee has made an important contribution relative to the prior art.[74]

This approach has a number of weaknesses. First, a test with five or more factors would be nearly impossible to administer consistently. Any test with so many unranked factors requires judges to resolve competing factors. With a highly divided Federal Circuit, this resolution would likely become as panel-dependent as the current approach.

Second, the proposed factors are all broad and would necessitate significant discovery and trial time to implement. This would be a step in the wrong direction because the Federal Circuit is already deeply concerned about the cost and length of patent cases.[75]

Third, and similarly to the Risch usefulness test, this inquiry depends on factual issues. To their credit, the authors advocate shifting § 101 to a back-end consideration that is resolved only if the patent first survives the other requirements of the Patent Act. However, this change would likely require the Supreme Court's blessing after *Prometheus*, making it difficult for the Federal Circuit to effectively adopt.

Finally, and perhaps most importantly, the "Life After Bilski" test is simply not helpful in close cases, especially for computer-implemented inventions. Most close § 101 cases consider claims that are narrowly tailored to a practical end. For example, the claim in *Cybersource* is limited to detecting credit card fraud online, and is solely directed to that purpose. Although it is defensible to conclude that all of the relevant post-*Bilski* Federal Circuit cases should pass the § 101 test, this test may not even reject the *Bilski* claims that restrict the hedging method to energy commodity markets. Those claims are narrowly tailored to energy commodity markets and strive to protect buyers and sellers from rapid price changes—certainly a practical end. Nearly any invention can be said to have a practical end; so as long as the claims are limited to a particular field, they will likely pass the "Life After Bilski" test.

Further, several of the "Life After Bilski" factors focus on the industry or technological field, so the test fails to distinguish between inventions within each field.[76] The main remaining factor is whether the patentee has "made an important contribution relative to the prior art," an exceptionally subjective consideration. Clearly, a more specific test is needed.

---

74. *Id*. at 1339-41.

75. Randall Rader, Chief Judge, Fed. Cir., *The State of Patent Litigation*, Address at a Joint Meeting of the Fed. Circuit Bar Ass'n and the Eastern Dist. of Tex. Bar Ass'n (Sept. 27, 2011), *available at* http://www.patentlyo.com/files/raderstateofpatentlit.pdf.

76. See, for example, factors two—whether the industry relies heavily on cumulative invention—and three—whether the technological field is fast moving. The first factor, while examining the invention itself, still somewhat looks at the industry, as it asks whether the claimed invention is potentially generative of many kinds of new inventions—a highly industry-dependent factor. *See generally* Lemley, *supra* note 62, at 1339-41.

III. A TEST FOR COMPUTER-IMPLEMENTED INVENTIONS

A more specific test should focus on the computer's role in the invention. Some inventions merely use a computer to perform a great number of calculations or to transfer data from point A to B. In these cases, the computer is necessary to perform the claimed method, but it does not act as any further limitation on the process. For example, in *Benson*, the claim is to a mathematical formula only useful in the context of that calculation or data transfer, which is an unpatentable abstract idea.[77]

If a computer is to confer patentability on a claim, it should be because the invention interacts with the computer in a particular way. Any other methodology skirts the basic question: how does the computer affect the invention's patentability?

Therefore, rather than search for a poor proxy test, courts should confront the issue and explicitly scrutinize the computer's role in the process. As computers manipulate data by definition, the manipulation of data is a good focal point. Therefore, this Note proposes the following test:

> A claim is not directed to an abstract idea if it contains a computer element such that:
>
> 1) The computer manipulates data, rather than merely being present,
> 2) The data being manipulated is inherent to the computer, and
> 3) The data manipulation is directed to one or more particular applications.

The main potential ambiguity is in defining "inherent to the computer." Data is inherent to the computer if it does not directly represent information about the world.[78] For example, the *Cybersource* process manipulates data on credit card numbers. While that data is stored in the computer in an artificial, computer-readable format, that data directly represents information about the world—credit card numbers—so the *Cybersource* claim fails the test.

Take the *RCT* method by contrast. That process involved a comparison of pixels and a blue noise mask. While the pixels make up an image, which is potentially a direct representation of the world, the process operates one level lower, on the pixels themselves. This particular data, as *Bancorp* explained,[79] is

---

77. *Benson*, 409 U.S. at 71-72 ("The mathematical formula involved here has no substantial practical application except in connection with a digital computer, which means that if the judgment below is affirmed, the patent would wholly pre-empt the mathematical formula and in practical effect would be a patent on the algorithm itself.").

78. In some ways, this is similar to a test of "advances in computer technology" or "advances in programming." However, this language would not constitute a workable test, as it is highly subjective. Any new test must involve little freedom in its application because there is such disagreement within the Federal Circuit. As the post-*Bilski* jurisprudence demonstrates, vague tests like "advances in computer technology" would only fuel further debate.

79. *Bancorp*, 687 F.3d at 1279.

inherent to the computer, so the claim is not directed to an abstract idea.[80]

A. *Applying the Data Manipulation Test: Some Examples*

The following subparts apply to the data manipulation test to the remaining post-*Bilski* Federal Circuit cases, several other cases, and one hypothetical.

1. *Inventions Not Manipulating Data*

Ultramercial v. Hulu is one example of a case impacted by part (1) of the test. The *Ultramercial* claims recite a process that (a) shows the consumer an advertisement, (b) registers a response to the advertisement or verifies that the advertisement was shown, (c) gives the consumer access to a media product, and (d) updates records to reflect that the advertisement was shown. Although it is a close call, *Ultramercial* probably does not pass part (1). The process requires a computer because it takes place online, but the computer does not truly manipulate data. While online video streaming technology involves "complex computer programming,"[81] the claim does not recite a method applying that technology. Rather, the patent claims the idea of showing an advertisement before giving access to a media product. The patent does not provide an implementation of the "giving access" step, nor does it say how the computer can do something like stream video online. Thus, while the method requires the computer to manipulate data, the claim is not to this manipulation.

After that observation, the only possible data manipulation left in the process happens when the computer registers that the user has viewed the advertisement (by listening for a mouse, click for example). This manipulation is so de minimis that it should not count as data manipulation under part (1). A court applying the test, however, would not have to consider that question because the data is obviously not inherent to the computer. The data represents whether the user viewed the advertisement, or, put simply, whether the user clicked the mouse. Therefore, the outcome in the district court was correct—*Ultramercial* claims the abstract idea of using an advertisement as currency, and the computer element does nothing to save the claim.

2. *Inventions Not Involving Data Inherent to the Computer*

Like the *Cybersource* claims, the inventions in *Dealertrack*, *Fort*

---

80. To some extent, this is an inversion of how *In re Bilski* read *In re Abele*, 684 F.2d 902 (C.C.P.A. 1982) *abrogated by In re Bilski*, 545 F.3d 943. *In re Bilski*, 545 F.3d at 963 ("the electronic transformation of the data itself into a visual depiction in *Abele* was sufficient" to render the claim patentable). However, the data manipulation test proposed here arrives at the same result as the *Abele* court. *See* discussion *infra* Part IV.1.e.

81. This quote is a reference to the court's opinion. *Ultramercial*, 657 F.3d at 1328.

Properties, *CLS Bank*, and *Bancorp* would all classify as inventions not involving data inherent to the computer. While the computer is necessary to the claims in these cases—in fact, the computers in some must perform relatively complex functions—the data manipulated is not inherent to the computer. Taking *Dealertrack* as an example, the computers in the *Dealertrack* method manipulate credit card application and credit report data, transferring it between various parties. This data—credit scores, personal information, and whatever else is included in the applications—directly represents information about the world. Although the computer programmers may use computer science-specific data structures to represent the data, this does not change the fact that the data directly reflect some fact about the world. Thus, unlike *RCT*, in which the method operates on pixels, which is a step removed from a real-world image, the *Dealertrack* method operates on the real-world data itself.

The *Fort Properties* (real estate portfolio data), *CLS Bank* (financial transaction data), and *Bancorp* (life insurance policy data) claims are the same: these inventions manipulate data that directly represents information about the world. These claims would therefore be invalid under the data manipulation test.

3. *Inventions Not Directed to a Particular Application*

a. In re Abele

The *In re Abele*[82] claims demonstrate the importance of part (3) of the proposed test. *Abele*, decided prior to *Bilski*, principally involved two claims: the first was to an algorithm for computer tomography (claim 5), and the second applied that algorithm in the context of x-ray imaging (claim 6).[83] While the *Abele* court found claim 5 invalid under § 101, claim 6 survived. The test proposed in this Note would reach the same result.

Both claim 5 and claim 6 pass the first two parts of the test. The data manipulation in these claims is similar to the *RCT* process—an algorithm is used to arrange computer data into a meaningful image. Here, the data manipulated are data points in a data field. Claim 6 applies this algorithm to "X-ray attenuation data produced in a two dimensional field by a computed

---

82. 684 F.2d 902 (C.C.P.A. 1982) *abrogated by In re Bilski*, 545 F.3d 943.

83. *Id*. at 908. For reference, the claims are reproduced in full below:

> 5. A method of displaying data in a field comprising the steps of calculating the difference between the local value of the data at a data point in the field and the average value of the data in a region of the field which surrounds said point for each point in said field, and displaying the value of said difference as a signed gray scale at a point in a picture which corresponds to said data point.
>
> 6. The method of claim 5 wherein said data is X-ray attenuation data produced in a two dimensional field by a computed tomography scanner.

U.S. Patent No. 850,892 (filed Nov. 15, 1977).

tomography scanner."[84]

Part (3) of the test asks whether the "data manipulation is directed to one or more particular applications."[85] This language sets a low bar for claims, so tomography is a particular application under part (3), and claim 6 is patentable. On the other hand, claim 5, which is not restricted to tomography, clearly fails part (3). Claim 5 expounds the algorithm in the abstract, speaking only of a "data field" and a "data point." At no point is there any context for this data, nor is there any concrete discussion of the data's use. Despite part (3)'s relative leniency, claim 5 would not meet this basic condition.

b. *Linked List*

To further explore the bounds of part (3), take the linked list data structure in computer science. A linked list is a "list implemented by each item having a link to the next item."[86] In other words, each item holds certain information, which is ordered by the item also storing the memory location of the next item in the list. The linked list is a basic concept in computer science that is used in a multitude of programs.

Assume that a programmer at one moment in time invented the linked list. Could he patent the idea? Under *Benson*, the answer is no because the patent would preempt the whole mathematical formula, as it "has no substantial practical application except in connection with a digital computer."[87] However, the linked list patent, like claim 5 in *Abele*, easily passes the first two parts of the data manipulation test. To stay true to *Benson*, the test requires a third prong relating to the invention's application. As discussed earlier, applied algorithms are patentable, but algorithms limited to a particular technological field are not.[88] Thus, all that part (3) requires is that the algorithm be directed to some concrete use. As in *Abele*, the algorithm itself is unpatentable, but the algorithm applied in a particular process is patentable subject matter.

Turning back to the linked list, it is only patentable in the context of the particular program in which it was first invented. This is directed to a particular application, and since it is such a basic concept with simple application to other programs, it is likely obvious in any future software. To be clear, however, the

---

84. U.S. Patent Application No. 850,892 claim 6 (filed Nov. 15, 1977). This application issued after the *Abele* decision as U.S. Patent No. 4,433,380 on Feb. 21, 1984.

85. *See supra* Part III.

86. Paul E. Black, *Linked List*, in DICTIONARY OF ALGORITHMS AND DATA STRUCTURES (National Institute of Standards and Technology 2011), *available at* http://xlinus.nist.gov/dads/HTML/linkedList.html.

87. *Benson*, 409 U.S. at 71-72 ("The mathematical formula involved here has no substantial practical application except in connection with a digital computer, which means that if the judgment below is affirmed, the patent would wholly pre-empt the mathematical formula and in practical effect would be a patent on the algorithm itself.").

88. *See supra* note 32.

linked list does impart § 101 eligibility in future programs. The concept itself may not be patented. Patentability for concepts is thus narrowly defined, allowing for downstream innovation, while preserving the patentability of computing breakthroughs, as in *RCT* or claim 6 in *Abele*.

4. *Inventions Passing the Data Manipulation Test*

Another instructive case is *SiRF Technology, Inc. v. International Trade Commission* (*SiRF*),[89] which the Federal Circuit decided between its *In re Bilski* decision and the Supreme Court's *Bilski v. Kappos*.[90] *SiRF* involved patents on improved GPS (Global Positioning System) technology. The method claims at issue claimed a process for calculating a GPS receiver's position based on satellite signals. The court held that the claims recited patentable subject matter,[91] a correct outcome according to the test proposed here.

Applying the test, the claims easily pass part (1), as the process interprets data. The claims satisfy part (2) as well, because the process manipulates satellite signals, which is data inherent to the satellites. This case is quite difficult to decide in the post-*Bilski* world, where the machine or transformation test is a useful and important clue, but not the sole test. First, the court would have to decide, without any further guidance, whether a GPS receiver is a "special purpose computer," satisfying the machine or transformation test.[92] Then, even if it does pass the machine or transformation test, the court would have to blindly decide whether this claim is otherwise an abstract idea. However, under the data manipulation test, this case is quite easy. The computer manipulates satellite signals (parts 1 and 3), data that is inherent to the computer (part 2).

5. *New Computer Systems Distinguished From Computer-Implemented Inventions*

One final example illustrates the important difference between machine and process claims. A company called Paice recently filed suit against Hyundai and Kia for infringing three of its patents on hybrid vehicles.[93] Although the three patents all claim entire hybrid vehicles, the invention is really a hybrid car engine controller, which is a computer system that regulates the various electric and combustion motors in a hybrid car.[94] Unlike a process claim, which

89. SiRF Tech., Inc. v. Int'l Trade Comm'n, 601 F.3d 1319 (Fed. Cir. 2010).

90. *SiRF* was thus decided purely on the machine or transformation test.

91. *SiRF*, 601 F.3d at 1333.

92. *See, e.g.*, *Cybersource*, 654 F.3d at 1374-76.

93. Paice LLC v. Hyundai Motor Co., No. 1:12-cv-00499-WDQ (D.Md. Feb. 16, 2012).

94. *See, e.g.*, U.S. Patent No. 7,104,347 col. 58 l. 13-37 (filed Mar. 7, 2003); U.S. Patent No. 7,237,634 col. 58 l. 2-27 (filed Jan. 13, 2006); U.S. Patent No. 7,559,388 col. 56

describes the steps performed, these claims characterize the controller in terms of its functions.[95]

Claims to a hybrid car engine controller (or, technically, an entire hybrid vehicle) presumptively fall within the "machine" or "manufacture" categories of § 101. However, as shown by *Cybersource*, it is possible for a claim that was drafted as a machine claim to be analyzed as a process claim.[96] When the claim is really a process claim written as a system claim, courts sometimes refuse to "exalt form over substance."[97] Indeed, to prevent claim drafters from circumventing § 101 limitations, the Federal Circuit has recently scrutinized system claims to ensure that they are not disguised processes.

The data manipulation test, in line with the questions presented in the *CLS Bank en banc* grant, solves this problem by characterizing all such claims as "computer-implemented inventions" and refusing to treat otherwise identical system and process claims differently. While there is sometimes a fine line between new computer systems and computer-implemented inventions, courts should be trusted to discern whether the invention is a new computer system or merely implemented by a computer. They have already proven adept at doing so in *Cybersource* and *Fort Properties*. Moreover, posing the initial question, "is this invention a computer or is it merely implemented by a computer?" quickly exposes claims to processes that are written as system claims.

When the hybrid car engine controller claims are scrutinized, it is clear that they are directed to machines rather than software. As the controller is described in terms of its components each performing certain functions, the patent claims the computer rather than the computer's software.[98] These claims are materially different than claims to a "computer-readable medium"—so-called *Beauregard* claims[99]—and are therefore system claims. As true machine

---

l. 42-67 to col. 57 l. 1-5 (filed May 8, 2006).

95. *See, e.g.*, *supra* note 94, '634 Patent col. 58 l. 19-27 ("wherein the controller is operable to operate the engine when torque required from the engine to propel the hybrid vehicle and/or to drive one or more of the first or the second motors to charge the battery is at least equal to a setpoint (SP) above which the torque produced by the engine is efficiently produced, and wherein the torque produced by the engine when operated at the SP is substantially less than the maximum torque output (MTO) of the engine.").

96. *Cybersource*, 654 F.3d at 1373-75.

97. *Id.* at 1374; *Abele*, 684 F.2d at 909.

98. Of course, it is possible to claim software as a computer and its components. However, these claims, such as the system claims in *CLS Bank*, are easily identified as software claims because the computer components are defined in terms of the steps they perform in the method.

99. Claims to a "computer-readable medium," often containing program instructions (software) to perform a function stem from *In re Beauregard*, 53 F.3d 1583 (Fed. Cir. 1995). Along with *Alappat*, 33 F.3d 1526, the Federal Circuit essentially allowed all software claims to become patentable if they were written as system claims. Recent decisions have called these cases into question, and *Alappat* has even been abrogated. *See Cybersource*, 654 F.3d at 1373-74; Ex Parte Jodi L. Coppinger, David P. Delay, Brian J. Levine, & Frank A. Pavelski, No. 2009-009934, 2011 WL 798170, at *2 (Bd. Pat. App. & Interf. Mar. 2, 2011)

claims, § 101 presents no real obstacle, so these patents claim patentable subject matter. In contrast, the software used in the hybrid car engine controller would be subject to the data manipulation test. The result of this inquiry would likely turn on whether the processes manipulate data inherent to the computer.

