

Production: Produced by members of the *Berkeley Technology Law Journal*.
All editing and layout done using Microsoft Word.

Printer: Joe Christensen, Inc., Lincoln, Nebraska.
Printed in the U.S.A.

The paper used in this publication meets the minimum requirements of American National Standard for Information Sciences—Permanence of Paper for Library Materials, ANSI Z39.48—1984.

Copyright © 2013 Regents of the University of California.
All Rights Reserved.

Berkeley Technology Law Journal
University of California
School of Law
3 Boalt Hall
Berkeley, California 94720-7200
btlj@law.berkeley.edu
<http://www.btlj.org>



BERKELEY TECHNOLOGY LAW JOURNAL

VOLUME 28

NUMBER 1

SPRING 2013

TABLE OF CONTENTS

ARTICLES

DO BAD THINGS HAPPEN WHEN WORKS ENTER THE PUBLIC DOMAIN?: EMPIRICAL TESTS OF COPYRIGHT TERM EXTENSION.....	1
<i>Christopher Buccafusco & Paul J. Heald</i>	
STATE PATENT LAWS IN THE AGE OF <i>Laissez Faire</i>	45
<i>Camilla A. Hrdy</i>	
THE BACKGROUND OF OUR BEING: INTERNET BACKGROUND CHECKS IN THE HIRING PROCESS.....	115
<i>Alexander Reicher</i>	
THE LAW OF THE ZEBRA	155
<i>Andrea M. Matmyshyn</i>	
EXACTITUDE IN DEFINING RIGHTS: RADIO SPECTRUM AND THE “HARMFUL INTERFERENCE” CONUNDRUM.....	227
<i>Thomas W. Hazlett & Sarah Oh</i>	

SUBSCRIBER INFORMATION

The *Berkeley Technology Law Journal* (ISSN1086-3818), a continuation of the *High Technology Law Journal* effective Volume 11, is edited by the students of the University of California, Berkeley, School of Law (Boalt Hall) and is published in print three times each year (March, September, December), with a fourth issue published online only (July), by the Regents of the University of California, Berkeley. Periodicals Postage Rate Paid at Berkeley, CA 94704-9998, and at additional mailing offices. POSTMASTER: Send address changes to Journal Publications, University of California, Berkeley Law—Library, LL123 Boalt Hall—South Addition, Berkeley, CA 94720-7210.

Correspondence. Address all correspondence regarding subscriptions, address changes, claims for non-receipt, single copies, advertising, and permission to reprint to Journal Publications, University of California, Berkeley Law—Library, LL123 Boalt Hall—South Addition, Berkeley, CA 94705-7210; (510) 643-6600; JournalPublications@law.berkeley.edu. *Authors:* see section titled Information for Authors.

Subscriptions. Annual subscriptions are \$65.00 for individuals and \$85.00 for organizations. Single issues are \$30.00. Please allow two months for receipt of the first issue. Payment may be made by check, international money order, or credit card (MasterCard/Visa). Domestic claims for non-receipt of issues should be made within 90 days of the month of publication; overseas claims should be made within 180 days. Thereafter, the regular back issue rate (\$30.00) will be charged for replacement. Overseas delivery is not guaranteed.

Form. The text and citations in the *Journal* conform generally to the THE CHICAGO MANUAL OF STYLE (16th ed. 2010) and to THE BLUEBOOK: A UNIFORM SYSTEM OF CITATION (Columbia Law Review Ass'n et al. eds., 19th ed. 2010). Please cite this issue of the *Berkeley Technology Law Journal* as 28 BERKELEY TECH. L.J. ____ (2013).

BTLJ ONLINE

The full text and abstracts of many previously published *Berkeley Technology Law Journal* articles can be found at <http://www.btlj.org>. Our site also contains a cumulative index, general information about the *Journal*, and the Bolt, a collection of short comments and updates about new developments in law and technology written by members of BTLJ.

INFORMATION FOR AUTHORS

The Editorial Board of the *Berkeley Technology Law Journal* invites the submission of unsolicited manuscripts. Submissions may include previously unpublished articles, essays, book reviews, case notes, or comments concerning any aspect of the relationship between technology and the law. If any portion of a manuscript has been previously published, the author should so indicate.

Format. Submissions are accepted in electronic format through the ExpressO online submission system. Authors should include a curriculum vitae and resume when submitting articles, including his or her full name, credentials, degrees earned, academic or professional affiliations, and citations to all previously published legal articles. The ExpressO submission website can be found at <http://law.bepress.com/expresso>.

Citations. All citations should conform to THE BLUEBOOK: A UNIFORM SYSTEM OF CITATION (Columbia Law Review Ass'n et al. eds., 19th ed. 2010).

Copyrighted Material. If a manuscript contains any copyrighted table, chart, graph, illustration, photograph, or more than eight lines of text, the author must obtain written permission from the copyright holder for use of the material.

DONORS

The *Berkeley Technology Law Journal* and the Berkeley Center for Law & Technology acknowledge the following generous donors to Berkeley Law's Law and Technology Program:

Partners

COOLEY LLP

FENWICK & WEST LLP

COVINGTON & BURLING LLP

ORRICK, HERRINGTON &
SUTCLIFFE LLP

Benefactors

FISH & RICHARDSON P.C.

SKADDEN, ARPS, SLATE, MEAGHER
& FLOM LLP & AFFILIATES

KASOWITZ BENSON TORRES &
FRIEDMAN LLP

WEIL, GOTSHAL & MANGES LLP

KIRKLAND & ELLIS LLP

WHITE & CASE LLP

LATHAM & WATKINS LLP

WILMER CUTLER PICKERING HALE
AND DORR LLP

MCDERMOTT WILL & EMERY

WILSON SONSINI

MORRISON & FOERSTER LLP

GOODRICH & ROSATI

WINSTON & STRAWN LLP

Members

ALSTON & BIRD LLP	HOGAN LOVELLS LLP
BAKER BOTTS LLP	IRELL & MANELLA LLP
BAKER & MCKENZIE LLP	KILPATRICK TOWNSEND & STOCKTON LLP
BINGHAM MCCUTCHEN LLP	KNOBBE MARTENS OLSON & BEAR LLP
DURIE TANGRI LLP	MORGAN, LEWIS & BOCKIUS LLP
FINNEGAN, HENDERSON, FARABOW, GARRETT & DUNNER, LLP	MUNGER, TOLLES & OLSON LLP
GTC LAW GROUP & AFFILIATES	ROPES & GRAY LLP
GUNDERSON DETTMER STOUGH VILLENEUVE FRANKLIN & HACHIGIAN, LLP	SIDLEY AUSTIN LLP
HAYNES AND BOONE, LLP	SIMPSON THACHER & BARTLETT LLP
HICKMAN PALERMO TRUONG BECKER BINGHAM WONG, LLP	TURNER BOYD LLP
KEKER & VAN NEST LLP	VAN PELT, YI & JAMES LLP
	WEAVER AUSTIN VILLENEUVE & SAMPSON, LLP

BOARD OF EDITORS

2012–2013

Executive Committee

Editor-in-Chief

DAVID ROSEN

Managing Editor
LAUREN ESCHER

Senior Articles Editors
YVONNE LEE
HANNAH MINKEVITCH

Senior Executive Editor
JOE SEXTON

Senior Annual Review Editors
WINNIE HUNG
JANE LEVICH

Senior Scholarship Editor
COURTNEY BOWMAN

Editorial Board

Submissions Editors
JULIA KOLIBACHUK
PENNI TAKADE

Production Editor
JARAD BROWN

Bluebook Editors
GABRIEL MILITELLO
YUE WANG
ZACH WOOD

Annual Review Editors
ANGEL DIAZ
LALI MADDURI

Notes & Comments Editors
ELISE EDLIN
BRITT HARWOOD

Symposium Editors
KAREN KOPEL
KILEY WONG

Publishing Editor
DINA ROUMIANTSEVA

Web Content Editor
CHRIS CIVIL

Web & Technology Editors
EDUARD MELESHINSKY
ANDREA YANKOVSKY

External Relations Editor
GAVIN LIU

Member Relations Editor
CASEY HULTIN

Articles Editors
BRADY BLASCO
JULIE BYREN
EMILY CHEN
ROSS COHEN
CHRISTINA FARMER

Articles Editors
GRANT GARBER
ROBIN KUNTZ
NICHOLAS LAMPROS
MARK LANGER

MARIENNA MURCH
JUSTIN ORR
SARAH ORRICK
LAUREN SMITH
ETHAN WEINER

MEMBERSHIP

Vol. 28 No. 1

Associate Editors

ASHWIN ANAND	JEFFREY HIRSCHHEY	ROSALIND SCHONWALD
NATASHA ARORA	JULIA HOFFMAN	DONGBIAO SHEN
JASON BELL	SAMANTHA HUANG	DANIELA SPENCER
LAUREN BLAKELY	ALEX LI	MARTIN STIERLE
MICHAEL BURSHTEYN	GI-KUEN LI	JENNA STOKES
LAUREN CARDINET	KASSANDRA MALDONADO	CHRISTOPHER SUNG
ANDREW COLLINS	JOSEPH MORNIN	STEVEN SWANSON
RAFAEL CRUZ	PURBA MUKERJEE	TUYET TU
BEVAN DOWD	VIABHAW RANGARAJAN	BONNIE WATSON
MISA EIRITZ	LEXI RUBOW	ANDREW YOCOPIS
JONATHAN FRANCIS	NAZLI SAKA	MARK ZAMBARDA
REBECCA FRIEDMAN	EMILY SANGI	OGNJEN ZIVOJNVIC
	ALLISON SCHMITT	

Members

ADAM AMIR	CHRISTOPHER HAMMOND	LEA MEKNECHE
CRAIG ARMSTRONG	VIKRAM IYENGAR	WADE ORBELIAN
ERIK BAUMAN	DAMION JURENS	JOHN OWEN
MARION BERGERET	ROHAN KADAM	JILLIAN PESIN-FULOP
DALLAS BULLARD	DROFN KAERNSTED	MARIE PFAMMATTER
DOMINIQUE CAAMANO	TIGIST KASSAHUN	MEREDITH REED
SIENA CARUSO	SEHYUN KIM	JEANNE REYES
ANTHONY CHANRASMI	ALEX KIM	ARAM ROSTOMYAN
VICTOR CHIU	KAMOLA KOBILDJADNOVA	GLADYS STEPHENS
ALEX COLEY	JESSE KOEHLER	MICHAEL LIU SU
LEON DIJKMAN	LEON KOTLYAR	BEN TAKEMOTO
MATT DONOHUE	ANAGHA KRISHNAN	JONATHAN TAMIMI
ALAN ENRIQUEZ	YI LI	EMMANUELLE TANG
PAUL FETTER	DEBBIE LI	JENNIFER TRUONG
NATALIE FLEICHSIG	EMILY LY	KATHARINE WARREN
MICHAEL GAFFNEY	SAHAR MAALI	MATTHEW WASSERMAN
NYDIA GALARZA	ANGELIKA MAREK	ADAM WHITAKER
S. ZUBIN GAUTAM	DARIUS MASSOUDI	KEVIN XU
DANNY GOLDSTEIN	STEVEN S. MCCARTY-SNEAD	ZELIN YANG
MARK GRAY		LUIS ZAMBRANO

BTLJ ADVISORY BOARD

ROBERT BARR

*Executive Director of the
Berkeley Center for Law & Technology*
U.C. Berkeley School of Law
Berkeley, California

ROBERT C. BERRING, JR.

Walter Perry Johnson Professor of Law
U.C. Berkeley School of Law
Berkeley, California

JESSE H. CHOPER

Earl Warren Professor of Public Law
U.C. Berkeley School of Law
Berkeley, California

PETER S. MENELL

*Professor of Law and Faculty Director of the
Berkeley Center for Law & Technology*
U.C. Berkeley School of Law
Berkeley, California

ROBERT P. MERGES

*Wilson Sonsini Goodrich & Rosati Professor
of Law and Technology and Faculty Director of
the Berkeley Center for Law & Technology*
U.C. Berkeley School of Law
Berkeley, California

REGIS MCKENNA

Chairman and CEO
Regis McKenna, Inc.
Palo Alto, California

DEIRDRE K. MULLIGAN

*Clinical Professor and Faculty Director of the
Berkeley Center for Law and Technology*
U.C. Berkeley School of Information
Berkeley, California

JAMES POOLEY

*Deputy Director General of the
World Intellectual Property Organization*
Washington, D.C.

MATTHEW D. POWERS

Tensegrity Law Group, LLP
Redwood Shores, California

PAMELA SAMUELSON

*Professor of Law & Information
and Faculty Director of the
Berkeley Center for Law & Technology*
U.C. Berkeley School of Law
Berkeley, California

LIONEL S. SOBEL

*Professor of Law and Director of the
International Entertainment & Media Law
Summer Program in London, England*
Southwestern University School of Law
Los Angeles, California

LARRY W. SONSINI

Wilson Sonsini Goodrich & Rosati
Palo Alto, California

MICHAEL STERN

Cooley LLP
Palo Alto, California

MICHAEL TRAYNOR

Cobalt LLP
Berkeley, California

THOMAS F. VILLENEUVE

Gunderson Dettmer Stough Villeneuve
Franklin & Hachigian LLP
Redwood City, California

BERKELEY CENTER FOR LAW & TECHNOLOGY 2012–2013

Executive Director

ROBERT BARR

Faculty Directors

KENNETH BAMBERGER

PETER MENELL

ROBERT MERGES

DEIRDRE MULLIGAN

PAMELA SAMUELSON

PAUL SCHWARTZ

SUZANNE SCOTCHMER

MOLLY VAN HOUWELING

Associate Director

LOUISE LEE

Affiliated Faculty and Scholars

AARON EDLIN

JOSEPH FARRELL

RICHARD GILBERT

BRONWYN HALL

THOMAS JORDE

MICHAEL KATZ

DAVID MOWERY

DAVID NIMMER

DANIEL RUBINFELD

ANNALEE SAXENIAN

JASON SCHULTZ

HOWARD SHELANSKI

CARL SHAPIRO

MARJORIE SHULTZ

LON SOBEL

TALHA SYED

DAVID TEECE

JENNIFER M. URBAN

HAL R. VARIAN

DAVID WINICKOFF

DO BAD THINGS HAPPEN WHEN WORKS ENTER THE PUBLIC DOMAIN?: EMPIRICAL TESTS OF COPYRIGHT TERM EXTENSION

Christopher Buccafusco[†] & Paul J. Heald^{††}

ABSTRACT

According to the current copyright statute, copyrighted works of music, film, and literature will begin to transition into the public domain in 2018. While this will prove a boon for users and creators, it could be disastrous for the owners of these valuable copyrights. Therefore, the next few years will likely witness another round of aggressive lobbying by the film, music, and publishing industries to extend the terms of already-existing works. These industries, and a number of prominent scholars, claim that when works enter the public domain, bad things will happen to them. They worry that works in the public domain will be underused, overused, or tarnished in ways that will undermine the works' economic and cultural value. Although the validity of their assertions turns on empirically testable hypotheses, very little effort has been made to study them.

This Article attempts to fill that gap by studying the market for audiobook recordings of bestselling novels, a multi-million dollar industry. Data from this study, which includes a novel human-subjects experiment, suggest that term-extension proponents' claims about the public domain are suspect. Audiobooks made from public domain bestsellers (1913–22) are significantly more available than those made from copyrighted bestsellers (1923–32). In addition, the experimental evidence suggests that professionally made recordings of public domain and copyrighted books are of similar quality. Finally, while a low quality recording seems to lower a listener's valuation of the underlying work, the data do not suggest any correlation between that valuation and the legal status of the underlying work. Accordingly, this research indicates that the significant costs of additional copyright protection for already-existing works may not be justified. These findings will be relevant to the inevitable congressional and judicial debate over copyright term extension in the next few years.

© 2013 Christopher Buccafusco & Paul J. Heald.

[†] Assistant Professor, Chicago-Kent College of Law; Co-Director, Center for Empirical Studies of Intellectual Property.

^{††} Professor of Law, University of Illinois, Urbana-Champaign; Professorial Fellow, Centre for Intellectual Property Policy and Management, Bournemouth University (UK). We wish to thank Emily Barney, Steven Benethen, Jessica Bregant, Kacy King, Bin Li, Megan Nolan, Tyler Slack, and Alex Wilgus for their assistance with researching, recording, designing, and analyzing these studies. We are also grateful to the following people for comments on earlier versions of this Article: Richard Watt and participants at the Canterbury University (Christchurch, N.Z.) Department of Economics Workshop; Suzy Frankel and the University of Wellington (Wellington, N.Z.) School of Law colloquium series; participants at the Conference on Empirical Legal Studies; and participants at the annual meeting of the Society for Economic Research on Copyright Issues.

TABLE OF CONTENTS

I.	INTRODUCTION	2
II.	THE COPYRIGHT TERM EXTENSION DEBATE	6
A.	THE UNITED STATES: THE 1998 SONNY BONO COPYRIGHT TERM EXTENSION ACT AND LOOKING AHEAD TO 2018	6
B.	INTERNATIONAL LOBBYING EFFORTS.....	10
C.	ECONOMIC JUSTIFICATIONS OF TERM EXTENSION: TESTABLE HYPOTHESES.....	12
1.	<i>The Underuse Hypothesis</i>	13
2.	<i>The Overuse Hypothesis</i>	15
3.	<i>The Tarnishment Hypothesis</i>	16
III.	EMPIRICALLY TESTING THE ECONOMIC ASSUMPTIONS: THE CASE OF AUDIOBOOKS	17
A.	EXISTING EMPIRICAL STUDIES OF COPYRIGHT TERM EXTENSION AND THE PUBLIC DOMAIN	18
B.	STUDY 1: THE EXPLOITATION OF POPULAR FICTION IN AUDIOBOOKS	19
1.	<i>Methods: Study 1</i>	21
2.	<i>Results: Study 1</i>	22
C.	STUDY 2: AUDIOBOOK QUALITY AND TARNISHMENT	23
1.	<i>Methods: Study 2</i>	24
2.	<i>Results: Study 2</i>	26
D.	DATA LIMITATIONS.....	28
IV.	IMPLICATIONS FOR IP LAW AND POLICY: THE NEXT TIME DISNEY COMES KNOCKING	29
A.	ADDRESSING THE UNDERUSE HYPOTHESIS	29
B.	ADDRESSING THE OVERUSE HYPOTHESIS	30
C.	ADDRESSING THE TARNISHMENT HYPOTHESIS	33
D.	TARNISHMENT BEYOND TERM EXTENSION	36
V.	CONCLUSION	36
	APPENDIX A: FULL SAMPLE OF BESTSELLING NOVELS, 1913–1932	38
	APPENDIX B: DURABLE FICTION BESTSELLERS, 1913–1932	43

I. INTRODUCTION

In 2018, for the first time in two decades, copyrighted works of art, music, film, and literature are scheduled to enter the public domain. This

promises to be a huge boon to both the public, who will be able to access these works freely, and to creative artists, who wish to perform, adapt, copy, or otherwise make use of them. Of course, to the owners of some of these copyrighted works, their transition into the public domain means the loss of millions of dollars of revenue.¹ Book publishers and movie studios will face a world where their creations are available for unauthorized copying and adaptation by anyone.² It seems inevitable that, just as they did in the 1990s, the copyright industries will engage in another round of congressional lobbying to extend the term of protection for an additional period.

The standard justification for intellectual property (“IP”) protection is that the exclusive rights of copyright law provide incentives for creators to invest in creating new works.³ Theoretically, without IP protection, creators would not be able to recoup the costs of their investment if their creations could be freely copied. The primary argument in favor of extending the copyright term for future works is based on this incentive-to-create rationale: a longer term means that the author will be able to generate more money from her work, thereby increasing the ex ante incentive to create the work in the first place.⁴

The incentive-to-create rationale fails entirely, however, in the case of extending the copyright term for already existing books, music, and movies. The extension of protection for *The Sun Also Rises* does not increase the incentives for Hemingway to produce more or better work.⁵ He is, after all, dead.⁶ Accordingly, proponents of term extension have had to offer other

1. The brief for the Petitioners in *Eldred v. Ashcroft* estimated these losses and noted that “because of CTEA [Copyright Term Extension Act], the public will . . . have to pay an additional \$317 million annually in royalties.” Brief for Petitioner, *Eldred v. Ashcroft*, 537 U.S. 186 (2003) (No. 01-618), 2002 WL 1041928, at *7.

2. Trademark law will provide Disney some relief against unauthorized uses, such as third-party production of a Mickey Mouse doll, that are likely to confuse consumers as to the source of goods or services. *See* 15 U.S.C. § 1125(a) (2011).

3. WILLIAM M. LANDES & RICHARD A. POSNER, *ECONOMIC STRUCTURE OF INTELLECTUAL PROPERTY LAW* 1 (2003).

4. *Id.* at 212.

5. Some argue that a potential author today seeing an extension of Hemingway’s copyright will perceive a signal that Congress will give the potential author’s works similarly gracious treatment in the future, thereby stimulating the potential author to produce more now. With the present copyright term already at life-of-the-author plus seventy years, the “added incentive” argument has not been taken very seriously. *See* Lawrence B. Solum, *Congress’s Power to Promote the Progress of Science: Eldred v. Ashcroft*, 36 *LOY. L.A. L. REV.* 1, 77 (2002).

6. *Hemingway Dead of Shotgun Wound; Wife Says He Was Cleaning Weapon*, *N.Y. TIMES*, July 3, 1961, at 1, available at <http://www.nytimes.com/books/99/07/04/specials/hemingway-obit.html>.

reasons why longer copyrights will increase social welfare. During the adoption of the last copyright term extension legislation and the litigation surrounding it, the copyright industries and some leading scholars asserted three justifications for increasing the term of protection for already existing works.⁷

First, proponents of term extension have argued that, without additional protection, the publishing industries will not have sufficient incentives to preserve, protect, and commercialize old works. They claim that without the protections that copyright provides, works that fall into the public domain will be underused.⁸ This is a version of the classic “public goods” problem in economics that asserts some intangible goods will be under-provided if their producers lack the power to exclude unauthorized users.⁹ Second, and in some ways the inverse of the first argument, proponents of term extension claim that freely available works will be overused, thereby undermining the works’ economic and cultural value.¹⁰ This is a version of the “tragedy of the commons”¹¹ (e.g., once anybody can use “Rhapsody in Blue” in a movie or a commercial, the song will be overused and lose its appeal). The proponents’ third argument claims that uncontrolled uses of culturally valuable works will tarnish or debase those works, because poorly made or “inappropriate” versions of the works will affect the public’s judgments about the works’ quality and meaning and therefore their underlying value.¹² Audiences who see a substandard production of Eugene O’Neill’s *The Iceman Cometh* performed by the Evans Elementary School Drama Club may not wish to read the play or see another performance of it afterward, and thereby never fully grasp the play’s treatment of anarchy and socialism. As with the incentive-to-create rationale for new works, these three justifications for extending the term of protection for already existing works have a theoretical appeal. The important question, however, is whether these justifications stand up to empirical scrutiny. This Article attempts to answer that question.

In recent years, legal scholars have turned increasingly to empirical and experimental methods to test longstanding assumptions about how laws operate. These methods have been particularly successful when applied to IP

7. See *infra* notes 25–31.

8. See *infra* notes 25–31.

9. Theodore Groves & John Ledyard, *Optimal Allocation of Public Goods: A Solution to the “Free Rider” Problem*, 45 *ECONOMETRICA* 783 (1977).

10. See *infra* notes 25–31.

11. See Garrett Hardin, *The Tragedy of the Commons*, 162 *SCI.* 1243 (1968).

12. See Justin Hughes, “Recoding” *Intellectual Property and Overlooked Audience Interests*, 77 *TEX. L. REV.* 923, 926 (1999) (“[N]on-owners commonly benefit from owner control that is used to keep a cultural object ‘stable.’”).

law, because, unlike some areas of the law, IP law's assumptions about markets, incentives, and human behavior are explicit.¹³ This Article continues the Authors' previous research applying empirical and experimental methods to IP issues. It reports data from two studies that test the validity of proponents' arguments for extending the copyright term. In short, *our study finds almost no evidence to support the claims made in favor of copyright term extension.*

Part II describes the debate over copyright term extension and the rationales in favor of it. This Part examines how these rationales affected the last term-extension act and the litigation following it, and how such rationales will likely come up again in renewed calls for extension. Part III reports on empirical tests of the extension rationales. These tests rely on an interesting and understudied creative industry: the market for audiobook recordings of novels. Audiobooks are "derivative works" within the definition of copyright law,¹⁴ and they present a number of opportunities for studying claims about the exploitation and commercialization of works. Our data compare the markets for audiobook recordings of popular novels on either side of the public domain divide: the decade of public domain works from 1913 to 1922 and the decade of copyrighted works from 1923 to 1932. Part IV applies the empirical findings to the debate about copyright term extension. Although

13. See Christopher Buccafusco & Christopher Jon Sprigman, *The Creativity Effect*, 78 U. CHI. L. REV. 31 (2011) [hereinafter Buccafusco & Sprigman, *Creativity Effect*]; Christopher Buccafusco & Christopher Sprigman, *Valuing Intellectual Property: An Experiment*, 96 CORNELL L. REV. 1 (2010); Deborah R. Gerhardt, *Copyright Publication: An Empirical Study*, 87 NOTRE DAME L. REV. 135 (2011); Paul J. Heald, *Property Rights and the Efficient Exploitation of Copyrighted Works: An Empirical Analysis of Public Domain and Copyrighted Fiction Bestsellers*, 92 MINN. L. REV. 1031, 1046–50 (2008) [hereinafter Heald, *Fiction Bestsellers*]; Paul J. Heald, *Does the Song Remain the Same? An Empirical Study of Bestselling Musical Compositions (1913–32) and Their Use in Cinema (1968–2007)*, 60 CASE. W. RES. L. REV. 1 (2009) [hereinafter Heald, *Musical Compositions*] (songs are just as likely to be used in films after they fall into the public domain); Paul J. Heald & Robert Brauneis, *The Myth of Buick Aspirin: An Empirical Study of Trademark Dilution by Product and Trade Names*, 32 CARDOZO L. REV. 2533 (2011); Raymond Shih Ray Ku et al., *Does Copyright Law Promote Creativity? An Empirical Analysis of Copyright's Bounty*, 62 VAND. L. REV. 1669 (2009); Thomas R. Lee et al., *An Empirical and Consumer Psychology Analysis of Trademark Distinctiveness*, 41 ARIZ. ST. L.J. 1033 (2009).

14. The Copyright Act defines a derivative work as follows:

[A] work based upon one or more preexisting works, such as a translation, musical arrangement, dramatization, fictionalization, motion picture version, sound recording, art reproduction, abridgment, condensation, or any other form in which a work may be recast, transformed, or adapted. A work consisting of editorial revisions, annotations, elaborations, or other modifications which, as a whole, represent an original work of authorship, is a "derivative work".

17 U.S.C. § 101 (2011).

this research is in no way conclusive on the issue, it strongly suggests that all three arguments in favor of copyright term extension are mistaken.

II. THE COPYRIGHT TERM EXTENSION DEBATE

The primary salience of the data analyzed in Part II relates to the ongoing and vociferous debate over the retroactive extension of copyright protection to existing creative works. Fifteen years ago powerful players in the copyright industries (primarily film, music, and book publishing) lobbied extensively to encourage Congress to pass legislation to prevent their works from falling into the public domain.¹⁵ Following the success of those efforts in the United States, the copyright industries have pushed for term extensions internationally.¹⁶ This Part briefly charts the history of the lobbying efforts in both the United States and abroad. It then presents the three primary economic justifications offered in favor of copyright term extension, all of which assert that bad things happen when works fall into the public domain. The data presented in Part III tend to refute the attempts prominent economists and the copyright industries have made to justify extending the term of protection to existing works.

A. THE UNITED STATES: THE 1998 SONNY BONO COPYRIGHT TERM EXTENSION ACT AND LOOKING AHEAD TO 2018

The U.S. Constitution provides Congress with the power to “promote the Progress of Science and the Useful Arts, by securing for limited Times to Authors and Inventors the exclusive Right to their respective Writings and Discoveries.”¹⁷ In 1790, one year after the Constitution went into effect, Congress passed the first copyright statute, providing protection for maps, charts, and books.¹⁸ This first Act provided authors with a fourteen-year term of protection that could be renewed for an additional fourteen years.¹⁹ Since the eighteenth century, Congress has extended the copyright term for existing works several times. In 1831, Congress extended the initial term of protection to twenty-eight years with a fourteen-year renewal term,²⁰ and the 1909 Copyright Act extended the renewal term to twenty-eight years.²¹

15. *See infra* notes 17–26.

16. *See infra* notes 33–40.

17. U.S. CONST. art. I, § 8, cl. 8.

18. *See* Act of May 31, 1790, ch. 15, 1 Stat. 124 (repealed 1831).

19. *Id.*

20. *See* Act of Feb. 3, 1831, ch. 16, 4 Stat. 436 (repealed 1870).

21. *See* Act of Mar. 4, 1909, ch. 320, 35 Stat. 1075 (repealed 1976).

The last major revision of the copyright statute, the 1976 Act, further lengthened the period of copyright protection.²² For existing works that had not yet entered the public domain, the Act added forty-seven years of protection to the twenty-eight-year term resulting in a total of seventy-five years of protection. The Act, which went into effect in 1978, did not retroactively revive copyright protection for works that had already entered the public domain; consequentially, all works published prior to 1923 remain in the public domain. The oldest works still subject to copyright were those published in 1923, and their copyrights were set to expire at the end of 1998. The possibility of these valuable works falling into the public domain seemed disastrous to the copyright owners, who turned to Congress for another extension.

By the time Americans had begun to debate the merits of another copyright term extension, Congress had already passed legislation doing so. The 1998 Sonny Bono Copyright Term Extension Act (“CTEA”) added an additional twenty years of protection to the copyright term for all existing works.²³ Works created between 1923 and 1978 now receive ninety-five years of protection, while works created since 1978 would be protected for the duration of the lives of their authors plus seventy years, with anonymous works, pseudonymous works, and works made for hire receiving a defined ninety-five-year term of protection.²⁴

The intense, well-documented lobbying efforts of Disney²⁵ and other copyright owners²⁶ resulted in the passage by voice vote of the CTEA.²⁷ In

22. See Copyright Act of 1976, Pub. L. No. 94-553, 90 Stat. 2541.

23. See Sonny Bono Copyright Term Extension Act, Pub. L. No. 105-298, 112 Stat. 2827 (1998).

24. 17 U.S.C. §§ 302–304 (2011).