This correlates with *Cybersource*'s distinction between a general-purpose computer and a special-purpose computer. The addition of a general-purpose computer does not render an otherwise unpatentable software claim patentable under *Cybersource*. A process involving a special-purpose computer is often patentable.[100] Put simply, if the computer is designed for a special-purpose function, it can be patented using machine claim language. If it is not, the data manipulation test does not allow an otherwise unpatentable process involving a general-purpose computer to be patentable as a system because it does not distinguish between system and process claims directed to identical subject matter. Therefore, if a new general-purpose computer is developed, the inventor can claim it using machine claim language. The data manipulation test will not allow unpatentable processes to be patentable under the fiction that they constitute a new machine.[101]

B. *Advantages of the Data Manipulation Test*

This Note now will discuss four advantages of the proposed test: 1) it is simple to administer and avoids confusion, 2) it is difficult to draft around, 3) it promotes good policy, and 4) it is drawn from current precedent, so it is straightforward to adopt.

1. *Simple to Administer*

As discussed previously in this Note, the current state of the law is highly confusing.[102] *Bilski v. Kappos* directs courts to consider the machine or transformation test, but it also tells them to look at other unnamed factors.[103] Perhaps the Supreme Court hoped that the Federal Circuit would develop new standards in the wake of *Bilski*, but the earlier survey of post-*Bilski* Federal

---

(distinguishing *Beauregard*); *In re Bilski*, 545 F.3d 943 (abrogating *Alappat*).

100. *See SiRF*, 601 F.3d at 1332-33.

101. As is fairly obvious, this statement directly contradicts the Federal Circuit's *en banc* ruling in In re Alappat, 33 F.3d 1526 (Fed. Cir. 1994) *abrogated by In re Bilski*, 545 F.3d 943. Although *Alappat* may still technically be valid precedent after the Supreme Court affirmed *In re Bilski* on different reasoning, it is extinct in practice. This is especially true now that adherents to the manifestly abstract test in *CLS Bank* blessed the idea that the addition of a general-purpose computer does not render an otherwise unpatentable process patentable. The Federal Circuit should bury *Alappat* with finality and abrogate it in the *CLS Bank* rehearing.

102. *See supra* p. 1.

103. *Bilski v. Kappos*, 130 S. Ct. at 3227 ("The machine-or-transformation test is not the sole test for deciding whether an invention is a patent-eligible 'process.'").

Circuit cases demonstrates that no progress has been made. Sometimes courts use the machine or transformation test and sometimes they do not, but all opinions include some wandering through various abstract ideas concepts before arriving at a conclusion. Without any guiding principles, it should come as no surprise that the current split has arisen. Moreover, for all judges, who are firmly committed to coming to the correct decision by the rules, this rule-less wasteland of post-*Bilski* jurisprudence is excruciatingly difficult to maneuver. Thus, especially given the variety of opposing views among members of the Federal Circuit on the issue, the need for a clear test that courts can consistently apply is obvious.

The test proposed in this Note provides a solution. Although the machine or transformation test is relatively definitive, serious debate arose over what constitutes a machine or a transformation. The data manipulation test is much less ambiguous. Most claims will pass part (1), which is fairly straightforward—either the computer manipulates data or it does not. For part (2), lawyers will be able to make superficial arguments (e.g. "the process manipulates the individual bits"), but ultimately it is a simple application of the facts as to whether the computer data represents something about the real world (e.g. credit card numbers) or whether the process works on data inherent to the computer (e.g. the pixels of a computer image). This is a workable test that requires minimal briefing and few judicial resources to apply.

2. *Difficult to Draft Around*

The distinction between data inherent to the computer and data not inherent to the computer is easy to administer, but at first glance it also appears easy to draft around. Upon closer inspection, however, such an attempt would be fruitless. If a claim were drafted to add an extra level of indirection, making the data appear inherent to the computer, it would fail to catch infringers, who have no incentive to inefficiently use that indirection. Consider *Cybersource* again: if we modified the *Cybersource* claim so that the process acted on data representing the credit card data, what would result? *Cybersource* may claim patentable subject matter, but competitors would never infringe, as they would manipulate the credit card data directly.

Cybersource could claim infringement under the doctrine of equivalents, but courts can reject this argument in one of three ways. First, the doctrine of equivalents is limited by prosecution history estoppel.[104] If the patentee originally claimed a process without the extra level of indirection, and later added the indirection to maneuver around § 101, estoppel would prevent the patentee from asserting that the claim covers the process without indirection. Second, courts can apply the ensnarement defense to limit the patent to

104. Festo Corp. v. Shoketsu Kinzoku Kogyo Kabushiki Co., Ltd., 535 U.S. 722, 732-40 (2002).

patentable subject matter.[105] The doctrine of equivalents is meant to give patents a fair scope. If the patentee never could have obtained a patent on a process that it claims infringes under the doctrine of equivalents, the doctrine of equivalents should not apply. Finally, and most simply, courts can find that an extra level of indirection does not perform the same function "in substantially the same way," as the doctrine of equivalents requires.[106]

Part (1) is similarly difficult to draft around. Patent drafters can add some data manipulation to the process, but two problems result. First, this requires an extra step in the process, one that potential infringers may not take. Second, the added data manipulation must pass part (2) of the test. In all, it is nearly impossible to artificially include data manipulation steps in the process that infringers will take and that avoids all the pitfalls of drafting around part (2) mentioned above.

Take *Ultramercial* as an example. Suppose that the *Ultramercial* claim included a database of information related to the process, and the process was written in terms of that database (i.e. A method comprising . . . "recalling the number of times the advertisement has played" . . .). This may allow the *Ultramercial* claims to get past part (1), but it still fails part (2). That information, such as the number of times an advertisement is shown, is not inherent to the computer. Potential infringers may also be able to cheaply design around the claim by creating a different system for storing and interacting with the data.

Finally, drafters can maneuver around part (3), but it forces claims to be limited, freeing up downstream innovation. Part (3) thus accomplishes its purpose of preventing broad patents on pure algorithms or ideas.

3. *Fosters Good Policy*

As a policy matter, using the extra level of indirection in part (2) to divide patentable and unpatentable subject matter is important because it successfully distinguishes between bland applications of computers as efficient information processors and real advances in computer technology. This Note does not intend to enter the broader policy debates around § 101, but as it proposes a new rule, it would be incomplete without a brief mention of the policies it furthers.

New contributions to computer technology should be given robust patent protection.[107] The data manipulation test clearly provides that protection, as

---

105. Depuy Spine, Inc. v. Medtronic, Inc., 567 F.3d 1314, 1323 (Fed. Cir. 2009).

106. Graver Tank & Mfg. Co. v. Linde Air Prods. Co., 339 U.S. 605, 608 (1950).

107. *See* Menell, *supra* note 7, at 1312-13, which argues that patent protection should be restricted to contributions to the technological arts, such as improvements in computers themselves (e.g. increasing processor speed or memory capacity). This Note does not go that far, as the test includes inventions like the halftoning process in *RCT*.

new inventions in computing technology involve data inherent to the computer. If the invention does something new with internal computer data, it very likely enhances computer or computer-related technology.[108] As the data manipulation test is not a gatekeeping test, it does not categorically exclude any particular kind of patent. However, claims that merely use computers to quickly process large data sets, such as the claims in *Cybersource* and *Fort Properties*, will have a difficult time surviving.

4. *Drawn from Current Precedent*

The data manipulation test follows smoothly from current precedent, so it would be easy for courts to adopt. The Supreme Court's *Bilski* opinion leaves the door wide open for a new rule to supplant the machine or transformation test, as it rejects the machine or transformation test as the sole test and invites the Federal Circuit to develop "other limiting criteria."[109] Moreover, the data manipulation test expounds on language in several post-*Bilski* decisions requiring that a computer element "impose a meaningful limitation on the claim's scope."[110]

The data manipulation test also squares with earlier Supreme Court doctrine on abstract ideas. Like *Benson*, it rejects claims to the idea itself. *Flook* introduces the concept of "post-solution activity," which indicates that the addition of a computer to efficiently process data is insufficient to transform an unpatentable claim into a patentable one. The data manipulation test internalizes this doctrine in part (2), which invalidates claims like the ones in *Cybersource* and *Fort Properties*. Moreover, this Note's test squares with *Flook* and *Prometheus*' admonition to assume that the discovered law of nature is known. The test examines each step of the process for the manipulation of data inherent to the computer. As applied in any process that passes part (3) and satisfies §§ 102 and 103, such data manipulation is unconventional activity beyond the background algorithm. Finally, as in *Diehr*, the test prevents claims from preempting the abstract idea by requiring them to recite specific, practical applications of data manipulation.

One common but misguided criticism is that, by requiring that the data manipulated be inherent to the computer, the passing claims will be more

---

108. Practical considerations prevent something like "enhancing computer or computer-related technology" from being the test. Courts cannot consistently administer this test, just as the *Life After Bilski* test fails in its application because it cannot be clearly defined. One main reason the case law surrounding abstract ideas is so convoluted is the fact that "abstract idea" is nearly impossible to define. Merely defining "abstract idea" as something that is also difficult to define only perpetuates the problem.

109. *Bilski v. Kappos*, 130 S. Ct. at 3231.

110. *E.g.*, *Cybersource*, 654 F.3d at 1375; *Dealertrack*, 674 F.3d at 1333; *Fort Properties*, 671 F.3d at 1323. Note that this language was introduced in *In re Bilski*, but the Supreme Court never disapproved of this part of the opinion in *Bilski v. Kappos*.

abstract—not less, as the abstract ideas exception contemplates. This misunderstanding conflates the meaning of abstraction in patent law with its meaning in computer science. In computer science, data inherent to the computer is sometimes referred to as one "level of abstraction" away from the real world. In contrast, the abstract ideas concept in patent law refers to fundamental ideas or algorithms. The data manipulation test, by requiring a particular application in part (3), prevents basic ideas from being patented, and, by requiring the manipulation of data inherent to a computer, moves further into the world of technology. The resulting inventions, rather than being algorithms or routine applications of computers to large data sets, will principally be advances in computer technology—core subject matter that patent law intends to protect.

Lastly, the test strikes a balance between all sides of the debate. It refuses to invalidate all business method patents per se; it allows processes that do not improve computer hardware, so long as they further the broader improvement of computer technology; and it excludes claims that merely run a large data set through a computer-implemented algorithm. The data manipulation test borrows concepts from all three of the current approaches: part (1) from the scope test, part (2) from the integral test, and part (3) from the manifestly abstract test. Of the seven post-*Bilski* Federal Circuit decisions, only *Ultramercial* and *CLS Bank*, decisions since vacated, would reach a different result. Furthermore, the test is friendly to other claims that courts have upheld in cases like *SiRF* and *Nazomi Communications v. Samsung Telecommuncations*.[111] Thus, rather than reject one view entirely, the data manipulation test presents a realistic compromise that the Federal Circuit could adopt in *en banc* consideration of *CLS Bank*.

CONCLUSION

Something must be done about § 101 and computer-implemented inventions. The Federal Circuit is panel-dependent on the issue, and the Supreme Court has not provided any useful guidance. Thankfully, the Federal Circuit has recognized the problem and granted an *en banc* petition in *CLS Bank*.

This Note presents a solution. The data manipulation test avoids vague legal standards in favor of a simple factual inquiry. It also ensures greater accuracy by examining the process itself, refusing to use a distant proxy like the machine or transformation test. Finally, and perhaps most importantly, it stakes out middle ground between the various positions and upsets little precedent, making it a viable compromise answer for the court to adopt in *CLS Bank*.

---

111. No. C-10-05545 RMW, 2012 WL 967968 (N.D. Cal. Mar. 21, 2012).

STANFORD TECHNOLOGY LAW REVIEW
Volume 16, Number 1 Fall 2012

NOTE

ANTICOMPETITIVE TYING AND BUNDLING ARRANGEMENTS IN THE SMARTPHONE INDUSTRY

Thomas H. Au*

CITE AS: 16 STAN. TECH. L. REV. 188 (2012)
http://stlr.stanford.edu/pdf/smartphonetying.pdf

ABSTRACT

As smart technologies become more prevalent and mobile devices become the digital platform of choice, how will antitrust law adapt? While current tying law has been criticized for its reliance on dated physical product precedents, the principles of tying and bundling doctrines are well-suited to address the next technology-based product combinations and integrations. Smartphones are an ideal foil for emerging antitrust issues, as these devices stand at the crossroads of tying and bundling inquiries. Integration of a smartphone mobile operating system ("mOS") with an application clearinghouse exemplifies the types of challenges that will test current tying and bundling doctrines. These dynamic high technology products exhibit a dual nature, as they are easily cognizable as separate products, yet they can also be characterized as a single product. This Note examines several permutations of tying tests to draw out the consequences of treating smartphone technologies as single or multiple products under the law.

To ensure the viability and increase the efficacy of current tying and bundling applications by courts, this Note recommends that (1) non-equilibrium-based assertions by either plaintiffs or defendants should be recognized as the

---

* B.A., Kenyon College (2008); J.D., Case Western Reserve University School of Law (expected 2013); U.S. Fulbright Fellow (U.S.-P.R.C. Comparative Competition Law, 2009-2010). For helpful inspiration, discussions, and suggestions, I thank Peter Gerhart, Raymond Ku, Casey R. Triggs, Kristian Rogers, Justina M. Wasicek, and Thomas Y. Au.

exception, not the norm; (2) flexible rule of reason analyses should be utilized in high-technology industries; and (3) evidentiary tools, such as natural experiments and price-indexes, should be considered a permissible inference of anticompetitive arrangements. Although it is unorthodox to suggest that current tying and bundling law does not need to be significantly altered in order to address the issues raised by dynamic high technology industries, effective integration of evidentiary and econometric analyses rises to the challenge—while maintaining a modest role for antitrust in the economy.

INTRODUCTION

There are over half a million apps and counting, on the iPhone. Apps that can take you anywhere, do anything. You might say there's no limit to what this amazing device can do. *So the question to ask is, why would anyone want*

to limit the iPhone?[1]

An anonymous hand holds a phone in front of a camera. With a couple of light taps, the user accesses applications displaying everything from the current skiing conditions to walking directions to a parked car.[2] Advertisements for smartphones frequently emphasize the phone's access to third party applications—applications accessible through a proprietary application market and which may only be available for an additional cost.[3] Consumers and critics have acclaimed products, like smartphones, for generating new secondary markets and adding value by integrating the functions of many different products that were, until recently, only available separately.[4] While integrated consumer technology products may produce procompetitive efficiencies, not all product combinations are procompetitive. By integrating products, producers may be presented with greater opportunities to evade rate regulation in their respective technology markets or increase barriers to entry for firms. Essentially, products can be 'tied' or 'bundled'[5] in a manner that generates

---

1. Sprint, "*Unlimited" Commercial*, YOUTUBE (Nov. 27, 2011), http://www.youtube.com/watch?v=GOazR24VNkM (emphasis added); this Note does not single out any particular company or manufacturer; the same antitrust theories could be applied in regards to any company that produces both a smartphone operating system and an application market, except where specifically stated otherwise. This Note only uses smartphones and related examples as hypothetical illustrations of the challenges facing antitrust law. Commercial asserting the utility of applications on a particular mobile operating system exist for all major manufacturers. *See also* Googlenexus, *Galaxy Nexus*: *Calling All*, YOUTUBE (Nov. 16, 2011), http://www.youtube.com/watch?v=CdD8s0jFJYo&feature=BFa&list=PL3FD7B7B5B (Google's Android); AT&T, *"Moby Dick" Commercial*, YOUTUBE (Sep. 26, 2011), http://www.youtube.com/watch?v=91EynpbLrg8; Research in Motion, *"Super Apps – Personal Assistant" Commercial*, YOUTUBE (Jan. 10, 2011), http://www.youtube.com/watch?v=grdJfW7BWZE (Research in Motion's Blackberry).

2. Apple, *"There's An App For That" Commercial*, YOUTUBE (Feb. 4, 2009), http://www.youtube.com/watch?v=szrsfeyLzyg.

3. It is not disputed that many applications are free of charge. However, free applications may also serve as a competitive barrier or may reflect a zero-pricing phenomena that reinforces a company's monopoly. *See infra* Part V.B.ii.

4. *See generally* Janna Anderson & Lee Rainie, *The Future of Smart Systems*, PEW INTERNET & AMERICAN LIFE PROJECT (Jun. 29, 2012), *available at* http://pewinternet.org/Reports/2012/Future-of-Smart-Systems.aspx?utm_source=Mailing+List&utm_campaign=5976672c14-Newsletter_07132012&utm_medium=email ("By 2020, experts think tech-enhanced homes, appliances, and utilities will spread, but many of the analysts believe we still won't likely be living in the long-envisioned 'Homes of the Future.' Hundreds of tech analysts foresee a future with "smart" devices and environments that make people's lives more efficient.").

5. Franklin M. Fisher, *Innovation and Monopoly Leveraging*, *in* DYNAMIC COMPETITION AND PUBLIC POLICY: TECHNOLOGY, INNOVATION, AND ANTITRUST ISSUES 139, 139 (Jerry Ellig ed., 2001) (Although similar, tying and bundling are not technically synonymous. "Tying occurs when a seller of product A requires all purchasers of A also to purchase product B from it. Sometimes this means that purchasers are required to purchase B from the same seller if they purchase A (e.g., block booking in movie exhibition)[sic]; sometimes it means that purchasers of A, *if they wish to buy B*, must do so from the seller of

higher consumer prices. Inklings of these anticompetitive concerns are already percolating in a number of dynamic technology markets. Due to the complexity of consumer products and the difficulty in ascertaining follow-on costs, consumers may not even be aware[6] they are paying higher prices than they should, if the markets were more competitive.[7] One such market that may exemplify these economic dynamics is the market for applications that market and sell other applications, or "application clearinghouses". Because consumers typically only have one such application clearinghouse on a device (which is normally the application clearinghouse produced by the mobile operating system producer), price transparency decreases and costs for third-party applications can rise. However, not all such arrangements are anticompetitive. Accurately assessing whether such arrangements are procompetitive or anticompetitive is a major function of antitrust law.[8]

Antitrust law exists to maintain competitive markets, prevent market failures, and protect consumers.[9] However, application of antitrust law to emerging, layered technologies remains controversial.[10] Overextension of antitrust law to emerging, integrated products can stymie innovation—decreasing consumer welfare.[11] By the same token, refusing to extend antitrust

---

A." Whereas, "[i]n bundling, the seller of A automatically includes B as part of the sold package, but does so at no separately stated charge." In other words, the price of a tied product is explicit while the price of a bundled product is not).

6. *Compare* Jefferson Parish Hosp. Dist. No. 2 v. Hyde, 466 U.S. 2, 14, n.24 ("Especially where market imperfections exist, purchasers may not be fully sensitive to the price or quality implications of a tying arrangement, and hence it may impede competition on the merits.") *with* Town Sound and Custom Tops, Inc. v. Chrysler Motors Corp., 959 F.2d 468, 489 (3d Cir. 1992) ('consumer surprise' argument: "[t]he plaintiffs' other primary theory of causation rests on the premise that many car buyers do not comparison shop for autosound systems at the time of their car purchase . . . [t]hey therefore may be surprised after purchasing the car to find that the . . . system was not as good as they expected. In contrast . . . if consumers purchased (or at least could purchase) . . . separately, they would be more likely to get what they wanted at a fair price. To put it bluntly, . . . [the] tie-in may well take advantage of consumer laziness.").

7. In fact, all one can actually know is that consumers are paying at or below their individual 'reservation' prices. *See generally* Kamel Jedidi, Sharan Jagpal & Puneet Manchanda, *Measuring Heterogeneous Reservation Prices for Product Bundles*, in 22 MARKETING SCI. 107 (2007) (for a model of heterogeneous reservation prices across consumers for low priced products and bundles).

8. *See* Town Sound and Custom Tops, Inc. v. Chrysler Motors Corp., 959 F.2d at 492 ("The antitrust laws were designed to deal with markets that are actually or potentially uncompetitive (in the sense of having higher prices and lower quantities produced than in a competitive market).").

9. Orrin G. Hatch, *Antitrust in the Digital Age*, *in* COMPETITION, INNOVATION AND THE MICROSOFT MONOPOLY: ANTITRUST IN THE DIGITAL MARKETPLACE 20, 20 (Jeffery A. Eisenach & Thomas M. Lenard eds. 1999) ("[T]he role of our antitrust laws is to maximize consumer welfare . . . . The basic premise is that antitrust protects 'competition' in the marketplace, and that a competitive marketplace enhances consumer welfare.").

10. *See id.*

11. *See id.*

laws may lead to higher prices—which also results in a decrease in consumer welfare.[12] Yet, by striking an appropriate balance, consumers can benefit both from enhanced innovation and lower prices.[13]

Although such a goal[14] is laudable, determining when product combinations shift from procompetitive offerings to anticompetitive ties or bundles is challenging in the real world. Gone are the days of *AT&T*,[15] when the competitive concerns over technologies, such as telephone services, were well defined and the injuries on consumer prices were more easily perceived. Clear perception of competitive concerns in modern technology markets, such as the smartphone market, are obscured by rapid cycles of innovation and opaque pricing strategies.[16] However, simply because anticompetitive conduct is difficult to detect does not mean that it is not present.[17] Rather, obscurity increases the likelihood that anticompetitive actions will be taken to increase market share or even short-term profits.[18] Teasing out the true nature of competition for multilayered technology products and applying antitrust concepts, such as tying and bundling, to anticompetitive behaviors is a critical challenge as antitrust law adapts to evolving markets.

This Note illustrates the dual nature of dynamic high technology markets by evaluating the integration of a smartphone mobile operating system ("mOS") with an application clearinghouse. An application clearinghouse is a middleware-market software application that identifies, sells, and distributes downstream third-party applications (*e.g.*: the iTunes Store, Blackberry's

---

12. *See id.*

13. *See id.* ("Proper antitrust enforcement plays an important role in protecting free markets. From Adam Smith to Robert Bork, free market, free enterprise proponents have long recognized as much. So let me debunk the myth that economic conservatives do not believe in antitrust. To the contrary, we believe strongly in antitrust – so long as the role of antitrust is understood properly and not overextended.").

14. Frank H. Easterbrook, *When Is It Worthwhile to Use Courts to Search for Exclusionary Conduct?*, 2003 COLUM. BUS. L. REV. 345, 346 (2003) ("The goal of antitrust, to be more precise, is preventing the allocative loss that comes about when firms raise price over long run marginal cost, and thus deprive consumers of goods for which they are willing to pay more than the cost of production.").

15. United States v. AT&T, 552 F. Supp. 131 (D.D.C. 1982), *aff'd sub nom.* Maryland v. United States, 460 U.S. 1001 (1983).

16. *See infra* note 199 and accompanying text.

17. *See* Easterbrook, *supra* note 14, at 345 ("Aggressive, competitive conduct by any firm, even one with market power, is beneficial to consumers. Courts should prize and encourage it. Aggressive, exclusionary conduct is deleterious to consumers, and courts should condemn it. The big problem lies in this: competitive and exclusionary conduct look alike.").

18. *See* Transamerica Computer Co., Inc. v. Int'l Bus. Machines Corp., 698 F.2d 1377, 1387 (9th Cir. 1983) ("[T]his court has already recognized that prices exceeding average total cost might nevertheless be predatory in some circumstances. The specific example we discussed was 'limit pricing,' in which a monopolist sets prices above average total cost but below the short-term profit-maximizing level so as to discourage new entrants and thereby maximize profits over the long run.").

Appworld, Google's Play[19], or the Amazon Appstore are all applications that sell other applications, like Angry Birds). Obviously, there are many ways to understand the market for these software products. However, if a mOS (such as the Android mOS) ties or bundles the mOS with an application clearinghouse (like Google Play), then it is possible for the producer to exert market power on the market for third-party application clearinghouses. The result is increased application prices for end-user consumers. Ties like this could also be used, in an anticompetitive way, to buttress a producer's mOS market share by controlling or manipulating the market for applications that run on that mOS. These types of claims are similar to those raised in the *Microsoft* cases[20]—yet the viability of such claims remains largely unresolved.

In order for tying and bundling doctrines to remain effective in high technology industries, tying and bundling must be adapted to address: (1) non-equilibrium-based economic theories; (2) dynamic shifts in technology markets that may suggest a multiple products at one timeframe and single products in a later time frame, or vice-versa; and (3) flexible, sub-market-specific timeframes for analyzing competitive effects.

To illustrate these challenges, this Note first identifies the emerging issues coming to the fore in tying and bundling inquiries. Part I describes the nature of the competitive problem plausibly resulting from the integration of smartphone mOS and application clearinghouses. Part II discusses the impact of the *Microsoft* antitrust cases. Finally, in Part III this Note applies tying and bundling legal doctrines to a hypothetical challenge of the smartphone mOSs and application clearinghouses. This illustration identifies several potential pitfalls in evidentiary and economic analyses that can distract courts from effective and consistent application of tying law to high technology markets.

This Note concludes that current tying and bundling law is sufficiently flexible to address future high-technology antitrust issues. However, to ensure the viability and increase the efficacy of current tying and bundling applications by courts, this Note recommends that (1) non-equilibrium-based assertions by either plaintiffs or defendants should be recognized as the exception, not the norm; (2) flexible rule of reason analyses should be utilized in high-technology industries; and (3) evidentiary tools, such as natural experiments and price-indexes, should be treated as permissible inferences of anticompetitive arrangements. Although it is unorthodox to suggest that current tying and bundling law does not need to be significantly altered in order to address the issues raised by dynamic high technology industries, effective integration of evidentiary and econometric analyses rises to the challenge while maintaining antitrust's modest role in the economy.

---

19. Previously known as Android Market.
20. *See infra* Part II.

I. THE FUTURE OF TYING AND BUNDLING IN ANTITRUST LAW: SEARCHING FOR A ROLE IN HIGH TECHNOLOGY

As high technology cases increasingly challenge economic perceptions of how markets function and antitrust litigation becomes more time-intensive and expensive, antitrust law faces an existential crisis.[21] Does antitrust have a role to play in evolving high technology markets?[22] Exactly how should antitrust law, and specifically doctrines like tying and bundling, be applied in high technology industries?[23] Moreover, even if there is a role for antitrust to play, are high technology markets so inherently dynamic[24] or functionally isolated[25] that they simply do not raise the kinds of competitive concerns addressed in traditional antitrust inquiries?

These issues cannot be resolved through abstract discussion alone. Abstract discussion of antitrust principles does little to advance the understanding of the limitations and potential of antitrust doctrines that depend entirely on the characteristics of the market and the behaviors involved. Likewise, current precedent provides little insight into the outer boundaries of tying and bundling doctrines. Instead, this Note considers the hypothetical application of tying and bundling doctrines to an emerging high technology industry—smartphones—as a means of drawing out the challenges, limitations and foreseeable extensions in the future of antitrust law.

A. *Why Would Anyone Want to Limit 'Smart' Technologies?*

Given the utility and versatility of 'smart' technologies like smartphones,

---

21. *See* Daniel A. Crane, *Antitrust Modesty*, 105 MICH. L. REV. 1193, 1196 (2007) ("[T]he spirit of this particular moment in U.S. antitrust law both for the present generation and historians to come . . . . [is that] the antitrust enterprise has become, in a word, modest—existentially, procedurally, and substantively."); *but cf.* Alan Devlin, *Antitrust in an Era of Market Failure*, 33 HARV. J. L. & PUB. POL'Y 557, 561 (2010) ("Although the global financial meltdown demonstrates that unqualified support for the free market was dogmatic, it has revealed no systemic market failure that suggests or supports a shift in substantive antitrust policy . . . . If the banking crisis has taught us anything, it is that financial actors are myopic in their avid pursuit of short-run gains. This practice highlights the presence of incentives that justify the pre-crisis approach to competition law." (footnote omitted)).

22. *Id.* at 1196-97.

23. Michael L. Katz & Carl Shapiro, *Antitrust in Software Markets*, *in* COMPETITION, INNOVATION AND THE MICROSOFT MONOPOLY: ANTITRUST IN THE DIGITAL MARKETPLACE 29, 29 (Jeffrey A. Eisenach & Thomas M. Lenard eds., 1999); *see also* Jonathan B. Baker, *Editor's Note* to Symposium, *Antitrust at the Millennium (Part I)*, 68 ANTITRUST L. J. 1, 1-2 (2000).

24. *See generally* Daniel L. Rubinfeld and John Hoven, *Innovation and Antitrust Enforcement*, *in* COMPETITION, INNOVATION AND THE MICROSOFT MONOPOLY: ANTITRUST IN THE DIGITAL MARKETPLACE 65, 65 (Jeffery A. Eisenach & Thomas M. Lenard eds., 1999).

25. *See* Crane, *supra* note 21 at 1197 ("A seller's own brand can never be a market unto itself, even if the seller prevents competitor from offering customers replacement parts and services for the brand and, thereby, is able to extract monopoly rents from customers.").

one may question why smartphones are an appropriate industry for the hypothetical application of tying and bundling law. The answer is simple—because smartphones are increasingly ubiquitous[26] (thereby magnifying the implications of even relatively small consumer harms) and because the primary market for smartphones and the secondary markets for applications[27] exhibit the representative economic characteristics of many dynamic smart technology industries. These characteristics include high barriers to entry (*i.e.* significant start-up capital costs),[28] heterogeneous product offerings,[29] and rapid innovation cycles.[30] Thus, smartphones are the ideal foil for evaluating the utility of current tying and bundling law principles. Furthermore, smartphone market characteristics naturally result in highly concentrated markets—an antitrust indicator[31] that naturally raises the eyebrows of trustbusters and consumer advocates alike. Yet, it is axiomatic that a market is not anticompetitive simply because it exhibits characteristics correlated with monopolies or oligopolies.[32] Rather, these may be indicators that the competitors are aggressively seeking "the spectacular prize"—namely large pecuniary rewards—through innovation and rapid growth or output.[33] Only if firms engage in conduct that is designed to artificially constrain competition, resulting in higher prices or lower levels of innovation, do antitrust laws become relevant.[34]

---

26. Aaron Smith, *Americans and Their Cell Phones*, PEW INTERNET & AMERICAN LIFE PROJECT (Aug. 15, 2011).

27. Kristen Purcell, *Half of Adult Cell Phone Owners Have Apps on Their Phones*, PEW INTERNET & AMERICAN LIFE PROJECT (Nov. 2, 2011) ("The share of adult cell phone owners who have downloaded an app to their phone nearly doubled in the past two years – rising from 22% in September 2009 to 38% in August 2011 . . . . The share of U.S. adults who purchased a phone already equipped with apps also increased five percentage points in the past year, from 38% in May 2010 to 43% in the current survey.").

28. Broadcom Corp. v. Qualcomm, Inc., 501 F.3d 297, 307 (3d Cir. 2007) ("Barriers to entry are factors, such as regulatory requirements, high capital costs, or technological obstacles, that prevent new competition from entering a market in response to a monopolist's supracompetitive prices. . . ." (citing United States v. Microsoft Corp., 253 F.3d 34, 51 (D.C. Cir. 2001))).

29. *Cf.* Harrison Aire, Inc. v. Aerostar Int'l., Inc., 423 F.3d 374, 381 (3d Cir. 2005) ("[W]hen dealing with a heterogeneous product or service . . . a reasonable finder of fact cannot infer monopoly power just from higher prices." (citing Blue Cross & Blue Shield United v. Marshfield Clinic, 65 F.3d 1406, 1412 (7th Cir. 1995)) (citations omitted)).

30. *In re* IBM Peripheral EDP Devices Antitrust Litig., 481 F. Supp. 965, 990 (N.D. Cal. 1979) *aff'd sub nom.* Transamerica Computer Co., Inc. v. Int'l Bus. Machines Corp., 698 F.2d 1377 (9th Cir. 1983) ("This state of perfect competition will continue until another innovator starts the innovation-monopoly-imitator-competition cycle over again.").

31. U.S. DEP'T OF JUSTICE AND THE U.S. FED. TRADE COMM'N, HORIZONTAL MERGER GUIDELINES, § 5.3 (2010).

32. *See In re IBM Peripheral EDP Devices Antitrust Litig.*, 481 F. Supp. at 990.

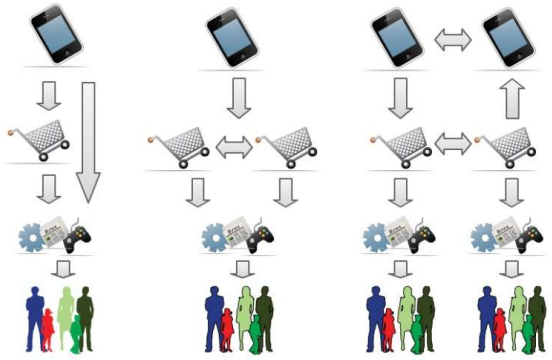
33. Anne Mayhew, *Schumpeterian Capitalism Versus the "Schumpeterian Thesis,"* 14 J. ECON. ISSUES 583, 586-87 n.2 (1980).

34. *See In re IBM Peripheral EDP Devices Antitrust Litig.*, *supra* note 32.

B. *The Smartphone "Problem"*

For the purposes of this smartphone illustration, it is important to note that there are different perspectives on how competition between smartphones relates to competition for application clearinghouses. There are three basic perspectives, each represented graphically in Figure 1. Based on one's understanding of the market, competitive harm can arise on different market levels. First, a smartphone mOS can be viewed as an individual market (See Figure 1, Perspective A). Within that market, the mOS producer creates a single application clearinghouse that provides the primary interface for obtaining third party applications. If the market is not viewed as a feature of the mOS, then tying may foreclose competitors from entering the application clearinghouse market. Furthermore, the only constraint on the separate market for clearinghouse services are mobile Internet browsers, (addressed in Part III.B.ii. on the competitive 'fix') which allow individuals to access third party applications directly. Tying, or pre-installation of application clearinghouses can result in higher prices charged by the application clearinghouse directly to consumers who are a captive market for the clearinghouse.