25. See Bill McAllister, *A Capital Way to Stop a Headache*, WASH. POST, Oct. 15, 1998, at A21 (“Hill staff members said that other Disney representatives, along with other movie industry representatives, had made strong pleas for a 20-year extension to all copyrights.”), available at <http://www.public.asu.edu/~dkarjala/commentary/WashPost10-15-98.html>.

26. See John L. Fialka, *Songwriters’ Heirs Mourn Copyright Loss*, WALL ST. J., Oct. 30, 1997, at B1, available at <http://news.google.com/newspapers?id=-5NGAAAIBAJ&sjid=DvgMAAAAIBAJ&dq=songwriters-heirs-mourn-copyright-loss&pg=5468%2C3903721>.

27. See Keith Pocaro, *Private Ordering and Orphan Works: Our Least Worst Hope?*, 2010 DUKE L. & TECH. REV. 15, 15 (2010) (“The current state of copyright law, with wildly longer term limits and automatic protection, is a result of continuous content-industry lobbying to protect their valuable, aging intellectual property.”); Alan K. Ota, *Disney in Washington: The Mouse That Roared*, CQ WEEKLY, Aug. 8, 1998, available at <http://www.cnn.com/ALLPOLITICS/1998/08/10/cq/disney.html> (describing the Disney lobbying strategy); Dennis Karjala, OPPOSING COPYRIGHT EXTENSION (Jan. 23, 2012), <http://homepages.law.asu.edu/~dkarjala/OpposingCopyrightExtension/> (collecting documents related to term extension efforts).

fact, Disney Chairman Michael Eisner lobbied Senate Majority Leader Trent Lott directly.²⁸ The bill sailed through both houses, with eighteen of twenty-five sponsors receiving Disney money, including Lott on the very day he signed up as a co-sponsor.²⁹ Opponents to term extension had very little opportunity to participate in discussion of the bill; according to Professor Dennis Karjala:

The hearings [on term extension] were combined with some other bills, so they were not publicized under the bill numbers for those trying to follow the legislation. The proponents of extension—surprise, surprise!—knew about the House hearings and of course testified in favor. The opponents did not even know the hearings took place until several months later!³⁰

With significant royalty streams at stake,³¹ copyright owners and the sponsors of their bill were taking no chances on a full-blown debate over the wisdom of extending the term of protection for valuable works that were about to fall into the public domain.

The failure of Congress to seriously consider arguments that term extension was a hidden tax on consumers, a drag on follow-on creators, and an unconstitutional failure to “promote the Progress of Science”³² suggests that any rationale offered in the legislative history of CTEA was merely

28. See *Disney Lobbying for Term Extension No Mickey Mouse Effort*, CHI. TRIB., Oct. 17, 1998, at 22.

29. See *id.*; see also WILLIAM M. LANDES & RICHARD A. POSNER, *THE POLITICAL ECONOMY OF INTELLECTUAL PROPERTY LAW* 16 (2004) (noting that the Center for Responsive Politics showed that in 1996 media interests donated \$1.5 million to six of the sponsors of the Copyright Term Extension Act); John Solomon, *Rhapsody in Green*, BOSTON GLOBE, Jan. 3, 1999, at E2. John Solomon wrote:

Behind the scenes, however, [Disney] has been active. Congressional Quarterly reported that Disney chairman Michael Eisner personally lobbied Senate Majority Leader Trent Lott, a Republican from Mississippi. That day, according to the Center for Responsive Politics, Disney gave Lott a \$1,000 contribution, following up two weeks later with a \$20,000 donation to the National Republican Senatorial Committee.

Id.

30. Dennis Karjala, *About Copyright Term Extension*, OPPOSING COPYRIGHT EXTENSION, <http://homepages.law.asu.edu/~dkarjala/OpposingCopyrightExtension/what.html> (last visited Oct. 29, 2012).

31. See Marvin Ammori, *The Uneasy Case for Copyright Extension*, 16 HARV. J.L. & TECH. 287, 292 (2002) (noting that Disney in particular stood to lose control of billions of dollars' worth of copyrights—Mickey Mouse and Winnie-the-Pooh alone were valued at nearly \$8 billion dollars in revenue each—if the CTEA was not passed).

32. U.S. CONST. art. I, § 8, cl. 8.

make-weight.³³ Nonetheless, the House Report stated that retroactive extension would “provide copyright owners generally with the incentive to restore older works and further disseminate them to the public.”³⁴ In the brief debate over the legislation, Representative Howard Coble adopted this rationale and stated that “[w]hen works are protected by copyright, they attract investors who can exploit the work for profit.”³⁵ Bruce Lehman, former Commissioner of Patents and Trademarks, put the case most strongly in his statement before Congress:

[T]here is ample evidence that shows that once a work falls into the public domain it is neither cheaper nor more widely available than most works protected by copyright. One reason quality copies of public domain works are not widely available may be because publishers will not publish a work that is in the public domain for fear that they will not be able to recoup their investment or earn enough profit.³⁶

Whether worries over the lack of availability of older works actually motivated Congress or not, the Supreme Court picked up on the argument in the failed constitutional challenge to the CTEA in *Eldred v. Ashcroft*.³⁷ The Court found that Congress “rationally credited projections that longer terms would encourage copyright holders to invest in . . . public distribution of their works.”³⁸ The *Eldred* litigation forced copyright owners to articulate neutral, public interest rationales to justify retroactively protecting copyrights in existing works. The primary arguments defending term extension enlarged upon the brief statements in the legislative history—that works would be less available to the public if they fell into the public domain.

The lobbying effort for term extension in the late 1990s began as an ordinary—and wildly successful—plea to Congress to maintain the flow of various copyright-fueled income streams without serious consideration of issues involving the public domain. The debate that peaked in *Eldred* five

33. See generally Dennis Karjala, *Value of the Public Domain*, OPPOSING COPYRIGHT EXTENSION, <http://homepages.law.asu.edu/~dkarjala/OpposingCopyrightExtension/publicdomain.html> (last visited Nov. 9, 2012).

34. H.R. REP. NO. 105-452, at 4 (1998).

35. 144 CONG. REC. H1458 (daily ed. Mar. 25, 1998) (statement of Rep. Howard Coble).

36. *Copyright Term, Film Labeling, and Film Preservation Legislation: Hearing on H.R. 989, H.R. 1248, and H.R. 1734 Before the Subcomm. on Courts and Intellectual Property of the H. Comm. on the Judiciary*, 104th Cong. 217–18 (1995) (statement of Bruce Lehman, Assistant Secretary of Commerce and Commissioner of Patents and Trademarks).

37. *Eldred v. Ashcroft*, 537 U.S. 186 (2003).

38. See *id.* at 207; see also Lawrence B. Solum, *The Future of Copyright*, 83 TEX. L. REV. 1137, 1165–68 (2005).

years later had evolved into a full frontal assault on the public domain by copyright owners. In need of a public interest rationale to defend their monetary objectives, rights holders argued that myriad bad things would happen if works were allowed to fall into the public domain,³⁹ and thus they asserted term extension was necessary to protect the public interest.

Because the present copyright term extension expires in 2018, Congress will soon decide whether to acquiesce to the next round of lobbying by copyright owners.⁴⁰ In the meantime, other jurisdictions are actively considering U.S.-style term extension. With significant royalty streams at stake in other jurisdictions, the pro-extension lobbying effort has gone global, with mixed success.

B. INTERNATIONAL LOBBYING EFFORTS

U.S. copyright owners, whose interests are well represented by U.S. trade negotiators, have also poured considerable effort and money into securing term extensions in other countries. These copyright owners have already been successful in imposing term extension on Australia as part of the Australia-U.S. Free Trade Agreement.⁴¹ Japan⁴² is currently under similarly intense pressure, as are Jamaica⁴³ and other developing countries.⁴⁴ The

39. See, e.g., Scott M. Martin, *The Mythology of The Public Domain: Exploring The Myths Behind Attacks on the Duration of Copyright Protection*, 36 LOY. L.A. L. REV. 253 (2002).

40. Joseph P. Liu has already looked ahead to 2018 in his recent article. See Joseph P. Liu, *The New Public Domain* (Sept. 12, 2011) (unpublished manuscript), available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1926381.

41. See Matthew Rimmer, *Robbery Under Arms: Copyright Law and the Australia-United States Free Trade Agreement*, 11 FIRST MONDAY, NO. 3 (Mar. 6, 2006), <http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/1316/1236> (“In the trade negotiations, [the U.S. Trade Representative] demanded that Australia ratify the World Intellectual Property Organization (WIPO) Copyright Treaty and Performances and Phonograms Treaty. He supported an extension of the copyright term, so that Australia adopted the standards set by the Sonny Bono Copyright Term Extension Act.”) (emphasis removed); see also Maree Sainsbury, *Governance and the Process of Law Reform: The Copyright Term Extension in Australia*, 9 CANBERRA L. REV. 1 (2006) (detailing lobbying effort in Australia to ratify the Free Trade Agreement).

42. See Mike Masnick, *Copyright Extension Moves to Japan*, TECHDIRT (Nov. 20, 2009, 7:15 AM), <http://www.techdirt.com/articles/20091119/1840217016.shtml>; CPB Netherlands Bureau for Economic Policy Analysis, *Copyright Protection; Not More But Different* (The Hague, Working Paper No. 122, 2000), available at <http://www.cpb.nl/en/publication/copyright-protection-not-more-different> (describing an industry “call for additional copyright legislation and enforcement” in the Netherlands).

43. See Mike Masnick, *Jamaica the Latest to Embrace Retroactive Term Extension and Screw the Public Domain*, TECHDIRT (Oct. 21, 2011, 4:59 AM), <http://www.techdirt.com/articles/20111014/00471816347/jamaica-latest-to-embrace-retroactive-copyright-term-extension-screw-public-domain.shtml>.

European Union recently acceded to retroactive extension for sound recordings,⁴⁵ as has Argentina.⁴⁶ A leaked first draft of the proposed Transpacific Partnership between New Zealand, Japan, and Canada would require retroactive extension for all copyrighted works.⁴⁷ But other jurisdictions have not been so easy to convince. Although pressure is constant from the copyright lobby, both the United Kingdom⁴⁸ and Japan⁴⁹ have refused to extend the term of protection for existing works other than sound recordings. One major political party in Brazil has even proposed a reduction in the copyright term.⁵⁰

The United Kingdom seems particularly reluctant to adopt the proposed changes in the absence of supporting empirical data. In fact, the recent government report by Ian Hargreaves urges that

the IP System [be] driven as far as possible by objective evidence. Policy should balance measurable economic objectives against social goals and potential benefits for rights holders against impacts on consumers and other interests. These concerns will be of

44. See Andrew Rens & Lawrence Lessig, *Forever Minus A Day: A Consideration Of Copyright Term Extension In South Africa*, 7 S. AFR. J. INFO. & COMM. 22 (2006); *Mexico—Copyright Law Amended*, LADAS & PARRY LLP (Mar. 21, 2004), http://www.ladas.com/BULLETINS/2004/0304Bulletin/Mexico_CopyrightLaw.html.

45. See Martin Kretschmer, *Creativity Stifled? A Joined Academic Statement on the Proposed Copyright Term Extension for Sound Recordings*, 9 EUR. INTEL. PROP. REV. 314 (2008) (statement of sixty-one law professors opposing extension).

46. Mike Masnick, *Here We Go Again: Argentina Extends Copyright*, TECHDIRT (Dec. 22, 2009, 5:27 AM), <http://www.techdirt.com/articles/20091221/1756577455.shtml>.

47. See Michael Geist, *TPP Copyright Extension Would Keep Some of Canada's Top Authors Out of Public Domain For Decades*, MICHAEL GEIST'S BLOG (Jan. 9, 2012), <http://www.michaelgeist.ca/content/view/6226/125/>.

48. See ANDREW GOWERS, GOWERS REVIEW OF INTELLECTUAL PROPERTY 56–57 (2006), available at http://www.hm-treasury.gov.uk/media/6/E/pbr06_gowers_report_755.pdf (study commissioned by the British Treasury department rejecting ex post justifications for extending copyright protection for existing works).

49. See Mike Masnick, *Copyright Extension Moves to Japan*, TECHDIRT (Nov. 29, 2009, 7:15 AM), <http://www.techdirt.com/articles/20091119/1840217016.shtml> (reporting on efforts to extend copyright in Japan); Andreas Bovens, *Japan to Extend Posthumous Copyright Term to 70 Years?*, CHOSAQ, (July 24, 2006), <http://chosaq.net/archives/2006/07/japan-to-extend-posthumous-copyright-term-to-70-years.html>.

50. *No National Leeway? Copyright Reform Proposals in Brazil and the Czech Republic*, GOVERNANCE ACROSS BORDERS (Sept. 3, 2010) <http://governancexborders.com/2010/09/03/no-national-leeway-copyright-reform-proposals-in-brazil-and-the-czech-republic/#more-1095>.

particular importance in assessing future claims to extend rights or in determining desirable limits to rights.⁵¹

Consistent with the Hargreaves approach, the earlier commissioned *Gowers Review of Intellectual Property* examined existing empirical evidence and rejected arguments that retroactive term extension was necessary.⁵² Although the United Kingdom had no choice but to accede to the new E.U. directive retroactively extending protection to sound recordings,⁵³ the level of skepticism from U.K. officials was significant.⁵⁴

The debate over the economic wisdom of term extension around the world turns on the validity of the same factual assumptions asserted to justify term extension in the United States.⁵⁵ Before explaining how our data bear on the validity of those assumptions, we provide a fuller account of the pro-extension arguments below.

C. ECONOMIC JUSTIFICATIONS OF TERM EXTENSION: TESTABLE HYPOTHESES

Jack Valenti, the President of the Motion Picture Association of America, once testified derisively to Congress that public domain works were “orphan[s],”⁵⁶ meaning that without parents (owners) they would be subject

51. IAN HARGREAVES, *DIGITAL OPPORTUNITY: A REVIEW OF INTELLECTUAL PROPERTY AND GROWTH 20* (2011), available at <http://www.ipo.gov.uk/ipreview-finalreport.pdf>.

52. See GOWERS, *supra* note 48, at 56–57.

53. See Directive 2011/77/EU of the European Parliament and of the Council of 27 September 2011 amending Directive 2006/116/EC on the term of protection of copyright and certain related rights, 2011 O.J. (L 265) 1, available at <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2011:265:0001:0005:en:pdf>.

54. See *Commission Staff Working Document: Impact Assessment on the Legal and Economic Situation of Performers and Record Producers in the European Union*, COM (2008) 464 final (Apr. 23, 2008), available at http://ec.europa.eu/internal_market/copyright/docs/term/ia_term_en.pdf (analyzing EU proposal to extend copyright term in sound recordings from fifty to ninety-five years); Eric Bangeman, *U.K. Government Resists Music Industry Pressure, Caps Copyrights at 50 Years*, ARS TECHNICA (July 24, 2007, 8:59 AM), <http://arstechnica.com/tech-policy/news/2007/07/uk-government-resists-music-industry-pressure-caps-copyrights-at-50-years.ars>.

55. See Laura Bradford, *A Closer Look at the Public Domain*, 13 GREEN BAG 343, 344–45 (2010) (“Currently a debate exists globally about the scope of protections for IP. . . . Proponents of the current strong rules protecting intellectual property argue that a failure to reward innovation curtails investment.”); Kretschmer, *supra* note 45.

56. See *Copyright Term, Film Labeling, and Film Preservation Legislation: Hearing on Copyright Term Extension, H.R. 989 Before the Subcomm. on Courts and Intellectual Property of the H. Comm. on the Judiciary*, 104th Cong. 55 (1995) (statement of Jack Valenti, President and CEO, Motion Picture Association of America).

to distressing abuse.⁵⁷ Sophisticated commentators in support of copyright term extension have offered more detailed and theory-driven arguments in support of their position. These arguments fall into three categories: underuse, overuse, and tarnishment. All three primary arguments rely on factual assertions about what happens when works fall into the public domain. Our study of the market for audiobooks, discussed in Part III, *infra*, tests all three assertions.

1. *The Underuse Hypothesis*

The most prominent justification for term extension asserts that works become less available to consumers when they fall into the public domain. In their influential article arguing for indefinitely renewable copyright for valuable works, law and economics scholars William Landes and Richard Posner reason that “an absence of copyright protection for intangible works may lead to inefficiencies because of impaired incentives to invest in maintaining and exploiting these works.”⁵⁸ Landes and Posner’s argument is a version of the classic “public goods” problem in economics. IP is expensive to create, but once it has been created, it can be cheaply copied and used by others. Because creators of IP cannot easily exclude others from using it, theory implies that they will not be able to recoup their investment costs and will never engage in creating the work in the first place. Thus, the law has to step in to create legal boundaries allowing creators the chance to recover their investments.⁵⁹

This argument can be applied not only to new works but to already created works as well. Some works require costly investments to maintain, produce, and distribute them over time. For example, when audio formats changed, someone had to spend money to transfer recordings on old vinyl disks to a digital format or the old music would not be accessible to most listeners. In theory, because those who would invest resources in the conversion cannot prevent others from free riding on their efforts, they will

57. On the role of metaphors in copyright law, see WILLIAM PATRY, *MORAL PANICS AND THE COPYRIGHT WARS* 44 (2009). Regarding orphan works, Patry writes: “Use of the term ‘orphan’ inaccurately conjures up an emotional need to protect these works against those who would use them with the copyright owner’s permission, even though the ‘parents’ long ago dropped any interest in them.” *Id.* at 77.

58. William M. Landes & Richard A. Posner, *Indefinitely Renewable Copyright*, 70 U. CHI. L. REV. 471, 475 (2003); *see also* LIOR ZEMER, *THE IDEA OF AUTHORSHIP IN COPYRIGHT* (2007) (arguing for indefinitely renewable copyright based on five-year renewal terms).

59. Another commentator explains: “If [works enter] the public domain, they [become] obscure and thus no one [will] invest in them due to the problem of free riding. Items which retain enough value for future use should be given indefinite copyrights to maintain their value.” Miriam Bitton, *Modernizing Copyright Law*, 20 TEX. INTELL. PROP. L.J. 65, 77 (2011).

not be able to recoup their investment and, thus, will never bother to invest in the first place. Although this appears not to be the case with digital recordings,⁶⁰ without a method for recouping the cost of conversion, preservation, or reproduction, the underuse hypothesis maintains that commercializers will have inadequate incentives to continue production and distribution of older works. This was the primary worry that Congress expressed when passing CTEA in 1998.⁶¹

Professor Arthur Miller adds a related concern about the underuse of copyrighted works. He worries that new works deriving from and based on materials in the public domain would be under-produced.⁶² Copyright law gives owners the exclusive right to make or license derivative works like adaptations, sequels, and translations that are based on the original work.⁶³ Miller argues that these derivative works will not be made without longer copyright terms.⁶⁴ He reasons that “you have to provide incentives for [producers] to produce the derivatives, the motion picture, the TV series, the documentary, whatever it may be—perhaps even a musical! . . . We must incentivize the dissemination industries, the preservation industries, and the derivative work industries.”⁶⁵ According to Miller’s argument, without the ability to prevent copiers, no one will be willing to invest the resources in creating a musical version of *A Passage to India*, because, if it proved successful, others would be able to prepare their own musicals of the book. These competing versions would drive down the value of the first musical, thereby undermining the incentives to create it in the first place. A staunch advocate of term extension, Miller believes that works need owners in order to be adequately exploited in derivative forms.⁶⁶

60. See *infra* note 85 and accompanying text.

61. See H.R. REP. NO. 105-452, at 4 (1998).

62. Symposium, *The Constitutionality of Copyright Term Extension: How Long is Too Long?*, 18 CARDOZO ARTS & ENT. L.J. 651, 693 (2000) (panel comments of Arthur Miller).

63. 17 U.S.C. § 106(2) (2011).

64. *The Constitutionality of Copyright Term Extension: How Long is Too Long?*, *supra* note 62, at 693.

65. *Id.*; cf. Lee Anne Fennell, *Common Interest Tragedies*, 98 NW. U. L. REV. 907, 919 (2004) (“The tendency towards overgrazing could thus reinforce one towards underinvestment, leading to a commons featuring too few, and too intensively exploited, intellectual products—at least in the absence of legal rules or norms designed to cabin these tendencies.”).

66. See *The Constitutionality of Copyright Term Extension: How Long is Too Long?*, *supra* note 62, at 692–94.

2. *The Overuse Hypothesis*

The “tragedy of the commons,” whereby common ownership leads to the degradation of a shared resource, forms the basis of the second primary theoretical justification for copyright term extension.⁶⁷ The tragedy of the commons can occur when a group of people collectively owns some resource, like a pasture. Each person has the incentive to maximize his use of the pasture before others can do so. This leads to overuse and depletion of the pasture through overgrazing. Similarly, if no one has the exclusive right to a creative work, then it might be overused (imagine dozens of advertisers all using the same song).⁶⁸ In such situations, the typical economic solution is to assign individual ownership of the resource so that a single control structure can efficiently manage use.⁶⁹

Landes and Posner make the tragedy of the commons analogy to copyright term extension explicit: “a novel or a movie or a comic book character or a piece of music or a painting” could be depleted like “unlimited drilling from a common pool of oil or gas would deplete the pool prematurely.”⁷⁰ Similarly, Stan Liebowitz and Stephan Margolis conclude:

Firms producing copies or derivatives of creative works after the copyright expires may be in the position of fisherman [sic] on an open access lake. They produce at their own private optima, not taking into account the effects that they have on other producers. Ownership can effectively manage these interactions, and copyright provides that ownership.⁷¹

67. See Hardin, *supra* note 11.

68. At least one commentator asserts that this was the fate of the classic film *It's a Wonderful Life* before it was rescued from the public domain. See Scott M. Martin, *The Mythology of the Public Domain: Exploring the Myths Behind Attacks on the Duration of Copyright Protection*, 36 LOY. L.A. L. REV. 253, 274–75 (2002). Martin explained:

By the 1980s, there were multiple versions of [*It's a Wonderful Life*], all in horrid condition. The film was “often sliced and diced by local stations who stuffed it with commercials.” There was no quality control over home video copies of the film—consumers had no way of knowing whether the tape they were purchasing was a poor quality bootleg version (which most were).

Id. (citations omitted).

69. See Michael J. Madison, Brett M. Frischmann & Katherine J. Strandburg, *Constructing Commons in the Cultural Environment*, 95 CORNELL L. REV. 657 (2010).

70. Landes & Posner, *supra* note 58, at 487.

71. Stan J. Liebowitz & Stephen Margolis, *Seventeen Famous Economists Weigh in on Copyright: The Role of Theory, Empirics, and Network Effects*, 18 HARV. J.L. & TECH. 435, 451 (2005).

In other words, without owners to police the frequency with which a work is used, others may wear it out and reduce its value.

The overuse hypothesis rests on the assumption that the value of creative works, like the value of a pasture, is finite and exhaustible.⁷² Each work has an optimal level at which it should be exploited and each use beyond that level decreases the work's value to others. While an individual owner of the copyright has the incentive to maintain the value of a work over time by preventing it from being overused, once the work falls into the public domain, others will rush to exploit the work's value immediately.⁷³ According to this theory a creative work such as a song has increasing social and economic value up to a certain number of uses in a given time period (e.g., in commercials during a year). Once that usage level is met, however, its value diminishes. Individual copyright owners are incentivized to exploit their works at the socially optimal maximum, but if works fall into the public domain, others will overuse the works and diminish their value.⁷⁴

3. *The Tarnishment Hypothesis*

The third rationale for extending copyright protection to already existing works is based on the fear that creative works will lose their value not through overuse but through misuse or tarnishment. A number of commentators have expressed concern that inappropriate uses of works will debase them and reduce their value.⁷⁵ Karjala, a leading opponent of term

72. The overuse hypothesis also assumes that people will exploit the resource in such a way that its value will be diminished. But considerable social science evidence, including from the field of behavioral game theory, demonstrates that this kind of overexploitation does not always take place. *See, e.g.*, ELINOR OSTROM, GOVERNING THE COMMONS: THE EVOLUTION OF INSTITUTIONS FOR COLLECTIVE ACTION (James E. Alt & Douglass C. North eds., 1990).

73. Of course, the assumption that creative works have finite and exhaustible value is itself open to empirical testing and may, in fact, be false. Psychological studies suggest that repeated exposure to things may actually increase their attractiveness. *See* Robert B. Zajonc, *Attitudinal Effects Of Mere Exposure*, 9 J. PERSONALITY & SOC. PSYCHOL. MONOGRAPH SUPPLEMENT 1, 23 (1968).

74. *See* Harold Demsetz, *Toward a Theory of Property Rights*, 57 AM. ECON. REV. 347, 355 (1967).

75. *See* Michael Steven Green, *Copyrighting Facts*, 78 IND. L.J. 919, 925 (2003). Green argues:

In addition to encouraging authors to create new works, copyrights also encourage authors to efficiently *utilize* constituents of works that already exist. For example, if no one had a property right in the character Superman, authors could freely create works in which Superman appeared as a character without concern for the effect their works had on the value of actual and potential Superman-based works.

extension, has coined a phrase to explain what is allegedly lacking when a work falls into the public domain: “proper husbandry by the copyright owner.”⁷⁶ The idea behind this hypothesis is that creative works can lose their value not just through overuse but also through the wrong kinds of uses. While the creation of some kinds of derivative works from an original work will be valuable and increase social welfare, other kinds of derivative works will actually decrease the value of the original and harm social welfare.

The most commonly expressed concern here involves the specter of unauthorized pornographic use that dots the literature on the subject.⁷⁷ As Karjala notes, “Rowling, Disney and other creative authors have at least some justification for being outraged when their characters are used in contexts wholly different from the original, such as pornography”⁷⁸ If viewers see a pornographic poster of *Harry Potter*, for example, they may tend to dislike and avoid the original movie. Presumably, though, other uses of the original work could harm it through the feedback effects of an audience’s reaction to the low quality derivative work as well. As noted above, poor quality productions of plays could undermine people’s sense of the value of the drama and its author. Or a poor movie version of a novel might reduce the public’s interest in the book. Hence, the asserted need for “proper husbandry,” and thus continued ownership of the work.

III. EMPIRICALLY TESTING THE ECONOMIC ASSUMPTIONS: THE CASE OF AUDIOBOOKS

Several years ago Professors Liebowitz and Margolis provided an invitation that the present study accepts:

There are, of course, many expensive derivative works that are based upon creations entirely in the public domain. The question is whether they are produced as regularly or as well as they would be

Id.; see also Alex Kozinski, *Mickey & Me*, 11 U. MIAMI ENT. & SPORTS L. REV. 465, 469 (1994) (arguing that unauthorized uses “end up diminishing the value of the product, not just to the creator, but to the general public as well”); Liebowitz & Margolis, *supra* note 71, at 449 (“Malicious or offensive derivative uses of some creative works might seriously diminish their value without a sufficient offset in the form of public benefit.”). *Cf.* Hughes, *supra* note 12, at 926 (arguing that “non-owners commonly benefit from owner control that is used to keep a cultural object ‘stable’”).

76. Dennis S. Karjala, *Harry Potter, Tanya Grotter, and the Copyright Derivative Work*, 38 ARIZ. ST. L.J. 17, 37 (2006).

77. See, e.g., Liebowitz & Margolis, *supra* note 71, at 449 n.24 (“The existence of a ‘Madeline Does Dallas’ might lead to some awkward questions during bedtime stories.”); Heald, *Musical Compositions*, *supra* note 13, at 25 (“The entire debate seems to turn on the effect of having unauthorized porn movies starring Mickey Mouse or Superman.”).

78. Karjala, *supra* note 76, at 36.

if they were protected by copyright. . . . This is an empirical question to which economists do not yet have the answer.⁷⁹

The Article uses the audiobook market to answer both questions about the quantity and quality of derivative works posed by these two prominent economists.

Audiobooks—audio recordings of fiction and nonfiction books—have become increasingly popular. Originally known as “books on tape,” sales of audiobooks have skyrocketed in recent years as technological changes in storage capacity, accessibility, and the ubiquity of smart phones have made listening to recorded versions of books incredibly convenient. The market for audiobooks is estimated to take in \$1 billion per year, and it is growing at over 10% per year.⁸⁰ This growth has been led by more than 300% growth in sales of downloaded audiobooks over a five-year period starting in 2005.⁸¹ Despite the significance of the audiobook market, no previous research has studied it with an eye towards IP law.

This Part reports two empirical studies of the audiobook market that test the economic assumptions supporting copyright term extension. Study 1 tests the underuse and overuse hypotheses by comparing the availability of audiobook recordings of popular fiction works from the decades on either side of the copyright-public domain divide. Study 2 applies a novel experimental technique to test the misuse hypothesis. Before describing those studies, Section III.A discusses some of the existing research that bears on these questions.

A. EXISTING EMPIRICAL STUDIES OF COPYRIGHT TERM EXTENSION AND THE PUBLIC DOMAIN

One of us has previously tracked the availability of 166 fiction bestsellers from 1913 to 1922 while they were still in copyright and after they fell out of copyright between 1988 and 1997.⁸² Heald measured the percentage of best sellers in print and the average number of publishers per work in a given year and found that, until 2001, public domain books were as available as their copyrighted counterparts.⁸³ After 2001, the percentage of in-print public domain bestsellers was significantly higher, as was the number of publishers

79. Liebowitz & Margolis, *supra* note 71, at 449.

80. See *Industry Data*, AUDIO PUBLISHERS ASS'N, <http://audiopub.org/resources-industry-data.asp> (last visited Oct. 15, 2012).

81. *Id.* The report also notes: “The CD format still represents the largest single source of dollars but showed slight declines overall in 2010—58% of revenue (down from 65%) and 43% of unit sales (down from 46%).” *Id.*

82. See Heald, *Fiction Bestsellers*, *supra* note 13, at 1039–43.

83. See *id.* at 1040–41.

per work. By 2006, 98% of the public-domain bestsellers from 1913 to 1922 were in print compared to only 74% of the copyrighted bestsellers from 1923 to 1932.⁸⁴ These data indicate that the fears about underuse may be inflated.