Figure 1: Different Smartphone mOS and Application Clearinghouse Market Dynamics



A. B. C.

Second, within a single mOS market, multiple application clearinghouses, created by the mOS producer or third-party competitors compete on price to sell the same applications to consumers See Figure 1, Perspective B). This is a more competitive market than the first example, but competitive harm can still result if the mOS producer, ties its mOS to its application clearinghouse and uses its market power to illegally constrain competition from third party application clearinghouses (a § 1 violation) or to buttress the mOS producer's monopoly (a § 2 violation).

Third, different smartphone mOS platforms compete with each other partially by proxy wars in application clearinghouses (See Figure 1, Perspective C). Customers may be 'locked in' to a mOS platform with higher prices for

applications in tied application clearinghouses, even though they would prefer to switch to a different mOS platform. As Justice Scalia's dissent in *Eastman Kodak* observes, cross-elasticity between mOS platforms can constrain anticompetitive prices for third party applications.[35]

Thus, tying between the same two products can result in different forms of competitive harms. How one views the competitive interactions between markets affects whether courts perceive redressable antitrust injury or what type of injury has occurred.

C. *Refining the Tying and Bundling Inquiry*

Industries that are naturally dynamic (meaning they are characterized by rapid cycles of innovation) are not, by default, procompetitive.[36] Firms may still hurt overall consumer welfare by extracting "rent", or charging prices above those that would be charged in a competitive market.[37] These anticompetitive behaviors are more difficult to detect[38], as the anticompetitive pricing action or the depression in innovation may not last for the protracted period found in historical antitrust actions.[39] Yet, that does not mean that economic harms fail to accrue to consumers.[40] In order for tying and bundling theories to remain relevant in high technology industries, then identifying and distinguishing when market failures occur and when competitors are engaging in anticompetitive conduct is critical.

Tying and bundling causes of action are the right legal "tools" for the antitrust "job," particularly in high technology markets[41] In these markets, the vast majority of potential consumer harm emanates from a competitor's leveraging[42] a strong market position in a primary market to constrain

---

35. Eastman Kodak Co. v. Image Tech. Svcs., 504 U.S. 451, 486 (1992) (6-3 decision) (Scalia, J., dissenting).

36. *See* Frank H. Easterbrook, *When Is It Worthwhile to Use Courts to Search for Exclusionary Conduct?*, 2003 COLUM. BUS. L. REV. 345, 345 n.13 (2003).

37. Michael Billiel, *Fine-Tuning Deregulation: The Interstate Commerce Commission's Use of Its General Rail-Exemption Power*, 53 GEO. WASH. L. REV. 827, 850, n.187 (1985) ("Economic rents, or monopoly profits, are amounts earned by a business in excess of costs plus a normal return on investment.").

38. *See* Easterbrook, *supra* note 13, at 345.

39. *See generally* Alan J. Meese, *Tying Meets the New Institutional Economics: Farewell to the Chimera of Forcing*, 146 U. PA. L. REV. 1, 1-3 (1997).

40. *See* Easterbrook, *supra* note 13, at 345.

41. *See generally* Allan M. Soobert, *Antitrust Implications of Bundling Software and Support Services: Unfit to Be Tied?*, 21 U. DAYTON L. REV. 63, 64 (1995).

42. Lawrence A. Sullivan, *Is Competition Policy Possible in High Tech Markets?: An Inquiry into Antitrust, Intellectual Property, and Broadband Regulation As Applied to "The New Economy,"* 52 CASE W. RES. L. REV. 41, 50 (2001) ("Antitrust has two significant tasks: to monitor the contest to become the standard, and to inhibit the winner either from stifling such continuing dynamism as might eventually unseat it or from leveraging its power into other markets.").

competition in a secondary market (that buttresses the competitor's dominance in the primary market). Alternatively, harm can occur when competitors extract higher prices in a secondary/downstream market.[43] Simply, control of a product like a smartphone mOS may give a competitor the ability to constrain competition in a secondary market, such as the intermediary market for the purchasing and downloading applications. However, in order to understand whether or not a competitor has the necessary market power and, if so, whether that competitor is using that market power in an anticompetitive manner, courts must be able to look to an economic theory that supports a theory of competitive harm.

II. THE *UNITED STATES V. MICROSOFT* TRILOGY

Perhaps the most prominent tying case involving technology products remains *United States v. Microsoft*, which was predicated on and applied the two Supreme Court decisions in *Jefferson Parish Hospital District No. 2 v. Hyde*[44] and *Eastman Kodak v. Image Technical Services*.[45] While these cases have been dissected and dissected, much of the academic discussion on this dispute was generated concurrent with the litigation and much less has been said since the Department of Justice settled the dispute, ending the federal component of the litigation.[46] Given the fundamental shifts since these cases were decided in how technology is used, it is surprising that a greater discussion of whether these precedents still remain viable guideposts for modern technologies has not ensued. This Note provides a more complete retrospective analysis of the *Microsoft* litigation, including the implications of the consent decree, in order to provide a stronger baseline for modern applications.

A. Microsoft I[47]

In district court, the Government alleged that Microsoft had illegally combined its proprietary operating system with its proprietary Internet browser.[48] Applying the Supreme Court's per se test,[49] the court determined

---

43. David S. Evans & Michael Salinger, *Why Do Firms Bundle and Tie? Evidence from Competitive Markets and Implications for Tying Law*, 22 YALE J. ON REG. 37 (2005).

44. Jefferson Parish Hosp. Dist. No. 2 v. Hyde, 466 U.S. 2 (1984).

45. Eastman Kodak Co. v. Image Technical Services, Inc., 504 U.S. 451 (1992).

46. Litigation by the States continued, although the resolution of that dispute is largely germane to the discussion of how tying and bundling law applies to modern, dynamic technological markets; therefore the State litigation will not be discussed in this Note.

47. United States v. Microsoft, 87 F. Supp. 2d 30 (D.D.C. 2000).

48. *Id.* at 47.

49. *Id.* ("Liability for tying under §1 exists where (1) two separate 'products' are involved; (2) the defendant affords its customers no choice but to take the tied product in

that Microsoft had indeed formed an illegal tying arrangement.[50] Proceeding through the four elements of the per se test, the court first found that the operating system and Internet browser were separate products, as "consumers today perceive operating systems and browsers as separate 'products,' for which there is separate demand."[51] Second, the sale of Windows was conditioned, automatically, on the purchase of Internet Explorer. Next, the court determined that Microsoft possessed market power in the tying market sufficient to force purchasers to change their product choices.[52] Interestingly, the court did not discuss the causality at length, but rather premised its conclusion on the court's earlier finding that "Microsoft possesses monopoly power in the worldwide market for Intel-compatible PC operating systems (i.e., the tying product market)."[53] Finally, the court identified that a "'not insubstantial' amount of commerce [had been] foreclosed to competitors as a result of . . . [the] decision to bundle Internet Explorer with Windows."[54]

The district court opinion also considered one of the unique elements of software bundling: zero pricing. Zero pricing presents a unique challenge in tying law, because it is difficult to assert that a purchaser is financially or economically harmed when there is no cost associated with the tied product. However, the district court posited that "licensees, including consumers, are forced to take, and pay for, the entire package of software and that any value to be ascribed to Internet Explorer is built into this single price."[55] Therefore, even free products or services could constitute forcing to consumers when zero pricing raised the costs to competitors to prohibitively high levels.[56]

Despite recognizing that the D.C. Circuit was concerned with "the perils associated with a rigid application of the traditional "separate products" test to

---

order to obtain the tying product; (3) the arrangement affects a substantial volume of interstate commerce; and (4) the defendant has 'market power' in the tying product market.") (citing Jefferson Parish, 466 U.S. at 12-8).

50. *Id.* at 49.

51. *Id.* ("[P]roduct and market definitions [are] to be ascertained by reference to evidence of consumers' perception of the nature of the products and the markets for them, rather than to abstract or metaphysical assumptions as to the configuration of the 'product' and the 'market.'").

52. *Id.*

53. *Id.*

54. *Id.* at 49-50 ("Although the Court's Findings do not specify a dollar amount of business that has been foreclosed to any particular present or potential competitor of Microsoft in the relevant market, including Netscape, the Court did find that Microsoft's bundling practices caused Navigator's usage share to drop substantially from 1995 to 1998, and that as a direct result Netscape suffered a severe drop in revenues from lost advertisers, Web traffic and purchases of server products. It is thus obvious that the foreclosure achieved by Microsoft's refusal to offer Internet Explorer separately from Windows exceeds the Supreme Court's de minimis threshold.").

55. *Id.* at 50.

56. *Id.*

computer software design,"[57] the district court held that Microsoft's tying arrangement was illegal because "it was the result of a deliberate and purposeful choice to quell incipient competition before it reached truly minatory proportions."[58] This holding would come under fire on appeal.

B. Microsoft II[59]

On appeal, Microsoft contested both the legal conclusions of the district court and the resulting remedial order.[60] With respect to the district court's findings on tying,[61] the D.C. Circuit determined that the separate products test and the per se analysis were inappropriate given the characteristics of the market, and that the tying of Microsoft's operating system and Internet browser should be evaluated under the rule of reason. Adopting a Schumpeterian economic perspective,[62] the court was careful not to conclude that Microsoft's "integration [was] welfare-enhancing or that [Microsoft] should be absolved of tying liability,"[63] the court nevertheless suggested that "integration of new functionality into platform software is a common practice and . . . wooden application of per se rules . . . may cast a cloud over . . . innovation."[64] The court then remanded the case with instructions that if the government renewed its tying claim, that claim would be subject to the rule of reason.[65] In other words, while it was possible that Microsoft could be liable for illegally tying their operating system and Internet browser together in one package, the standard for demonstrating the existence of an illegal tie would be much higher, and Microsoft would be able to present efficiency arguments in their defense.

In order to arrive at this decision, the D.C. Circuit first considered whether the separate products test truly shed any light on the competitive concerns for newly integrated products, such as with operating systems and Internet

---

57. *Id.* at 51.

58. *Id.*

59. United States v. Microsoft, 253 F.3d 34 (*Microsoft II*) (D.C. Cir. 2001) (per curiam).

60. *Id.* at 46.

61. *Id.* at 84-85 ("The key District Court findings are that: (1) Microsoft required licensees of Windows 95 and 98 also to license IE as a bundle at a single price . . . ; (2) Microsoft refused to allow OEMs to uninstall or remove IE from the Windows desktop . . . ; (3) Microsoft designed Windows 98 in a way that withheld from consumers the ability to remove IE by use of the Add/Remove Programs utility . . . ; and (4) Microsoft designed Windows 98 to override the user's choice of default web brow[s]er in certain circumstances.").

62. *Id.* at 49 ("In technologically dynamic markets, however, such entrenchment may be temporary, because innovation may alter the field altogether.") (citing JOSEPH A. SCHUMPETER, CAPITALISM, SOCIALISM AND DEMOCRACY 81-90 (1976)).

63. *Microsoft II*, 253 F.3d at 89.

64. *Id.* at 95.

65. *Id.*

browsers.[66] Beginning with the basic rationale "that tying prevents goods from competing directly for consumer choice on their merits,"[67] and adding the coordinate truism that "not all ties are bad,"[68] the court identified the goals underlying prohibited tying arrangements. Reasoning from these first principles, the D.C. Circuit observed that "[t]he per se rule's direct consumer demand and indirect industry custom inquiries are . . . backward-looking and therefore [are] systematically poor proxies for overall efficiency in the presence of new and innovative integration."[69]

Next, the court challenged the per se tying analysis, opting instead to apply the more flexible rule of reason approach.[70] When the products were physically and technologically integrated and when the combination improved the value of the tying product to users and producers of complementary goods,[71] the court reasoned that "[a]pplying per se analysis . . . creates undue risks of error and of deterring welfare-enhancing innovation."[72] In contrast, the rule of reason approach would permit the consideration of efficiencies arguments never before presented to courts.[73]

Finally, the D.C. Circuit laid out three explicit conditions for any further tying claims on remand. First, "plaintiffs must show that Microsoft's conduct unreasonably restrained competition."[74] Second, that "[i]n order for the District Court to conclude [that Microsoft's] practices[75] also constituted §1 violations, plaintiffs must demonstrate that their benefits . . . are outweighed by the harms in the *tied product* market."[76] Third, "the District Court must . . . consider an alleged tying violation . . . price bundling."[77]

Although nearly a decade old—the current precedential value of the D.C. Circuit's decision should not be belittled. Before remanding the case, the D.C.

---

66. *Id.* at 85-89.

67. *Id.* at 87.

68. *Id.* The court also presented the classic economic justification, that "[b]undling obviously saves distribution and consumer transaction costs." *Id.*

69. *Id.* at 89. While questioning the value of the analysis, the court never directly engaged in an analysis of whether operating systems and internet browsers should be considered a single product for the purposes of assessing a per se illegal tie. In fact, the D.C. Circuit's reference to the District Court findings suggested that operating systems and internet browsers were separate products, especially in light of Microsoft's concession that "many consumers demand alternative browsers." *Id.*

70. *Id.* at 89-90.

71. *Id.* at 88-89 ("[T]here are strong reasons to doubt that the integration of additional software functionality into an OS falls among [impermissible per se] arrangements.").

72. *Id.* at 89-90.

73. *Id.* at 93.

74. *Id.* at 95.

75. *Id.* at 96 ("Microsoft's refusal to allow OEMs to uninstall IE . . . and [the] removal of the IE entry from the Add/Remove Programs utility in Windows 98.").

76. *Id.* (emphasis in original).

77. *Id.* at 96.

Circuit explicitly narrowed their holding to cases where "the tying product is software whose major purpose is to serve as a platform for third-party applications and the tied product is complementary software functionality."[78] Additionally, the court's remark that procompetitive justifications "appl[y] with distinct force when the tying product is *platform* software" may be easily extended to new technological platforms that share many of the same competitive dynamics, such as smartphones.

C. Microsoft III[79]

On remand, the United States District Court for the District of Columbia ordered settlement negotiations.[80] The Department of Justice entered a consent decree and the court reviewed for compliance with the Tunney Act.[81] Part (e) of the Tunney Act requires that the court "make an independent determination as to whether or not entry of a proposed consent decree is in the public interest."[82] When evaluating the public interest, the court considered the competitive impact of the judgment as well as its effect on the public.[83]

In regards to tying, the consent decree focused on behaviors and remedies that would prevent or limit the effect of illegal tying arrangements. The pertinent parts included anti-retaliation provisions,[84] remedies for commingling Internet browser and non-Internet browser code,[85] and disclosure of technical information that would ensure the interoperability of 'middleware.'[86] Essentially, these provisions were designed to promote a competitive market by removing the barriers faced by downstream producers. Although these provisions were expansive, "[s]uch expansion is appropriately forward-looking and accords with the general law of remedies in antitrust law."[87]

Additionally, Part III.A of the consent decree also applied to the tying issue—although its relevance was not directly stated. This section "bar[red] retaliation where Microsoft knows the OEM 'is or is contemplating' . . . 'developing, distributing, promoting, using, selling or licensing any software

78. *Id*. at 95.

79. United States v. Microsoft, 215 F. Supp. 2d 1 (D.D.C. 2002) *and* United States v. Microsoft (*Microsoft III*), 231 F. Supp. 2d 144 (D.D.C. 2002).

80. *Microsoft III*, 215 F. Supp. 2d at 3.

81. The Antitrust Procedures and Penalties Act, 15 U.S.C. § 16(b)-(h) (2006) (requiring the government to publish a competitive impact statement in the Federal Register sixty days prior to the effective date of the final judgment and solicit public comments).

82. *Microsoft III*, 231 F. Supp. 2d at 152 (quoting 15 U.S.C. § 16(e) (2006)); *see also* S. Rep. 93-298 at *5.

83. *Id.* (citing United States v. Bechtel, 648 F.2d 660, 666 (9th Cir. 1981)).

84. *Id.* at 168-69.

85. *Id.* at 179-81.

86. *Id.* at 186-92.

87. *Id.* at 188.

the competes with Microsoft Platform Software or any product or service that distributes or promotes any Non-Microsoft Middleware.'"[88] Although the precise, technical definitions obscured the products' common names, the term 'platform software' product referred to Microsoft's operating system and the term 'middleware product' referred to Internet browsers, among a wide variety of other products.[89] It was important that middleware products were defined broadly, so as to "enable[] inclusion . . . of future technologies."[90]

Thus, the litigation between the Department of Justice and Microsoft yielded no clear victors and no concrete precedent for judging future software tying arrangements. Nevertheless, the litigation highlighted tests, standards, and issues that may be applied to analysis of the smartphone industry.

III. COMPETITIVE ISSUES ARISING FROM SMARTPHONE MOBILE OPERATING SYSTEMS AND CLEARINGHOUSES

A. *Section 1 Actions and Applications*

Tying and bundling actions can be brought as either per se violations of § 1[91] or as violations based on the rule of reason.[92] Assessing the combination of smartphone mobile operating systems and application clearinghouses illustrates the robust value of current tying and bundling law as applied to high technology industries as well as the potential pitfalls courts must avoid when reviewing inherently dynamic technologies. First, this Note will address the traditional per se tying approach. Second, this Note will evaluate a hypothetical tying and bundling claim under a hybrid O'Connor/*Microsoft II* rule of reason analysis.

1. *Per se violations*

Traditional per se violations are the oldest forms of tying and bundling actions. However, their strict requirements force litigants to work under awkward legal fictions in order to mold modern high-technology industries into pre-information age precedents. The awkward fit between the per se elements and modern information technologies is well illustrated by the mOS/application clearinghouse arrangement.

In *Eastman Kodak*, the Court defined tying as "an agreement by a party to sell one product but only on the condition that the buyer also purchase a different (or tied) product, or at least agrees that he will not purchase that

88. *Id.* at 164.
89. *Id.* at 165-68.
90. *Id.* at 167.
91. The Sherman Antitrust Act, 15 U.S.C. § 1 (2006).
92. *Town Sound*, 959 F.2d at 482.

product from any other supplier."[93] The Court explicitly stated that "[s]uch an arrangement violates § 1 of the Sherman Act if the seller has 'appreciable economic power' in the tying product market and if the arrangement affects a substantial volume of commerce in the tied market."[94] Lower courts have combined these statements into a three-part test[95] or a four-part test[96] that establishes the criteria that the plaintiff must prove.[97]

Unfortunately, the three-part formulation can result in a significant amount of confusion, as the first element actually subsumes two different inquiries: (1)(a) whether two distinct, separate products exist and (1)(b) whether the sales of those two products are tied together. Stated more clearly, the criteria is whether the sale of the tying product is conditioned on the purchase of the tied product.[98] For this reason, the four-part test will be used in the following mOS/application clearinghouse illustration. Thus, a plaintiff alleging an illegal tying relationship between a smartphone mOS and an application clearinghouse would have to prove that:

- two separate products exist,[99] *i.e.* that smartphone mOS is a separate product from the application clearinghouse,
- the sale of the tying product is conditioned on the purchase of the tied product, [100] *i.e.* that the 'sale' (or pre-installation) of the application clearinghouse is conditioned on the 'purchase' of the mOS,
- the defendant has market power in the tying product,[101] *i.e.* that the

---

93. *Eastman Kodak*, 504 U.S. at 461-62 (1992) (citing Northern Pacific Ry. Co. v. United States*,* 356 U.S. 1, 5-6 (1958)).

94. *Id.* at 461-62.

95. These are: "(1) that there exist two distinct products or services in different markets whose sales are tied together; (2) that the seller possesses appreciable economic power in the tying product market sufficient to coerce acceptance of the tied product; and (3) that the tying arrangement affects a 'not insubstantial volume of commerce' in the tied product market." *See, e.g.*, Paladin Assoc., Inc. v. Montana Power Co*.,* 328 F.3d 1145, 1159 (9th Cir. 2003).

96. BookLocker.com, Inc. v. Amazon.com, Inc., 650 F. Supp. 2d 89, 98 (D. Me. 2009) ("There are four elements of a per se tying claim: (1) the tying and tied products are actually two distinct products; (2) there is an agreement or condition, express or implied, that establishes a tie; (3) the entity accused of tying has sufficient economic power in the market for the tying product to distort consumers' choices with respect to the tied product; and (4) the tie forecloses a substantial amount of commerce in the market for the tied product." *citing* Data Gen. Corp. v. Grumman Sys. Support Corp., 36 F.3d 1147, 1179 (1st Cir. 1994)).

97. *Id.*

98. *See Eastman Kodak*, 504 U.S. at 451; BookLocker.com, 650 F. Supp. 2d at 89; *see also* J. Gregory Sidak, *An Antitrust Rule For Software Integration*, 18 YALE J. ON REG. 1, 20 (2001).

99. *See Eastman Kodak*, 504 U.S. at 451; Booklocker.com, 650 F. Supp. 2d at 89; *see also* Sidak, *supra* note 99.

100. *See Eastman Kodak*, 504 U.S. at 451; Booklocker.com, 650 F. Supp. 2d at 89; *see also* Sidak, *supra* note 99.

101. *See Eastman Kodak*, 504 U.S. at 451; Booklocker.com, 650 F. Supp. 2d at 89; *see also* Sidak, *supra* note 99.

mOS producer has 'market power' that can be exercised in the market for application clearinghouses, and

- the tie-in forecloses a substantial amount of potential sales of the tied product"[102], *i.e.* that the combination (or bundle, or pre-installation) of the mOS and the application clearinghouse forecloses potential 'sales' of third-party application clearinghouses. In essence, the mOS/application clearinghouse tie-in harms consumers by reducing consumer choice.[103]

The first element requires the plaintiff to show that two separate products exist. While this legal rule may appear to merely incorporate a truism,[104] it is also the element that has received the vast majority of academic attention and scrutiny.[105] This is because this 'separate products test' puts the cart before the horse, requiring a definitive finding that two separate products exist when a tying or bundling action may be raised specifically for the purpose of showing that the appearance of a single product is in fact an illusion.[106] Such an illusion can be created when one competitor has virtually eliminated competitors for the secondary product, thereby creating the illusion of a single product market.

---

102. *See Eastman Kodak*, 504 U.S. at 451; Booklocker.com, 650 F. Supp. 2d at 89; *see also* Sidak, *supra* note 99.

103. *See Eastman Kodak*, 504 U.S. at 451; Booklocker.com, 650 F. Supp. 2d at 89; *see also* Sidak, *supra* note 99; *see also* Sam Grobart & Jenna Wortham, *Which Apps Are Threatened by Apple's Upgrades* (June 6, 2011 3:23 PM), http://bits.blogs.nytimes.com/2011/06/06/which-apps-are-threatened-by-apples-upgrades (last visited Mar. 14, 2012) ("How do you know if you've created a really great, useful iPhone app? Apple tries to put you out of business. That may be overstating it, but a number of new features for Apple's operating systems that it announced at its Worldwide Developers Conference have been available through existing apps and services for some time. Some of those apps are quite popular, and have been lucrative for the people who developed them."); c*f.* Saul Hansell, *Apple Denies It Rejected Google Application for iPhone*, NY Times (Aug 21, 2009), http://www.nytimes.com/2009/08/22/technology/companies/22apple.html?scp=1&sq=iphone%20app%20denied&st=cse.

104. If two products are not involved, how can you tie or bundle a product to itself?

105. *See, e.g.*, Aryeh S. Friedman, *Law and the Innovative Process: Preliminary Reflections*, 1986 COLUM. BUS. L. REV. 1, 21 (1986); Robert F. Goff, *Emphasizing Conduct over Context and Market Definition over Market Power: Short-Term Strategic Anticompetitive Behavior Absolved in Blue Cross v. Marshfield Clinic*, 3 CONN. INS. L.J. 381, 402 (1997); Renato Mariotti, *Rethinking Software Tying*, 17 YALE J. ON REG. 367, 389 (2000); Erik B. Wulff & Scott A. McIntosh, *The Separate Product Test in Franchise Tying Cases: Through the Microsoft Lens of Reason*, 21 FRANCHISE L.J. 70, 71 (2001).

106. *See, e.g.*, *BookLocker*, 650 F. Supp. 2d at 99 ("[While] the First Circuit expressed skepticism about whether a tie existed where a tying product did not take the conventional form of a product or service bought by a consumer . . . . has also accepted that commercial transactions lacking an outright "sale" of a tying product might constitute unlawful ties."); see also Marts v. Xerox, Inc., 77 F.3d 1109, 1112 (8th Cir. 1996) (leaving open the possibility that a tie could involve a product neither paid for nor leased when finding that a warranty provided at no additional charge with the purchase of a copier might be a tying product, although such a claim "does not fit easily into the existing structure of antitrust law").

For most physical products, the specter of illusory single product markets is not at issue. For example, when purchasing a car, there is no question that you will also purchase an engine as part of that transaction.[107] For these types of products, it is sufficient that "[t]he relevant market must be a product market, whose boundaries are defined not by the consumers but by the products or producers themselves."[108] More practically stated: "in determining whether one or two products is involved, the focus should be 'not on the functional relation between them, but rather on the character of the demand for the two items.'"[109] On its face, this "demand test"[110] appears to create a reasonable, mechanical, and viable mechanism for discerning whether a particular market is for a single product or multiple products.

However, high technology markets, particularly software markets, are plagued by the question of whether a plaintiff's (or a defendant's) market definition is illusory—*i.e.* whether it is facially sustainable.[111] Likewise, traditional tests based on market definition that are used to establish separate products, such as the demand test, do not address the quintessential functionality issue, nor are demand tests necessarily sensitive enough to identify many anticompetitive arrangements in software or other high technology industries.[112] Thus, while the other elements may be more dispositive on the merits of the tying or bundling question, any competitive effects analysis can be foreclosed by a viable assertion that the technology product(s) at issue constitute only a single product. Yet, market power and competitive effects analysis may prove more informative than market

---

107. *But cf.* ToyotaUSA, *Built – Reinvented 2012 Toyota Camry,* YOUTUBE, http://www.youtube.com/watch?v=X3bhbYXiwYk (last visited Mar. 10, 2012) (raising an interesting question with respect to the integration of mobile applications into cars—should they be considered merely features of the car because they are 'part' of the car, *i.e.*: installed on the car's computer, or does the fact that the same applications can be found on a car computer and on a smartphone actually *prove* that they are distinct, separate products?) .

108. *In re* Webkinz Antitrust Litig., 695 F. Supp. 2d 987, 994 (N.D. Cal. 2010).

109. PSI Repair Svcs., Inc. v. Honeywell, Inc., 104 F.3d 811, 815 (6th Cir. 1997).

110. *See* Rick-Mik Enter. v. Equilon Enter., 532 F.3d 963, 975 (2008) ("[T]he 'purchaser demand' test of *Jefferson Parish* 'examine[s] direct and indirect evidence of consumer demand for the tied product separate from the tying product.'").

111. *See* Brown Shoe Co. v. United States, 370 U.S. 294, 325 (1962) ("The outer boundaries of a product market are determined by the reasonable interchangeability of use or the cross-elasticity of demand between the product itself and substitutes for it."); Queen City Pizza v. Domino's Pizza, 124 F.3d 430, 436 (3d Cir. 1997) ("Where the plaintiff fails to define its proposed relevant market with reference to the rule of reasonable interchangeability and cross-elasticity of demand, or alleges a proposed relevant market that clearly does not encompass all interchangeable substitute products even when all factual inferences are granted in plaintiff's favor, the relevant market is legally insufficient and a motion to dismiss may be granted.").

112. However, where tying or bundling actions in software product markets create the illusion of a single product market, then it is difficult, if not impossible to demonstrate a separate demand function for a product that consumers may not even be aware exists separately.

definition. For this reason, a plaintiff may seek to reverse-engineer their claim based on the other elements of the tying inquiry. This Note does not seek to resolve this existential legal debate. However, it is important to be cognizant of the unique characteristics of tying and bundling law in dynamic, non-physical, high technology product markets.

For the purposes of this illustration, it is sufficient to say that it is legally permissible for a plaintiff to resolve the relevant product market inquiry of the first element in two ways, either by (1) alleging the tied product exists in an entirely different product market,[113] or (2) alleging that the tied product exists in a "legally cognizable" submarket.[114] However, regardless of the approach, the underlying single product or two-product question must be addressed.

The mOS/application clearinghouse arrangement provides an informative illustration of this issue, as the product market can accurately be described as a single product market or multiple product markets. Conceptualized as a single product market, an application clearinghouse is merely a 'feature' of a smartphone mOS. Like the QUERTY keyboard[115] or the integration of spellcheck into word processing software,[116] what may initially appear to be multiple product markets may practically morph into a single product market. This is frequently the case in dynamic technology markets, as new technologies increase the utility of a single product by combining functions previously available only through multiple products into a single product.[117] A smartphone mOS is the epitome of this dynamic—the utility of a smartphone mOS is predicated on the mOS's ability to function as a platform for features developed by third-parties. Thus, a smartphone mOS combined with any integrated application—especially an application clearinghouse, which allows consumers to access features developed by third parties—is the quintessential single product.

However, in markets where there are no significant competitors for application clearinghouses, such as in the Apple iOS market, there is no separate demand function because the proprietary application clearinghouse is the presumed means of accessing downstream third-party applications. In fact, consumers may be largely unaware that they may be suffering higher prices for third-party applications by being restricted to a single application clearinghouse. Even if consumers were aware of an alternative product and

---

113. *In re Webkinz Antitrust Litig*., 695 F. Supp. 2d at 994.

114. *Id.*

115. *See* S.J. Liebowitz & Stephen E. Margolis, *Should Technology Choice Be a Concern of Antitrust Policy?*, 9 HARV. J.L. & TECH. 283, 312-14 (1996).

116. Richard Schmalensee, *Antitrust Issues in Schumpeterian Industries*, 90(2) AM. ECON. REV. 192 (2000) *citing* JOSEPH A. SCHUMPETER, CAPITALISM, SOCIALISM AND DEMOCRACY 84 (3d ed. 1950).

117. *See* David A. Heiner, *Assessing Tying Claims in the Context of Software Integration: A Suggested Framework for Applying the Rule of Reason Analysis*, 72 U. CHI. L. REV. 123, 126-129 (2005).

there was real demand, there is still no separate demand function because consumers do not know to 'demand' an alternative, competing product in this market.[118] Essentially, the single product test runs the very real risk of conflating the tying product market, *i.e.* the smartphone mOS market, with the tied product market, *i.e.* the application clearinghouse market. Thus, without analysis of how market power can be leveraged between two, at least plausible, product markets, per se restrictions put the cart before the horse by restricting per se inquiries to the prerequisite multiple product showing.

Alternatively viewed as two products, an entirely different array of issues appear. The most obvious challenge is the 'slippery slope' concern, as courts are reluctant to substitute their judgment regarding product differentiation for that of the market's judgment (normally the *status quo*).[119] Less obvious, however, is the damning and legitimate affirmative defense that even if a smartphone mOS and an application clearinghouses are separate products, their 'hand-in-hand' relationship has the effect of creating new markets that never existed before.[120] Such an effect is overwhelmingly procompetitive. The latter argument is especially appealing in the smartphone industry, as the downstream market for third-party applications has grown exponentially[121] and is widely considered one of the most dynamic nascent markets of the new millennia.[122] Moreover, the existence of a dominant application clearinghouse can be considered a 'public good,'[123] which reduces information costs,[124] allowing

---

118. For a classic description of the 'alternative sources' problem, which is fundamentally based on geographic considerations, *see* Matt Koehler, Comment, *The Importance of Correctly Identifying the Consumer for an Antitrust Relevant Market Analysis*, 67 UMKC L. REV. 521, 532 (1999) (This forces antitrust parties to continually pose the question, 'where else can the customer go' . . . Evidence of actual consumer patterns is not as important as evidence of possible consumer patterns. . . . A rare product or service might conceivably include national or even international practicable alternatives.").

119. *Cf.* Frank H. Easterbrook, *Allocating Antitrust Decisionmaking Tasks*, 76 GEO. L.J. 305, 306 (1987-1988) (suggesting that judges lack the comparative advantage over markets when determining anticompetitive conduct) ("So we do not trust judges to make business decisions in corporate law. Yet antitrust law now calls for the same sorts of economic judgments. We ought to be skeptical of judges' and juries' ability to give good answers.").

120. *See* Liebowitz, *supra* note 116, at 286-90.

121. *Compare* Robin Wauters, *Smartphone App Market Reached More than $2.2 Billion in First Half of 2010*, TECHCRUNCH.COM (Aug. 20, 2010) http://techcrunch.com/2010/08/20/smartphone-applications-market-size/, *with Smartphone App Market to Hit $15 B in a Year*, TECHJOURNAL (Oct. 13, 2011), http://www.techjournalsouth.com/2011/10/smartphone-app-market-to-hit-15b-in-a-year/.

122. *Id.*

123. For methods to measure the public good, *see* David S. Brookshire & Don. L. Coursey, *Measuring the Value of a Public Good: An Empirical Comparison of Elicitation Procedures*, 77 AM. ECON. REV. 554, 555, 565 (1987) ("Although individuals may initially exaggerate their preferences for the public good, they modify their stated values as a function of the incentives, feedback, interactions, and other experiences.").

124. *See* Yoram Barzel, *Some Fallacies in the Interpretation of Information Costs*, 20(2) J. L. & ECON. 291, 306-307 (1977) (*"*[I]n each situation people will first assess the

consumers to efficiently access the third-party application market.[125] Thus, the single product test of the per se approach provides little insight into the anticompetitive economic dynamic addressed by tying and bundling law.

Second, the traditional tying inquiry requires that the sale of the tying product be conditioned on the purchase of the tied product.[126] To reiterate, in this illustration the second tying[127] element requires that the 'sale' of the smartphone mOS be conditioned on the 'purchase' of the application clearinghouse. This element is particularly unwieldy in the software context, as a smartphone mOS producer will generally pre-install its own application clearinghouse as part of the entire software package, which also includes other software programs or features that are loaded onto a smartphone. What to do (and how to think) about pre-installation is the critical piece of guidance that the district court provided in *Microsoft I*[128] until its reversal by the D.C. Circuit in *Microsoft II*.[129] While the issue remains unresolved, zero-pricing and bundling does not necessarily negate the second element of the per se tying inquiry.