A second study, tracking the use of public domain songs in movies, showed that public domain songs were exploited at a rate equal to that of their copyrighted counterparts.⁸⁵ Heald measured the rate at which songs from 1913 to 1932 appeared in movies and accounted for the number of moviegoers who attended each movie the year of its release. He found no difference in the rates at which moviegoers were exposed to public domain and copyrighted songs.⁸⁶ Also, the study took on the overuse claim directly and found that copyright owners were willing to license their songs for use in movies at an equal or higher rate than public domain songs were used.⁸⁷ In other words, ownership did not function as a relative constraint on comparative use rates in that market.

Finally, at the request of the Library of Congress, Tim Brooks studied the rate at which copyright owners were making old vinyl audio recordings of popular music available to the public. He found that non-owners had converted more music from vinyl to digital format than copyright owners had.⁸⁸

B. STUDY 1: THE EXPLOITATION OF POPULAR FICTION IN AUDIOBOOKS

While the research discussed above has cast doubt on the hypotheses offered by some economists and proponents of term extension, the present audiobook studies enable us to more directly ascertain what happens to works when they fall into the public domain. Studying the audiobook market suggests a number of distinct advantages. First, audiobooks constitute derivative works under U.S. copyright law because they are transformations of other copyrighted works.⁸⁹ Many of the arguments concerning term extension discuss the public domain's presumed ill effects on the production of derivative works, so, unlike the research discussed above, this study can

84. *Id.* at 1040.

85. See Heald, *Musical Compositions*, *supra* note 13, at 10–12 (demonstrating songs in the public domain were used in film once every 3.8 years while songs protected by copyright were used in film once every 3.3 years, with little difference in the relative popularity of films using public domain or copyrighted songs).

86. *Id.*

87. *Id.* at 14–15.

88. See TIM BROOKS, NAT'L RECORDING PRES. BD., LIBRARY OF CONG., SURVEY OF REISSUES OF U.S. RECORDINGS 7–8 & 7 tbl. 4 (2005) (demonstrating that copyright owners have made only an average of 14% of popular recordings from 1890 to 1964 available on CD, while non-owners have made 22% of them available to the public on CD).

89. 17 U.S.C. § 101 (2011).

help explore the public domain's effect on different versions of the same work. Second, derivative works require a capital investment to create, thus studying audiobooks provides an opportunity to measure whether investors are deterred from exploiting public domain works.

The market for audiobooks is distinctive in its heterogeneity. Many audiobooks, of both public domain and copyrighted works, are produced at significant expense by firms that use professional actors working on sound stages. The production and distribution of these audiobooks may cost thousands of dollars per book.⁹⁰ With improvements in computing, however, private individuals may also make their own audiobooks with nothing more than a copy of the book, a computer, and some free software. For example, the website Librivox.org collects, organizes, and distributes thousands of privately recorded audiobooks produced by lay readers.⁹¹ The site encourages members of the public to submit their own recordings of public domain works; the Librivox staff then reviews the submissions to ensure accuracy and comprehensibility. Librivox, however, makes no effort to judge the quality of recordings or to limit its listings only to those of high quality.⁹² Accordingly, while many of the audiobooks available on its website rival professional recordings in quality, many others are quite poorly made. Study 2 takes advantage of this heterogeneity in audiobook quality to test the hypothesis that low quality derivative works affect the value of the underlying work.

90. Celebrity readers are often paid \$4000 to \$6000 for standard six-hour recordings. See *Why Celebrities are Lending Their Voices to Audiobooks*, PUBLISHING CENTRAL, <http://publishingcentral.com/articles/20061126-32-06e6.html?si=4> (last visited Nov. 4, 2012).

91. LIBRIVOX, <http://www.librivox.org> (last visited Nov. 4, 2012).

92. See hugh, Comment to *COMPLETE FAQ – Everything You Need to Know (Almost)*, LIBRIVOX (Nov. 13, 2005, 2:45 PM), <https://forum.librivox.org/viewtopic.php?f=18&t=219&sid=ce01d19d7a0c0bf3fc0d2fb30171548c>. In response to the frequently asked question, “Don’t you have any standards?” user hugh stated:

Our feeling is this: in order for LibriVox to be successful we must welcome anyone who wishes to honour a work of literature by lending their voice to it. Some readers are better than others, and the quality of reading will change from book to book and sometimes from chapter to chapter. But we will not judge your reading, though we may give you some advice if you ask for it. This is not Hollywood, and LibriVox has nothing to do with commercial media’s values, production or otherwise. However: we think almost all of our readings are excellent, and we DO try to catch technical problems (like repeated text *etc.*) with our Listeners Wanted/prooflistening stage.

Id.

1. *Methods: Study 1*

The underuse and overuse hypotheses make empirically testable assertions about the availability of works once they enter the public domain. These hypotheses assert that works will be either under-exploited or diluted, respectively, after they lose copyright protection. Many works subject to copyright, however, have no significant remaining value when they fall into the public domain. Accordingly, the hypotheses are only relevant to those works that have retained significant value at the time when they would enter the public domain. Our study focuses on just these works.

Following the methodology used in one of the studies discussed above,⁹³ we derived a list of bestselling novels that were published in the decade before (1913–1922) and the decade after (1923–1932) the copyright-public domain divide (i.e., all of the novels published between 1913 and 1922 have entered the public domain, while all of those published in or after 1923 are still subject to copyright protection).⁹⁴ The list includes 171 public domain novels and 174 copyrighted novels.⁹⁵ Our goal was to collect a sample of fiction from the same period large enough to support statistically meaningful analyses.

Of course, many books that were bestsellers when published may no longer have significant value. Accordingly, we derived a second, smaller list of novels that have shown enduring popularity. This list—generated by examining the number of editions in print and consulting with experts in the literature of the period⁹⁶—includes twenty public domain novels and twenty copyrighted novels.⁹⁷ These books, like James Joyce’s *The Portrait of the Artist as a Young Man* (1916) and William Faulkner’s *The Sound and the Fury* (1929), are still widely read and retain significant cultural and economic value. This Article refers to these novels as “durable.”

To test the underuse and overuse hypotheses, we collected data on the availability and prices of audiobook versions of all 375 works. We searched the most widely used online retailers of audiobooks, Audible.com,⁹⁸

93. See Heald, *Fiction Bestsellers*, *supra* note 13.

94. We discarded a handful of post-1922 bestsellers that had not been renewed after the expiration of their initial twenty-eight year copyright term. Such works fell into the public domain and were not eligible for the 1976 or 1998 term extensions.

95. See *infra* Appendix A.

96. See Heald, *Fiction Bestsellers*, *supra* note 13, at 1038–39 (describing the selection methodology for “durable” novels).

97. See *infra* Appendix B.

98. AUDIBLE.COM, www.audible.com (last visited Nov. 4, 2012). Audible.com is owned and operated by Amazon. *Id.*

Amazon,⁹⁹ and Barnes and Noble,¹⁰⁰ all of which sell versions in either CD or downloadable mp3 format. We also cross-referenced our results against the online listing *Books in Print*, published by Bowker.¹⁰¹ In addition to noting the availability of titles, we computed the average prices of professional recordings across the different retailers. Finally, using Librivox, we collected data on the availability of free recordings of public domain novels.

2. Results: Study 1

A comparison of the full samples of 171 public domain novels and 174 copyrighted novels shows some similarities and some differences. Of the public domain novels, 58 of the 171 titles (33%) have at least one available recording. Of those, 17 only exist in a Librivox recording. There are a total of 193 complete recordings of the recorded works (67 on CD and 126 on mp3), for an average of 3.3 recordings per recorded title. For the 174 copyrighted titles, 27 are available in audiobook format (16%). Of these, there are a total of 80 complete recordings (44 on CD and 36 on mp3), for an average of 3.0 recordings per recorded title. Interestingly, the average price for the available recordings is fairly similar for public domain and copyrighted titles (Public Domain: CD = \$26, mp3 = \$22; Copyrighted: CD = \$28, mp3 = \$19).

Table 1: Full Sample of Novels

	# in Sample	# Recorded	% Recorded	Total Recordings	Recordings/Recorded Title	Avg. Price CD	Avg. Price mp3
Public Domain	171	58	33	193	3.3	\$26	\$22
Copyrighted	174	27	16	80	3.0	\$28	\$19

For bestselling novels from 1913 to 1932, the data suggest that being in the public domain roughly doubles the likelihood that the work will be available in audiobook format. Despite this increase, however, the fact that a work is in the public domain—and is thus free to be used without licensing—does not ensure that it will be made into an audiobook. Including the versions available on Librivox, fewer than half of the public domain titles are available as audiobooks. Moreover, the similarity in prices between professionally read public domain and copyrighted audiobooks at least suggests that the public domain titles are not being produced in appreciably lower quality versions.

99. AMAZON, www.amazon.com (last visited Nov. 4, 2012).

100. BARNES & NOBLE, www.bn.com (last visited Nov. 4, 2012).

101. See Bowker, BOOKS IN PRINT, www.booksinprint.com (last visited Nov. 4, 2012).

For the titles of enduring popularity, the story is similar. All of the twenty public domain titles are currently available in an audiobook version, and there are 6.25 recordings per title. Of the enduringly popular copyrighted works, however, only 17 are currently available in audiobook format (85%), and there are only 3.25 versions per recorded title. The data on pricing are consistent with Heald's earlier study finding that the 20 copyrighted durable books were significantly more expensive on a price-per-page basis than the 20 public domain durable books.¹⁰² This study concludes that durable copyrighted audiobooks cost \$0.050 per minute for CDs and \$0.036 per minute for mp3 downloads. The corresponding price for the durable public domain audiobooks was significantly lower: \$0.038 per minute for CDs and \$0.028 for mp3 downloads.

Table 2: Enduringly Popular Novels

	# in Sample	# Recorded	% Recorded	Total Recordings	Recordings/Recorded Title	Avg. Price/min. CD	Avg. Price/min. mp3
Public Domain	20	20	100	134	6.25	\$0.038	\$0.028
Copyrighted	20	16	80	62	3.25	\$0.050	\$0.036

As with the full sample, being in the public domain increases the likelihood that a work of enduring popularity will be available in audiobook format, and it increases the number of recordings of the title that are likely to be available when compared to similar copyrighted works. For these works, there is full exploitation of public domain novels in audiobook format. Part IV, *infra*, analyzes whether the number of recordings per title constitutes overexploitation.

C. STUDY 2: AUDIOBOOK QUALITY AND TARNISHMENT

Study 2 addresses the tarnishment hypothesis put forward by economists and proponents of copyright term extension. According to this hypothesis, once works enter the public domain and are free to be used by anyone, the works will be subjected to a variety of inappropriate and poor quality uses that will undermine the works' cultural and economic value. Without copyright ownership, so the argument goes, valuable works will not be properly husbanded. This study focuses only on the durable works described

102. Heald, *Fiction Bestsellers*, *supra* note 13, at 1048–49.

in Study 1. Using a novel experimental methodology, we tested: (1) whether public domain works are produced in poorer quality audiobook versions than are copyrighted works, and (2) whether poorer quality audiobook versions affect the perceived value of the novels from which they are made.¹⁰³

1. *Methods: Study 2*

To test these questions, we relied on the heterogeneity of available audiobook recordings from multiple sources. As mentioned above, audiobooks are available from both professional and amateur sources. If the tarnishment hypothesis is correct, we would expect that: (1) the quality of audiobook recordings of copyrighted works would be higher than that of audiobooks based on public domain works (because the copyrighted works have an owner to husband them); and (2) the lower quality of the public domain audiobooks would be reflected in a lower perceived value of the underlying novel.

To test these assumptions, we recruited subjects through Amazon Mechanical Turk¹⁰⁴ to listen to selections of audiobook recordings and to provide feedback on them. After agreeing to participate, the subjects were directed to the survey instrument, hosted on a standard survey platform.¹⁰⁵ The subjects were told that the survey was being conducted by researchers who were testing the quality of different people as potential audiobook readers. The subjects were then presented with five alternating five-minute recordings taken from the beginning of the fifth chapter of the selected novels.¹⁰⁶ After listening to each selection, the subjects were asked a series of questions:

1. They were asked two comprehension questions to ensure that they were paying attention.
2. They were asked to rate the quality of the reader's readiness for commercial distribution on a scale of one to six.¹⁰⁷

103. See *supra* notes 71–75 and accompanying text.

104. *Amazon Mechanical Turk*, WIKIPEDIA.ORG, http://en.wikipedia.org/wiki/Amazon_Mechanical_Turk (last visited Nov. 9, 2012) (“The Amazon Mechanical Turk (MTurk) is a crowdsourcing Internet marketplace that enables computer programmers (known as Requesters) to coordinate the use of human intelligence to perform tasks that computers are currently unable to do.”); see also AMAZON MECHANICAL TURK, <https://www.mturk.com/mturk/welcome> (last visited Nov. 9, 2012).

105. QUALTRICS, www.qualtrics.com (last visited Nov. 9, 2012).

106. We selected the fifth chapter to avoid biases associated with particularly well-known or interesting first chapters.

107. The points on the scale were labeled:

- 1) This reader could never produce a commercially acceptable audiobook.
- 2) With great improvement this reader could produce an acceptable audiobook.

3. They were asked if they had read or seen other versions of the work and, if so, how much they liked them.
4. After being told that the surveyors would have multiple paperback copies of the book left over after completing the survey, they were asked to indicate the price for which the surveyors should sell the extra copies. They were instructed that paperback copies typically sell for \$8 to \$12.¹⁰⁸

Finally, after listening to the five different recordings, the subjects were asked a series of questions about their own audiobook usage and demographic background.

We chose the recordings from works on our list of the most enduringly popular novels on either side of the copyright-public domain divide. The sample recordings came from several different sources. Since at the time of the study there were only sixteen professional recordings of the twenty most durable copyrighted works, we selected all sixteen of them. In addition, we randomly selected sixteen of the twenty professionally recorded public domain audiobooks. Comparing the subjects' responses to these sets of recordings enabled us to test whether the professional versions of the public domain works were being produced at the same standards as professional versions of the copyrighted works.

In addition, this study included versions of the works produced by non-professionals. Accordingly, we selected recordings of the same sixteen public domain works that are downloadable on the website Librivox. These recordings had been made by private parties using their own equipment. Of course, because the copyrighted works are still under copyright protection, non-professional recordings of these works are not available publicly. To complete the sample and to provide a control for the comparative attractiveness of the content of all the underlying works, we employed a non-professional reader to record copies of the sixteen copyrighted works. This ensured that particularly exciting or interesting prose did not bias the evaluation of the reader.

This strategy gave us a 2 x 2 matrix of recordings (Legal Status: Copyrighted vs. Public Domain; Source: Professional vs. Non-Professional).

3) This reader is close to good enough, but still needs some improvement.
 4) The reader was acceptable for commercial distribution.
 5) The reader was very good, clearly ready for commercial distribution.
 6) The reader was excellent.

Figure 1

LEGAL STATUS

		<i>Copyrighted</i>	<i>Public Domain</i>
RECORDING SOURCE	<i>Professional</i>		
	<i>Non-professional</i>		

2. Results: Study 2

Our data provide almost no support for the arguments made by proponents of copyright term extension that, once works fall into the public domain, the works will be produced in poor quality versions that will undermine their cultural or economic value.¹⁰⁹ Our data indicate no statistically significant difference, for example, between the listeners' judgments of the quality of professional audiobook readers of copyrighted and public domain texts.¹¹⁰ Study 2 also found no significant difference between the recommended prices for which the paperback copies should be sold.¹¹¹ This suggests, as we discuss in more detail below, that the producers

109. In addition to the data reported here, we reran the study with a sample of subjects recruited from the general population by Qualtrics. The results of that study are identical to those reported here, and we chose to report the mTurk data because the quality of the responses that we received were higher in the mTurk sample. The results are available online at <http://www.kentlaw.iit.edu/institutes-centers/center-for-empirical-studies-of-intellectual-property/cesip-projects/copyright-term-extension>.

110. Two sample t test, $p = 0.4452$. To indicate a statistically significant difference, the "p value" should be less than 0.05.

111. Two sample t test, $p = 0.9203$.

of professional audiobook recordings of public domain works are not using poorer quality readers than are the producers of copyrighted works.

The data do reveal, however, that the amateur recordings of both copyrighted and public domain works are perceived to be of lower quality than are the professional versions.¹¹² Librivox recordings of public domain works were perceived to be significantly worse than professional recordings (3.54 versus 4.30, on a scale of 1 to 6, respectively),¹¹³ and the recordings of our assistant were perceived to be significantly worse than professional recordings of both copyrighted and public domain works.¹¹⁴ This difference is not surprising—the resources that go into professional recordings will tend to be much greater than those that go into amateur recordings.

The important question, however, is whether the perceived difference in quality between amateur and professional recordings resulted in corresponding value judgments of the underlying work. The answer is a qualified “no,” but the data are not entirely unambiguous. In general, we found a positive and statistically significant relationship between the perceived quality of a recording and the amount for which subjects thought copies should be sold. This is important for two reasons. First, it suggests that our metric for studying the underlying value of a work (i.e., asking how much we should sell copies for) is sensitive to changes in quality of the recording and, thus, indicates the validity of the measure. Second, it suggests that people who listen to poor quality recordings of audiobooks are likely to attribute some of their dissatisfaction to the underlying work. Accordingly, there appears to be some feedback effect between the quality of a given version of a work and the value of the underlying work.

Table 3: Quality and Price of Recordings

	Avg. Quality (1–6)	Avg. Price
Public Domain Professional	4.30	\$8.30
Copyright Professional	4.17	\$8.26
Public Domain Librivox	3.54	\$8.00
Copyright Research Asst.	3.56	\$8.40
Public Domain Research Asst.	3.55	\$7.78

112. See *infra* Table 3.

113. Two sample t test, $p = 0.0002$.

114. Assistant vs. Copyrighted: two sample t test, $p = 0.0027$; Assistant vs. Public Domain: two sample t test, $p = 0.0001$. We detected no significant difference between our assistant’s recordings of public domain works and his recordings of copyrighted works.

Importantly, the correlation between recording quality and price did not manifest itself in the manner predicted by proponents of copyright term extension. The data indicated no statistically significant differences in book price between any of the paired conditions.¹¹⁵ Thus, for example, although the Librivox recordings of public domain works were judged to be of lower quality than professional recordings of public domain works, we detected no significant difference between the prices that subjects indicated for the paperback books.¹¹⁶ Moreover, although we detected a positive correlation between quality and price for the entire sample of works, we found no such correlation within any of the subsamples. These results suggest that although there may be a modest feedback effect associated with poor quality versions of creative works, that effect is not related to whether a work is protected by copyright or not.

D. DATA LIMITATIONS

We do not and cannot claim to have established all the precise effects of works falling into the public domain. There may be effects that were not measured or that apply to other industries left unexplored.

One limitation, of course, is that audiobooks are not necessarily representative of copyright-eligible works as a whole. But perhaps the biggest limitation of our data involves the difficulty of scientifically proving the lack of a difference. Social scientific research and statistical methods are normally aimed at demonstrating the existence of a difference between a treatment group and a control group. When such a difference is shown, there is reason to believe that it is the result of true differences between the groups. When no difference is detected, however, the inverse inference is not necessarily true. Failure to find an effect may be the result of an insufficiently sensitive experimental design or inadequate statistical power.

While it is possible that some such problem accounts for our failure to detect a difference between the quality of copyrighted or public domain professional readings, the findings still likely track reality. First, this study included hundreds of subjects sourced via multiple methods, which should have provided the statistical power necessary to detect a difference. Recall, also, that this study detected a significant difference between the quality of Librivox recordings and the quality of professional recordings and a positive correlation between the quality of a recording and the valuation of the

115. In addition, we found no meaningful effects based on prior exposure to the works, although this likely was the result of the small sample of subjects who had prior experience with the works.

116. Two sample t test, $p = 0.3203$.

underlying work, although that correlation did not map on to differences between the source of the reading (professional vs. Librivox).

Future research should continue to study the effects of the public domain on the value of works. Perhaps other methods can be devised that overcome some of these limitations. In the meantime, however, our data suggest that anxieties about the public domain are substantially overblown.

IV. IMPLICATIONS FOR IP LAW AND POLICY: THE NEXT TIME DISNEY COMES KNOCKING

Our audiobook study has obvious implications for the ongoing worldwide debate over the extension of copyright terms in existing works. That debate has centered on factual assumptions about what happens to works when they fall into the public domain—assumptions that are contradicted by our data. In addition, our data on audiobook pricing, in conjunction with similar data on book pricing,¹¹⁷ illustrate one important reason why the copyright term extension debate should matter to consumers: we found higher prices for recordings of the most popular older works.

A. ADDRESSING THE UNDERUSE HYPOTHESIS

Lack of availability has been the most prominent concern expressed by Congress and commentators about works falling into the public domain. If works tended to disappear after their copyright terms expired, a plausible argument could be made for term extension because these lost works would be unavailable for future readers, users, and creators. Consistent with several previous studies,¹¹⁸ however, we found that audiobooks were significantly more likely to be made from older bestselling public domain works than from bestselling copyrighted works from the same era. Even excluding audiobooks available for free at the Librivox website, the public domain works were more available to consumers in audiobook form. For the full sample, public domain works were twice as likely to be available, and for the sample of enduringly popular works, public domain titles were 20% more likely to be available. *These data suggest that copyright status, in fact, seems to reduce availability, even for the most popular books.* Even today, there are no unabridged audio recordings for three of the most popular copyrighted novels in our study: *Magnificent Obsession* by Lloyd Douglas, *Mutiny on the Bounty* by Nordoff

117. See Heald, *Fiction Bestsellers*, *supra* note 13, at 1048–49; Petra Mosse, et al., *Dead Poets' Property – The Copyright Act of 1814 and the Price of Books in the Romantic Period* (Nov. 14, 2012) (unpublished manuscript), available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2170447.

118. See *supra* notes 76–82 and accompanying text.

and Hall, and *Death Comes for the Archbishop* by Willa Cather. Further, D.H. Lawrence's *Lady Chatterley's Lover* (1930) did not appear as an unabridged audiobook until 2011.

The finding that there is a greater availability of audiobooks made from public domain works represents a significant advance over a previous study's finding that bestselling public domain novels are more likely to be in print and in more editions than the bestselling copyrighted novels from the same era.¹¹⁹ Unlike reprinted novels, audiobooks are derivative works that require time and effort to produce. Professional versions of audiobooks can cost substantial sums to record, produce, and market; yet, producers do not hesitate to expend significant capital resources to produce their own versions of the work.¹²⁰ Producers of audiobooks are clearly not deterred by their inability to exclude competitors from making competing products. Of 171 public domain titles, 75 had audiobook versions (including 17 Librivox-only versions), while of 173 copyright titles, only 27 had been made into audiobooks. Producers appear more interested in expending capital to produce public domain titles. As our data suggest, the market for public domain audiobooks thrives even though multiple competing versions are often available of the same work. A right to exclude is clearly not needed to incentivize the production of audiobooks made from older works.

If the argument for copyright term extension turns on the need for incentives to reproduce older works or create derivative works from them, then existing empirical evidence suggests that term extensions are not needed and are probably counter-productive.

B. ADDRESSING THE OVERUSE HYPOTHESIS

As discussed in Part II, *supra*, economists not only worry about the underuse of public domain works, but they also are concerned that some works will be over-exploited if no single owner has the right to exclude others. This tragedy-of-the-commons argument suggests that because no individual has the right to exclude others, everyone has the incentive to rush to exploit the resource while it has value. According to the argument, the public will encounter public domain works so frequently that their value will

119. See Heald, *Fiction Bestsellers*, *supra* note 13, at 1046–47.

120. See Liebowitz & Margolis, *supra* note 71, at 451 (“Firms producing copies or derivatives of creative works after the copyright expires may be in the position of fisherman on an open access lake. They produce at their own private optima, not taking into account the effects that they have on other producers.”).

be lost.¹²¹ Our data contradicts this alternative ground for copyright term extension.

For the whole data set, we found an average of 3.3 recordings made for each recorded public domain work and 3.0 recordings for each recorded copyrighted book, an insignificant difference that provides no indication that public domain books are over-exploited and worn out due to their unprotected legal status. In addition, the average price of recorded books in the full public domain data set and the full copyrighted data set was virtually the same, suggesting that the value of the public domain works in comparison to their copyrighted counterparts had not been destroyed by overuse.

This study uncovered no evidence of over-exploitation even when we consider only the most enduringly popular public domain and copyrighted works. In that regard, this study revealed a significant difference in exploitation rates, although the sample size was small. Within their respective groups, the 20 most enduringly popular public domain books had an average 6.25 audiobook recordings per title, while the 16 most popular copyrighted works had only 3.25 audiobook recordings per title. While this is evidence of a higher level of exploitation, further data suggest that there is not evidence of harmful overuse.

One clue that the increased availability of public domain works is not a signal of over-exploitation comes from the pricing data that we accumulated.¹²² Although audiobooks made from the durable public domain works do not command as high a price, the price is still fairly high and close to that for copyrighted works. Even with the competition that professional public domain versions face from free recordings on Librivox, these versions are still able to command market prices that are reasonably close to those obtained by copyrighted works.

While professionally produced public domain audiobooks are priced lower than copyrighted versions, there is little reason to believe that this price difference is due to over-exploitation and the “wearing out” phenomenon. Several compelling explanations for the price difference that are unrelated to an overuse effect also exist. First, the producers of the audio recordings from copyrighted books have to pay a royalty to the copyright owner that may increase the cost of producing the work and raise its price in relation to the public domain works, which require no such payment. Just as likely, the

121. For a succinct expression of this concern in the publicity rights context, see Bitton, *supra* note 59, at 78 (“[I]f everyone uses the likeness of Humphrey Bogart in advertising, it will eventually become worthless.”).

122. See *supra* Table 3.

“intradbrand” competition between the multiple editions of the audiobooks based on the same public domain work will drive down their prices even in the absence of any “wearing out” phenomenon. Note, however, that despite this competition and the competition from free Librivox recordings, the price for professionally produced public domain audiobooks is still fairly high. Finally, data presented in a prior study suggest a significant disparity in the popularity and appeal of the public domain and copyrighted titles at issue.¹²³ If the copyrighted works are indeed more iconic, then we would expect versions of them to be sold at a higher price. Overall, the pricing disparity between audiobooks based on public domain and copyrighted works does not convince us that the public has seen its most valuable public domain works dangerously over-recorded.

In addition, as a practical matter, it is difficult to see how the availability of multiple versions of an audiobook would diminish the value of the underlying work. No one is forced to consume an audiobook, so multiple copies are not flung in the face of the consuming public who then become tired of hearing the story. If audiobooks were played in the background of commercials or department stores, perhaps repetitive, choice-less consumption might negatively affect consumer attitudes, but audiobooks are not used that way. And even with music, which does appear in commercials and in the background ambience of shopping areas, we suspect that businesses try not to alienate their customers by overusing the same music.

123. See Heald, *Fiction Bestsellers*, *supra* note 13, at 1046–47. Heald wrote:

[A]s of 1965, when all of the forty durable books were still protected by copyright, only five of the twenty books (1913–1922) that have since fallen into the public domain had sold 1,000,000 copies. As of the same date, eleven of the twenty books (1923–1932) still protected by copyright today had sold 1,000,000 copies, despite having on the average ten fewer years to accomplish that feat. Even more tellingly, the top five books from the public domain set (1913–1922) had sold a total of only 7,381,709 volumes as of 1965, while the top five sellers from the copyrighted set (1923–1932) had sold 20,289,943 volumes. And as of 1965, the top five books still protected by copyright had fifteen fewer years to sell than those that have since fallen into the public domain. Sales data for books selling fewer than 1,000,000 copies as of 1965 is not publicly available. An update on books that had sold over 2,000,000 volumes by 1975 reemphasizes the comparative popularity of the books published from 1923–1932. Only one of the durable books published from 1913–1922 is on the list (*Of Human Bondage*, with sales of 2,609,236), while seven from 1923–1932 are on the list. Sales of those seven books, as of 1975, totaled 28,732,714.

Id. (citing ALICE PAYNE HACKETT, 70 YEARS OF BEST SELLERS, 1895–1965, at 111–45 (1967); ALICE PAYNE HACKETT & JAMES HENRY BURKE, 80 YEARS OF BEST SELLERS, 1895–1975 (1977)).

Market discipline should make over-exploitation highly unlikely—it is just bad business. It is difficult to imagine how any harm flows from the higher exploitation rate that we measured in the set of the twenty most enduringly popular public domain works.

C. ADDRESSING THE TARNISHMENT HYPOTHESIS

Although many legal analysts are skeptical of the claim that “inappropriate” uses of a work can negatively affect its value,¹²⁴ the present study is the first to evaluate empirically the claim that a work will be tarnished by unconstrained uses in the absence of a copyright owner to “husband” the work and protect it from the ravages of the free market. One of us has earlier argued that even pornographic version of works are unlikely to affect value,¹²⁵ but one could imagine, for example, that a truly horrible movie made from a book might have an effect on the sales of the book. If the *Howard the Duck* comic book had still been regularly in print at the time of the release of its famously awful movie version,¹²⁶ perhaps sales would have dropped (although such a fate would also serve as an example of how copyright ownership does not prevent debasement). By the same token, one could imagine that someone listening to an inferior recording of an audiobook might become less likely to consume the underlying written work, thereby diminishing its value.

Given how easily supporters of copyright term extension can assert the claim of misuse of works in the public domain, it was critical to take the debasement argument seriously. The audiobook context provided an attractive opportunity for study because the claim of tarnishment caused by a poor audiobook reading seems more credible than the claim that Santa Claus

124. See Richard A. Epstein, *Liberty versus Property? Cracks in the Foundations of Copyright Law*, 42 SAN DIEGO L. REV. 1, 26 (2005) (“Anyone is hard pressed to believe that Shakespeare’s star has been dimmed by the calamities committed in his name”); cf. Yi Qian, *Counterfeiters: Foes or Friends?* (NBER Working Paper No. 16785, 2011) (finding that the advertising benefits of counterfeits in the market outweigh any substitution losses); Renée Ann Richardson Gosline, *The Real Value of Fakes: Dynamic Symbolic Boundaries in Socially Embedded Consumption* (May 5, 2009) (unpublished Ph.D. dissertation, Harvard Business School) (on file with authors) (finding “social networks enable counterfeit consumers to develop relationships with the authentic brand”).