The third criterion is that the defendant possesses market power in the tying product.[130] Market power, as defined by the Supreme Court, is the power "to force a purchaser to do something that he would not do in a competitive market."[131] As a practical matter, market power is "ordinarily . . . inferred from the seller's possession of a predominant share of the market."[132] While market share is the focal heuristic in determining market power, there is no 'magic number'[133] at which a party is endowed with the ability to unduly exercise its

---

resource cost of obtaining a particular piece of information which subsequently will have little effect on resource allocation, and will then compare its assessment with the cost of various alternatives that would preempt the use of information. The lowest cost method will prevail.").

125. *Cf*. *infra* Part III.B.i (for further discussion of market forces may result in a dominant product, the continued existence of which may ultimately be beneficial).

126. Sidak, *supra* note 99, at 20.

127. 'Bundling' is the more appropriate word in this illustration, however, for consistency, 'tying' will continue to be used.

128. *Cf*. Microsoft I, *supra* note 48.

129. *Cf*. Microsoft II, *supra* note 60.

130. For practical purposes, this Note will not distinguish between 'properly alleging' and 'proving' market power in the tied product market, although this distinction is significant for summary judgment purposes. *See, e.g.*, Rick-Mik Enterprises v. Equilon Enterprises, 532 F.3d 963, 971-72 (9th Cir. 2008) ("[I]n all cases involving a tying arrangement, the plaintiff must prove that the defendant has market power in the tying product. And to prove it, it must first be properly alleged.") (citation omitted); *see also* Sidak, *supra* note 99, at 20.

131. Eastman Kodak Co. v. Image Tech. Servs., 504 U.S. 451, 464 (1992) (citing Jefferson Parish, 466 U.S. at 14).

132. *Id.* at 464 (citing Jefferson Parish, 466 U.S. at 17).

133. It is a foolhardy expedition to attempt to place a presumptive number at which a market share indicates market power. Every market is unique. *See* BookLocker.com v.

will in the tying product market.[134]

Applied to this illustration, the third criterion would require the mOS producer to have sufficient 'market power' in the market for application clearinghouses in order to raise price and restrict output. The question is: in software markets, how should a court determine whether a competitor possesses sufficient market power in the tied product market? There are two possible approaches: (1) simply draw inferences based on market share statistics alone; or (2) make a finding as to whether a competitor has distorted, or has the potential to distort, consumers' choices with respect to the tied product.[135]

Under the first approach, market share statistics are dispositive of market power.[136] In order to simplify the smartphone illustration, the relevant market share statistics for current smartphone operating system producers and their respective Herfindahl-Hirschman Index (HHI) calculations are provided in Figure 2. With the caveat that HHI figures cannot be fully calculated since in tying assertions there is no merger or acquisition of firms, HHI figures may still provide a useful heuristic for identifying moderately or highly concentrated markets.[137] As a practical matter, the market share heuristic is only valuable—and viable—in two situations: (a) when abundant corroborative but no conclusive evidence of market power exists,[138] and (b) when a competitor has such an overwhelming market share (in a fundamentally related market) that no other conclusion is possible. The latter usage was the approach appropriately used by Judge Jackson in *Microsoft I*.[139] Likewise in smartphone mOS markets, the latter approach would be appropriate in mOS markets where the same competitor produces both a mobile operating system and application clearinghouse.

Amazon.com, 650 F. Supp. 2d 89, 103-04 (D. Me. 2009) (for a longer discussion of sufficient market shares indicative of market power in tied products).

134. *Id.*

135. Data Gen. Corp. v. Grumman Sys. Support Corp., 36 F.3d 1147, 1178-79 (1st Cir. 1994) *rev'd on different grounds*, Reed Elsevier, Inc. v. Muchnick, 130 S. Ct. 1237 (U.S. 2010) (jurisdictional requirements of copyright suits).

136. *See BookLocker.com*, 650 F. Supp. 2d at 103-04.

137. U.S. DEP'T OF JUSTICE AND THE U.S. FED. TRADE COMM'N, HORIZONTAL MERGER GUIDELINES § 5.3 (2010).

138. The principle behind the permission for plaintiffs to establish market power by market share *is only to prevent exclusion* of tying or bundling cases when abundant corroborative evidence but no conclusive evidence of market power exists. *See, e.g.*, *Town Sound*, 959 F.2d at 482 ("[T]ying cases have primarily been concerned with the leveraging of economic power from the market for one product into the market for another, with the resulting forcing of consumers to take unwanted products and foreclosure of competitors from the tied product market.").

139. United States v. Microsoft Corp., 87 F. Supp. 2d 30, 36 (D.D.C. 2000).

Figure 2: Market Share Statistics & HHI Calculations for Smartphone mOS

OS Name	Pew Internet Market Share[140]	Herfindahl-Hirschman Index (HHI) (Pew)	Nielsen Total Market Share[141]	Nielsen Recent Adaptors Market Share[142]	Herfindahl-Hirschman Index (HHI) (Nielsen Total)
Android	35	1225	51	54.6	2601
Apple iOS	24	576	34	36.3	1151
RIM Blackberry	24	576	9	4	81
Win. Mobile (Win. Phone 7)	4	16	3 (1.3)	N/A	9
Palm / HP WebOS	6	36	0.6	N/A	0.36
Symbian OS	†	†	0.9	N/A	0.81
(Other)	(7)	†	N/A	5	N/A
(Total)	100	2429	99.8	99.9	3843.17

However, the formulaic character of the first approach does not addresses the underlying 'forcing' issue highlighted in the Supreme Court's description of the tying requirements.[143] In most cases, a finding (as to whether a competitor has distorted, or has the potential to distort, consumers' choices with respect to the tied product) more appropriately addresses the third criterion of per se tying.[144] This is particularly true in high-technology markets—particularly

140. Aaron Smith, Smartphone Adoption and Usage, PEW INTERNET & AMERICAN LIFE PROJECT (July 11, 2011), *available at* http://pewinternet.org/Reports/2011/Smartphones.aspx.

141. *Two Thirds of New Mobile Buyers Now Opting for Smartphones*, NIELSENWIRE (July 12, 2012), http://blog.nielsen.com/nielsenwire/?p=32494.

142. *Id.*

143. Justice Blackmun's lengthy quote of *Jefferson Parish* is not just an homage to tying precedent. Rather, the reiteration of the Court's statement in *Jefferson Parish* that:

> [T]he *essential characteristic* of an invalid tying arrangement lies in the seller's exploitation of its control over the tying product to *force* the buyer into the purchase of a tied product that the buyer either did not want at all, or might have preferred to purchase elsewhere on different terms. When such 'forcing' is present, competition on the merits in the market for the tied item is restrained and the Sherman Act is violated.

serves as a reminder that the it is conduct, not merely size, that is in question. Eastman Kodak Co. v. Image Tech. Servs., 504 U.S. 451, 464, n.9 (1992) (citing *Jefferson Parish* 466 U.S. at 12) (emphasis added). This is also an important distinction from Schumpeterian economics, which also looks at firm size and presumes that "firms compete not only on the margins of price and output, but by offering new products, new technologies and new forms of organization." JERRY ELLIG & DANIEL LIN, DYNAMIC COMPETITION AND PUBLIC POLICY: TEHNOLOGY INNOVATION, AND ANTITRUST ISSUES 6 (2001).

144. Data Gen. Corp. v. Grumman Sys. Support Corp., 36 F.3d 1147, 1178-79 (1st Cir.

software markets—where market share statistics are constantly in flux but a competitor's reputation, or another non-tangible characteristic, may still permit distortion in tied product markets. In the smartphone hypothetical, a smartphone's mOS's dominant market share could be used to create artificial distortions in the application clearinghouse market. These distortions can be created by even simple factors, like pre-installation of an application clearinghouse (with no other obvious information or means of obtaining a viable, competing application clearinghouse). Pre-installation can 'nudge'[145] consumers into exclusive use of the mOS producer's application clearinghouse—thereby distorting the tied product market with no obvious relationship to their market share.

The fourth element of per se tying is the 'threshold', requiring that the tie-in forecloses a substantial amount of potential sales of the tied product"[146] As a practical matter, this is almost never an issue.[147] Applied in the smartphone context, this would mean that an alleged tie between a competitor's smartphone mOS and that same competitor's application clearinghouse (by means of bundling or pre-installation) plausibly foreclosed potential 'sales' of third-party application clearinghouses. While not a difficult threshold to meet, the fourth element is ill-suited to many software markets, which are not monetized by sales of the actual product, but rather by 'follow-on' or 'follow-through' purchases.[148] In fact, many such software products are 'zero-priced'[149]

This makes it all too easy to raise a strict textual defense that a purported tie or bundle fails to meet the threshold of foreclosed commerce because there is no 'real' money at stake. The real issue is that this defense side-steps the heart of the tying or bundling claim—the assertion that a threshold-level of consumer harm is obtained by a reduction in consumer choices or prices paid by downstream consumers.[150] In the smartphone context, it is much easier to see the potential for consumer harm (in higher prices paid for third-party applications or the reduction in the choice of available applications) as a result of tying between a mOS and an application clearinghouse than it is to legally prove such an occurrence.

Of course, there are many plausible defenses that deserve a court's attention and consideration, including: business decisions regarding brand

---

1994).

145. *See generally* Richard A. Thaler & Cass R. Sunstein, NUDGE: IMPROVING DECISIONS ABOUT HEALTH, WEALTH, AND HAPPINESS (2008).

146. Sidak, *supra* note 99, at 20.

147. *See, e.g.*, Audell Petroleum Corp. v. Suburban Paraco Corp., 903 F. Supp. 364, 371 (E.D.N.Y. 1995).

148. *See infra* Part III.B.i.

149. *See infra* Part III.B.ii.

150. A competitor does have standing to challenge based on downstream effects. *See, e.g.*, Free FreeHand Corp. v. Adobe Sys. Inc., 852 F. Supp. 2d 1171, 1188-90 (N.D. Cal. 2012).

defense (as was contended by Kodak in *Eastman Kodak*)[151] and other efficiency[152] arguments. However, the problem is it is all too easy for courts to confound the assertion of these defenses with a lack of plausibility in plaintiff claims.[153] The per se title and deceptively specific criteria of tying and bundling claims make per se claims ill-suited to high technology cases. Rather, these criteria mask areas of substantive dispute between parties that go to the core of tying and bundling claims—and can act either as artificial bars to legitimate complaints or unnatural burdens to defenses in concentrated (but competitive) markets.

2. *The Dysfunctional Functional Rule of Reason*

A completely distinct alternative to a per se approach is the "rule of reason."[154] Simply put, a rule of reason analysis in tying claims "focuses directly on the challenged restraint's impact on competitive conditions."[155] In 1992, Circuit Judge Becker noted that despite the fact that "[per se] tying claims and the rule of reason have coexisted in antitrust law for eighty years, the case law on tying claims under the rule of reason is amazingly sparse."[156] Twenty years later, the Third Circuit's statement holds true—especially in the high technology context.

Partially, the dearth of rule of reason cases is due to divergent guidance from what little precedent exists. As described by Justice O'Connor in *Jefferson Parish*,[157] a tying claim that met the following threshold criteria would be evaluated under a balancing test between the economic harms and benefits of the alleged trying arrangement[158]: (1) "the seller must have power in the tying product market,"[159] (2) "there must be a substantial threat that the tying seller will acquire market power in the tied-product market,"[160] and (3) "there must be a coherent economic basis for treating the tying and tied products as distinct."[161]

---

151. *See* Eastman Kodak Co. v. Image Tech. Servs*.*, 504 U.S. 451, 464 (1992).

152. *Cf.* U.S. DEP'T OF JUSTICE AND U.S. FED. TRADE COMM'N, HORIZONTAL MERGER GUIDELINES, § 10 (2010).

153. *In re* Webkinz Antitrust Litig., 695 F. Supp. 2d 987, 994 (N.D. Cal. 2010).

154. *Town Sound*, 959 F.2d at 482.

155. Nat'l Soc'y of Prof'l Eng'rs v. United States, 435 U.S. 679, 688 (1978).

156. *Town Sound*, 959 F.2d at 482.

157. *Jefferson Parish*, 466 U.S. at 34-37 (O'Connor, J., concurring) ("The 'per se' doctrine . . . . incurs the costs of a rule-of-reason approach without achieving its benefits: the doctrine calls for the extensive and time-consuming economic analysis characteristic of the rule of reason, but then may be interpreted to prohibit arrangements that economic analysis would show to be beneficial.").

158. *Id.* at 4.

159. *Id.* at 37-38.

160. *Id.* at 38-39.

161. *Id.* (O'Connor, J., concurring).

However, when applied by the Circuits without any binding precedent, the waters become even more muddied. In an early—and direct—application of Justice O'Connor's concurring opinion,[162] the Third Circuit expressly stated that there was no requirement that the seller have market power in the *tying* product market.[163] In place of the three threshold criteria, the Third Circuit required only one prerequisite—that the plaintiff present a viable theory of causation of antitrust injury.[164] Adding to the confusion, the Fifth Circuit held that a tying arrangement violates § 1 under the rule of reason approach only if it has an "actual adverse effect on competition."[165] The Ninth Circuit's discussion of tying claims under § 1 provide no more clarity, as the court required only the most basic elements necessary for *any* § 1 violation before withdrawing the opinion.[166] Lower court attempts have largely only added to the cacophony, as illustrated by the Southern District of New York, which articulated a burden-shifting approach to the rule of reason in tying cases.[167] Addressing this lack of uniformity, the Eastern District of New York highlighted that "certain core elements appear in the various 'rule of reason' discussions, including the need for allegations identifying the relevant market, as well as the adverse effect on that market caused by the defendant's anticompetitive conduct."[168]

Nevertheless, the D.C. Circuit's opinion in *Microsoft II* provides the most clairvoyant discussion of the rule of reason in dynamic high technology

---

162. *Town Sound*, 959 F.2d at 482.

163. *Id.* at 484 ("[W]e decline to read *Jefferson Parish* as . . . deciding a rule of reason claim may ignore proffered evidence of actual conduct and economic performance in the *tied* product market simply because of a finding on *tying* product market structure.").

164. *Id.* at 486.

165. Breaux Bros. Farms, Inc. v. Teche Sugar Co., Inc., 21 F.3d 83, 86 (5th Cir. 1994) (quoting *Jefferson Parish*, 466 U.S. at 29).

166. Kendall v. Visa U.S.A., Inc., 518 F.3d 1042, 1047 (9th Cir. 2008) ("(1) a contract, combination or conspiracy among two or more persons or distinct business entities; (2) by which the persons or entities intended to harm or restrain trade or commerce among the several States, or with foreign nations; (3) which actually injures competition.") (these requirements are so broad that it does not provide a framework to distinguish anticompetitive actions from parallel conduct, as the court found in this case).

167. Although the district court takes into consideration both sides of Justice O'Connor's rule of reason analysis, the district court changed the basic nature of the inquiry by considering only one side of the scale at a time through a burden shifting approach:

> [P]laintiffs bear an initial burden to demonstrate the defendants' challenged behavior had an *actual* adverse effect on competition as a whole in the [tied product] market . . . . If the plaintiff fulfills this preliminary burden, however, the burden shifts to the defendants to offer evidence of the pro-competitive effects of their agreement. Assuming defendants can provide such proof, the burden shifts back to the plaintiffs to prove that any legitimate competitive benefits offered by defendants could have been achieved through less restrictive means.

In re Wireless Tel. Serv. Antitrust Litig., 385 F. Supp. 2d 403, 415 (S.D.N.Y. 2005) (citations omitted).

168. Audell Petroleum Corp. v. Suburban Paraco Corp., 903 F. Supp. 364, 372 (E.D.N.Y. 1995).

markets, while remaining largely consistent with Justice O'Connor's concurring opinion in *Jefferson Parish*. The most fascinating aspect of the smartphone illustration is that it tests the precedential value of the D.C. Circuit's rule of reason analysis. Bundling an mOS with an application clearinghouse is just similar enough to fall within the D.C. Circuit's explicitly narrow holding—which applied only to cases where "the tying product is software whose major purpose is to serve as a platform for third-party applications and the tied product is complementary software functionality"[169]—but also sufficiently different due to externalities (such as mobility, network effects, and data limitations). Thus the question becomes: was *Microsoft II* a sort of legal *hapax legomenon*,[170] or was it a prescient standard for technology claims to come?

Application of the D.C. Circuit's rule of reason analysis to smartphone mOS and application clearinghouses answers this question. As a preliminary matter, *Microsoft II* explicitly finds price bundling to be a legally cognizable form of tying under the rule of reason.[171] Even though bundling actions are generally permitted,[172] this explicit permission incorporates the understanding that the separate product test may not always be an effective threshold and that zero pricing can be just as insidious as it is facially procompetitive. Applied to a mOS/application clearinghouse bundle, this means that a plaintiff is not barred from bringing a § 1 tying action under the rule of reason simply because they lack sufficient "demand test" econometric evidence. *Microsoft II* also requires plaintiffs to "show that [the seller's] conduct unreasonably restrained competition."[173] As phrased, such a showing is broad enough to accommodate allegations of market distortions. [174] In markets where the tied product is complementary software functionality, competitive effects are most likely to take the form of market distortions. In the context of application clearinghouses, distortions can take the form of extended periods of higher application prices, an absence of price competition (or sale prices) for applications,[175] or even more difficult-to-measure effects like the time it takes an application to become available in the market (beyond delays by the

169. *Microsoft II*, 253 F.3d at 95.

170. A word or form occurring only once in a document or corpus or something said only once. *See* Hapax Legomenon Definition, MERRIAM-WEBSTER.COM, http://www.merriam-webster.com/dictionary/hapax%20legomenon (last visited Nov. 15, 2012).

171. *Microsoft II*, 253 F.3d at 96.

172. *See generally* Thomas A. Lambert, *Appropriate Liability Rules for Tying and Bundled Discounting*, 72 OHIO ST. L.J. 909 (2011).

173. *Microsoft II*, 253 F.3d at 95.

174. Data Gen. Corp. v. Grumman Sys. Support Corp., 36 F.3d 1147, 1179 (1st Cir. 1994).

175. Hendrik Koekkoek, *The Amazon Appstore: Show Me the Money*, DISTIMO (Feb. 21, 2012), http://www.distimo.com/blog/2012_02_the-amazon-appstore-show-me-the-money/.

application's developer). Such a broad standard is appropriate in dynamic high technology markets since, from an Austrian economic perspective, courts cannot possibly anticipate developments in technological innovation. Likewise, courts cannot possibly anticipate new forms of competitive harm that can result from those innovations. The fair response, as implicitly recognized by the D.C. Circuit, is to allow new claims to be tried but to hold plaintiffs accountable through the economic balancing test advocated by Justice O'Connor. Thus, the D.C. Circuit requires plaintiffs to "demonstrate that their benefits . . . are outweighed by the harms in the *tied product* market."[176] Under this approach, the burden[177] is on the plaintiff to show that anticompetitive consequences, such as marginally higher prices for third-party applications, are significant enough to outweigh procompetitive considerations, including the "public good" and "efficiency" addressed in Part III.A. Thus, even though the plaintiffs may have a "pass" in the types of allegations that may be brought, equity is maintained by placing the initial burden on the plaintiff, thereby shielding defendants from onerous tactical suits.

Despite plaintiffs' general lack of success with tying under the rule of reason, the D.C. Circuit's opinion in *Microsoft II* should be recognized as moving beyond the gyre of other rule of reason opinions, especially when applied to dynamic high technology markets. What distinguishes the D.C. Circuit's approach as a viable solution is the court's focus on flexibility and economic balancing, which addresses the insipid flaws of the per se approach when applied to new technologies, yet does not go so far as to disincentivize innovation.

3. *Section 2 actions and applications*

A § 2 theory has a certain inherent appeal in highly concentrated technology markets, as high levels of concentration may earmark a market lacking in vigorous competition[178] due to a competitor's attempt to maintain a monopoly. Normally, these types of claims fall beyond the purview of tying—except when a plaintiff asserts that a defendant competitor has exercised market power in the *tied* product market for the purpose of maintaining a monopoly in the *tying* product market. There is a unique draw for these types of claims in software antitrust actions, since advances in modern software occur primarily by integrating functions or features previously found in multiple software or hardware products.[179] However, § 2 claims have the somewhat contradictory

176. *Microsoft II*, 253 F.3d at 96 (emphasis in original).

177. *In re* Wireless Tel. Serv. Antitrust Litig., 385 F. Supp. 2d 403, 415 (S.D.N.Y. 2005).

178. U.S. DEP'T OF JUSTICE AND THE U.S. FED. TRADE COMM'N, HORIZONTAL MERGER GUIDELINES § 1.51 (2010).

179. Boualem Benatallah & Hamid R. Motahari Nezhad, *Service Oriented*

result in which the complaint arises out of competitive effects in the *tied* product market, but the claim is premised on an intent to maintain a monopoly in the *tying* product market. More importantly, § 2 tying claims are fundamentally flawed mechanisms for relief, as they disregard critical cross-elasticity concerns and the Supreme Court's explicit acceptance of monopoly prices as "an important element of the free market system."[180] Application of the mOS/application clearinghouse exemplifies these two critical concerns, especially in light of Justice Scalia's majority opinion in *Trinko*.[181]

As cited by the District Court for the District of Columbia in *Microsoft I*[182] and most recently reiterated by the Supreme Court in *Trinko*,[183] a "monopoly maintenance" violation of § 2 of the Sherman Act has two elements:(1) "the possession of monopoly power in the relevant market"; and (2) "the willful acquisition or maintenance of that power as distinguished from growth or development as a consequence of a superior product, business acumen, or historic accident."[184]In the smartphone mOS/application clearinghouse context, plaintiffs would have to allege that (1) smartphone mOS producers posses *monopoly power*, not just market power, in the relevant market for application clearinghouses; and (2) mOS producers are *willfully* maintaining that power in an anticompetitive manner that cannot be confused with innovation, business judgment, or random chance. The first element raises a much higher standard for economic power than is required by either the per se or rule of reason approaches under § 1. When the largest market shares for an mOS producer are less than 50% and as low as 30%, any contention of monopoly power is suspect.[185] Moreover, the only other way to overcome this hurdle is to assert

---

Architecture: Overview and Directions, *in* ADVANCES IN SOFTWARE ENGINEERING (Egon Börger & Antonio Cisternino eds., 2008) ("Integration has been one of the main drivers in the software market during the late nineties and into the new millennium. It has led to a large body of research and development in areas such as data integration, software components integration (EAI), and recently service integration and composition.").

180. Verizon.com, Inc. v. Trinko, 540 U.S. 398, 407 (2004) ("The mere possession of monopoly power, and the concomitant charging of monopoly prices, is not only not unlawful; it is an important element of the free-market system. The opportunity to charge monopoly prices-at least for a short period-is what attracts "business acumen" in the first place; it induces risk taking that produces innovation and economic growth. To safeguard the incentive to innovate, the possession of monopoly power will not be found unlawful unless it is accompanied by an element of anticompetitive *conduct.*").

181. *Id.* at 407.

182. *Microsoft I*., 87 F. Supp. 2d at 35.

183. *Trinko*, 540 U.S. at 407.

184. United States v. Grinnell Corp., 384 U.S. 563, 570-71 (1966).

185. *See* BookLocker.com, Inc. v. Amazon.com, Inc., 650 F. Supp. 2d 89, 103-04 (D. Me. 2009); *Cf.* Illinois Tool Works Inc. v. Indep. Ink, Inc., 547 U.S. 28, 44 (2006) ("[T]he vast majority of academic literature recognizes that a patent does not necessarily confer market power."); CIBA Vision Corp. v. De Spirito, No. 109CV01343JOF, 2010 WL 553233 (N.D. Ga. 2010).

that a branded mOS is a market unto itself.[186]

The second element requires the plaintiff to show that mOS producers are *willfully* maintaining anticompetitive monopoly power. This element has two component parts: (a) that the maintenance of the monopoly power is occurring in an anticompetitive fashion and (b) that it is being done in a *willful* manner. In the smartphone illustration, the theory that zero pricing of applications creates barriers to entry in the application clearinghouse market meets the first component part. Furthermore, because the price of the application market is 'rolled' into the price of the mOS, it is impossible for a would-be competitor to develop a profitable alternative product. Thus, zero pricing of application markets would be illegally used to maintain a company's market share in the mobile operating system market. Simply stated, proprietary application clearinghouses can be used to reinforce the market positions of smartphone mobile operating systems, which would support a monopoly maintenance claim.

The second component part would require that the plaintiff show that mOS producers are *willfully* using anticompetitive means to maintain their monopoly position. The problem with such a position is that zero prices for application clearinghouse services may not be the result of anticompetitive bundling, but the result of a different mechanism for monetization,[187] a fundamental business consideration. Additionally, it is plausible that a single application market is dominant simply because fringe competitors have not developed a better product.[188] Moreover, the ultracompetitive price for application clearinghouses is fundamentally a benefit to consumers.

The second significant consideration in high technology tying cases is the cross-elasticity of demand, highlighted by Justice Scalia in his *Eastman Kodak* dissent.[189] Cross-elasticity concerns are particularly relevant in dynamic high

---

186. Such an assumption may also be faulty from a traditional economic perspective. If, hypothetically, a particular smartphone was such a unique product that there were no other close substitutes, then the constraining economic choice would be between purchasing the product or not. However, more realistically, smartphones are constrained by other similar products such as feature phones. While the purchase of feature phones has declined in comparison to smartphones, the array of products that may constitute the combined amenities of a smartphone may still act as a price constraint. *See generally*, Aaron Smith, *Nearly Half of American Adults Are Smartphone Owners*, PEW INTERNET & AMERICAN LIFE PROJECT (March 1, 2012), http://pewinternet.org/Reports/2012/Smartphone-Update-2012/Findings.aspx.

187. *See* Koekkoek, *supra* note 177, at *5.

188. *Cf*. Priya Ganapati, *Independent App Stores Take On Google's Android Market*, WIRED.COM (June 11, 2010), http://www.wired.com/gadgetlab/2010/06/independent-app-stores-take-on-googles-android-market/.

189. Eastman Kodak Co. v. Image Tech. Servs, 504 U.S. 451, 491 (1992) (Scalia, J., dissenting) ("Interbrand competition would render Kodak powerless to gain economic power over an additional class of consumers, to price discriminate by charging each customer a 'system' price equal to the system's economic value to that customer, or to raise barriers to entry in the interbrand equipment markets.").

technology markets because feature integrations can merge markets and diminish competitive concerns. For example, if the market for smartphone mobile operating systems is viewed as cross-platform, as Scalia perceived the market for photocopiers, then the market share for any particular mOS producer would be too small to support the first element of a § 2 claim. Furthermore, competition between mOS producers, like cross-brand competition between photocopier OEMs, would constrain downstream anticompetitive price effects. This contention still has merit in modern technology markets, as a mOS competes partially based on the number of applications available on each system. Essentially, a consumer can choose to switch between mOS platforms if the third-party application selection and prices available through an application clearinghouse are not competitive. In fact, commentators have contended that is exactly the reason that Blackberry lost market share in 2011. Thus, if a mOS producer increases third-party application prices at the application clearinghouse 'bottleneck', then mOS producers would actually be weakening their competitive position.

Tying claims under section § 2 claims remain highly contentious. Although there are instances where such claims may be warranted,[190] these claims may suffer from significant conceptual deficiencies—especially in regards to cross-elasticity defenses. Nevertheless, § 2 claims capture forms of tying not easily addressed by § 1, yet are endemic in software and other high technology industries.

B. *Common High-Technology Economic Complexities in the Smartphone Paradigm*

1. *Evidence of competitive effects through "natural" experiments: price indexes, Google Play, and Amazon Appstore*

Consistently, one of the most significant challenges in antitrust litigation is establishing (and, more to the point, proving) anticompetitive effects. This challenge is particularly acute in tying and bundling disputes—where plaintiffs are frequently left in the undesirable position of asserting that prices would be lower had the defendant not exercised market power to anticompetitive ends in a second market. In short, tying and bundling claims may depend heavily on the counterfactual.[191] In a perfectly competitive world, one may assume that

190. *See, e.g.*, *Free FreeHand Corp. v. Adobe Sys. Inc*., 852 F.Supp.2d 1171 (N.D.Cal. 2012) at * 1 ("Plaintiffs have plausibly alleged that Adobe willfully acquired monopoly power and maintained that power through anticompetitive conduct. If, as alleged, Adobe ceased the development of FreeHand while steering existing FreeHand users to a bundled product, thereby further raising already high barriers to entry, it is plausible to infer that this conduct tended 'to impair the opportunities of rivals' and 'did not further competition on the merits.'").

191. Granted, similar problems are not entirely uncommon–consider the debate over

the market could support multiple clearinghouse vendors who would compete to sell applications to downstream customers. Competition between buyers and sellers would drive the price of the product(s) down to the equilibrium—a price point lower than would be if there were only a single vendor (producer). However, real world markets rarely reflect these characteristics. In many markets, especially high-technology software markets, there may be only a single producer. And there may be good reasons that only a single producer exists. For example, the market may not be large enough to profitably support multiple producers. Moreover, simply because a producer has market power does not mean that competitor is using that power anticompetitively. High-technology producers regularly assert these types of arguments—and rightly so. Both Economists and courts assume that economic actors act rationally.[192] Accordingly, the lack of product diversity is most likely the result of natural market forces, not anticompetitive tying or bundling arrangements.[193]

Perhaps the most legally functional and ultimately practical means of demonstrating anticompetitive effects is to show that the market can be, and would be, different by way of a 'natural experiment.'[194] A natural experiment is "evidence that the posited harm has occurred under circumstances similar to the [conduct in question]."[195] Of course, it is rare that a suitably similar situation has occurred in virtually the same market, with the same relevant competitors, and under the same operative market dynamics. Nevertheless, evidence in the form of a natural experiment is both economically and legally poignant, not to mention easily accessible by judges and juries.

The launch of Amazon Appstore, available on the same platform (mOS) as

---

negative evidence. *See*, *e.g.*, George Fisher, *The Jury's Rise As Lie Detector*, 107 YALE L.J. 575, 648 (1997).

192. *See* Richard A. Posner, *Rational Choice, Behavioral Economics, and the Law*, 50 STAN. L. REV. 1551 (1998).

193. *In re* Flat Glass Antitrust Litig., 191 F.R.D. 472, 484 (W.D.Pa. 1999) ("[C]ontentions of infinite diversity of product, marketing practices, and pricing have been made in numerous cases and rejected.").

194. F.T.C. v. Foster, CIV 07-352 JBACT, 2007 WL 1793441, at *38 (D.N.M. May 29, 2007) ("[A]ntitrust agencies rely extensively on natural market experiments to provide relevant evidence to show whether or not a transaction is likely to lessen competition.").

195. *Id. But cf.* F.T.C. v. Church & Dwight Co., Inc., 747 F. Supp. 2d 3, 7 (D.D.C. 2010), *aff'd,* 665 F.3d 1312 (D.C. Cir. 2011) (a description in a very different circumstance of a general concept does not create a legal standard). *See also* F.T.C. v. ProMedica Health System, Inc., 3:11 CV 47, 2011 WL 1219281, at *14 (N.D. Ohio March 29, 2011) ("Real-world natural experiments in the marketplace confirm that St. Luke's successfully competed with ProMedica for a significant number of patients. For example, ProMedica estimated that St. Luke's readmission to X's network in 2009, after being excluded since 2005, would cost ProMedica X dollars in gross margin annually. . . . After St. Luke's was readmitted in July 2009, St. Luke's market share in its core service area rose from 36 percent to 43.1 percent in 2010, while ProMedica's market share in the same area declined. . . . Mercy's and UTMC's shares during this period changed little in comparison.") (citations omitted).

Google Play, provides just such an example of a 'natural experiment.'[196] Essentially, price and overlap data between two significant competitors (Amazon Appstore and the Google Play/the Android Market)[197] demonstrates how the world might be different if there were other producers who could compete vigorously. In fact, analysts have found that the average price of applications is lower in the Amazon Appstore (the independent competitor application clearinghouse) than in Google Play (the primary mOS manufacturer produced application clearinghouse).[198] One analyst noted that "[w]hile all available paid applications [average a price of] . . . $3.13 in the Google Android Market, these applications [average a price of] . . . $2.77 in the Amazon Appstore."[199] What makes this comparison even more interesting is the fact that "50% of all apps in the Amazon Appstore are also directly available in the Google Android Market."[200] While part of the price differential can be explained by special sale prices[201] and different mechanisms for monetizing the intermediary application market[202], this does not negate the procompetitive effects for consumers. Moreover, the price differentials and alternatives monetization systems may be 'innovations' that are indicative of a resource-based or evolutionary economic dynamics.[203] In sum, the competition between

196. *See* Koekkoek, *supra* note 176, at *5 ("The Amazon Appstore is very similar to the Google Android Market in [many] ways. Both stores have applications on the same platform, and 50% of all apps in the Amazon Appstore are also directly available in the Google Android Market. However, there are several important differences. Due to the way Amazon handles the pricing of applications, the prices of the same apps can be very different in these two stores.").

197. The Amazon Appstore is not the only third party application clearinghouse available on the Android mOS platform; although the alternatives, GetJar, Opera Mobile App Store, and Handango are best described as "fringe competitors" whose pricing and availability have no statistically significant effect on market prices or dynamics. *See id.* at 3.

198. *Id.* at 4.

199. *Id.* at 4 ("Looking at only the 100 most popular paid applications, the difference is even larger. The average price of [the] top 100 applications is $3.76 in the Google Android Market and $2.24 in the Amazon Appstore.").

200. *Id.* at 4, 6 ("The Google Android Market had 368,985 available apps in January in the US . . . while the Amazon Appstore has only 26,826 available applications. A large number of these applications – 13,432 – are available in Google Android Market as well.").

201. *Id.* at 5 ("In the Google Android Market, these [top 100] applications had an average price $3.47. In the Amazon Appstore, the average price of these same applications was $2.89 during the period they were ranked in the top 100. Applications that cause this difference are, for example, Monopoly, which cost $0.99 for a limited time in the Amazon Appstore (normal price $4.99) while it cost $4.99 for the whole month in the Google Android Market. Another one is Splashtop Remote Desktop, which been free for a day in the Amazon Appstore, while it cost $4.99 in the Google Android Market. Especially when looking at these kinds of temporary price reductions of top applications in January, we found that the reductions are generally larger in the Amazon Appstore than in the Google Android Market.").

202. *Id.* at 5 ("One of the reasons for the difference is that Amazon is responsible for setting the price in its store, while in the Google Android Market, the developer is.").