125. See Heald, *Musical Compositions*, *supra* note 13, at 25–26.

126. See *Howard the Duck (film)*, WIKIPEDIA.ORG, [http://en.wikipedia.org/wiki/Howard_the_Duck_\(film\)](http://en.wikipedia.org/wiki/Howard_the_Duck_(film)) (last visited Nov. 4, 2012) (“The film frequently ranks among the worst films of all time.”).

has been debased by the numerous pornographic movies with “Santa” appearing in the title.¹²⁷

Part II explained that any claim of debasement in the audiobook market would be predicated on two underlying factual assumptions. First, readers of public domain audiobooks are inferior to readers of copyrighted audiobooks; and second, the inferior versions of the audiobooks negatively affect consumers’ valuation of the underlying work. We found little support for either assumption.

Regarding the first prong of the tarnishment hypothesis, professional readers of audiobooks made from public domain works were rated just as highly as professional readers of copyrighted books. The companies that produce public domain audiobooks appear to be selecting readers who are as talented as those selected for copyrighted titles. According to the results of our study, when consumers go to the three main sources for audiobooks (www.audible.com, Amazon, and Barnes & Noble), they will likely find that the public domain books are as well produced as the copyrighted books. These data substantially undermine any claim of debasement in the most important market for audiobooks. Market discipline is apparently sufficient to ensure that the producer of an audiobook for commercial sale will hire a competent reader. Producers of audiobooks would like to establish a positive reputation and make a steady profit in the market.¹²⁸ It should be no surprise that such producers take adequate care in the selection of readers, whether the underlying work chosen for exploitation is copyrighted or in the public domain. Amateur readers who distributed audio versions of public domain books on Librivox were, not surprisingly, rated significantly lower than professional readers of the same books. Non-professionals using their own equipment produce significantly lower quality recordings than professional readers in recording studios.

The question for the second prong of the tarnishment hypothesis, then, is whether these lower quality recordings resulted in lower valuations of the underlying works. Although we did find a positive correlation between the quality of readings and the subjects’ valuation of the underlying work, that effect did not correlate with the source of the recording. In other words,

127. See Search Results for “Santa”, INTERNET ADULT FILM DATABASE, www.iafd.com/results.asp?searchtype=title&searchstring=santa (searching for “Santa” under the title criterion) (last visited May 23, 2012).

128. On the value of attribution and reputation in intellectual property, see Christopher Jon Sprigman, Christopher Buccafusco & Zachary Burns, *Valuing Attribution and Publication in Intellectual Property*, 93 BOSTON U. L. REV. (forthcoming 2013), available at <http://ssrn.com/abstract=2011403> (finding that creators significantly value opportunities for attribution).

quality correlates with valuation whether the subject listened to an amateur recording, a professional recording of a public domain book, or a professional recording of a copyrighted book. However, the absolute values assigned to the underlying works by subjects who listened to audiobooks from all three sources were not significantly different. So, the tarnishment thesis has some force, but ownership does not prevent tarnishment in this particular market. Of course, this is contrary to what proponents of term extension argue: that ownership prevents tarnishment in a way that free market discipline does not.

We do not and cannot claim that our data conclusively prove that falling into the public domain has no effect on the value of a work. But even if works are theoretically harmed by falling into the public domain, proponents of term extension should be expected to establish such losses empirically because term extension comes with considerable costs that must be justified. One such cost, already noted, involves pricing. The exclusive rights granted by copyright can sometimes allow owners to charge above-market rates for their products. Imposing such costs on consumers is only worthwhile if the public is getting something valuable in return. Another cost of copyright to consumers is diminished availability. We show that twice as many public domain books have audio versions compared to copyrighted books. If proponents feel that imposing these costs are justified, then they should support their arguments with more than bare assertions.

Perhaps more important than the cost to consumers, other creators must bear higher costs when existing works continue to remain subject to copyright protection. Creators may wish to perform these works, adapt them for new uses, or incorporate them into other kinds of works.¹²⁹ When works are protected by copyright, however, creators must obtain a license or face stiff legal penalties. This license requirement creates multiple problems for new creators and, thus, the public. Copyright owners may demand more in licensing fees than creators are willing or able to pay, resulting in works not getting made.¹³⁰ In other cases, the copyright owners may be impossible to locate and contact. For these “orphan works,” the opportunity for bargaining over their use is impossible, and again, derivative works remain uncreated. If the public is going to be asked to bear costs for an additional period of years,

129. See LAWRENCE LESSIG, *FREE CULTURE: HOW BIG MEDIA USES TECHNOLOGY AND THE LAW TO LOCK DOWN CULTURE AND CONTROL CREATIVITY* (2004).

130. See Buccafusco & Sprigman, *Creativity Effect*, *supra* note 13 (showing that owners of IP rights often demand substantially more money to license their works than others are willing to pay, leading to inefficiencies in IP markets).

it is incumbent upon term-extension proponents to establish that those costs are worth bearing.

D. TARNISHMENT BEYOND TERM EXTENSION

In addition, our audiobook quality and valuation data may be relevant in multiple contexts outside the copyright-term-extension debate. First, some copyright fair use disputes seem to turn on the argument that inappropriate uses will devalue a copyrighted work. For example, those who oppose the publication of fan fiction (for example, new Harry Potter tales concocted by enthusiastic fans on the internet¹³¹) often allege that the copyrighted characters will be tarnished by unconstrained storytelling on the web.¹³² The data may suggest that amateur fan fiction is unlikely to negatively affect the value of the underlying character franchise.

Second, outside of the realm of copyright law, this study might provide support for those who applaud the judiciary's continuing reluctance to vigorously implement the Federal Trademark Anti-Dilution Act.¹³³ The tarnishment prong of dilution doctrine asserts that a trademark loses some of its intrinsic value when consumers encounter the mark used in an inappropriate context, such as when the mark is placed on goods of inferior quality. The data show that listeners to Librivox recordings find the readers to be inferior but do not translate that sentiment to a significantly lower valuation of the associated work. Similarly, the doctrine of post-sale confusion in trademark law rests on the assumption that a trademark owner is harmed when a bystander merely observes a trademark on an inferior product (imagine someone who sees a poor quality Chicago Bears sweat shirt without knowing that it is a knock off). The data suggest that the assumption of such harm is unrealistic and comport with previous studies suggesting that lower quality counterfeits may increase the value of the authentic brand.¹³⁴

V. CONCLUSION

The copyright term extension debate, as it once again begins to heat up, will have substantial consequences for the creative industries and the

131. See HARRYPOTTERFANFICTION.COM, <http://www.harrypotterfanfiction.com/> (last visited Nov. 4, 2012) (containing over 78,000 Harry Potter stories written by fans).

132. See Karjala, *supra* note 76; Rebecca Tushnet, *Payment in Credit: Copyright Law and Subcultural Creativity*, 70 L. & CONTEMP. PROBS. 135 (2007).

133. See 15 U.S.C. § 1125(c) (2011) (creating a cause of action against diluting and tarnishing uses of a famous trademark).

134. See Qian, *supra* note 124 (finding that the advertising benefits of counterfeits in the market outweigh any substitution losses); Gosline, *supra* note 124, at iii (finding "social networks enable counterfeit consumers to develop relationships with the authentic brand").

consuming public. If copyrighted works begin to enter the public domain, their owners will stand to lose millions of dollars in revenue. On the other hand, that revenue comes directly from consumers' pockets, and the expiration of valuable copyrights would save consumers considerable costs. Perhaps more importantly, those works will be available to an army of creative artists who will be able to use them in their works in ways that were impossible while the works were copyrighted. Whether this use will be a good thing if and when it happens is an empirical question that is susceptible to quantitative measurement. This Article has attempted to address that question.

The data suggest that the three principal arguments in favor of copyright term extension—underuse, overuse, and tarnishment—are unsupported. There seems little reason to fear that once works fall into the public domain, their value will be substantially reduced by the amount or manner of their use. Although there may be costs associated with movement into the public domain, allowing open access to public domain works will yield considerable benefits. These benefits should dramatically outweigh the costs.

APPENDIX A: FULL SAMPLE OF BESTSELLING NOVELS, 1913–1932

Public Domain Works (1913–1922)

John Fox, *Heart of the Hills* (1913); Robert Herrick, *His Great Adventure* (1913); Jack London, *John Barleycorn* (1913); Gene Porter, *Laddie* (1913); Willa Cather, *O Pioneers* (1913); Eleanor Porter, *Pollyanna* (1913); O. Henry, *Rolling Stones* (1913); D.H. Lawrence, *Sons and Lovers* (1913); Frances Burnett, *T. Tembarom* (1913); Jeffrey Farnol, *The Amateur Gentleman* (1913); Winston Churchill, *The Inside of the Cup* (1913); Rex Beach, *The Iron Trail* (1913); Gilbert Parker, *The Judgment House* (1913); W.B. Maxwell, *The Devil's Garden* (1913); Jack London, *The Valley of the Moon* (1913); Hall Caine, *The Woman Thou Gavest Me* (1913); Henry Harrison, *V.V.'s Eyes* (1913); Ellen Glasgow, *Virginia* (1913); Robert Herrick, *Clark's Field* (1914); James Joyce, *Dubliners* (1914); Leona Dalrymple, *Diane of the Green Van* (1914); Booth Tarkington, *Penrod* (1914); Edgar Burroughs, *Tarzan of the Apes* (1914); Rex Beach, *The Auction Block* (1914); Harold Wright, *The Eyes of the World* (1914); William Locke, *The Fortunate Youth* (1914); George Barr McCutcheon, *The Prince of Graustark* (1914); Mary Watts, *The Rise of Jennie Cushing* (1914); Owen Johnson, *The Salamander* (1914); Frank Norris, *Vandover and the Brute* (1914); Winston Churchill, *A Far Country* (1915); Henry Harrison, *Angela's Business* (1915); Jean Webster, *Dear Enemy* (1915); F. Hopkinson Smith, *Felix O'Day* (1915); William Locke, *Jaffery* (1915); Mary Roberts Rinehart, *K* (1915); Gene Stratton Porter, *Michael O'Halloran* (1915); Somerset Maugham, *Of Human Bondage* (1915); Irving Cobb, *Old Judge Priest* (1915); Eleanor Porter, *Pollyanna Grows Up* (1915); Harry Leon Wilson, *Ruggles of Red Gap* (1915); Dorothy Canfield, *The Bent Twig* (1915); Theodore Dreiser, *The Genius* (1915); Stewart White, *The Gray Dawn* (1915); Ernest Poole, *The Harbor* (1915); Raphael Sabatini, *The Sea-Hawk* (1915); Zane Grey, *The Lone Star Ranger* (1915); Willa Cather, *The Song of the Lark* (1915); Booth Tarkington, *The Turmoil* (1915); James Joyce, *A Portrait of the Artist* (1916); Ethel Dell, *Bars of Iron* (1916); Peter Bernard Kyne, *Cappy Ricks* (1916); William McFee, *Casuals of the Sea* (1916); Eleanor Porter, *Just David* (1916); Ellen Glasgow, *Life and Gabriella* (1916); H.G. Wells, *Mr. Britling Sees it Through* (1916); Frank Spearman, *Nan of Music Mountain* (1916); Booth Tarkington, *Seventeen* (1916); Winston Churchill, *The Dwelling Place of Light* (1916); Kathleen Norris, *The Heart of Rachael* (1916); William Dean Howells, *The Leatherstocking God* (1916); Henry Kitchell Webster, *The Real Adventure* (1916); Harold Wright, *When a Man's a Man* (1916); Edith Wharton, *Xingu* (1916); Ring Lardner, *You Know Me, Al* (1916); Alice Cholmondeley, *Christine* (1917); Edna Ferber, *Fanny Herself* (1917); Ring

Lardner, *Gullible's Travels* (1917); Ernest Poole, *His Family* (1917); Robert Hichens, *In the Wilderness* (1917); Christopher Morley, *Parnassus on Wheels* (1917); David Graham Phillips, *Susan Lennox: Her Rise and Fall* (1917); James Branch Cabell, *The Cream of the Jest* (1917); Jeffrey Farnol, *The Definite Object* (1917); Ethel Dell, *The Hundredth Chance* (1917); Ralph Connor, *The Major* (1917); Irving Bacheller, *The Light in the Clearing* (1917); William Locke, *The Red Planet* (1917); Stephen McKenna, *Sonia* (1917); Eleanor Porter, *The Road to Understanding* (1917); May Sinclair, *The Tree of Heaven* (1917); Joseph Hergesheimer, *The Three Black Pennys* (1917); Zane Grey, *Wildfire* (1917); Gene Porter, *A Daughter of the Land* (1918); Thorne Smith, *Biltmore Oswald* (1918); Zona Gale, *Birth* (1918); Zane Grey, *The Desert of Wheat* (1918); Edward Streeter, *Dere Mable* (1918); V. Blasco Ibanez, *The Four Horsemen of the Apocalypse* (1918); Joseph Hergesheimer, *Java Head* (1918); Willa Cather, *My Antonia* (1918); Eleanor Porter, *Oh, Money! Oh, Money* (1918); Mary Roberts Rinehart, *The Amazing Interlude* (1918); Booth Tarkington, *The Magnificent Ambersons* (1918); Emerson Hough, *The Passing of the Frontier* (1918); E. Phillips Oppenheim, *The Pawns Count* (1918); Robert Chambers, *The Restless Sex* (1918); Temple Bailey, *The Tin Soldier* (1918); Zane Grey, *The U.P. Trail* (1918); Treat 'Em Rough, *Ring Lardner* (1918); Margaret Atherton, *The Avalanche* (1919); Elizabeth von Arnim, *Christopher and Columbus* (1919); Mary Roberts Rinehart, *Dangerous Days* (1919); Gene Porter, *Dawn* (1919); Winston Churchill, *Dr. Jonathan* (1919); Frannie Hurst, *Humoresque* (1919); Robert Chambers, *In Secret* (1919); James Cabell, *Jurgen* (1919); Albert Terhune, *Lad, A Dog* (1919); Ethel Dell, *The Lamp in the Desert* (1919); Joseph Hergesheimer, *Linda Condon* (1919); Joseph Conrad, *The Arrow of Gold* (1919); Irving Bacheller, *A Man for the Ages* (1919); Ellen Glasgow, *The Builders* (1919); Harold Wright, *The Re-Creation of Brian Kent* (1919); James Curwood, *The River's End* (1919); Emerson Hough, *The Sagebrusher* (1919); Ralph Connor, *The Sky Pilot in No Man's Land* (1919); Sherwood Anderson, *Winesburg, Ohio* (1919); Edith Wharton, *The Age of Innocence* (1920); Kathleen Norris, *Harriet and the Piper* (1920); Peter Kyne, *Kindred of the Dust* (1920); Sinclair Lewis, *Main Street* (1920); Eleanor Porter, *Mary-Marie* (1920); Zona Gale, *Miss Lulu Bett* (1920); Floyd Dell, *Moon Calf* (1920); James Huneker, *Painted Veils* (1920); Sherwood Anderson, *Poor White* (1920); Mary Roberts Rinehart, *A Poor Wise Man* (1920); E. Phillips Oppenheim, *The Great Impersonation* (1920); Zane Grey, *The Man of the Forest* (1920); Joseph Lincoln, *The Portygee* (1920); Anne Sedgwick, *The Third Window* (1920); Francis Fitzgerald, *This Side of Paradise* (1920); James Curwood, *The Valley of Silent Men* (1920); Booth Tarkington, *Alice Adams* (1921); Ben Hecht, *Erik Dorn* (1921); Harold Bell Wright, *Helen of the Old House* (1921); Gene Porter, *Her Father's Daughter* (1921); A.S.M. Hutchinson, *If Winter Comes* (1921); Brian Donne-

Byrne, *Messer Marco Polo* (1921); Rafael Sabatini, *Saramouche* (1921); Ring Lardner, *The Big Town* (1921); Dorothy Fisher, *The Brimming Cup* (1921); Eden Phillpotts, *The Grey Room* (1921); Coningsby Dawson, *The Kingdom Round the Corner* (1921); Louis Hemon, *Maria Chapdelaine* (1921); Zane Grey, *The Mysterious Rider* (1921); Don Marquis, *The Old Soak* (1921); Willa Cather, *One of Ours* (1921); Edith Hull, *The Sheik* (1921); Gertrude Atherton, *The Sisters in Law* (1921); Sherwood Anderson, *The Triumph of the Egg* (1921); John Passos, *Three Soldiers* (1921); Sinclair Lewis, *Babitt* (1922); Thomas Stribling, *Birthright* (1922); Booth Tarkington, *Gentle Julia* (1922); Carl Vechten, *Peter Whiffle* (1922); Robert Keable, *Simon Called Peter* (1922); Francis Fitzgerald, *The Beautiful and the Damned* (1922); Mary Roberts Rinehart, *The Breaking Point* (1922); Raphael Sabatini, *Captain Blood* (1922); Emerson Hough, *The Covered Wagon* (1922); Temple Bailey, *The Dim Lantern* (1922); Elizabeth von Arnim, *The Enchanted April* (1922); Edward Cummings, *The Enormous Room* (1922); Frances Burnett, *The Head of the House of Coombe* (1922); A.S.M. Hutchinson, *This Freedom* (1922); James Joyce, *Ulysses* (1922); Herbert Quick, *Vandermark's Folly* (1922); Christopher Morley, *Where the Blue Begins* (1922).

Copyrighted Works (1923–1932)

Willa Cather, *A Lost Lady* (1923); Gertrude Atherton, *Black Oxen* (1923); Phillip Gibbs, *The Heirs Apparent* (1923); Arthur Train, *His Children's Children* (1923); Elliot Paul, *Impromptu* (1923); Mazo de la Roche, *Jalna* (1923); John Dos Passos, *Streets of Night* (1923); Margaret Wilson, *The Able McLaughlins* (1923); Robert Chambers, *The Hijackers* (1923); Harold Bell Wright, *The Mine with the Iron Door* (1923); Zane Grey, *The Wanderer of the Wasteland* (1923); James Oliver Curwood, *A Gentleman of Courage* (1924); Margaret Kennedy, *The Constant Nymph* (1924); Will James, *Cowboys, North and South* (1924); Michael Arlen, *The Green Hat* (1924); Clarence Mulford, *Hopalong Cassidy Returns* (1924); Ernest Hemingway, *In Our Time* (1924); Emerson Hough, *Mother of Gold* (1924); Edith Wharton, *Old New York* (1924); Edna Ferber, *So Big* (1924); Coningsby Dawson, *The Coast of Folly* (1924); Louis Bromfield, *The Green Bay Tree* (1924); Dorothy Fisher (1924); Anne Douglas Sedgwick, *The Little French Girl* (1924); Booth Tarkington, *The Midlander* (1924); Percy Marks, *The Plastic Age* (1924); Robert Herrick, *Waste* (1924); Theodore Dreiser, *An American Tragedy* (1925); Sinclair Lewis, *Arrowsmith* (1925); Ellen Glasgow, *Barren Ground* (1925); PC Wren, *Beau Geste* (1924); Sherwood Anderson, *Dark Laughter* (1925); James Boyd, *Drums* (1925); Anita Loos, *Gentlemen Prefer Blondes* (1925); E. Barrington, *Glorious Apollo* (1925); John Dos Passos, *Manhattan Transfer* (1925); ASM Hutchinson, *One Increasing Purpose* (1925); Robert Benchley, *Pluck and Luck* (1925); DuBose Heyward, *Porgy* (1925); John Erskine, *The Private Life of Helen of Troy* (1925); A. Hamilton

Gibbs, *Soundings* (1925); Rafael Sabatini, *The Carolinian* (1925); Francis Scott Fitzgerald, *The Great Gatsby* (1925); Gene Stratton Porter, *The Keeper of the Bees* (1925); Gertrude Stein, *The Making of Americans* (1925); Anne Parrish, *The Perennial Bachelor* (1925); Willa Cather, *The Professor's House* (1925); Christopher Morley, *Thunder on the Left* (1925); Susan Ertz, *After Noon* (1925); PC Wren, *Beau Sabreur* (1926); Louis Bromfield, *Early Autumn* (1926); Dorothy Canfield, *Her Son's Wife* (1926); Carl Van Vechten, *Nigger Heaven* (1926); Zona Gale, *Preface to a Life* (1926); Edna Ferber, *Show Boat* (1926); William Faulkner, *Soldier's Pay* (1926); Warwick Deeping, *Sorrell and Son* (1926); Thomas Stribling, *Teeftallow* (1926); Temple Bailey, *The Blue Window* (1926); Sylvia Thompson, *The Hounds of Spring* (1926); Ellen Glasgow, *The Romantic Comedians* (1926); John Galsworthy, *The Silver Spoon* (1926); James Branch Cabell, *The Silver Stallion* (1926); Ernest Hemingway, *The Sun Also Rises* (1926); Elizabeth Roberts, *The Time of Man* (1926); Thorne Smith, *Topper* (1926); A.A. Milne, *Winnie-the-Pooh* (1926); Louis Bromfield, *A Good Woman* (1927); Julia Peterkin, *Black April* (1927); Conrad Aiken, *Blue Voyage* (1927); Thornton Wilder, *The Bridge of San Luis Rey* (1927); Willa Cather, *Death Comes for the Archbishop* (1927); Warwick Deeping, *Doomsday* (1927); Sinclair Lewis, *Elmer Gantry* (1927); Honore Willies Morrow, *Forever Free* (1927); Ole Rolvaag, *Giants in the Earth* (1927); Mary Roberts Rinehart, *Lost Ecstasy* (1927); James Boyd, *Marching On* (1927); Ernest Hemingway, *Men Without Women* (1927); Glenway Westcott, *The Grandmothers* (1927); Don Marquis, *The Lives and Times of Archy Mehitabel* (1927); Anne Douglas Sedgwick, *The Old Countess* (1927); Booth Tarkington, *The Plutocrat* (1927); Anne Parrish, *Tomorrow Morning* (1927); Edith Wharton, *Twilight Sleep* (1927); Fannie Hurst, *A President is Born* (1928); Anne Parrish, *All Kneeling* (1928); Vina Delmar, *Bad Girl* (1928); Booth Tarkington, *Claire Ambler* (1928); H.W. Freeman, *Joseph and his Brethren* (1928); Honore Willies Morrow, *Mary Todd Lincoln* (1928); Roark Bradford, *Ol Man Adam n His Chillun* (1928); Warwick Deeping, *Old Pybus* (1928); Julia Peterkin, *Scarlet Sister Mary* (1928); John Galsworthy, *Swan Song* (1928); S.S. Van Dine, *The Greene Murder Case* (1928); Louis Bromfield Stokes, *The Strange Case of Miss Annie Spragg* (1928); Hugh Walpole, *Wintersmoon* (1928); Ernest Hemingway, *A Farewell to Arms* (1929); Erich Maria Remarque, *All Quiet on the Western Front* (1929); Anne Douglas Sedgwick, *Dark Hester* (1929); Sinclair Lewis, *Dodsworth* (1929); James Thurber, *Is Sex Necessary?* (1929); Oliver LaFarge, *Laughing Boy* (1929); Thomas Wolfe, *Look Homeward, Angel* (1929); Lloyd Douglas, *Magnificent Obsession* (1929); DuBose Heyward, *Mamba's Daughters* (1929); O.E. Rolvaag, *Peder Victorious* (1929); Warwick Deeping, *Reaper's Row* (1929); Ellen Glasgow, *They Stooped to Folly* (1929); SS Van Dine, *The Bishop Murder Case* (1929); Susan Glaspell, *The Fugitives Return* (1929); Susan Ertz, *The Galaxy* (1929); Ellery Queen, *The Roman Hat Mystery* (1929);

William Faulkner, *The Sound and the Fury* (1929); Susan Glaspell, *Alison's House* (1930); J.B. Priestly, *Angel Pavement* (1930); Kenneth Roberts, *Arundel* (1930); A. Hamilton Gibbs, *Chances* (1930); Edna Ferber, *Cimarron* (1930); Warwick Deeping, *Exile* (1930); Katherine Anne Porter, *Flowering Judas* (1930); Vicki Baum, *Grand Hotel* (1930); Michael Gold, *Jews Without Money* (1930); D.H. Lawrence, *Lady Chatterley's Lover* (1930); William McFee, *North of Suez* (1930); Hugh Walpole, *Rogue Herries* (1930); John Dos Passos, *The 42nd Parallel* (1930); Arthur Train, *The Adventures of Ephraim Tutt* (1930); Mary Roberts Rinehart, *The Door* (1930); Elizabeth Madox Roberts, *The Great Meadow* (1930); Dashiell Hammett, *The Maltese Falcon* (1930); Thornton Wilder, *The Woman of Andros* (1930); Honore Willsie Morrow, *Tiger! Tiger!* (1930); Louis Bromfield, *Twenty-Four Hours* (1930); Margaret Ayer Barnes, *Years of Grace* (1930); Katharine Brush, *Young Man of Manhattan* (1930); Bess Streeter Aldrich, *A White Bird Flying* (1931); Susan Glaspell, *Ambrose Holt and Family* (1931); Fannie Hurst, *Back Street* (1931); Honore Willsie Morrow, *Black Daniel* (1931); Mazo de la Roche, *Finch's Fortune* (1931); William McFee, *The Harbourmaster* (1931); John Galsworthy, *Maid in Waiting* (1931); William Faulkner, *Sanctuary* (1931); Willa Cather, *Shadows on the Rock* (1931); Warwick Deeping, *The Bridge of Desire* (1931); Thomas Stribling, *The Forge* (1931); Pearl Buck *The Good Earth* (1931); Ernest Hergesheimer, *The Limestone Tree* (1931); Thorne Smith, *The Night Life of the Gods* (1931); Erich Maria Remarque, *The Road Back* (1931); Ole Rolvaag, *Their Fathers' God* (1931); Sherwood Anderson, *Beyond Desire* (1932); Aldous Huxley, *Brave New World* (1932); Julia Peterkin, *Bright Skin* (1932); Vardis Fisher, *In Tragic Life* (1932); Phyllis Bentley, *Inheritance* (1932); Louis Golding, *Magnolia Street* (1932); Booth Tarkington, *Mary's Neck* (1932); Charles Barnard Nordoff, *Mutiny on the Bounty* (1932); Warwick Deeping, *Old Wine and New* (1932); Pearl Buck, *Sons* (1932); Phillip Stong, *State Fair* (1932); Thorne Smith, *The Bishop's Jaegers* (1932); Robert Herrick, *The End of Desire* (1932); Charles Morgan, *The Fountain* (1932); Ellen Glasgow, *The Sheltered Life* (1932); Thomas Stribling, *The Store* (1932); AJ Cronin, *Three Loves* (1932); Erskine Caldwell, *Tobacco Road* (1932); Sinclair Lewis, *Ann Vickers* (1933); Hervey Allen, *Anthony Adverse* (1933); Gladys Carroll, *As the Earth Turns* (1933); Lloyd Douglas, *Forgive us our Trespasses* (1933); Erskine Caldwell, *God's Little Acre* (1933); Caroline Miller, *Lamb in his Bosom* (1933); Hans Fallada, *Little Man, What Now?* (1933); Bess Streeter Aldrich, *Miss Bishop* (1933); William McFee, *No Castle in Spain* (1933); John Galsworthy, *One More River* (1933); Robert Herrick, *One More Spring* (1933); Philip Stong, *Stranger's Return* (1933); Louis Bromfield, *The Farm* (1933); Mazo de la Roche, *The Master of Jalna* (1933); Ernest Hemingway, *Winner Take Nothing* (1933).

APPENDIX B: DURABLE FICTION BESTSELLERS, 1913–1932

Public Domain Durable Works

Sherwood Anderson, *Winesburg, Ohio* (1919); Edgar Burroughs, *Tarzan of the Apes* (1914); Willa Cather, *My Ántonia* (1918); Willa Cather, *O Pioneers!* (1913); Willa Cather, *The Song of the Lark* (1915); F. Scott Fitzgerald, *The Beautiful and the Damned* (1922); F. Scott Fitzgerald, *This Side of Paradise* (1920); Zane Grey, *The Lone Star Ranger* (1915); James Joyce, *A Portrait of the Artist as a Young Man* (1916); James Joyce, *Dubliners* (1914); James Joyce, *Ulysses* (1922); D.H. Lawrence, *Sons and Lovers* (1913); Sinclair Lewis, *Main Street* (1920); Sinclair Lewis, *Babbitt* (1922); W. Somerset Maugham, *Of Human Bondage* (1915); Eleanor H. Porter, *Pollyanna* (1913); Rafael Sabatini, *Captain Blood* (1922); Rafael Sabatini, *Scaramouche* (1921); Booth Tarkington, *The Magnificent Ambersons* (1918); Edith Wharton, *The Age of Innocence* (1920).

Copyrighted Durable Works

Pearl S. Buck, *The Good Earth* (1931); Willa Cather, *Death Comes for the Archbishop* (1927); John Dos Passos, *Manhattan Transfer* (1925); Theodore Dreiser, *An American Tragedy* (1925); William Faulkner, *Sanctuary* (1931); William Faulkner, *The Sound and the Fury* (1929); F. Scott Fitzgerald, *The Great Gatsby* (1925); Dashiell Hammett, *The Maltese Falcon* (1930); Ernest Hemingway, *A Farewell to Arms* (1929); Ernest Hemingway, *The Sun Also Rises* (1926); Aldous Huxley, *Brave New World* (1932); D.H. Lawrence, *Lady Chatterley's Lover* (1930); Sinclair Lewis, *Arrowsmith* (1925); Sinclair Lewis, *Elmer Gantry* (1927); A.A. Milne, *Winnie-The-Pooh* (1926); Charles Nordhoff, *Mutiny on the Bounty* (1932); Erich Maria Remarque, *All Quiet on the Western Front* (1929); Thornton Wilder, *The Bridge of San Luis Rey* (1927); Thomas Wolfe, *Look Homeward, Angel* (1929); Percival Christopher Wren, *Beau Geste* (1925).