203. *See* Koekkok, *supra* note 177 at *5; *see also* J.S. Metcalfe, *Evolutionary*

Google's Android Market and the Amazon Appstore suggests that (1) the intermediary market for applications is robust enough to support multiple competitors, (2) competition is likely to drive prices down, and (3) competitors are likely to develop 'innovations' regarding access and monetization that increase efficiency and have procompetitive effects downstream (*e.g.*, increasing the revenue returned to application producers).

One of the most serious pitfalls of the 'natural experiment' is that it is essentially anecdotal evidence. However, certain economic analyses can change the nature of the inferences gained from natural experiments. Primarily, price indexes[204] can be used to assess the approximate value of increased competition in the analogous market.[205] A price index is used to "compare[] the prices of a set of products at different points in time" [206] which allows for a numerical measurement of "price changes or price differentials rather than price levels."[207] Thus, natural experiments can be used to demonstrate both the *existence*[208] and the *magnitude*[209] of anticompetitive effects.

Economics and Technology Policy, 104 ECON. J. 931, 933 (1994) (stated alternatively, "the principal themes of evolutionary economics are twofold: the processes which determine the range of actual innovations (variety) introduced into the economy; and, the processes which alter the relative economic importance of the competing alternatives (selection). The fundamental issues are dynamic, they are related to the nature of competition as a process of endogenous change, and the relation between variety and selection is two-way.").

204. There are three major variations of price indexes: the Fisher price index (also known as the 'ideal' price index), the Laspeyres price index and the Paasche price index. Kenneth Flamm advocates a form of the Fisher price index:

$$\sqrt{\sum_i \left(\frac{p_i^0 q_i^0}{\sum_i p_i^0 q_i^0}\right)\frac{p_i^1}{p_i^0} \div \sum_i \left(\frac{p_i^1 q_i^1}{\sum_i p_i^1 q_i^1}\right)\frac{p_i^0}{p_i^1}}$$

although this equation requires complete cross-elasticity data (which is frequently difficult to obtain). *See* Kenneth Flamm, *Digital Convergence? The Set-Top Box and the Network Computer*, *in* COMPETITION, INNOVATION AND THE MICROSOFT MONOPOLY: ANTITRUST IN THE DIGITAL MARKETPLACE 259-60 (Jeffery A. Eisenach & Thomas M. Lenard eds., 1999); *see also Glossary:Fisher price index*, EUROSTAT EUROPEAN COMMISSION, http://epp.eurostat.ec.europa.eu/statistics_explained/index.php/Glossary:Fisher_price_index; Bert M. Balk, *Axiomatic Price Index Theory: A Survey*, 63 INT'L STATISTICAL R. 69 (1995) (regarding Laspreyres & Paasche price indexes).

205. Flamm, *supra* note 209, at 259-60.

206. EUROSTAT EUROPEAN COMMISSION, *European Price Statistics An Overview, in* EUROSTAT EUROPEAN COMMISSION 15, (Gunter Schäfer ed., 2008), *available at* http://epp.eurostat.ec.europa.eu/cache/ITY_OFFPUB/KS-70-07-038/EN/KS-70-07-038-EN.PDF, ("A price index shows how much must be paid for a set of products at some point in time relative to what would have been paid for the same set of products at another point in time, which latter is taken as the reference of the comparison. This is done by setting the index value for the reference . . . to 100 so-called index points.").

207. *Id.* at 15.

208. The variation proposed in this Note is analogous to the 'during or after' or 'before and after' methods of analysis. *Cf. In re* Scrap Metal Antitrust Litig., 527 F.3d 517 (6th Cir. 2008) (although this case was a conspiracy case—not a tying or bundling case—the Sixth Circuit's dicta on expert testimony illustrate how a price index can transform anecdotal

Furthermore, price indexes measure the price differential relative to time. Because the unit of time can be changed to conform with the circumstances, price index-based calculations of competitive effects are especially attractive in dynamic high-technology markets, where the period of competitive harm may vary from product to product and the era in which the harm occurred. These dynamics make any hard-and-fast rule, such as a two-year or three-year rule, for evaluating the competitive effects practically and conceptually untenable.[210]

An inference can be drawn from this natural experiment that if vigorous price competition within intermediary market for applications on the Android mOS platform is possible, then more competition in other mOS platform markets (*e.g.*, iOS or RIM/BlackBerry) may also be possible. Accordingly, if more vigorous competition is possible, then the question becomes why *isn't* there more competition in the other mOS platform markets? In response, the correlation between firms producing a mOS and the (same) firms producing the dominant application clearinghouse for that mOS cannot be ignored. Although not dispositive in-and-of itself, a natural experiment can be persuasive evidence of anticompetitive conduct.[211] In the smartphone mOS & clearinghouse context, pre-installation of a mOS producer-made application clearinghouse may suggest an illegal bundling arrangement. In this illustration, lower prices in the Andrioid mOS market resulting from competition between Google Play and the Amazon Appstore could be used as a 'natural experiment' to demonstrate the existence and magnitude of anticompetitive effects in other mOS producer markets, such as the iOS market or the Blackberry mOS market. While natural experiments are not prima facie evidence of § 1 violations, they

---

evidence into more concrete competitive effects: "[e]mploying this method, the profits made by antitrust defendants *during* the alleged conspiracy are compared with the profits made by the defendants in the period *after* the alleged conspiracy. In simple terms, by analyzing this difference, an expert can determine the amount of profit during the conspiracy period had the antitrust violation not occurred. Presumably, the data would show that, but for the anticompetitive conduct, the defendants' profit margin would have been lower and the plaintiffs' profit margin would have been higher.").

209. *See id.* Although the analytical method used is not exactly parallel to the method proposed in this Note, the types of conclusions that can be reached are the same. For example, "[t]he 'price spread'—the difference between (1) the price a dealer paid to a generator for unprocessed scrap, and (2) the comparative figure from a price index, representing the amount the dealer earned from selling the processed scrap to a user—was the dealer's profit. By comparing the price spread during the conspiracy period with the price spread after the conspiracy period, Leitzinger concluded that the results were consistent with anticompetitive conduct: Defendants' profits declined after the conspiracy, while the generators' profits rose," *id.* at 517.

210. This Note recognized the argument that for the very reason that high-technology markets *are* dynamic, and that because consumer harm can dissipate quickly within a short period of time due to new innovations, tying and bundling actions are not appropriate mechanisms to restore competition. However, that argument sidesteps an essential goal of antitrust law—to provide relief to consumers who have been economically harmed by anticompetitive arrangements, like tying or bundling.

211. *See* cases cited *supra* note 197.

are the best possible evidence of an anticompetitive tying or bundling arrangement when no 'smoking gun'[212] emerges in discovery.

2. *The competitive "fix": cross-platform solutions and mobile Internet browsers*

One of the most common defenses in software tying actions is that any plausible competitive concerns are 'fixed' by the existence of cross-platform products or complementary software functionalities that restrain a competitor's ability to charge anticompetitive prices. This issue, of whether secondary market software applications can act as a competitive relief valve, was never definitively resolved in the *Microsoft* cases and the issue continues to elicit strong opinions in the academic literature.[213] As illustrated by the smartphone mobile operating systems and application clearinghouses, both the economic rationales supporting and rejecting secondary market competitive fixes remain viable with no clear resolution in sight.

Under this theory, any competitive problem that arises because of a tie between mobile operating systems and application clearinghouses can be solved by potential use of mobile Internet browsers, which give consumers an alternate means of accessing the end-user products (third-party smartphone applications) directly. Therefore, it does not matter whether application clearinghouses are tied to mobile operating systems, because the producer cannot charge anticompetitive prices for applications in the clearinghouse without losing sales to direct purchases via mobile Internet browsers. Essentially, if prices for applications get too high in the application clearinghouse, then consumers can bypass the clearinghouse via mobile Internet browsers and get the application they want directly from the third-party producer at a lower price. Such a competitive 'fix' is possible because application clearinghouses are zero-priced (thus there is no explicit cost associated with the means of obtaining the third party application).

Economically, the incentives to zero-price application clearinghouses on smartphones are the same as was Microsoft's incentive to zero-price Internet browsers with operating systems. Application clearinghouses may carry a zero price because "that . . . software carries with it the potential to steer . . . users to a particular [application] and thereby generate significant advertising revenues

---

212. *Cf.* InterVest, Inc. v. Bloomberg, L.P., 340 F.3d 144, 159 (3d Cir. 2003) ("Because direct evidence, the proverbial 'smoking gun,' is difficult to come by, plaintiffs have been permitted to rely solely on circumstantial evidence (and the reasonable inferences that may be drawn therefrom) to prove a conspiracy.").

213. *See, e.g.*, Benjamin Klein, *Microsoft's Use of Zero Price Bundling to Fight the "Browser Wars," in* COMPETITION, INNOVATION AND THE MICROSOFT MONOPOLY: ANTITRUST IN THE DIGITAL MARKETPLACE 251 (Jeffery A. Eisenach & Thomas M. Lenard eds., 1999).

and on-line shopping commissions."[214] This is certainly not illegal behavior. However, there is "[t]he contrary view, that . . . [a competitor may] set a zero . . . price as a temporary predatory tactic to drive out competitors before planning to recoup lost profits by raising the . . . price to a monopoly level."[215] This would constitute a Sherman Act violation.

Conclusive evidence of either economic explanation is rarely possible. In the smartphone paradigm, defendants advocating the competitive fix may easily contend that the ultracompetitive price is not intended to foreclose competition between application clearinghouses, but is merely a residual effect of natural competitive behavior. As is with pricing for most "other Internet-related software . . . a zero . . . price profitably maximizes penetration and more closely reflects the essentially negative marginal cost to the . . . supplier of additional software sales."[216]

Figure 3: Cross-Platform Market Share and Mobile Browser Market Share Comparisons

Browser	Mobile Browser Market Share[217]	Mobile OS	Moblie OS Market Share[218]	Browser Usage as a Percentage of Mobile OS
Safari	65.8	iOS	65.3[219]	1.007%
Android Browser	19.2	Android	19.7 (46.3)[220]	97.461% (38.401%)
Opera Mini/Mobile	10.6	Java ME (Android)	10.2[221] (46.3)[222]	103.921% (42.289)[223]
BlackBerry	1.5	BlackBerry	1.9	78.947%

214. *Id.* at 219.

215. *Id.*

216. *Id.* at 223.

217. Net Applications, *Mobile/Table Browser Market Share*, NETMARKETSHARE.COM (July, 2012), http://www.netmarketshare.com/browser-market-share.aspx?qprid=0&qpcustomd=1.

218. *Id.*

219. Statistic includes other devices, such as the iPad and iPod.

220. *More US Consumers Choosing Smartphones as Apple Closes the Gap on Android,* NIELSEN WIRE (Jan. 19, 2012), http://blog.nielsen.com/nielsenwire/consumer/more-us-consumers-choosing-smartphones-as-apple-closes-the-gap-on-android/.

221. This statistic substitutes the Java ME market share for the Opera Mini/Mobile market share because Opera Mini is based on Java ME code. *See Mobile/Tablet Operating System Market Share*, NETMARKETSHARE.COM (July. 2012), http://www.netmarketshare.com/mobile-market-share.

222. *See* NIELSEN WIRE, *supra* note 225.

223. Using the Nielsen data, this number expresses the Opera browser as a percentage of the Android Market share based on Nielsen. This is not an equivalent expression, as there is a global versus domestic difference—but it probably more accurately reflects reality, as Opera is reported to be nearly exclusively used on Android phones. *See id.*

However, the competitive 'fix' does not address the underlying issues of information costs related to seeking out a particular end-product (such as a specific third-party application) without the use of an application clearinghouse, nor does the 'fix' address the dampening effects of bundling on competition from nascent cross-platform application clearinghouses. Such a result is even more troubling in a world where the utility of the Internet itself, versus the utility of applications is in question.[224] In other words, the competitive 'fix' may be obscuring foreclosure of competitors from creating a cross-platform application clearinghouse that would allow smartphone application producers to produce just one version of the product.[225]

What is clear from the competitive 'fix' application to the smartphone paradigm is that both economic and legal rationales remain viable because they address basic conceptions of the market. From a resource-based economic perspective, the mere existence of a competitive 'fix' opens the door for competitors to fundamentally change the market dynamics by reallocating their resources in more efficient, competitive combinations that bypass any 'bottleneck' held by a current competitor. However, from a path-dependence perspective, competitive fixes may be magnifying the utility of complementary software functionalities. Thus, competitive 'fixes' should not be presumed to support a procompetitive outcome, but should be evaluated closely based on the structure and dynamics of a particular market.

CONCLUSION & RECOMMENDATIONS

The bundling of smartphone operating systems and application markets is a modern microcosm of tying's age old question: will declaring product combinations illegal tying arrangements, especially in nascent technology markets, benefit consumers or will it stymie innovation? And what does the smartphone illustration tell us about the utility of current tying law in dynamic high technology markets? One simple point: the principles of current tying and bundling law can be effectively applied to address new technologies and their concordant antitrust concerns. While the fit may not be perfect, there is no reason to adopt new tests when current tests can effectively address plaintiff's concerns and protect defendants from dilatory suits. Nevertheless, current tying and bundling law can be tailored to address critics by (1) recognizing that antitrust law is based on the assumption that a competitive equilibrium is achievable, and that non-equilibrium-based assertions and defenses, while not

224. *See generally*, Janna Anderson & Lee Raine, *The Future of Apps and Web*, PEW INTERNET & AMERICAN LIFE PROJECT (Mar. 23, 2012), *available at* http://pewinternet.org/Reports/2012/Future-of-Apps-and-Web/Overview.aspx.

225. *See* Peter Wayner, *The Cross-Platform Option: Web Apps for Smartphones*, INFOWORLD, http://www.infoworld.com/d/developer-world/cross-platform-option-web-apps-smartphones-485 (last visited Mar. 14, 2012).

excluded, should be the exception; (2) making greater use of the rule of reason in high technology tying cases; and (3) viewing natural experiments buttressed by econometric evidence as indicative of anticompetitive conduct.

A. *Non-Equilibrium-Based Assertions Should be Recognized as the Exception, Not the Norm*

Antitrust law is premised on the idea that markets do have competitive equilibriums. Anticompetitive behavior is only conceptually possible when competitors' conduct creates market distortions. Non-equilibrium-based assertions are legitimate explanations of market behavior. However, legal interpretations of tying and bundling arrangements based on non-equilibrium characterizations of market behavior cut the antitrust inquiry off at the pass. In effect, this denies plaintiffs a true opportunity to investigate, detect, and seek redress for truly anticompetitive tying arrangements.

B. *Flexible Rule of Reason Analyses Should Be Utilized in High Technology Industries*

As discussed in Part III.A.2, the rule of reason focuses directly on the balancing implicit in all tying and bundling cases; whether the tying arrangement's restraints on competition were severe enough to warrant relief or even considering all its faults, or whether the tying arrangement was essentially procompetitive. Although plaintiffs have been wary to bring § 1 tying claims under the rule of reason, the rule of reason provides a better forum to address anticompetitive conduct in high technology markets than either § 1 per se tying claims or § 2 monopoly maintenance claims. While variations on the rule of reason have arisen among the circuits, the D.C. Circuit's approach to the rule of reason in *Microsoft II* provides the best template for evaluating tying claims in dynamic high technology markets.

C. *Evidentiary Tools, Such as Natural Experiments and Price-Indexes Used to Evaluate Short-Term Competitive Effects, Should Be Considered a Permissible Inference of Anticompetitive Tying or Bundling Arrangements*

While natural experiments are not prima facie evidence of § 1 violations, they are the best possible evidence of an anticompetitive tying or bundling arrangement when no 'smoking gun' emerges in discovery. In order for tying and bundling law to maintain its flexibility and ensure equity, courts must evaluate new forms of evidence, like natural experiments, in a more favorable light.[226] When a natural experiment is internally valid, adequately explained

226. *Cf.* Easterbrook, *supra* note 121, at 319 ("We need not think that things can get

through econometric analysis, and not overextended to apply to the instant claim, then it should be presumed valid to infer anticompetitive effects. However, in order to maintain equity, defendants must be given an adequate opportunity to rebut the underlying assumptions of the natural experiment and to reframe the natural experiment through arguments and evidence that may not have been addressed in the original description of the experiment.

Dynamic high technology markets, like smartphones, are likely to receive greater antitrust scrutiny. As new innovations in mobile technologies appear, new competitive harms may arise. Current tying and bundling law is sufficiently flexible to address these claims, but courts must also be aware of the dangers posed by non-equilibrium-based arguments and econometric analyses, while remaining open to new means of proving competitive harm. Tailoring current law with the rule of reason and natural experiments improves the utility of current tying and bundling inquiries and reflects the complexity of dynamic high technology markets.

---

worse to be skeptical of the comparative advantage of courts in helping inframarginal buyers, however . . . How many people think that courts could produce a net improvement in social welfare if they were given statutory power to examine manufacturers' decisions . . . ? The answer must be 'no one.' Yet since judicial control of design decisions is economically identical to judicial control of the retail service decisions, there is no reason to believe that judges have a comparative (or absolute) advantage at that task either. If we would not want courts to have power over design, why should we entrust courts with power over functionally identical economic matters?").