STATE PATENT LAWS IN THE AGE OF *Laissez Faire*

Camilla A. Hrdy[†]

ABSTRACT

This Article brings to light the heretofore unstudied views of esteemed nineteenth century jurists, including Chief Justice of the New York Supreme Court James Kent (1763–1847), that states have concurrent constitutional authority to grant their own patents alongside Congress in order to stimulate innovation and economic development in their own territories. Based on arguments surrounding the constitutional validity of New York’s infamous steamboat monopoly, this Article reveals that in the height of America’s supposed commitment to *laissez faire* economics, concurrent state patent powers were justified by a fundamental concern that Congress’s new and uniquely “hands-off” patent system was not a sufficient replacement for the active patent policies of state and colonial governments prior to adoption of the Constitution. Therefore, in the tradition of Alexander Hamilton—who tempered his vision of a strong central government with a recognition of the importance of autonomous state policymaking and his vision of a vibrant free market with a recognition that targeted government incentives are sometimes necessary to stimulate investment in beneficial activities—state patents were seen as an important policy tool for encouraging the private sector to invest in developing costly technology of unproven value that states deemed worthy of support.

© 2013 Camilla A. Hrdy.

† Visiting Scholar at the Berkeley Center for Law and Technology at University of California Berkeley School of Law and Postdoctoral Fellow at the Information Society Project at Yale Law School. J.D., University of California Berkeley School of Law; M.Phil. in History and Philosophy of Science, University of Cambridge; B.A., Harvard University. I am immensely grateful to Robert Merges and John Yoo, who were essential sources of guidance and encouragement throughout this project. Thanks also to Mario Biagioli, Daniel Kazhdan, Elizabeth Perlman, Zorina Khan, Will Baude, Bryan Choi, Kathy Abrams, and the members of Berkeley Law’s legal scholarship seminar, 2011–12; David Rosen, Mark Langer, and the editors of the *Berkeley Technology Law Journal*; and Nilgun Gokur, Charlie Wyzanski, Marc Esserman, Ursula DeYoung, and my family. This Article is dedicated to the memory of Edward C. Walterscheid.

TABLE OF CONTENTS

I.	INTRODUCTION	47
A.	A WORLD WITHOUT STATE PATENT LAWS	47
B.	AN UNEXPECTED APPEAL FOR STATE PATENTS IN THE AGE OF <i>Laissez Faire</i>	52
C.	THE STRUCTURE OF THIS ARTICLE	55
II.	STATE AND COLONIAL PATENT LAWS	58
A.	LESS EMPHASIS ON NOVELTY; MORE EMPHASIS ON RESULTS	60
B.	LIMITING PRINCIPLES ON COLONIAL AND STATE PATENT LAWS.....	64
III.	THE INCOMPLETE TRANSITION TO A NATIONAL PATENT SYSTEM	67
A.	“THE STATES CANNOT SEPARATELY MAKE EFFECTUAL PROVISION”	67
B.	THE EVIDENCE FOR STATES’ REMAINING PATENT POWERS	70
1.	<i>The Text of the IP Clause</i>	71
2.	<i>The Patent Act’s “Relinquishment” Provision</i>	72
3.	<i>The Decision to Limit U.S. Patents to Universally Novel Inventions</i>	74
IV.	THE LAST STATE PATENT	76
A.	UNCOVERING A STATE PATENT AT THE HEART OF THE STEAMBOAT MONOPOLY	78
B.	THE FIRST COMMERCIAL STEAMBOAT ENTERPRISE IN AMERICA.....	79
V.	JUSTICE JAMES KENT’S CONCURRENT STATE PATENT SYSTEM	80
A.	CHANCELLOR LANSING’S DISCOMFORT WITH STATE- SANCTIONED MONOPOLIES	80
B.	THE NEW YORK SUPREME COURT’S ENDORSEMENT OF STATE PATENT POWERS	81
C.	STATE PATENT LAWS IN THE WAKE OF <i>GIBBONS</i>	88
VI.	JUSTIFYING CONCURRENT STATE PATENT POWERS	93
A.	THE WEAKNESS OF U.S. PATENTS IN THE EARLY NINETEENTH CENTURY	94
B.	U.S. PATENT LAW LEFT MAJOR GAPS IN AMERICAN INNOVATION POLICY	95
1.	<i>No Concern for the Social Utility of the Subject Matter</i>	95
2.	<i>No U.S. Patents for Previously Known or Used Technology</i>	96

3. <i>No Assurance of Local Working of Inventions</i>	98
C. JUSTIFYING CONCURRENT STATE PATENT POWER IN THE AGE OF <i>LAISSEZ FAIRE</i>	100
D. NEW YORK’S STEAMBOAT PATENT AS A MARKET CORRECTIVE.....	104
VII. CONCLUSION	108

I. INTRODUCTION

A. A WORLD WITHOUT STATE PATENT LAWS

In theory, by granting inventors exclusive rights to their original inventions for a limited period, patents give the private sector a needed incentive to invest in innovation and encourage public disclosure of previously unknown technological know-how.¹ Today patent law is purely a federal creature. Article I, Section 8, Clause 8 of the Constitution (“the IP Clause”) gives Congress 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.”² Under the Patent Act, Congress determines the terms and conditions under which inventors may apply for federal patents (“U.S. patents”).³ A single agency, the United States Patent and Trademark Office (“USPTO”), is responsible for granting and issuing those patents.⁴ Federal district courts have exclusive jurisdiction over patent cases, and a single court of appeals, the Federal Circuit, hears all patent appeals in the country.⁵ Meanwhile, states do not grant their own patents and have not done so for over two hundred years.

1. The primary utilitarian justifications for patents are the incentive-to-invent and incentive-to-disclose theories. See Rebecca Eisenberg, *Patents and the Progress of Science: Exclusive Rights and Experimental Use*, 56 U. CHI. L. REV. 1017, 1024–30 (1989). The incentive-to-innovate and prospect theory justifications add that the chance of obtaining a limited patent monopoly will entice patent holders to invest still more time and capital, after an invention has been made, in order to develop, perfect, and market it to the public. See *id.* at 1036–45 (explaining Joseph Schumpeter’s and Edmund Kitch’s famous theories). In today’s world of high financing, U.S. patents also may serve as a useful “signal” for capital-rich investors that a particular innovation or company is a good investment. See Clarisa Long, *Patent Signals*, 69 U. CHI. L. REV. 625, 653 (2002).

2. U.S. CONST. art. I, § 8, cl. 8.

3. 35 U.S.C. §§ 100–213 (2011).

4. *Id.* § 2(a)(1) (authority of USPTO to grant and issue patents). Until the recent reforms, which ordered the director to establish at least three satellite offices, there was only a single regional office of the USPTO in Alexandria, Virginia. See Leahy-Smith America Invents Act, Pub. L. No. 112-29, § 23, 125 Stat. 284, 336 (2011).

5. See 28 U.S.C. §§ 1295 (jurisdiction of Federal Circuit over patent appeals), 1338 (exclusive jurisdiction of federal courts in patent cases) (2011).

Historically, however, the states, and before that the American colonies, played an active role in encouraging inventors and entrepreneurs to invest time and labor in developing and implementing new and useful technology.⁶ One of their most valuable policy tools was individual statutes granting private actors exclusive rights to a particular innovation that promised local benefits.⁷ Although states often referred to these grants as “patents” after the English practice, these were extraordinarily different legal tools from modern U.S. patents.⁸ Universal novelty and original inventorship were not essential requirements for obtaining a state patent; instead, the patentee’s primary obligation was to establish a technology in the state that benefited the public.⁹ Also, unlike with U.S. patents, legislatures used various means, such as individualized term lengths, local working requirements, and compulsory licensing, to ensure that patents rights did not create unjustified monopolies or impede access to the building blocks of knowledge.¹⁰

Today it is essentially a unanimous assumption that states cannot grant their own patents. The Supreme Court has consistently held that the terms of

6. In this Article, I use “technology” in a loose utilitarian sense to mean applied science, such as manufacturing methods or machines that might be immediately profitable or of use to society, rather than basic research. On the gradual development of this modern conception of technology over the course of the nineteenth century, see DAVID NYE, *AMERICAN TECHNOLOGICAL SUBLIME* 45–46 (1994); see also Karl B. Lutz, *Patents and Science: A Clarification of the Patent Clause of the U.S. Constitution*, 18 GEO. WASH. L. REV. 50, 54 (1949) (arguing that “[t]he term ‘useful arts’ as used in the [IP Clause of the] Constitution . . . is best represented in modern language by the word ‘technology.’”).

7. On colonial and state patent laws, generally, see BRUCE BUGBEE, *GENESIS OF AMERICAN PATENT AND COPYRIGHT LAW* 57–103 (1967); Oren Bracha, *Owning Ideas: A History of Anglo-American Intellectual Property* 97–116 (June 2005) (unpublished Ph.D. dissertation, Harvard Law School), available at <http://www.utexas.edu/law/faculty/obracha/dissertation/>. For present purposes, the term “state patent” refers loosely to laws giving a person or entity the right to exclude others from making, using, selling, or practicing an invention or technological innovation, though it need not be entirely new or meet any of the other standards of patentability the U.S. patent laws impose today. See also discussion *infra* Part II.

8. See Edward Walterscheid, *The Early Evolution of the United States Patent Law: Antecedents (5 Part I)*, 78 J. PAT. & TRADEMARK OFF. SOC’Y 615, 616 n.4 (1996); see also notes to Part II, *infra*.

9. See Oren Bracha, *The Commodification of Patents 1600–1836: How Patents Became Rights and Why We Should Care*, 38 LOY. L.A. L. REV. 177, 243 (2004); see also Mario Biagioli, *Patent Republic: Representing Inventions, Constructing Rights and Authors*, 73 SOC. RES. 1129, 1138 (asserting that “[p]rivilege-granting authorities wanted to maximize local utility, not to disclose knowledge about the invention.”); see also *infra* notes and accompanying text in Part II.

10. On working clauses, compulsory licensing and price controls in state patents, see Section II.B, *infra*.

the IP Clause itself do not apply to the states¹¹ and that states are not “completely disabled from offering any form of protection to articles or processes which fall into the broad scope of patentable subject matter.”¹² But, in its leading patent preemption case, *Bonito Boats, Inc. v. Thundercraft Boats, Inc.*, the Court was nonetheless quite clear that states cannot grant patents or even “patent-like”¹³ rights because this would compete with inventors’ decision to apply for U.S. patents and interfere with Congress’s exclusive power to define the criteria and terms of patentability.¹⁴ Despite the

11. See *Goldstein v. California*, 412 U.S. 546, 560–62, 571 (1973) (holding that the IP Clause does not deprive states of their constitutional powers to grant intellectual property rights and upholding California’s authority to criminalize unauthorized copying of sound recordings left unprotected by the Copyright Act); see also *Kewanee Oil Co. v. Bicron Corp.*, 416 U.S. 470, 478–93 (1974). The *Kewanee Oil* Court discussed *Goldstein*, noting that:

[The Court has held that] the cl. 8 grant of power to Congress was not exclusive and that, at least in the case of writings, the States were not prohibited from encouraging and protecting the efforts of those within their borders by appropriate legislation This determination was premised on the great diversity of interests in our Nation—the essentially non-uniform character of the appreciation of intellectual achievements in the various States. Evidence for this came from patents granted by the States in the 18th century.

Id.

12. See *Bonito Boats, Inc. v. Thunder Craft Boats, Inc.*, 489 U.S. 141, 154 (1989) (finding that federal law preempted a Florida statute that prohibited using a controversial direct molding technique to copy boat hulls that had not been patented and did not appear to meet the federal criteria of patentability).

13. The term “patent-like” is not confined to state laws that create a right against the world to exclude others from inventive subject matter, as patent laws do. See *Bonito Boats*, 489 U.S. at 156–57. The “antimolding” statute that the Court struck in *Bonito Boats* did not grant boat hull manufacturers exclusive rights in their boat hull design, but simply prohibited using a particular method (a direct molding process) to copy the boat hulls. Thus, it did not give boat hull manufacturers the right to exclude competitors from making and selling the same boat hulls so long as they did not use the direct molding process to do so. This is why, in *Interpart Corporation v. Italia*, 777 F.2d 678, 684–85 (Fed. Cir. 1985), which was later overruled by the Supreme Court in *Bonito Boats*, 489 U.S. at 143–44, the Federal Circuit upheld California’s similar antimolding statute, basing its decision largely on the fact that the statute did not actually give the creator of the boat hull designs “the right to exclude others from making, using, or selling the product as does the patent law.” *Interpart*, 777 F.2d at 684–85.

14. See *Bonito Boats*, 489 U.S. at 151 (“The offer of federal protection from competitive exploitation of intellectual property would be rendered meaningless in a world where substantially similar state law protections were readily available.”); *Kewanee*, 416 U.S. at 489 (“If a State, through a system of protection, were to cause a substantial risk that holders of patentable inventions would not seek patents, but rather would rely on the state protection, we would be compelled to hold that such a system could not constitutionally continue to exist.”). In *Sears, Roebuck & Company*, the Court noted that Congress enacted the first Patent Act in 1790 and:

Court's contention that this prohibition is purely statutory, the Court has effectively read it into the Constitution, creating a so-called "Dormant Patent Clause."¹⁵ The Federal Circuit, for its part, has suggested that the Patent Act "occupies the field of patent law," if not unfair competition law.¹⁶

[E]ver since that time has fixed the conditions upon which patents . . . shall be granted. . . . Obviously a State could not, consistently with the Supremacy Clause of the Constitution, extend the life of a patent beyond its expiration date or give a patent on an article which lacked the level of invention required for federal patents. To do either would run counter to the policy of Congress of granting patents only to true inventions, and then only for a limited time.

376 U.S. 225, 229–31 (1964) (footnotes omitted); *see also* Daniel Laster, *The Secret is Out: Patent Law Preempts Mass Market License Terms Barring Reverse Engineering for Interoperability Purposes*, 58 BAYLOR L. REV. 621, 647 (2006). Laster summarized:

In a nutshell, although there is no preemption of the field of regulating IP, the Supremacy Clause of the Constitution requires preemption of any state law which either (1) provides a party (whether investor or inventor) a realistic and potentially preferable option to the limited term monopoly grant offered by patent law or (2) provides patent-like protection for functional matter which does not meet the novelty and nonobviousness requirements of patent law and thereby stifles competition in such matter without any concomitant advance in the "Progress of Science and useful Arts."

Id.

15. Ever since Chief Justice Marshall declined to decide the issue in *Gibbons v. Ogden*, 22 U.S. 1, 221 (1824), discussed below, courts have toyed with the idea of a "dormant patent clause." *See* Dan L. Burk, *Protection of Trade Secrets in Outer Space Activity: A Study in Federal Preemption*, 23 SETON HALL L. REV. 560, 610 (1992) ("Since *Gibbons*, the tantalizing prospect of a "dormant patent clause" has remained an unadjudicated possibility, but a possibility that informs a preemption analysis involving the Patent Clause."). In the middle of the twentieth century a highly influential judge on the Second Circuit, Learned Hand, posited the existence of a "Dormant IP Clause," which entirely extinguishes any state patent or copyright laws. *See* Miller, *infra* note 17, at 747. The Supreme Court has repeatedly claimed to reject this position. *See Goldstein*, 412 U.S. at 560; *Kewanee*, 416 U.S. at 479; *Bonito Boats*, 489 U.S. at 154. But the presumption that the IP Clause impliedly prevents states from granting their own patents pervades the Court's patent preemption cases. For example, the *Bonito Boats* Court stated:

The novelty and nonobviousness requirements of patentability embody a congressional understanding, *implicit in the Patent Clause itself*, that free exploitation of ideas will be the rule, to which the protection of a federal patent is *the exception* To a limited extent, the federal patent laws must determine not only what is protected, but also what is free for all to use.

Bonito Boats, 489 U.S. at 151 (emphasis added); *see also* *Compco Corp. v. Day-Brite Lighting, Inc.*, 376 U.S. 234, 237 (1964) The *Compco* Court held that:

[W]hen an article is unprotected by a patent or a copyright, state law may not forbid others to copy that article. To forbid copying would interfere with the federal policy, *found in Art. I, § 8, cl. 8, of the Constitution* and in the

Legal academia has also supported this assumption of federal supremacy.¹⁷ The most important illustration comes from Edward

implementing federal statutes, of allowing free access to copy whatever the federal patent and copyright laws leave in the public domain.

Id. (emphasis added); see also K. David Crockett, *The Salvaged Dissents of Bonito Boats v. Thunder Craft*, 13 GEO. MASON L. REV. 27, 28 (1990) (concluding that the Court in *Bonito* relied on theories of implied preemption, actual conflict preemption, and “a new ‘dormant patent clause.’”).

16. See *Hunter Douglas, Inc. v. Harmonic Design, Inc.*, 153 F.3d 1318, 1334 (Fed. Cir. 1998) (“With respect to field pre-emption, Title 35 occupies the field of patent law, not commercial law between buyers and sellers.”) (citing *Cover v. Hydramatic Packing Co.*, 83 F.3d 1390, 1393 (Fed. Cir. 1996)). An intent by Congress to preempt an entire field may be inferred from a “scheme of federal regulation . . . so pervasive as to make reasonable the inference that Congress left no room for the States to supplement it,” or when congressional legislation “touch[es] a field in which the federal interest is so dominant that the federal system will be assumed to preclude enforcement of state laws on the same subject.” *Rice v. Santa Fe Elevator Corp.*, 331 U.S. 218, 230 (1947).

17. Even those scholars who advocate for *less* stringent preemption of state IP laws do not suggest that states should be allowed to grant exclusive rights against the world like U.S. patents do. See, e.g., K. David Crockett, *The Salvaged Dissents of Bonito Boats v. Thunder Craft*, 13 GEO. MASON L. REV. 27, 57–59 (1990) (explaining that under “relational rights theory” the key distinction between state laws like trade secrets and trademarks and the federal patent laws is that the state laws only protect rights as between competitors and do not create intellectual property against the world like federal patents do); Douglas Gary Litchman, *The Economics of Innovation: Protecting Unpatentable Goods*, 81 MINN L. REV. 693, 731–32 (1997) (arguing from the law and economics perspective that states should be allowed to provide limited protections for unpatentable innovations, with the caveat that “[i]f a state law seems to do more than forbid a particularly egregious form of unauthorized duplication, or if a law leaves intact no tenable copying alternative, the court should and will strike the law down.”); Arthur Miller, *Common Law Protection for Products of the Mind: An Idea Whose Time Has Come*, 119 HARV. L. REV. 703, 752 (2006) (distinguishing state laws protecting against theft of undeveloped ideas from “a state right in something unpatented that is enforceable against the public” which would be preempted under the Patent Act). Other scholars have advocated abolishing state intellectual property laws entirely, including seemingly benign business incentives like trademarks and trade secrets. See, e.g., Joan E. Schaffner, *Patent Preemption Unlocked*, 1995 WIS. L. REV. 1081, 1085–86 (1995) (arguing that “any state protection of discoveries conflicts with the congressional patent scheme” and advocating for an even broader preemption theory under field preemption principles); Malla Pollack, *Unconstitutional Incontestability? The Intersection of the Intellectual Property and Commerce Clauses of the Constitution: Beyond A Critique of Shakespeare Co. v. Silstar Corp.*, 18 SEATTLE U. L. REV. 259, 270–326, 313–14 (1995) (noting historical support for the view that the IP Clause does not, on its own, extinguish states’ intellectual property powers, though going on to argue that various other theories place direct constitutional limits on states’ powers to protect intellectual property and that state trademarks can therefore only last for “limited times.”). Other scholars object in particular to state contract laws that allow IP owners to extend their exclusive rights beyond those afforded by U.S. patents. See, e.g., John E. Mauk, *The Slippery Slope of Secrecy: Why Patent Law Preempts Reverse-Engineering Clauses in Shrink-Wrap Licenses*, 43 WM. & MARY L. REV. 819 (2001) (patent law preempts shrink-wrap licenses that prohibit

Walterscheid, the preeminent historian of the IP Clause and American patent law.¹⁸ According to Walterscheid, “the enactment of federal patent and copyright laws in 1790 was largely viewed as removing the need for state patents and copyrights, because the advantages of uniformity and broader protection inherent in the federal system were obvious to almost everyone.”¹⁹ Walterscheid goes on to conclude that state patents are unconstitutional because under the IP Clause “Congress alone has the discretion to set the term of patents and copyrights”²⁰

B. AN UNEXPECTED APPEAL FOR STATE PATENTS IN THE AGE OF
LAISSEZ FAIRE

What the Supreme Court and modern scholars alike seem to have overlooked is that in 1798, ten years after ratification of the Constitution, the state of New York *did* grant a state patent, giving the wealthy and well-connected investor Robert Livingston (1746–1813) and his partner the inventor Robert Fulton (1765–1815) the exclusive right to make, use, and sell any vessels “propelled by the power of steam” within the territory of the state for up to thirty years.²¹ When Livingston and Fulton attempted to

reverse engineering); Laster, *supra* note 14, at 647 (patent law preempts mass market license terms that bar reverse engineering for interoperability purposes).

18. For an impressive review of all materials related to the IP Clause and the history of patent law in America, refer to EDWARD C. WALTERSCHEID, *THE NATURE OF THE INTELLECTUAL PROPERTY CLAUSE: A STUDY IN HISTORICAL PERSPECTIVE* (2002) as well as Walterscheid’s many articles cited throughout his book and this article.

19. *See* WALTERSCHEID, *THE NATURE OF THE INTELLECTUAL PROPERTY CLAUSE*, *supra* note 18, at 436–37; *see also* BUGBEE, *supra* note 7, at 102–03 (also suggesting federal patents made state patents obsolete); WALTERSCHEID, *THE NATURE OF THE INTELLECTUAL PROPERTY CLAUSE*, *supra* note 18, at 76–77 (highlighting the downsides of a patent system based on patents that were limited to the jurisdiction of one state), 438–42 (dismissing Justice Kent’s views that states could continue to grant their own patents alongside Congress).

20. WALTERSCHEID, *THE NATURE OF THE INTELLECTUAL PROPERTY CLAUSE*, *supra* note 18, at 442–43. To elaborate, Walterscheid wrote:

Under the clause, *Congress alone has the discretion to set the term of patents and copyrights* and once that term has expired, the clear intent of the clause is that the subject matter of the particular patent or copyright shall go into the public domain to be available for use or reproduction by all throughout the United States.

Id. at 442 (emphasis added).

21. *See* *Livingston v. Van Ingen*, 9 Johns. 507, 511 (N.Y. 1812). There has been a good deal of recent scholarship on *Gibbons v. Ogden*, 22 U.S. 1 (1824), and the controversy surrounding Livingston and Fulton’s monopoly. *See* THOMAS H. COX, *GIBBONS V. OGDEN, LAW, AND SOCIETY IN THE EARLY REPUBLIC* (2009); HERBERT A. JOHNSON, *GIBBONS V. OGDEN: JOHN MARSHALL, STEAMBOATS, AND THE COMMERCE CLAUSE* (2010); Norman R.

enforce their rights in *Livingston v. Van Ingen*, 9 Johns. 507 (N.Y. 1812), the court unanimously upheld the patent. In his lengthy opinion, the Chief Justice and respected American jurist, James Kent (1763–1847), included a remarkable piece of dicta stating that nothing in the Constitution prevented states from continuing to concurrently grant patents over a wide range of inventive subject matter, just as they had done prior to ratification.²²

Kent's position is entirely incommensurable with modern preemption doctrine, as is his decision to uphold Livingston and Fulton's thirty-year monopoly over a mode of navigation that was already, to some extent, in the public domain.²³ But although the U.S. Supreme Court famously struck down the steamboat monopoly in *Gibbons v. Ogden*, Chief Justice John Marshall did so on narrow grounds and did not decide or even address the patent preemption issue—leaving Kent's strongly worded dicta on states' concurrent patent powers entirely intact.²⁴ Less than ten years later, Justice Joseph Story, the father of many modern patent law doctrines and a justice on the Supreme Court during one of the most important periods in American legal development, endorsed Kent's opinion in *Livingston* in his seminal *Commentaries on the Constitution of the United States*.²⁵ Courts, including the Supreme Court, continued to cite favorably to Kent's opinion throughout the nineteenth century.²⁶ *Livingston* has never been overruled.

Despite these indicators that notable jurists in the period shared Kent's views on concurrent state patent powers, scholars have largely ignored them.²⁷ The reason is apparently their assumption that the availability of U.S.

Williams, *Gibbons*, 79 N.Y.U. L. REV. 1398 (2004). For a more complete historiography, see JOHNSON, *supra* note 21, at 1–3, 179–88.

22. *Livingston*, 9 Johns. at 581–82.

23. See *Bonito Boats, Inc. v. Thunder Craft Boats, Inc.*, 489 U.S. 141, 156–57 (1989).

24. See *Gibbons v. Ogden*, 22 U.S. 1, 221 (1824).

25. JOSEPH STORY, COMMENTARIES ON THE CONSTITUTION OF THE UNITED STATES 79 (1833) [hereinafter COMMENTARIES].

26. See, e.g., *Patterson v. Kentucky*, 97 U.S. 501, 508–09 (1878) (employing Kent's opinion in *Livingston* to uphold states' authority to pass laws requiring illuminating oils to be certified before sale in the state).

27. With the exception of Edward Walterscheid, whose views I discuss below, scholars have barely mentioned Kent's opinions on concurrent state patent powers. Cox and Johnson have both briefly discussed Kent's opinion in *Livingston*, but neither addresses Kent's views on a concurrent patent system in detail nor notes Kent's views on states' continuing powers to grant patents to inventors. See COX, *supra* note 21, at 60 (briefly summarizing Kent's opinion, though not commenting on Kent's dicta regarding a concurrent patent system); see also JOHNSON, *supra* note 21, at 33–34. Regarding Kent's holding on the patent power, Johnson stated:

While United States patent laws protected actual authors and inventors, state financial grants, prizes, and monopolies encouraged the importation

patents made state patents both obsolete and impractical, and that Kent's opinion was both wrong and irrelevant. For example, after noting the incommensurability between Kent's position in *Livingston* and modern preemption doctrine, Walterscheid concludes that Kent's position "was of doubtful validity in 1812 and most certainly would be deemed invalid today."²⁸ In Walterscheid's version of events, Kent's contemporaries believed that if states continued granting patents to inventors on their own terms and conditions this would lead to an unacceptable "patchwork of patent rights in the various states, a result which the Framers clearly intended to avoid by the constitutional authorization for federal patents and copyrights."²⁹

I do not disagree with Walterscheid that in the early nineteenth century, as now, the *risk* that concurrent state patents might undermine the U.S. patent system would have been obvious and might have led some people to disagree with Kent's conclusion. But I contend that we are still left with a perplexing question worth addressing in some depth: given that the perils of a concurrent patent system were so obvious, why then did nineteenth century jurists like James Kent, Joseph Story and perhaps John Marshall himself—all of whom were Federalists and adherents to Alexander Hamilton's vision of a strong central government, and were otherwise committed to Adam Smith's general aversion to state-sanctioned monopolies³⁰—nonetheless think there ought to be a concurrent patent system and that the two systems could effectively and beneficially coexist?

of foreign technology, augmenting the federal system. The Livingston-Fulton New York legislative grant was an outstanding example of this supplementary function of state legislative power, and therefore would not be in conflict with the interstate commerce powers should Congress elect to exercise those powers.

Id.; see also G. EDWARD WHITE, THE MARSHALL COURT AND CULTURAL CHANGE, 1815–1835, in 3–4 OLIVER WENDELL HOLMES DEVISE HISTORY OF THE SUPREME COURT OF THE UNITED STATES 569 (1988) (discussing only Kent's conclusion that states could regulate commerce unless state and federal powers in the field "[came] directly in contact"); Williams, *supra* note 21, at 1407 (stating regarding *Livingston* only that New York's highest court "rejected Van Ingen's Dormant Commerce Clause-based constitutional challenge to the monopoly, holding that the Commerce Clause did not divest the states of authority to regulate commercial activities such as steamship navigation."). Oren Bracha's dissertation on the history of the U.S. patent laws notes the historical question of whether state patents were preempted by the IP Clause in a brief footnote without addressing Kent's opinion. Bracha, *supra* note 7, at 115 n.268.

28. WALTERSCHEID, THE NATURE OF THE INTELLECTUAL PROPERTY CLAUSE, *supra* note 18, at 442.

29. *Id.* at 442–43.

30. On Hamilton and the Federalist agenda for a strong national government, both economically and politically, see GORDON S. WOOD, EMPIRE OF LIBERTY: A HISTORY OF THE EARLY REPUBLIC, 1789–1815, at 95–110 (2009).

C. THE STRUCTURE OF THIS ARTICLE

To answer this question, this Article proceeds as follows. In Part II, I introduce the patent practices of the colonies and the early states, showing how they differed from modern U.S. patents and emphasizing the unique features that made them effective tools for local policymakers, particularly their low cost as compared to outright subsidies and their utility for mitigating the risk of funding relatively new technology of unproven value.

In Part III, I explain the major reasons behind the decision to transition from a state to a national patent system. I then present evidence that the Framers of the Constitution and the drafters of the Patent Acts of 1790 and 1793—with participation from Secretary of State Thomas Jefferson—nonetheless deliberately left states with residual authority to grant their own patents.

In Part IV, I provide the factual background necessary for understanding New York's decision to grant a patent for the steamboat to Robert Livingston in 1798 and the eventual legal arguments challenging the patent's constitutionality.

In Part V, I present Justice Kent's opinion in *Livingston* upholding the steamboat patent and providing the fullest judicial expression of the argument that states retain broad concurrent powers to grant patents based on the principles of federalism espoused by Alexander Hamilton in *The Federalist*.³¹ I show that Kent's opinion was never overruled and received broad support throughout the nineteenth century, even in the wake of the Supreme Court's decision to strike down the steamboat monopoly on other grounds in *Gibbons*.³²

In Part VI, I seek to understand and justify the argument for concurrent state patent powers in the context of nineteenth-century economic and political theory. I argue that advocates for state patents were concerned that the national patent system was incapable of doing all the work of promoting innovation in America, let alone within individual states. Unlike their state

31. *Livingston v. Van Ingen*, 9 Johns. 507, 581–85 (N.Y. 1812). Hamilton famously explained that many state and federal powers, such as the power of taxation, are “concurrent,” not “repugnant,” and can generally be exercised side by side without conflict. See THE FEDERALIST NOS. 32, 33 (Alexander Hamilton); see also BERNARD BAILYN, TO BEGIN THE WORLD ANEW: THE GENIUS AND AMBIGUITIES OF THE AMERICAN FOUNDERS 120 (2003) (writing that Hamilton insisted that “[t]wo authorities with similar powers could concur, if their roles were clearly established—could divide their responsibilities into separate spheres—could even reinforce each other and clarify each other's role. The Constitution's federalist division of absolute powers was a structure of concurrence, he argued, not repugnancy.”).

32. *Gibbons v. Ogden*, 22 U.S. 1, 221 (1824).

counterparts, U.S. patents protected original inventors' exclusive rights in a relatively narrow range of universally novel subject matter and did not allow the government to play a role in selecting technologies or in tailoring patent rights to the investment incentive actually required. Therefore, in the tradition of Hamilton—who had recommended an array of national incentives to promote innovation and industrial development, including patents to “introducers” of foreign inventions³³—Federalists like Kent and Story believed states must be able to use patents to encourage local implementation of potentially invaluable technology that was otherwise too expensive or risky to attract private investment.³⁴ After all, as Joseph Schumpeter would famously put it over a century later, “[a]s long as they are not carried into practice, inventions are economically irrelevant.”³⁵

These early arguments resemble modern proposals for reforming *federal* patents to more effectively induce investment in costly innovations with uncertain chance of success.³⁶ However, I show that in Kent's time the more viable policy solution came from *state* governments, which were almost exclusively responsible for promoting economic development in their own territories and frequently granted legal monopolies to private developers of

33. See ALEXANDER HAMILTON, REPORT ON MANUFACTURERS (1791), reprinted in 2 ANNALS OF CONG. 971, 991–92, 1014, available at <http://memory.loc.gov/ammem/amlaw/lwaclink.html> [hereinafter REPORT ON MANUFACTURERS]. On Hamilton's support for patents of importation, see Edward Walterscheid, *Patents and Manufacturing in the Early Republic*, 80 J. PAT. & TRADEMARK OFF. SOC'Y 855, 860–78 (1998).

34. See *Livingston*, 9 Johns. at 581–85; STORY, COMMENTARIES, *supra* note 25, at 79.

35. JOSEPH SCHUMPETER, THE THEORY OF ECONOMIC DEVELOPMENT 88–89 (Redvers Opie trans., Harvard Univ. Press 1951).

36. The most popular strategy has been to focus on federal courts' role in crafting the standards of patentability. See Robert P. Merges, *Uncertainty and the Standard of Patentability*, 7 HIGH TECH. L.J. 1, 1–4, 43–55 (1992) (arguing that patent law's nonobviousness test “seeks to reward inventions that, viewed prospectively, have a low probability of success” and arguing for a moderate lowering of the patentability standards for inventions involving high-cost research and uncertain success); Dan L. Burk & Mark A. Lemley, *Policy Levers in Patent Law*, 89 VA. L. REV. 1575, 1582–83 (2003) (arguing that courts should use flexible standards to account for the incentive needs of different industries); Michael Abramowicz & John F. Duffy, *The Inducement Standard of Patentability*, 120 YALE L.J. 1590, 1593–96, 1625–26, 1676 (2011) (arguing that courts should revitalize an “inducement standard” of nonobviousness and apply a structured economic inquiry into the economic incentives that would exist without the patent, on the one hand, and the actual or expected economic costs of developing and commercializing the invention, on the other). Others suggest tailoring patents at the granting stage. See, e.g., Michael W. Carroll, *One Size Does Not Fit All: A Framework for Tailoring Intellectual Property Rights*, 70 OHIO ST. L.J. 1361 (2009); Eric E. Johnson, *Calibrating Patent Lifetimes*, 22 SANTA CLARA COMPUTER & HIGH TECH. L.J. 269 (2006); see also Gideon Parchomovsky & Michael Mattioli, *Partial Patents*, 111 COLUM. L. REV. 207, 219–23 (2011) (proposing a new option to obtain a “partial patent” or a “semi patent” and usefully summarizing extant proposals for patent law reforms).

expensive public works like bridges and roads.³⁷ As I will demonstrate, states' authority to grant patents on uncertain new technologies like the steamboat came from a direct analogy to their authority to grant public utilities monopolies.³⁸ Although political shifts in the 1830s subjected states' monopoly grants to stricter scrutiny, courts considered them constitutional and strictly enforced them against states under the Contracts Clause.³⁹

In conclusion, I draw together my major findings and introduce the obvious question left by all of this historical research: could states revive their patent practices today? From a strict Originalist perspective, I show in this Article that the answer is *yes*. For those unwilling to restrict government to a structure designed in the late eighteenth century, the answer may be different. My views on the modern policy implications of state patent laws are the subject of a work in progress.⁴⁰ I simply suggest here that, despite the predominance of the federal government in issuing patents for the past two centuries, patent law is not necessarily immune to the benefits of federalism and decentralized decision making.⁴¹ My purpose here is to lay the vital

37. See LAWRENCE M. FRIEDMAN, *A HISTORY OF AMERICAN LAW* 177–81 (2d ed. 1985) (describing states' nearly exclusive role in promoting local development, using various monopoly charters and other encouragements). On state monopolies in the nineteenth century, see Herbert Hovenkamp, *Technology, Politics, and Regulated Monopoly: An American Historical Perspective*, 62 *TEX. L. REV.* 1263, 1268 (1984).

38. See *Livingston*, 9 Johns. at 573 (analogizing New York's sovereign power to grant Livingston the exclusive right to operate steamboats to New York's power to grant monopolies to operators of ferries, bridges and roads). Herbert Hovenkamp argues that many of the monopolies states granted in the nineteenth century to operators of public works—including the steamboat monopoly—can be justified using natural monopoly theory; but Hovenkamp concedes that states did not rely on a structured economic analysis in granting these monopolies, and that governments' and courts' ideas of which technologies qualified for the protections of a legal monopoly changed over time with changing notions of which technologies the state had a duty to provide the public. Hovenkamp, *supra* note 37, at 1293–95.

39. See STANLEY I. KUTLER, *PRIVILEGE AND CREATIVE DESTRUCTION: THE CHARLES RIVER BRIDGE CASE* 18–21 (1971); see also cases cited *infra* note 272.

40. Camilla A. Hrdy, *State Patents as a Solution to Underinvestment in Innovation*, 62 *U. KAN. L. REV.* (forthcoming 2013), available at <http://ssrn.com/author=1687909>; see also Camilla A. Hrdy, *Dissenting State Patent Regimes*, 3 *IP THEORY* 78 (2013) (arguing that objectors to U.S. patents could seek state patents instead of U.S. patents, leading to more innovation spillovers and encouraging states to implement socially beneficial patent reforms).

41. The Supreme Court has recognized the benefits of allowing states to protect intellectual property in their own jurisdictions alongside the U.S. government, at least with respect to copyright, trade secret and unfair competition laws. See *Goldstein v. California*, 412 U.S. 546, 557–62 (1973); *Kewanee Oil Co. v. Bicron Corp.*, 416 U.S. 470, 478–79 (1974). More recently, Xuan-Thao Nguyen has recognized the benefits of using local patent rules and other local procedural reforms to improve the efficiency of the patent litigation process. See Xuan-Thao Nguyen, *Dynamic Federalism and Patent Law Reform*, 85 *IND. L.J.* 449, 451 (2010) (nonetheless concluding that local reform movements cannot change the

groundwork, showing that state patent laws have a vibrant history and that our predecessors believed the practice was too valuable to give up, in spite of Congress's national patent system.

II. STATE AND COLONIAL PATENT LAWS

Starting in the early seventeenth century, the American colonies, loosely following the British practice, frequently granted “patents” to inventors and importers of technology for the purpose of inducing costly projects with the prospect of a period of exclusivity.⁴² In England, the word “patent” had referred to special privileges granted by the monarch and recorded in “open letters” to become matters of public record.⁴³ Similarly, in the colonies the term generally implied a grant from the government of exclusive rights or “privileges”⁴⁴ and broadly included corporate charters and franchises given to

substance of patent laws due to accepted federal supremacy in patent law). Also, the economic literature generally favors *decentralized* rather than centralized policymaking structures “based on the observation that free-market economies perform better than planned, centralized economies.” Tim Wu, *Intellectual Property, Innovation, and Decentralized Decisions*, 92 VA. L. REV. 123, 126 (2006). For examples, see *id.* at 124 n.3. See also Richard Schragger, *Decentralization and Development*, 96 VA. L. REV. 1837, 1838–39 nn.1–3 (2010) (challenging studies showing a connection between decentralized policies and economic development, in part due to the possibility that divergence in results is due to fundamentally different economic geography, not governmental policy). On the benefits of decentralized innovation policies in particular, see, for example, CHARLES W. WESSNER, NATIONAL RESEARCH COUNCIL, GROWING INNOVATION CLUSTERS FOR AMERICAN PROSPERITY: SUMMARY OF A SYMPOSIUM, available at http://sites.nationalacademies.org/PGA/step/PGA_043846; ORG. FOR ECON. CO-OPERATION AND DEV., REGIONS AND INNOVATION POLICY 29–68 (2011), available at http://www.oecd-ilibrary.org/urban-rural-and-regional-development/regions-and-innovation-policy_9789264097803-en (arguing that regions cannot simply replicate national policies given the heterogeneity between and within regions and the importance of regional collaboration and local networks for producing innovation).

42. See BUGBEE, *supra* note 7, at 57–103; Bracha, *Ownning Ideas*, *supra* note 7, at 97–116; see also P.J. Federico, *Colonial Monopolies and Patents*, 11 J. PAT. OFF. SOC'Y 358 (1929); P.J. Federico, *State Patents*, 13 J. PAT. OFF. SOC'Y 166 (1931); Frank D. Prager, *A History of Intellectual Property from 1545 to 1787*, 26 J. PAT. OFF. SOC'Y 711 (1944).

43. Blackstone wrote:

The king's grants are also matter[s] of public record These grants, whether of lands, honors, liberties, franchises, or aught besides, are contained in charters, or letters patent, that is, open letters, *literae patentes*: so called because they are not sealed up, but exposed to open view, with the great seal pendant at the bottom; and are usually directed or addressed by the king to all his subjects at large.

2 WILLIAM BLACKSTONE, COMMENTARIES *346–47 (1768); see also ROBERT P. MERGES & JOHN F. DUFFY, PATENT LAW AND POLICY 3 n.6 (4th ed. 2007).

44. I use the terms “privilege” and “right” interchangeably mainly because the grants themselves often did. For example, a 1722 Connecticut grant conveyed the “the sole right, privilege and liberty to erect, use, maintain and support, a slitting mill” See 6 PUBLIC

operators of public utilities.⁴⁵ Although these patents frequently included a legal monopoly,⁴⁶ they might include other local privileges as well,⁴⁷ such as tax incentives, grants of land,⁴⁸ exclusive use of the colony's natural resources,⁴⁹ or authorization to establish a facility for production within the colony.⁵⁰ For instance, Massachusetts Bay's twenty-one-year monopoly grant

RECORDS OF THE COLONY OF CONNECTICUT (1636–1776) 312–13 (James H. Trumbell & Charles J. Hoadly eds., 1850–1890) [hereinafter CONNECTICUT RECORDS]; see also BUGBEE, *supra* note 7, at 181 n.34. However, Adam Mossoff has pointed out that the term “privilege” had a different meaning in the colonial period, sometimes implying that the grantee had a “natural right” in his invention, so it is important not to impose our modern definition of the term onto colonial grants. See Adam Mossoff, *Who Cares What Thomas Jefferson Thought About Patents? Reevaluating the Patent “Privilege” in Historical Context*, 92 CORNELL L. REV. 953, 968–69 (2007).

45. See BLACKSTONE, *supra* note 43, at *346. On colonial grants to operators of bridges and ferries, see KUTLER, *supra* note 39, at 6–8. For a history of franchises in America, see David Gurnick & Steve Vieux, *Case History of the American Business Franchise*, 24 OKLA. CITY U. L. REV. 37, 38 (1999). On early corporate charters, see RON HARRIS, *INDUSTRIALIZING ENGLISH LAW: ENTREPRENEURSHIP AND BUSINESS ORGANIZATION 1720–1844* (2000).

46. Modern patents are not necessarily the same thing as a monopoly, mainly because exclusive rights over a narrow range of subject matter do not necessarily confer market power over an invention or product that actually reaches the market. See Kenneth W. Dam, *The Economic Underpinnings of Patent Law*, 23 J. LEGAL. STUD. 247, 249–50 (1994). However, early colonial and state patents were likely to confer a monopoly since they were much broader in scope, likely covering an entire product or operation, had local working requirements, and often included other benefits besides exclusivity.

47. Biagioli has noted that, unlike modern U.S. patents, the exclusive privileges granted in early modern Europe “also [included] other benefits such as authorization to set up a business in a certain place . . . cash awards, free housing, capital investments in the invention, the permission to immigrate and assume citizenship, or the exemption from taxes [etc.]” and that they “were as much about excluding competitors as about providing resources and permissions to set up and operate a business based on that invention.” Mario Biagioli, *From Print to Patents: Living on Instruments in Early Modern Europe*, 44 HIST. SCI. 139, 147–48 (2006).

48. For example, in 1639, the General Court of Massachusetts Bay granted 500 acres of land at Pecoit to Edward Rawson for a gunpowder mill to make saltpeter. *Records of the General Court of Massachusetts*, quoted in J.L. BISHOP, *HISTORY OF AMERICAN MANUFACTURERS* II 23 (Philadelphia: Young, 1864).

49. For example, in 1641, the town of Plymouth granted John Jenney the use of a nearby island, “Clarke’s Hand,” in order to operate a salt works and supply the town with salt at two shillings a bushel, and also gave him exclusive use of the island’s wood for this purpose. 1 RECORDS OF THE TOWN OF PLYMOUTH (1636–1705), at 7 (Plymouth: Avery & Doten 1889–1903); see also BUGBEE, *supra* note 7, at 180 n.14.

50. For example, in 1718, in response to their request to set up a linseed oil mill, the assembly of Connecticut granted to John Prout, Moses Manfield, and Jerimiah Attwater “the sole and whole privilege of making linseed oyl [sic] within this Colony,” specifying “that no other person or persons shall set up any mill or other engines for that purpose within the county of New Haven during the space of twenty years next coming, nor in any other part of this Colony without the special leave of this Court.” To keep their rights, they were required to set up the mill within two years and keep it in “good repair at all times.” CONNECTICUT RECORDS, *supra* note 44, at 79–80; see also BUGBEE, *supra* note 7, at 181 n.34.

in 1644 to the “Company of the Undertakers of the Iron Works in New England” provided, along with the exclusive privilege to make iron and manage iron mines and works in the colony, various tax exemptions and free use of the colony’s natural resources, including “all manner of wood & timber” and “all manner of earth, stones, turfe, clay, & other materials for building & reparations of any of their workes, forges, mills, or houses, built or to be built”⁵¹

These early patents played a similar policy role to the patents the federal government eventually offered American inventors under the early Patent Acts.⁵² But they contained many features we do not associate with U.S. patents today; and they were also not directly analogous to those granted in Britain in the period.⁵³ Thus, to refer to them as “patents” at all is misleading *unless* we understand precisely what the term entailed in contemporary context.

A. LESS EMPHASIS ON NOVELTY; MORE EMPHASIS ON RESULTS

Unlike modern U.S. patents, early American patents were not uniform rights granted by a specialized agency with the sole job of reviewing and granting patents. Rather, they were individualized semi-contractual arrangements between the sovereign and the patentee created in statutes by the legislatures and assemblies. Like U.S. patents, the resulting exclusive rights were of limited term length; but the terms were not standardized, instead ranging anywhere from seven to around twenty years, depending on the estimated cost and risk of developing and implementing the technology

51. 2 RECORDS OF THE GOVERNOR AND COMPANY OF MASSACHUSETTS BAY IN NEW ENGLAND, 1642–1649, at 125–26 (Nathanial B. Shurtleff ed. 1853–1854); *see also* BUGBEE, *supra* note 7, at 62 (noting various other encouragements provided for the company).

52. *See* Patent Act of 1790, ch. 7, 1 Stat. 109 (repealed 1793); *see also* Patent Act of 1793, ch. 11, 1 Stat. 318 (repealed 1836).

53. *See* BUGBEE, *supra* note 7, at 57. Bugbee explained:

As in Elizabethan and Jacobean England, true patents of invention (in the present American sense, as involving originality) constituted only a small proportion of these colonial grants or awards. Unlike the English patents of invention, which were issued by royal grace and favor under the Prerogative of the sovereign, the American patents consisted almost entirely of private enactments of colonial legislatures in behalf of individual inventors, and included varying provisions and terms of effectiveness.

Id.; *see also* Bracha, *supra* note 7, at 97 n.222 (also noting a major difference from early British privileges was that colonial patents were created by the assemblies and legislatures, not the Crown).

involved.⁵⁴ The state legislature could extend the patent's term if further protection from competition was necessary to recoup costs or produce beneficial results,⁵⁵ or it could retract the patent entirely if the patentee failed to meet his end of the bargain.⁵⁶

Nor was novelty a necessary requirement for obtaining a patent. Whereas the U.S. Patent Act mandates full disclosure of an absolutely new invention in order to obtain a patent and employs various means for policing this requirement,⁵⁷ the colonies followed the English practice of defining "invention" more broadly as introducing a new trade or industry into the realm.⁵⁸ Although the legislatures did sometimes grant patents on ostensibly new inventions, these were not reliably distinguishable from other statutory monopolies given to entrepreneurs, developers, or inventors who imported inventions from abroad.⁵⁹ Likewise, disclosure to the public of the technological know-how required to practice an invention was not an essential element for obtaining a patent. Instead, the primary requirement was investing the time, money, and labor required to establish a working technology that produced beneficial results for the community.⁶⁰

Although this strategy is at odds with modern patent law's emphasis on universal novelty and the disclosure of new information, the colonies' emphasis on introduction rather than invention of known technology makes sense when viewed in historical context. The colonies were short on capital

54. See Bracha, *supra* note 7, at 101 (stating terms oscillated between seven and twenty years).

55. For instance, in 1652, Massachusetts Bay granted John Clarks a three-year monopoly over use of his invention saving firewood and heating rooms and after this term expired extended it for life. See BUGBEE, *supra* note 7, at 64.

56. See discussion *infra* Section II.B (discussing working clauses).

57. See 35 U.S.C. §§ 102 (novelty), 103 (non-obviousness) & 112 (disclosure and enablement) (2011).

58. Bracha, *supra* note 7, at 13, 19 (on English definition of invention), 99 (concept of invention employed in colonies that of the introduction of a new trade or industry).

59. In his survey of colonial and state patents, Bruce Bugbee has attempted to extract "true patents of invention" that appeared to claim original subject matter; but he concedes that there is no reliable way to tell whether the invention was truly new or had been used back in England. See BUGBEE, *supra* note 7, at 57–83; see also Bracha, *supra* note 7, at 99 ("Like the early English grants colonial patents for invention were not conceived of as forming a separate well differentiated channel for stimulating economic growth.").

60. Other scholars have noted that early patent privileges prioritized local establishment of a working invention or trade far more than disclosure of technological know-how; and that the shift in focus can be linked to the requirement that the inventor produce a specification designed to permit replication of the invention by one of ordinary skill in the art. See Biagioli, *Patent Republic*, *supra* note 47, at 1138; MERGES & DUFFY, *supra* note 43, at 6.

and in desperate need of industrial development.⁶¹ The colonists had access to plentiful resources but recognized that they would have to rely on technology, along with hard work, to convert this natural wilderness into a civilization comparable to that which they had enjoyed back home.⁶² They strove hard to transform America's plentiful natural resources into things that English people could use, like tobacco, flour, bread, furs, whale oil, and iron, and to develop the essential facilities that defined a civilized nation: agricultural fields, mills, sugar refineries, tanneries, salt works, and public utilities like ferries, bridges, and roads.⁶³

In order to accomplish these goals, the colonies used a variety of policies and institutions designed to give the colonists private incentives to invest and work hard.⁶⁴ Granting patents on technology was part of this strategy.⁶⁵ But colonial legislatures saw no clear reason to favor those who came up with entirely new ideas. Although they could obviously develop industry by encouraging original invention, the "safer road" was simply to copy what was already proven to work somewhere else.⁶⁶ We might object that, since these early patents did not primarily protect novel ideas from copyists or promote disclosure of new information, it would have been more effective for the government simply to offer subsidies or rewards to those who applied

61. For an empirical study addressing economic effects of the limited availability of credit in New England prior to the issuance of public currency, see Claire Priest, *Currency Policies and Legal Development in Colonial New England*, 110 YALE L.J. 1303, 1317–35 (2001).

62. See DARON ACEMOGLU & JAMES A. ROBINSON, *WHY NATIONS FAIL: THE ORIGINS OF POWER, PROSPERITY, AND POVERTY* 19–28 (2012) (arguing that the American colonies, unlike Spanish colonies in the same period, strove to become self-sufficient, rather than relying solely on plundering and indigenous labor, and developed economic policies to make the colonists themselves invest and work hard). *But see* MICHAEL ADAS, *DOMINANCE BY DESIGN* 35–45 (2006) (arguing that the colonists' mission to develop "the tools and skills needed to put the continent's abundant resources to proper use" went hand in hand with their mission to subdue and civilize the natives, whom they saw as undeserving of America's abundant resources).

63. For a list of the goods and facilities integral to the colonies' domestic economies, see Richard B. Sheridan, *The Domestic Economy*, in *COLONIAL BRITISH AMERICA: ESSAYS IN THE NEW HISTORY OF THE EARLY MODERN ERA* 43–85, 58 (Jack P. Greene & J.R. Pole, eds., 1984). In the early years, the colonists' end goal was subsistence, though in time they began to generate surplus and colonial economies became more export-driven. They came to rely on the income generated from shipping raw materials and semi-processed goods back to England. See Jacob M. Price, *The Transatlantic Economy*, in *COLONIAL BRITISH AMERICA: ESSAYS IN THE NEW HISTORY OF THE EARLY MODERN ERA* 18–42, 27–30 (Jack P. Greene & J.R. Pole, eds., 1984); see also Acemoglu & Robinson, *supra* note 62, at 26.

64. See Bracha, *Owning Ideas*, *supra* note 7, at 99; ACEMOGLU, *supra* note 62, at 19–28.

65. See Bracha, *Owning Ideas*, *supra* note 7, at 99 (explaining that patents in the American colonies were one of many tools used to promote economic growth).

66. See WALTERSCHEID, *THE NATURE OF THE INTELLECTUAL PROPERTY CLAUSE*, *supra* note 18, at 45–46.

known ideas to the task of implementing a particular local project.⁶⁷ And the colonial legislatures did use a “whole arsenal of tools and tactics” for promoting economic development.⁶⁸

But, historically, patents, not direct cash payments, have been the preferred means for governments with limited resources to reward investment in relatively new and unproven technology.⁶⁹ Patents are, at least initially, cheaper, costing taxpayers nothing up front and shifting the burden of financing the project to the private sector. If the technology fails or proves impossible to market, the state loses nothing since the patent will be completely valueless; if the technology shows promise but the patentee is an inefficient developer, the state legislature can simply retract the patent and transfer it to a more skilled operator under a local working clause. Though patents may ultimately tax the public in the form of monopoly prices, they are arguably a more accurate way to determine the appropriate amount for this tax than a sum supplied in a reward or contract, particularly where the

67. Subsidies like cash grants or tax incentives are generally given prior to development of a technology and cover the cost of development and deployment. The major downside is that it is difficult to value the technology prior to development, and the risk of failure is entirely on the government, not the innovator. See Cédric Schneidery, *Mixed R&D incentives: the effect of R&D subsidies on patented inventions 1* (Copenhagen Business School, Department of Economics, Working Paper No. 06-2008, 2008) (citing further secondary literature), available at <http://openarchive.cbs.dk/bitstream/handle/10398/7662/wp6-2008.pdf?sequence=1>; see also SUZANNE SCOTCHMER, *INNOVATION AND INCENTIVES* (2004). Rewards and prizes, in contrast to subsidies, are granted after the invention is developed. Major downsides are the requirement of paying cash directly and, again, the difficulty of valuing the invention, especially if the reward is given prior to full deployment or commercialization. See Steven Shavell & Tanguy Van Ypersele, *Rewards Versus Intellectual Property Rights*, 44 J.L. & ECON. 525 (2001).

68. As Bracha explained:

The colonies striving to promote the public good, to accommodate various interests, and in some periods to ensure their very survival and their basic needs employed a whole arsenal of tools and tactics . . . [including] bonuses, prizes, subsidies, payment of salaries to skilled artisans, loans, permissions to hold lotteries for raising funds, exemption from tax and military service and also grants covered all kinds of enterprises and manufactures, everything from mills to iron works or the operation of ferries.

Bracha, *supra* note 7, at 99–100.

69. Adam Smith, who otherwise generally opposed to state-sponsored monopolies, defended England’s policy of granting fourteen-year exclusive privileges to inventors. He believed it to be a more accurate and efficient way to value unknown technology than direct rewards. ADAM SMITH, *WEALTH OF NATIONS*, quoted in MERGES & DUFFY, *supra* note 43, at 8.

market value of the technology is entirely unknown before development.⁷⁰ Meanwhile, the patentee has a much stronger incentive to succeed than if the cash had already been paid or promised up front because they gain nothing from the prospect of a monopoly on a useless technology and their gains increase the more successful the technology becomes.⁷¹

B. LIMITING PRINCIPLES ON COLONIAL AND STATE PATENT LAWS

Obviously, state patents, like any legal monopoly, risked generating various social costs: the legal ability to charge monopoly prices, wasteful competition among different inventors vying for a valuable economic right from the state, lobbying by rights holders to protect or extend their rights (“rent-seeking”), and inhibition of future innovation by excluding others from experimenting and competing in the same intellectual space.⁷² This is why the federal government today limits patents by requiring that the patentee claim a novel and nonobvious invention and disclose it to the public.⁷³ The colonial and state legislatures also recognized the risks of patent rights.⁷⁴ They imposed various restrictions on patents to ensure that their

70. See Nancy Gallini & Suzanne Scotchmer, *Intellectual Property: When is it the best incentive system?*, in 2 INNOVATION POLICY AND THE ECONOMY 51, 70 (Adam Jaffe, Joshua Lerner & Scott Stern eds., 2002). Gallini and Scotchmer concluded:

IP is probably the best mechanism for screening projects when value and cost are not observable by the sponsor, since the private value of IP reflects the social value, and firms automatically compare some measure of value to the cost of innovation. In addition, IP encourages firms to accelerate progress, since the reward is conditional on success. Prizes could serve the same purposes if the size of the prize could be linked to the social value and without the deadweight loss of monopoly pricing.

Id.

71. A major utilitarian justification for patents is that they generate strong incentives not just to generate an invention but also to make further investments in developing and marketing it in order to enhance the value of their monopoly. See Eisenberg, *supra* note 1, at 1036–45 (citing Edmund Kitch, *The Nature and Function of the Patent System*, 20 J.L. & ECON. 265 (1977)); see also Gallini & Scotchmer, *supra* note 70, at 70.

72. See Dam, *supra* note 46, at 249–53 (enumerating three costs of a monopoly grant); see also Mark Lemley, *Property, Intellectual Property, and Free Riding*, 83 TEX. L. REV. 1031, 1058–64 (2005) (summarizing costs that an IP system imposes on a society).

73. See 35 USC §§ 102 (novelty), 103 (nonobviousness), 112 (U.S. patent law’s disclosure and enablement requirements) (2011).

74. At least two colonies had their own versions of England’s Statute of Monopolies (1624), which limited the Crown’s ability to grant monopolies to favored individuals. See Statutes at Large, 1624, 21 Jac., c.3 (Gr. Brit.). For instance, in 1641, the General Court of Massachusetts Bay added to its “Body of Liberties” a clause stating: “No monopolies shall be granted or allowed amongst us, but of such new inventions yt are pftable to ye Countrie, & yt [sic] for a short time.” See BUGBEE, *supra* note 7, at 61.

exclusive rights actually resulted in beneficial innovation and did not impede access to the building blocks of knowledge.

At the most general level, the state legislatures tailored the terms of each statutory patent grant based on the legislature's assessment of the benefits that the technology was likely to produce for the community, on the one hand, and the likely costs to the public and the patentee, on the other.⁷⁵ In addition, each patent came with significant *quid pro quos* for the patentee—and not just the requirement that he disclose his art to the public. First, to reduce the risk of patents entirely preempting competition in the same market, or all innovative activity in the same intellectual space, colonial and state patents nearly always specified a statutory maximum penalty for infringement.⁷⁶ This essentially compelled patentees to license their technology to anyone who wished to use it for a set fee.⁷⁷ Some patents even employed direct price controls, limiting the rates the patentee could charge for use or sale of the invention.⁷⁸

By far the most significant *quid pro quo* was the so-called “working clause,” providing that if the patentee did not develop and actually start practicing his invention within a certain number of years his grant would lapse.⁷⁹ For instance, in 1706, Massachusetts Bay gave Thomas Houghton “a Patent for the Improvement of the Whale Flesh,” giving him ten years’ exclusive privilege to make saltpeter out of the lean whale flesh discarded by

75. As Bracha puts it, the decision to grant a patent in any given case involved a “specific calculus of the public good”:

In a fashion similar to the early English grants applicants usually detailed specific tangible benefits offered by their inventions such as lower prices, the supply of a scarce commodity or the saving of labor. . . . The assembly was in charge of a specific calculus of the public good in each case, considering whether a patent was justified and what its specific terms should be.

Bracha, *supra* note 7, at 100–01.

76. See the state patents catalogued in BUGBEE, *supra* note 7, at 84–103.

77. Some states also used what Bugbee calls “abolition privileges” to achieve a similar result. BUGBEE, *supra* note 7, at 96. For instance, in 1785 Virginia gave a ten-year patent to James Rumsey for his steam engine, but included a provision that the assembly could cancel Rumsey’s patent at any time by paying £10,000 in gold or silver. *Id.* Bugbee also notes that “[s]everal state awards provided for their own repeal should it be known that someone other than the grantee was the first and true inventor. . . .” *Id.* at 103.

78. See BUGBEE, *supra* note 7, at 64, 90 (noting that “this principle of price control had already appeared in some early English awards and in Joseph Jenk’s Massachusetts patent of 1646”); see also *id.* at 89–90 (describing 1775 Pennsylvania Assembly grant to Donaldson on new machine for cleaning docks allowing the legislature to set the rates charged for the dredged material).

79. See Bracha, *supra* note 7, at 102.

Cape Code whalers.⁸⁰ In exchange, Houghton was required to “disclose and make publick his art” after the ten years elapsed; but he was also required to demonstrate within four years that he was working his method successfully in the colony.⁸¹ Even as states began granting more patents over truly new inventions in the 1780s, disclosure of new principles alone would not justify a patent. Legislatures still required the patentee, as a primary matter, to actually establish a working technology in the state in order to deserve a monopoly.⁸² Otherwise, the legislatures would retract or transfer the patent to a more deserving grantee.⁸³

80. BUGBEE, *supra* note 7, at 67.

81. *Id.*

82. For example, in 1780, New York granted Henry Guest a five-year patent over the making of the following:

[a] certain species of blubber and oyl proper for currying and dressing of leather made from ingredients that “abound[ed]” in the state of New York. In exchange, Guest was required to disclose his invention to the public, submitting “a writing containing the names and descriptions of the materials aforesaid, and the method and process of making such blubber and oyl” But the patent also specified that it would not take effect “until the said Henry Guest shall have a manufactory erected for the purpose, and shall have made blubber or oyl [sic], of the materials aforesaid, within this State.

Id. at 87–88. Another telling example of the longevity of this strict local working requirement is the patent that New Hampshire granted to John Young for a chimney design in 1791, after the first Patent Act of 1790 had gone into effect. *See id.* at 101–02.

83. The issue of whether a state could have retracted a patent after ratification under the Constitution and specifically the Contracts Clause was not actually tested. But New York’s retraction of Fitch’s 1787 grant in 1798 suggests that the state could have rescinded an inventor’s patent for failure to establish a working technology in state jurisdiction even when the patent contained no explicit working clause. *Livingston v. Van Ingen*, 9 Johns. 507 (N.Y. 1812). Cases from the nineteenth century involving public utilities monopolies also support that the state would have been free to retract a patent and transfer the exclusive rights to another. *See, e.g., Chenango Bridge Co. v. Paige*, 83 N.Y. 178 (1880) (holding that if a toll bridge with a monopoly grant is so far out of repair as not to afford a passage with ordinary convenience and safety, a rival bridge, even though within a previously prohibited distance, is not unlawful). Bracha reaches a similar conclusion:

Unlike the parallel issue of corporate charters there was no legal dispute that directly involved the question of whether the legislature could revoke a patent grant, but from a few instances in which such revocations happened it seems that at least during the eighteenth century the assumption was that the legislature could take away what it granted.

Bracha, *supra* note 7, at 110 n.251.

III. THE INCOMPLETE TRANSITION TO A NATIONAL PATENT SYSTEM

During the short period between independence in 1776 and ratification of the U.S. Constitution in 1788, during which the Articles of Confederation established a confederation of thirteen states of independent sovereignty, states continued to use limited monopolies as a cheap and efficient way to encourage development and implementation of useful technologies in their jurisdictions.⁸⁴ Although the state patents granted in the 1780s more often involved truly novel inventions and made disclosure a requirement, the same general features and underlying framework continued to guide their practices.⁸⁵ Nevertheless, despite potential benefits to states' local economies, the increasingly *interstate* nature of commerce heralded the end of state patents.

A. "THE STATES CANNOT SEPARATELY MAKE EFFECTUAL PROVISION . . ."

During the colonial period, trade had been intra-colonial or with foreign nations. There was no common currency among the different states.⁸⁶

84. On state patents generally, see Bracha, *supra* note 7, at 109–16. See also BUGBEE, *supra* note 7, at 84–103. South Carolina mentioned patents in its 1784 copyright statute, providing that “inventors of useful machines shall have a like exclusive privilege of making or vending their machines for the like term of fourteen years, under the same privileges and restrictions hereby granted to and imposed on the authors of books.” See Federico, *State Patents*, *supra* note 42, at 166.

85. As Bugbee explained:

A new outpouring of provincial patents of invention came with the 1780's, and was largely confined to that decade. In character it was essentially a continuation of the colonial practice of enacting private laws of varying provisions in favor of individual inventors, but to pre-Revolutionary patent features were added some significant improvements.

BUGBEE, *supra* note 7, at 84. Similarly, Bracha wrote:

On the eve of creating the American patent regime in 1790 the American patent grant practice was still rather similar on both the practical and the conceptual level to the traditional English framework. In fact, it was much closer to the English origin than its contemporaneous British counterpart. Colonial and state patents were individual privileges granted as a result of a case-specific policy-political decision by government in the name of the public good. Although this aspect was beginning to change late in the eighteenth century, patents were also conceived and practiced not as creating “ownership” in an intellectual-informational entity called an “invention;” but rather as commercial privileges to exercise a “trade.”

Bracha, *supra* note 7, at 109, 401.

86. See JOEL ACHENBACH, *THE GRAND IDEA: GEORGE WASHINGTON'S POTOMAC AND THE RACE TO THE WEST* 11–12 (2004). Colonial governments did begin issuing paper

Transportation of people and goods across state lines was slow and expensive. The country was full of impassable mountains, dense forests, and rivers that, for all practicable purposes, were not navigable. There were few bridges, and people relied on ferries.⁸⁷ Bodies of water were barriers to communication and trade between developed communities in distinct jurisdictions.⁸⁸ Under these conditions, a patent covering local use of an invention was a valuable asset. Even assuming the patentee actually discovered infringements outside his jurisdiction, this would not have significantly affected his ability to profit in local markets. If the invention was a method, such as a process for grinding corn using a water mill,⁸⁹ or an agricultural tool, such as an “engine” for husking rice,⁹⁰ it was likely to be connected to local resources and to require the establishment of a stable operation in the colony. Even if the invention was a product, traveling to another jurisdiction to find an infringing alternative would have been incredibly costly.

The development of new modes of production, better infrastructure, and the more rapid steam-powered transportation completely altered this situation.⁹¹ The steam engine allowed for cheaper and quicker production of goods, and also enabled vendors to more efficiently transport and market their wares across state borders. For the first time, developed economies of

money in their own jurisdictions during the period from 1690 to 1710. *See* Priest, *supra* note 61, at 1303.

87. Achenbach, *supra* note 86, at 14, 28.

88. *See* KUTLER, *supra* note 39, at 6 (stating that the river villages became “centers for intracolonial trade and depots for foreign commerce.”).

89. Many colonial grants involved mills. For instance, in 1646, the General Court of Massachusetts granted Peter Jenks a fourteen year patent to “[b]uild a mill for making of Sithes; and alsoe a new jnvented Saw Mill, and divers other Engines for making of divers sorts of Edge tooles” *See* BUGBEE, *supra* note 7, at 62. In 1722, Connecticut granted Ebinezar Fitch and company “the sole right, privilege and liberty to erect, use, maintain and support, a slitting mill or slitting mills within this Colony of Connecticut during [the space of fifteen years,]” provided that they “erect and set up a good, sufficient slitting mill” on the Stony Brook river or in “some place within this Colony,” within three years, and keep it in good repair. CONNECTICUT RECORDS, *supra* note 44, at 312–13; *see also* BUGBEE, *supra* note 7, at 181 n.34.

90. In 1691 South Carolina passed an act granting a two-year patent to “Mr. Peter Jacob Guerard” for a “pendulum engine” for husking rice “much better, and in lesse time and labour . . . than any other heretofore . . . used within this Province,” with a set infringement penalty of forty shillings. *See* BUGBEE, *supra* note 7, at 75.

91. *See* Nye, *supra* note 6, at 120–121; Adas, *supra* note 62, at 76–77; *see also* Achenbach, *supra* note 86, at 238 (“Commerce required better transportation networks, and better transportation networks led to increased commerce, and soon the whole system was heating up like a steam boiler.”), 238–39 (describing steamboats running on the Potomac).

different jurisdictions began interacting fluidly side-by-side.⁹² Infringing uses of patented inventions outside a state's jurisdiction would have been more readily visible. In many cases, such out-of-state infringing use would have immensely reduced the value of a state patent right because consumers could simply purchase infringing products or products made from infringing methods in other states. As a result, inventors in the 1780s started purchasing patents from adjoining states to consolidate the jurisdictional scope of their protection.⁹³

Inventors soon began lobbying for a nationwide patent system to more effectively protect their rights, for obvious reasons.⁹⁴ National patents provided superior protection against copyists and potentially entitled inventors to greater profits by enabling them to sell and license their inventions in multiple states. This was especially true for consumer goods, such as the clocks and watches proliferating in this period, which could be easily copied and sold in interstate markets.⁹⁵ Also, unlike a decentralized system of individualized state statutes, a uniform national system would be more likely to produce consistent decisions and, ideally, would put an end to

92. See BORIS BITTKER, BITTKER ON THE REGULATION OF INTERSTATE AND FOREIGN COMMERCE 1.02[A], at 1–8 (citing Albert S. Abel, *Commerce Regulation Before Gibbons v. Ogden: Interstate Transportation Facilities*, 25 N.C. L. REV. 121, 122 (1947)).

93. See Walterscheid, *Priority of Invention: How the United States Came to Have a "First to Invent" Patent System*, 23 AIPLA Q.J. 263, 277 (1995) (discussing priority battles between steam boat inventors attempting to buy up patents from various states and lobbying by inventors); see also MERGES & DUFFY, *supra* note 43, at 7 & nn.17–18.

94. See Michael Martin, *The End of the First-to-Invent Rule: A Concise History of Its Origin*, 49 IDEA 435, 461–67 (2009) (describing the influence of inventors like John Fitch, advocating for federal patents to protect their natural rights in their inventions). Walterscheid has succinctly summed up the weaknesses of the state-level system:

The most singular defect was that states could only legislate with respect to their own territory. Thus, state patents and copyrights could be infringed with impunity in adjoining states. Getting multiple patents or copyrights was time consuming, expensive, and frequently frustrating. Moreover, there was no certainty of consistency in terms and conditions from state to state.

WALTERSCHEID, THE NATURE OF THE INTELLECTUAL PROPERTY CLAUSE, *supra* note 18, at 76.

95. On the patent granted by Pennsylvania in 1789 to Robert Leslie for various improvements on clocks and watches, see BUGBEE, *supra* note 7, at 91; see also *id.* at 88 (describing a 1783 Connecticut patent over a wind-up clock preventing anyone from constructing, importing or selling such a clock in the state). For inventions like the steamboat that had to be operated within the boundaries of a particular state, the value of the patent was reduced, though not necessarily eliminated, by competitors' use of the invention in other states. Local public utilities monopolies are still valuable today.

never-ending rivalries over priority of invention and infringement.⁹⁶ The inventors found a captive audience for their plea in the nation's policymakers. According to lore, on August 20, 1787, two days after Madison first suggested adding the IP Clause to the Constitution, several members of the Constitutional Convention gathered on the banks of the Delaware River to watch a demonstration of Fitch's latest steamboat model.⁹⁷ No one appeared to disagree with James Madison's famous statement in the Federalist No. 43 that "[t]he States cannot separately make effectual provision for either of the cases [patent or copyright.]"⁹⁸ Thereafter, the Framers added the IP Clause to the Constitution with little discussion of the matter.⁹⁹

B. THE EVIDENCE FOR STATES' REMAINING PATENT POWERS

Over four years between 1787 and 1791, the states ratified the Constitution, thereby accepting the terms of Congress's new power of "securing for limited Times to . . . Inventors the exclusive Right to their respective . . . Discoveries."¹⁰⁰ On April 10, 1790, Fitch's prayers for a national patent on his steamboat were finally answered when Congress passed the first federal Patent Act, giving inventors the opportunity to obtain a fourteen-year period of exclusive rights in their inventions, nationwide.¹⁰¹ Once U.S. patents became available, inventors enthusiastically began applying

96. See Zorina B. Khan, *Property Rights and Patent Litigation in Early Nineteenth-Century America*, 55 J ECON HIST. 58, 62 (1995) ("Consistent regional decisions would serve to increase the value of holding a patent; first, by expanding the coverage of the patent to a much wider market; and second, by eliminating the uncertainty and costs of enforcement if litigation were governed by the laws of the individual states.").

97. See COX, *supra* note 21, at 1–2; see also JAMES THOMAS FLEXNER, *STEAMBOATS COME TRUE: AMERICAN INVENTORS IN ACTION* 126–27 (1944).

98. THE FEDERALIST NO. 43, at 279 (James Madison) (Earle edition, 1938). In a speech at the Pennsylvania ratifying convention, Thomas McKean echoed this concern, stating that "[t]he power of securing to authors . . . the exclusive right to their writings . . . could only with effect be exercised by the Congress." Thomas McKean, *Speech at Pennsylvania Ratifying Convention* (Dec. 10, 1787), *quoted in* WALTERSCHEID, *THE NATURE OF THE INTELLECTUAL PROPERTY CLAUSE*, *supra* note 18, at 10.

99. See *id.* at 110; see also Bracha, *Owning Ideas*, *supra* note 7, at 272 n.2 (and citations therein to literature on the history of the passage of the IP Clause).

100. U.S. CONST. art. I, § 8, cl. 8. On ratification, see THE DOCUMENTARY HISTORY OF THE RATIFICATION OF THE CONSTITUTION (John P. Kaminski and Gaspare J. Saladino eds., Wisconsin Historical Society Press 1981).

101. See Patent Act of 1790, ch. 7, 1 Stat. 109 (repealed 1793). The Act was revised three years later. See Patent Act of 1793, ch. 11, 1 Stat. 318 (repealed 1836).

for them.¹⁰² As noted in the introduction, the prevailing assumption is that the enactment of a national patent law “was largely viewed as removing the need for state patents”¹⁰³ However, in light of the evidence, it is apparent that at the time of ratification states were perceived as retaining significant constitutional authority to grant patents alongside Congress and that these views persisted throughout the nineteenth century.

1. *The Text of the IP Clause*

The IP Clause, which is included in the list of powers expressly granted to Congress in Article I, Section 8 of the Constitution, provides that: “The Congress shall have Power . . . [t]o 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”¹⁰⁴

The Clause gives Congress an independent power to “secur[e]” the “exclusive Right” of “Inventors” in their “Discoveries[]” for a limited period.¹⁰⁵ However, it does not suggest that this power is exclusive to the federal government or that it excludes states from granting patents within their own jurisdictions.¹⁰⁶ As Maya Pollack has noted, the IP Clause “gives Congress ‘power,’ not ‘*the* power,’ to promote science and the useful arts by granting exclusive rights—even though ‘*the*’ had appeared in earlier drafts.”¹⁰⁷ This is the case for all of the powers listed in Article I, Section 8, and some were exclusive grants to Congress.¹⁰⁸ But unlike Clause 4, granting Congress power to “establish an *uniform* Rule of Naturalization, and *uniform* Laws on the subject of Bankruptcies throughout the United States,”¹⁰⁹ the IP Clause does not state or imply that American patent laws must be

102. See Bracha, *Owning Ideas*, *supra* note 7, at 408 n.23 (stating that 57 patents were issued under the 1790 act but that at least 114 applications were received); see also COX, *supra* note 21, at 15 (further discussing inventors’ interest in federal patents).

103. See WALTERSCHEID, THE NATURE OF THE INTELLECTUAL PROPERTY CLAUSE, *supra* note 18, at 436–37; see also *id.* at 76–77, 438–42; BUGBEE, *supra* note 7, at 102–03 (also suggesting that federal patents made state patents obsolete).

104. U.S. CONST. art. I, § 8, cl. 8.

105. *Id.*

106. *Id.*; see also *Gibbons v. Ogden*, 22 U.S. 1, 45 (1824) (noting Thomas Oakley’s oral argument that that the patent power is “not granted exclusively to Congress. No exclusive terms are used. The grant is affirmative and general, like all the other powers.”); *Goldstein v. California*, 412 U.S. 546, 560 (1973) (“[T]he language of the Constitution neither explicitly precludes the States from granting copyrights nor grants such authority exclusively to the Federal Government.”).

107. Pollack, *supra* note 17, at 301 (emphasis in original).

108. U.S. CONST. art. I, § 8, cl. 4 (bankruptcy and naturalization powers).

109. *Id.* (emphasis added).

“uniform.”¹¹⁰ Therefore, the simple fact of non-uniformity does not necessarily suffice to divest states of their authority to concurrently grant patents on inventions practiced in the state—just as it does not deprive states of the power to concurrently levy their own taxes on state residents.¹¹¹

2. *The Patent Act’s “Relinquishment” Provision*

The Patent Act, unlike the Copyright Act, has no express preemption provision,¹¹² but Congress added a provision to the second Patent Act in 1793 that explicitly addressed state patents.¹¹³ The provision did not preempt state patents or preempt states from granting patents.¹¹⁴ Instead, the provision only required inventors who obtained U.S. patents to “relinquish” any patents they possessed that had been granted by a state “*before its adoption of the [Constitution.]*”¹¹⁵ Edward Walterscheid contends:

The limitation proposed by Jefferson and accepted by Congress can only be explained on the supposition that they believed that once a state ratified the Constitution it no longer had authority to

110. See THE FEDERALIST NO. 32, at 156 (Alexander Hamilton) (Earle ed. 1938) (noting that, unlike with the concurrent power to levy taxes, the federal government’s power “to establish an UNIFORM RULE of naturalization throughout the United States[] . . . must necessarily be exclusive; because if each State had power to prescribe a DISTINCT RULE, there could not be a UNIFORM RULE.”); see also U.S. CONST. art. I, § 8, cl. 8 (granting Congress power “[t]o 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.”).

111. See THE FEDERALIST NOS. 32 (Alexander Hamilton) (discussing the right of states to levy taxes).

112. The Copyright Act of 1976 has an express preemption provision. 17 U.S.C. § 301 (2011); see also 35 U.S.C. §§ 101–102 (2011) (Patent Act, lacking an express preemption provision).

113. Patent Act of 1793, Ch. 11, § 7, 1 Stat. 322 (Feb. 21, 1793). The provision was the brainchild of Secretary of State Thomas Jefferson, one of the trio of officials in charge of reviewing federal patent applications under the 1790 Act. Patent Act of 1790, ch. 7, 1 Stat. 109, 109–12 (repealed 1793) (naming three executive officers to review patents); see also WALTERSCHEID, THE NATURE OF THE INTELLECTUAL PROPERTY CLAUSE, *supra* note 19, at 437.

114. Patent Act of 1793, Ch. 11, § 7, 1 Stat. 322 (Feb. 21, 1793).

115. The provision stated:

Where any State, *before its adoption of the present form of government*, shall have granted an exclusive right to any invention, the party claiming that right shall not be capable of obtaining an exclusive right under this act, but on *relinquishing* his right under such particular State, and such *relinquishment*, his obtaining an exclusive right under this act, shall be sufficient evidence.

Patent Act of 1793, Ch. 11, § 7, 1 Stat. 322 (Feb. 21, 1793) (emphasis added); see also WALTERSCHEID, THE NATURE OF THE INTELLECTUAL PROPERTY CLAUSE, *supra* note 19, at 437 (discussing the 1793 provision).

issue its own state patents. That is, they believed that the language of the intellectual property clause preempted state authority to issue patents (and presumably copyrights as well). Otherwise, there would have been no good reason to limit the surrender requirement to state patents issued before ratification and not include those issued after ratification.¹¹⁶

This assumption cannot safely be made. In this period, many contested the issue of whether, and to what degree, states could continue their prior patent practices.¹¹⁷ The simpler and better explanation is that Jefferson did not address the constitutional or statutory viability of state patents issued after ratification because he knew states might want to keep granting their own patents. Instead, the relinquishment requirement had a very specific and practical purpose. In 1790, many inventors who had several state patents that were still in force also eagerly sought U.S. patents for the same inventions.¹¹⁸ Congress did not feel empowered to simply void the inventors' vested state rights without voluntary action by state patentees, but this created a problem because it would be confusing and unfair if applicants for U.S. patents were allowed to rely on both state and federal patent rights simultaneously.¹¹⁹ Thus, Jefferson sought to make it clear in the Patent Act that applicants must choose between federal or state patent rights.¹²⁰ Jefferson could have decided to expressly preempt any state patents granted in the future if he believed preemption was warranted. But his decision to simply require a choice between the two regimes easily supports the opposite conclusion: that he

116. *Id.* at 438 (citations omitted).

117. As Walterscheid himself goes on to note, the state of New York did grant a patent *five years after the provision went into effect*. *See id.*

118. *See COX, supra* note 21, at 15 (noting four steamboat inventors with several state patents who applied for patents).

119. Supporting this view is that Van Ingen's attorneys, in arguing to invalidate New York's steamboat patent, conceded that Congress could not simply revoke vested state patent rights issued prior to ratification, at least not without a voluntary act by a patentee who chose to exchange his state rights for federal ones. *See Livingston v. Van Ingen*, 9 Johns. 507, 540–41 (N.Y. 1812) (Messrs. Wells and Henry Van Vechten for the majority, and Messr. Van Vechten concurring, finding in favor of Van Ingen et al.).

120. An earlier draft of the provision submitted by Jefferson shows that his reason for requiring relinquishment of a state patent was indeed to ensure that patentees under the new federal Act would obtain "equal benefits" and also be "subject to equal restrictions with the other Citizens of the United States . . ." *See WALTERSCHEID, THE NATURE OF THE INTELLECTUAL PROPERTY CLAUSE, supra* note 18, at 437 (citing Jefferson's early draft of the provision).

believed states possessed independent authority to grant their own patents under the Tenth Amendment.¹²¹

3. *The Decision to Limit U.S. Patents to Universally Novel Inventions*

Further evidence that states retained concurrent patent powers comes from the fact that, when drafting the Patent Act of 1790, the First Congress deliberately limited U.S. patents to universally novel inventions.¹²² When the Second Congress revised the Patent Act in 1793, they kept this limitation in place.¹²³ It is still in place today.¹²⁴ But as Justice Story pointed out in 1825, Congress required that inventors do more than facilitate “the *mere importation*

121. See U.S. CONST. amend. X; see also *infra* note 214 (discussing George Tucker’s comments on this subject—which directly contradict Walterscheid’s contention that the relinquishment provision did not apply to state patents granted after ratification).

122. See Patent Act of 1790, ch. 7, § 1, 1 Stat. 109, 109–10 (repealed 1793). As Walterscheid has documented, an early draft of the first Patent Act was read into the House on February 16, 1790. The proposed bill allowed inventors to obtain patents for inventions “not before known or used”; that is, only universally novel inventions. The phrase was thereafter modified by “in the United States,” which would have allowed patents over imported inventions; and a new section was thereafter added providing that “the first importer of any art, machine, engine, device or invention, or any improvement thereon” should be treated the same as an “original inventor or improver within the United States.” However, with no recorded discussion on the matter, these revisions were deleted, and the original rule allowing only patents for universally novel inventions was reinstated in the final version of the Patent Act. See Walterscheid, *Patents and Manufacturing*, *supra* note 33, at 872–73.

123. Compare Patent Act of 1793, ch. 11, § 1, 1 Stat. 318, 318 (repealed 1836) (stating that the applicant must submit a petition “setting forth, that he, she, or they, hath or have invented or discovered any useful art, manufacture, engine, machine, or device, or any improvement therein *not before known or used*.”) (emphasis added), with Patent Act of 1790, ch. 7, § 1, 1 Stat. 109, 109–10 (repealed 1793) (limiting applications to the invention of “any new and useful art, machine, manufacture or composition of matter, or any new and useful improvement on any art, machine, manufacture or composition of matter, *not known or used before* the application”) (emphasis added).

124. There was a period where Congress gave some allowance for prior foreign activities. See, e.g., Patent Act of 1836, ch. 357, § 8, 5 Stat. 117, 121 (repealed 1861) (providing that an “original and true inventor” could obtain a U.S. patent despite having “previously taken out letters patent therefor in a foreign country, and the same having been published, at any time within six months next preceding the filing of his specification and drawings.”); see also *O’Reilly v. Morse*, 56 U.S. 62 (1853) (upholding the validity of Samuel Morse’s patent on the electro-magnetic telegraph despite allegations of prior inventorship abroad because “[a] previous discovery in a foreign country does not render a patent void, unless such discovery or some substantial part of it had been before patented or described in a printed publication.”); MERGES & DUFFY, *supra* note 43, at 494–95 (describing various forms of “domestic bias” in the Patent Act of 1952). However, as of September 16, 2011, a person is not entitled to a U.S. patent “if the claimed invention was patented, described in a printed publication, or in public use, on sale, or otherwise available to the public before the effective filing date of the claimed invention” 35 U.S.C. § 102(a)(1) (2011).

of a known machine” in order to obtain a patent—distinguishing U.S. patent law from both England’s and the states’ prior practices, which allowed inventors to obtain patents on inventions imported from other countries.¹²⁵ The reason why Congress decided to change policies is not entirely clear. Walterscheid suggests that James Madison probably insisted on the universal novelty requirement based on the IP Clause’s statement that Congress had power to grant exclusive rights to “*Inventors*” in their “*Discoveries*.”¹²⁶ Secretary of the Treasury Alexander Hamilton, for his part, had no qualms about patents for inventions that were already known abroad. In his *Report on Manufacturers* (1791), Hamilton famously proposed a variety of national incentives to promote investment, including national patents for “introducers” of foreign technology, “a policy which [had] been practiced with advantage in other countries.”¹²⁷ But even Hamilton reluctantly concluded that the “National Government” lacked authority to grant patents beyond the means specified in the IP Clause.¹²⁸

Why did Madison and Hamilton believe the Constitution imposed this limitation? One reason may be the belief that only original inventions, and only true inventors, were deserving of patent rights,¹²⁹ and that allowing

125. See *Earle v. Sawyer*, 8 F. Cas. 254, 256 (C.C. Mass. 1825); see also Walterscheid, *Patents and Manufacturing*, *supra* note 33, at 859 n.15 (noting that “[t]he British policy of obtaining foreign technology by this means was perceived to be highly successful.”); Bracha, *Onning Ideas*, *supra* note 7, at 492–93 (on continuing British practice of granting patents to importers), 493 (America’s unique decision to exclude patents of importation).

126. U.S. CONST. art. I, § 8, cl. 8 (emphases added); see also Walterscheid, *Patents and Manufacturing*, *supra* note 33, at 868–78.

127. See HAMILTON, REPORT ON MANUFACTURERS, *supra* note 33, at 1014; see also *id.* at 988–92, 1008–34 (recommending a variety of national incentives). On Hamilton’s policy agenda for promoting domestic industry and manufacturing through a variety of means including prizes for inventions and direct payments to businesses, see Wood, *supra* note 30, at 102; RON CHERNOW, ALEXANDER HAMILTON 377 (2004).

128. Hamilton stated that while “it is desirable, in regard to improvements and secrets of extraordinary value, to be able to extend the same benefit to introducers, as well as authors and inventors . . . there is cause to regret, that the competency of the authority of the National Government to the good which might be done, is not without question.” HAMILTON, REPORT ON MANUFACTURERS, *supra* note 33, at 1014; see also WALTERSCHEID, TO PROMOTE THE PROGRESS OF SCIENCE AND USEFUL ARTS: AMERICAN PATENT LAW AND ADMINISTRATION, 1798–1836, at 148–56 (1998) (discussing Hamilton’s proposal in his *Report on the Subject of Manufacturers* to grant patents of importation and Hamilton’s concerns about the constitutionality of such patents).

129. There are various lines of evidence supporting this hypothesis. First, scholars have observed that the purpose of U.S. patents was no longer solely to promote local working of technology, but also to promote disclosure of *new information* in a specification. See Mario Biagioli, *Patent Specification and Political Representation: How Patents Became Rights*, in MAKING AND UNMAKING INTELLECTUAL PROPERTY: CREATIVE PRODUCTION IN LEGAL AND CULTURAL PERSPECTIVE 25, 31–32 (Mario Biagioli, Peter Jaszi & Martha Woodmansee eds.,

patents on known foreign inventions would needlessly interfere with public access to know-how that was already disclosed to the public.¹³⁰ However, some research suggests an alternative explanation: that the decision to impose a unique requirement of universal novelty in U.S. patent law was not based exclusively on fear of government power to grant monopolies; rather, it was also based on fear of the federal government encroaching on the sovereign powers of the *states* to do so.¹³¹ The strength of this hypothesis will become clear once we examine the arguments and legal opinions generated by New York's infamous steamboat monopoly—which, I will show, can actually be considered the last state patent granted in America.

IV. THE LAST STATE PATENT

The most straightforward evidence that at least some policymakers believed states retained the power to grant patents is that several states, including Pennsylvania, New York, New Hampshire, Connecticut, and New Jersey, continued to do so.¹³² There is no denying that once U.S. patents became available in 1790, inventors enthusiastically began applying for them

2011); *see also* MERGES & DUFFY, *supra* note 43, at 6 (explaining that during the Industrial Revolution in England patent applicants were more frequently required to describe their inventions in exchange for a patent, and associating this development with Judge Mansfield's 1778 opinion in *Liardet v. Jobnson*, (1780) 62 Eng. Rep. 1000 (K.B.)). Second, scholars have noted a growing fixation in this period with the somewhat romantic notion of the invention and the inventor. *See, e.g.*, Bracha, *Owning Ideas*, *supra* note 7, at 430–31 (describing new preoccupation with invention and inventors and new focus on technological innovation, displacing traditional British definition of invention as a trade or industry not used before in the jurisdiction).

130. Walterscheid has unearthed evidence that some American businessmen feared that allowing patents on known foreign inventions would interfere with their ability to imitate important English inventions already disclosed in English patent specifications and publicly available to copy. *See* Walterscheid, *Patents and Manufacturing*, *supra* note 33, at 875–76 (discussing Richard Wells' apparently unused petition to Congress, submitted March 4, 1790, in opposition to a proposed amendment to the Patent Act that would have allowed "first importers" to be treated as original inventors).

131. Hamilton recognized that the power to grant patents to foreign introducers would theoretically remain with the states if Congress could not do so, writing that if federal patents of importation were allowed, "many aids might be given to industry, many internal improvements of primary magnitude might be promoted, by an authority operating throughout the Union; which cannot be effected as well, if at all, *by an authority confirmed within the limits of a single State.*" HAMILTON, REPORT ON MANUFACTURERS, *supra* note 33, at 1014 (emphasis added). As this Article later discusses, many believed states did have this power. *See* discussion *infra* Section V.B.

132. Crockett has also noted this fact. *See* Crockett, *supra* note 15, at 37 (using a hypothetical dissent by J. Frankfurter) (citing to Prager and Federico).

and rarely applied to state legislatures.¹³³ However, there are several reported cases of state patents issued after the Constitution was adopted,¹³⁴ and even after the Patent Act went into effect in 1790. For instance, in 1791, New Hampshire granted a patent to John Young for his chimney design, providing that should Young choose to obtain a U.S. patent instead, his state rights would be void.¹³⁵ In the same year, New Jersey issued a patent on a grist mill.¹³⁶

After Congress clarified in the 1793 Patent Act that state patentees had to give up their rights when they obtained a U.S. patent for the same invention, states' concurrent patent practices might have continued unnoticed. But then something happened that changed the trajectory of patent law in America. On March 27, 1798, the New York Legislature and Council of Revision, whose members included several participants in New York's federal ratifying convention, passed the first in a series of laws granting the exclusive right to navigate steamboats in New York waters to Robert Livingston and Robert Fulton.¹³⁷

133. See Bracha, *Owning Ideas*, *supra* note 7, at 408 n.23 (stating that 57 patents were issued under the 1790 act but that at least 114 applications were received). On the eagerness of inventors to apply for federal patent rights, see COX, *supra* note 21, at 15.

134. For instance, in 1789, Pennsylvania granted a patent to Robert Leslie for improvements in the mechanism of clocks and watches, which were also patented later under the first federal patent act. See Federico, *State Patents*, *supra* note 42, at 167–68. As Federico writes:

The patent to Leslie, it is to be noted, was granted after the Constitution was adopted. This was common among the states for a few years after the adoption. As instancing the overlapping of state and federal functions, the action of Samuel Briggs is cited. In 1789 he petitioned both Congress of the United States, and the General Assembly of Pennsylvania for a patent for his machine for making nails, screws and gimlets, and deposited with the executive of the state a model of the machine in a sealed box pending the outcome of either petition. He, together with his son, received the first patent for nail machinery under the federal patent law.

Id.; see also Crockett, *supra* note 15, at 37 (noting that some states, such as Pennsylvania, New Hampshire, New York, Connecticut, and New Jersey, continued to issue patents).

135. To avoid any confusion or injustice that might result if Young were to subsequently get a patent from Congress, the New Hampshire grant contained a clause requiring that Young relinquish his state rights if he should choose to obtain a U.S. patent. See BUGBEE, *supra* note 7, at 102 (quoting 5 LAWS OF NEW HAMPSHIRE 791 (Henry Harrison Metcalf ed., 1916)); see also Federico, *State Patents*, *supra* note 42, at 168; WALTERSCHEID, THE NATURE OF THE INTELLECTUAL PROPERTY CLAUSE, *supra* note 18, at 437 (describing the “voidance clause”).

136. Federico, *State Patents*, *supra* note 42, at 168.

137. Although a few members of the Council of Revision, including Governor John Jay and Chancellor Lansing, objected to the Act, none objected on constitutional grounds. See *Gibbons v. Ogden*, 22 U.S. 1, at 80; see also *infra* note 161 and accompanying text.

A. UNCOVERING A STATE PATENT AT THE HEART OF THE STEAMBOAT MONOPOLY

At first glance, the New York grant does not appear to be a patent, and few scholars have described it as one.¹³⁸ Livingston was not an inventor—he was a wealthy and well-connected businessmen and the former Chancellor of New York.¹³⁹ Fulton was a skilled engineer and would eventually develop original steamboat designs, but he copied liberally from the designs of John Fitch—who invented one of the first working steamboats in America and obtained one of the first U.S. patents on a steamboat in 1791.¹⁴⁰ On its face, the New York law was not styled as a “patent” but as an exclusive right granted to a mere “*possessor of a mode of applying the steam engine to propel a boat on new and advantageous principles.*”¹⁴¹ However, the law was a direct continuation of the patent New York had granted to John Fitch prior to ratification, which indicated that Fitch was the “inventor” of the steamboat.¹⁴² In 1798, New York simply transferred Fitch’s patent to Livingston based on an implied local working requirement after learning that Fitch had made no “attempt, in the space of more than ten years, of executing the plan for which he so obtained the exclusive privilege.”¹⁴³ Like Fitch’s prior patent, Livingston’s grant covered, quite broadly, “the sole and exclusive right and privilege” of making, using and navigating “all and every species or kinds of boats, or water craft, which might be urged or impelled through the water, *by the force of fire or steam*” and explicitly exempted any boat or watercraft “invented, or thereafter to be invented” that operated “by any other power,” such as wind powered sailboats.¹⁴⁴ Also, just like other state

138. *But see* Crockett, *supra* note 17, at 37–38 (arguing, in an invented dissent to *Bonito*, that the patent power was not meant to be exclusive and supporting this assertion with the fact that in *Gibbons* the Supreme Court ignored the patent preemption argument *even though* *New York’s grant to Livingston was based on a patent previously granted to Fitch*).

139. *See* Encyclopedia of World Biography – Robert R. Livingston, ENCYCLOPEDIA.COM, http://www.encyclopedia.com/topic/Robert_Livingston.aspx#2-1G2:3404703928-full (last visited May 1, 2013).

140. On John Fitch’s allegations of original inventorship and copying by Fulton, see THOMPSON WESTCOTT, LIFE OF JOHN FITCH: THE INVENTOR OF THE STEAM-BOAT 386–91 (1857); *see also* COX, *supra* note 21, at 1–16; JOHNSON, *supra* note 21, at 26–27. Fulton objected to these accusations. *See* COX, *supra* note 21, at 80.

141. *Livingston v. Van Ingen*, 9 Johns. 507, 583 (N.Y. 1812) (emphasis added).

142. *Id.* at 507 (“[O]n the 19th of March, 1787, the legislature of the state of New-York passed an act, entitled ‘An act for granting and securing to John Fitch the sole right and advantage of making and employing for a limited time, the steam-boat by him lately invented’”).

143. *Id.* at 509. The 1798 Act gave Livingston “privileges similar to those granted to John Fitch.” *Id.*

144. *Id.* at 507–08.

patents granted prior to ratification, this patent contained a fixed penalty for infringement, requiring operators of “offending” vessels to pay 100 pounds and forfeit their steamboats.¹⁴⁵

Thus, despite its comparably lax criteria for absolute novelty and original inventorship, everyone recognized that these exclusive rights were, in character if not in name, just like the patents states granted prior to ratification, and very similar to the patents Congress now granted to original inventors under the Patent Act.¹⁴⁶

B. THE FIRST COMMERCIAL STEAMBOAT ENTERPRISE IN AMERICA

The New York Legislature had learned from its mistake with Fitch and made quite clear in the 1798 law that, as a quid pro quo for the exchange, Livingston must actually implement functional steamboats in New York.¹⁴⁷ Specifically, he had to demonstrate to the Governor, Lieutenant-Governor, and Surveyor-General of New York, “or a majority of them,” that he had built a steamboat of at least twenty tons’ capacity that could travel not less than four miles an hour with or against the ordinary current of the Hudson.¹⁴⁸ Thereafter, for more than a year at a time, he had to have at least one of these boats travel between New York and Albany, lest his grant be retracted.¹⁴⁹ But thanks in part to Fulton’s efforts in experimenting with the designs of prior inventors and in part to Livingston’s wealth and political connections¹⁵⁰—and thanks to New York’s generous monopoly grant—Livingston and Fulton succeeded where Fitch failed.

After several extensions of their New York patent (and the inclusion of Fulton as a named grantee), they eventually satisfied the grant’s local working requirement. They dramatically showcased their boat on August 17, 1807, making the 120-mile trip from New York City to Albany in record time.¹⁵¹ Thereafter, they ensured that at least one steamboat was constantly

145. *Id.*

146. *See* *Gibbons v. Ogden*, 22 U.S. 1, 173–74 (1824) (discussing the argument that the New York law was simply a patent similar in character to U.S. patents).

147. *Livingston*, 9 Johns. at 509–11.

148. *Id.*

149. *Id.* at 509–10.

150. *See* JOHNSON, *supra* note 21, at 27 (noting the unique combination of political connections and financial patronage, which allowed Livingston and Fulton to succeed in turning their “discoveries into fully operational steamboats” as well as Fulton’s efforts in “por[ing] over recent U.S. patent filings” in 1806); *see also* COX, *supra* note 21, at 29 (“Together they succeeded where others had failed. Nonetheless, their work rested on the creativity and scientific knowledge of many who had come before.”).

151. On Fulton’s famous trial run up the Hudson, *see* COX, *supra* note 21, at 28; JOHNSON, *supra* note 21, at 27–28.

navigating the Hudson “except when the navigation of the river was interrupted by ice.”¹⁵² Their patent term was then extended, per its terms, to match the size of their fleet, increasing by five years for each additional boat they navigated on the river, but not to exceed thirty years.¹⁵³ Ultimately, as compensation for their willingness “to run the risk and hazard” of such an expensive experiment “which might prove so useful and beneficial to the community,”¹⁵⁴ Livingston and Fulton obtained the exclusive right to navigate steamboats on the Hudson for up to thirty years.¹⁵⁵

V. JUSTICE JAMES KENT’S CONCURRENT STATE PATENT SYSTEM

The constitutional validity of what may be the last state patent was tested when Livingston and Fulton attempted to enforce their rights against James Van Ingen and a group of twenty businessmen, who began operating their own steamboat called the “Hope” on the Hudson in contravention of Livingston and Fulton’s exclusive rights.¹⁵⁶

A. CHANCELLOR LANSING’S DISCOMFORT WITH STATE-SANCTIONED MONOPOLIES

In 1811, Livingston sued the prospective competitors in the New York Chancery court, only to have Chancellor John Lansing deny their application for an injunction.¹⁵⁷ As a member of New York’s Council of Revision, Lansing (along with John Jay) had previously objected to New York’s 1798 patent, though his reason for objecting was *not* that New York’s power to grant patents was extinguished by Congress’s powers to grant patents to inventors or to regulate interstate commerce.¹⁵⁸ Rather, Lansing asserted that New York simply lacked the sovereign authority to grant such a broad monopoly and had violated its duty to protect citizens’ common rights to navigate on public waterways.¹⁵⁹ Lansing recognized that monopolies might

152. *Livingston*, 9 Johns. at 511 (quoting Act of April 6, 1807 (5 LAWS OF THE STATE OF NEW YORK, ch. CLXV, 213–14)).

153. *Id.* at 511.

154. *Id.* at 508.

155. *Id.* at 510–12; *Gibbons v. Ogden*, 22 U.S. 1, 6–7 (1824).

156. *Livingston*, 9 Johns. at 512 (refusing to grant their *ex parte* application for an injunction, but entering an order that Van Ingen should provide a defense against an injunction).

157. *Id.* at 513; *see also* COX, *supra* note 21, at 55.

158. *See Gibbons*, 22 U.S. at 80 (explaining that Lansing and John Jay did not believe that laws passed by Congress had not eliminated state grants of exclusive rights).

159. *Livingston*, 9 Johns. at 519–20. Under contemporary notions of the “public trust doctrine,” the state held public waterways in trust for the people and had a duty to protect

sometimes be justified under the exceptional auspices of encouraging “the exertion of ingenuity and perseverance”; but he observed that in this case, steamboats had been known in America and in England for at least twenty-five years. “The combination of machinery, and the application of the power to give it effect, have been happily adapted to the propelling of vessels. It is a matter of public notoriety that they are now in a train of successful operation”¹⁶⁰ Therefore, Lansing concluded that, without any showing of true inventorship, or at least appropriate limits on the scope of the grant, New York lacked the power to transform commonly known subject matter into private property. This is very different from a declaration that states lacked the constitutional authority to grant patents. Indeed, some argued that Chancellor Lansing’s holding should be entirely disregarded as a precedent for determining the constitutionality of state patents.¹⁶¹

B. THE NEW YORK SUPREME COURT’S ENDORSEMENT OF STATE PATENT POWERS

Livingston and Fulton appealed Lansing’s decision to the New York Supreme Court (known in this time as the Court for the Trial of Impeachments and Correction of Errors, or simply the “Court of Errors”).¹⁶² The presiding Chief Justice was James Kent (1763–1847), appointed to the court in 1798 and made Chief Justice in 1804.¹⁶³ Kent was a self-proclaimed “zealous Federalist,” an admirer of Alexander Hamilton, who “got the

their rights to use and access them. See James L. Huffman, *Speaking of Inconvenient Truths—A History of the Public Trust Doctrine*, 18 DUKE ENVTL. L. & POL’Y F. 1, 96–98 (2007).

160. *Id.* at 520; see also COX, *supra* note 21, at 11, 15 (discussing U.S. steamboat patents), 16 (on steam engine publications); WESTCOTT, *supra* note 140, at 373–98 (discussing steamboat experiments in Europe and America).

161. Thomas Emmet argued in *Gibbons* that Lansing’s holding should be ignored because his prior objection at the Council of Revision (like the objection of Governor John Jay) was based “on principle” and showed that Lansing did not object to the law’s constitutionality. *Gibbons*, 22 U.S. at 80. Malla Pollack also asserts that Lansing’s objection to the steamboat monopoly was “on policy, not constitutional, grounds.” See Pollack, *supra* note 17, at 302 n.206 (citing Emmets’ comments in *Gibbons*, 22 U.S. at 80).

162. The Court of Errors at this time consisted of both state Supreme Court justices, who issued oral opinions explaining their reasoning, and New York senators, who had the final vote to affirm or reverse the judges’ decisions. It was the highest court of New York whose decisions could be appealed only to the U.S. Supreme Court. See WALTERSCHEID, *THE NATURE OF THE INTELLECTUAL PROPERTY CLAUSE*, *supra* note 18, at 438 n.11; see also COX, *supra* note 21, at 57–62.

163. On Kent’s life history, see JOHN THEODORE HORTON, *JAMES KENT: A STUDY IN CONSERVATISM 1763–1847* (1969). See also COX, *supra* note 21, at 57; Judith S. Kaye, *Commentaries on Chancellor Kent*, 74 CHI.-KENT L. REV. 11 (1998).

Federalist [Papers] almost by heart.”¹⁶⁴ Kent was also a well-read and meticulous legal scholar.¹⁶⁵ Before joining the New York Supreme Court, he had been a Professor of Law at Columbia College,¹⁶⁶ and his *Commentaries* (1826–1830) would become the foundational treatise on American law in the nineteenth century.¹⁶⁷ And Kent was no stranger to U.S. patent law: the *Commentaries* contain an entire section on the patent laws, which has been cited in patent disputes.¹⁶⁸ All sides recognized that Kent’s view of the case would be important. Not only was he the most distinguished of his colleagues and likely to be best prepared, but by his own account Kent dominated the New York court and wrote most of the opinions during stretches of his tenure.¹⁶⁹

In an erudite and thorough opinion, Kent reversed Chancellor Lansing and upheld the New York steamboat patent as a valid exercise of states’ sovereign powers under the Constitution.¹⁷⁰ In stark contrast to Lansing’s discomfort with legal monopolies, Kent began from the premise that “[t]he capacity to grant separate and exclusive privileges extends to every sovereign authority[]” and “is a necessary attribute of every independent government.”¹⁷¹ In England, the Crown had regularly granted monopolies to

164. See John Langbein, *Chancellor Kent and the History of Legal Literature*, 93 COLUM. L. REV. 547, 556 (1993) (quoting from Letter from James Kent to Thomas Washington, New York City (Oct. 6, 1828)); see also KUTLER, *supra* note 39, at 71 (describing Kent as “the unreconstructed Federalist” in opposition to Chief Justice Taney, a Democrat appointed to the Supreme Court by Jackson in 1836). On Kent’s great admiration for Hamilton and the lengthy letter he wrote to Hamilton’s wife upon Hamilton’s death, see MEMOIRS AND LETTERS OF JAMES KENT, LL.D. 227–29 (William Kent, ed. 1898).

165. On Kent’s preparation strategy, see FRIEDMAN, *supra* note 37, at 135–36; see also Kaye, *supra* note 163, at 13–15.

166. See COX, *supra* note 21, at 57; see also HORTON, *supra* note 163, at 95.

167. See Harold J. Krent, *Foreword: The Legacy of Chancellor Kent*, 74 CHI.-KENT L. REV. 3, 3 (1998) (noting that Kent’s “mammoth treatise on American law, Kent’s *Commentaries*, not only was the first full-fledged effort in this country, but also was the most extensively used throughout the nineteenth century, and is still widely cited today.”). As Cox writes, Kent’s “image as a brilliant and thorough jurist allowed him to become one of the foremost legal authorities in the early Republic, frequently using English common-law precedents to construct a distinctive American jurisprudence.” COX, *supra* note 21, at 57.

168. See Mossoff, *supra* note 44, at 982 n.136 (citing *Hotchkiss v. Greenwood*, 52 U.S. (11 How.) 248, 269 (1850) (Woodbury, J., dissenting)). Note that Kent wrote his COMMENTARIES after his opinion in *Livingston* in 1812.

169. Kent claimed to have written most of the opinions for certain years while he was on the bench, even those issued *per curiam*. See Kaye, *supra* note 163, at 19; see also FRIEDMAN, *supra* note 37, at 134.

170. See *Livingston v. Van Ingen*, 9 Johns. 507, 572–89 (N.Y. 1812).

171. *Id.* at 573.

those who invented or simply introduced new inventions into the realm.¹⁷² If the British Crown had these powers, then clearly so did the elected state governments, “for no one ever doubted (unless it be since the origin of this controversy) of the power of the Legislature to create an exclusive privilege.”¹⁷³ In fact, Kent noted, states exercised their power to grant exclusive rights all the time: “[a]ll our bank charters, turnpike, canal and bridge companies, ferries, markets, &c. are grants of exclusive privileges for beneficial public purposes.”¹⁷⁴ Surely, Kent reasoned, if states could grant a corporation the exclusive right to build a bridge, then states could grant a wealthy businessman like Robert Livingston the exclusive right to develop an expensive new invention like the steamboat for the sake of the public welfare.

Given states’ broad inherent authority to grant patents and other kinds of monopolies, the only question to address was whether states had ceded all or some of this authority to the federal government in ratifying the Constitution.¹⁷⁵ To answer this question, Kent drew on Alexander Hamilton’s famous test for constitutional preemption in *The Federalist No. 32*.¹⁷⁶ According to Hamilton, all powers not delegated to the federal

172. *Id.* at 584 (citing, among others, *Darcy v. Allen* (1602), 74 Eng. Rep. 1131, 1139 (K.B)).

173. *Id.* at 583–84. Indeed, Kent astutely noted that the Statute of Monopolies (1624), which otherwise limits the Crown’s powers to grant monopolies, exempted all grants by Parliament. *Id.*; see also Thomas Nachbar, *Monopoly, Mercantilism, and the Politics of Regulation*, 91 VA. L. REV. 1313, 1352 (2005) (explaining that Parliament retained power to grant monopolies).

174. *Livingston*, 9 Johns. at 573. In his role as Chancellor of New York, Kent would uphold these types of exclusive rights based on the theory that, without the prospect of a period of freedom from competition and the right to charge an “exclusive toll,” no one would “expend money upon great, and expensive, and hazardous public works, as roads and bridges,” or “become bound to keep them in constant and good repair.” See *Newburgh & C. Turnpike Road Co. v. Miller*, 5 Johns. Ch. 101, 106, 112, 116–17 (N.Y. Ch. 1821) (enjoining as a nuisance a rival bridge constructed close to a toll bridge that had been operating for over ten years); see also Joseph Dorfman, *Chancellor Kent and the Developing American Economy*, 61 COLUM. L. REV. 1290, 1293 n.16 (1961) (discussing Kent’s dissenting opinion in *Palmer v. Mulligan*, 3 Cai. R. 307, 318 (N.Y. Sup. Ct. 1805)).

175. *Livingston*, 9 Johns. at 573–74.

176. *Id.* at 576–77. See THE FEDERALIST NO. 32, at 155–58 (Alexander Hamilton) (Earle ed. 1938). The Tenth Amendment, added to the Constitution as part of the Bill of Rights in 1791, specifies that “[t]he powers not delegated to the United States by the Constitution, nor prohibited by it to the States, are reserved to the States respectively, or to the people.” U.S. CONST. amend. X. See WOOD, *supra* note 30, at 72. Kent did not apply the Tenth Amendment directly, instead employing Hamilton’s preemption test. This may not have been unusual for the period. The Tenth Amendment was added to the Bill of Rights mainly to appease Anti-Federalists who were concerned that otherwise valuable rights would be given up “by implication” upon ratification of the Constitution. See WOOD, *supra* note 30, at

government in the Constitution remained with the states unless the Constitution expressly made a certain power exclusive to the federal government or prohibited the states from exercising it, or where “a similar authority in the states would be absolutely and totally *contradictory* and *repugnant*.”¹⁷⁷

In his actual holding for purposes of the steamboat controversy, Kent avoided answering the difficult question of whether a state patent power would create a “repugnancy” with the federal government’s patent power. Instead, Kent relied on the gap in Congress’s own patent-granting powers, lamented by Hamilton in his *Report on Manufacturers*.¹⁷⁸ All parties conceded that Congress lacked the authority to grant exclusive rights to “importers from abroad of any useful invention or improvement.”¹⁷⁹ Therefore, according to Kent, “[s]uch persons must resort to the patronage of the state governments, in which the power to award their expensive and hazardous exertions was originally vested, and in which it still remains.”¹⁸⁰ Because Livingston and Fulton had not obtained their rights from New York as “inventors of the steam-boat,” but merely as “possessors,” the “privilege [was] totally unconnected with the patent power.”¹⁸¹

Kent must have recognized the weakness of this conclusion. First, as discussed above, many federal policymakers, including Hamilton, Washington, and Tench Coxe, believed Congress should be able to grant patents to importers; and it is not entirely clear that the choice to restrict U.S. patents to universally novel inventions was due solely to an adherence to federalism and deference to states’ rights. Second, Kent did not limit this category of “importers” to foreign inventors. Neither Robert Livingston nor Robert Fulton was a foreigner; and any steam technology they possessed had likely been copied from John Fitch and other true inventors. Finally, like Chancellor Lansing, Kent was surely aware that steamboats were not new to the United States and that American steamboat inventors had by this time published treatises on their designs and disclosed them in patent specifications under the new federal patent laws. Thus, New York’s steamboat monopoly was not “unconnected” to U.S. patent laws; it simply covered subject matter that could not have been patented under the Patent

70–71. Wood has observed that after it was finally ratified in 1791 “most Americans promptly forgot about the first ten amendments to the Constitution. The Bill of Rights remained judicially dormant until the twentieth century.” *Id.* at 72.

177. THE FEDERALIST NO. 32, at 155–56 (Alexander Hamilton) (Earle ed., 1938).

178. HAMILTON, REPORT ON MANUFACTURERS, *supra* note 33, at 1014.

179. *Livingston*, 9 Johns. at 583; *see also* discussion *supra* Section III.B.3.

180. *Livingston*, 9 Johns. at 583.

181. *Id.*

Act because it was already disclosed to the public in America as well as abroad, and because it was not actually invented by the patentees.¹⁸²

Therefore, for the sake of completeness, Kent went on to explain in dicta that the Constitution also left states with broad concurrent powers to grant patents to original inventors so long as state patents not interfere with U.S. patent rights or interstate commerce.¹⁸³ According to Kent's reading, Hamilton's test required preemption of state patent powers only if they came "*practically in collision with the actual exercise of some congressional power.*"¹⁸⁴ Kent took this to mean that a state patent would be preempted only when it covered precisely the same invention as a valid U.S. patent during the same period of time; and the U.S. patentee was able to prove in federal court that the state patent infringed his exclusive rights.¹⁸⁵ But the mere possibility of a future conflict between state and U.S. patent rights did not create a "repugnancy" and did not prohibit states from granting their own patents "in a variety of cases, without any infringement of the congressional power."¹⁸⁶

182. *See also* *Gibbons v. Ogden*, 22 U.S. 1, 173 (1824) (Attorney General Wirt's oral argument). Wirt argued:

The privilege here granted by the State, is to an American citizen, who claims to be the inventor. The privilege is the reward of invention, not of importation, and this it is which brings it in conflict with the act of Congress. It is true, the law does not call him the inventor; it calls him merely the "possessor." But, can the constitution and laws of the United States be evaded in this manner? If he was not the inventor, why this unjust tax which has been levied upon our admiration and gratitude?

Id.

183. In his opinion in *Gibbons*, Chief Justice Marshall would suggest that the Commerce Clause preempts a state patent that interferes with interstate commerce. *See* discussion *infra* Section V.C (discussing Marshall's opinion in *Gibbons*). That said, Kent rejected the Commerce Clause argument in *Gibbons*, even though the facts there involved a conflict between the New York monopolists and steamboat operators from out-of-state in possession of a federal coasting license. *See* Williams, *supra* note 21, at 1409. Furthermore, Kent generally disagreed with Marshall's holding in *Gibbons* striking the monopoly. *See* WHITE, *supra* note 27, at 578.

184. *Livingston*, 9 Johns. at 576 (emphasis added).

185. *Id.* at 582–83. Notably, Kent later had the case reporter clarify in a footnote that he was not specifically holding that "a state grant could . . . be questioned or controlled by a [federal] patent right[]" before *any* tribunal, whether state or federal. *Id.* at 582 n.(a). In this period judges rendered their opinions orally. They were written up by a reporter who tried his best to capture accurately what the judge had said. *See* Langbein, *supra* note 164, at 577.

186. *Livingston*, 9 Johns. at 581. Kent's most detailed description of states' concurrent patent powers is as follows:

A state cannot take away from an individual his patent right, and render it common to all the citizens. This would contravene the act of congress, and would be, therefore, unlawful. But if an author or inventor, instead of resorting to the act of congress, should apply to the legislature of this state

Therefore, if an inventor chose to apply to a state instead of Congress for an exclusive right to his invention, the state was free to grant it, creating a “complete and perfect” right within the state’s jurisdiction.¹⁸⁷

It is also clear from Kent’s opinion that he believed states had the power to independently determine the criteria for patentability and the terms and conditions of their own patents. Thus, state patents could last far longer than U.S. patents and extend to a much wider range of subject matter, including inventions like the steamboat that were to some extent known and free for all to copy and use.¹⁸⁸ At the same time, it is *not* clear from Kent’s opinion whether he believed these powers had any limits. Kent did state somewhat vaguely that a state’s power to grant monopolies was limited by “its own constitutional provisions, or by the fundamental principles of all government, and the unalienable rights of mankind.”¹⁸⁹ But his dicta on states’ concurrent patent powers, as well as his decision to uphold New York’s sweeping thirty-year monopoly grant to non-inventors, make clear that state patents would not be limited by the standards laid out by Congress in the Patent Act or by any restrictions that might be imposed upon Congress by the IP Clause.¹⁹⁰

Nor did Kent think states’ patent powers were limited by the general anti-monopoly principles that applied in England, where the Statute of

for an exclusive right to his production, I see nothing to hinder the state from granting it, and the operation of the grant would, of course, be confined to the limits of this state. Within our own jurisdiction, it would be complete and perfect. So a patentee under the act of congress may have the time of his monopoly extended by the legislature of any state, beyond the term of fourteen or twenty-eight years allowed by that law. Congress may secure, for a limited time, an exclusive right throughout the union; but there is nothing in the constitution to take away from the states the power to enlarge the privilege within their respective jurisdictions. The states are not entirely divested of their original sovereignty over the subject matter; and whatever power has not been clearly granted to the union, remains with them.

Id. at 581–82.

187. *See id.* at 581. Kent further opined that states could even grant patent extensions to federal patentees whose patents had expired. *Id.* Kent was likely motivated by a belief, apparently common in this period, that U.S. patents were too short—a belief that ultimately prompted Congress to begin granting patent extensions. *See, e.g.,* *Evans v. Jordan*, 13 U.S. (9 Cranch) 199, 202 (1815).

188. Thomas Oakley made this argument in *Gibbons v. Ogden*, 22 U.S. 1, 51 (1824). Oakley, like Attorney General Wirt, did not use the term “public domain” but used the terms “common rights” or “public property” to refer to formerly patented subject matter that should be free for all to use. *Id.* at 171. On the origins of the “public domain,” see Tyler Ochoa, *Origins and Meaning of the Public Domain*, 28 U. DAYTON. L. REV. 215, 217–22 (2002).

189. *Livingston*, 9 Johns. at 573.

190. *Id.* at 581–82.

Monopolies (1624) gave English courts powers to review the king's monopoly grants and check the monopolies of a free-wheeling, self-interested monarch.¹⁹¹ Instead, Kent believed state patents would be sufficiently limited by the democratic political process. In a brilliant turn of statutory construction, Kent pointed out that in America, state patents, like the colonial patents before them, were granted by the elected legislatures, not the Crown,¹⁹² and that under the Statute of Monopolies' express terms, "all grants of privileges by act of parliament were saved; for no one ever doubted (unless it be since the origin of this controversy) of the power of the legislature to create an exclusive privilege."¹⁹³ Unlike the royal privileges granted in England, each state patent had undergone review by various branches of state government and was enacted for the purpose of promoting the public good. It was therefore entitled to an "extremely strong" presumption of validity,¹⁹⁴ and should not be policed substantially by courts, even in those cases where they deemed the grant to be "inexpedient or unwise."¹⁹⁵

Ironically, in 1788 James Madison made exactly the same appeal to the protections of the democratic political process when defending the Framers' decision to give *Congress* the constitutional power to grant patents.¹⁹⁶ Why,

191. See Statutes at Large, 1624, 21 Jac., c.3 (Gr. Brit.), *supra* note 74; see also Bracha, *supra* note 7, at 47.

192. Bracha has also pointed out that it is notable that the colonial legislatures *rather than the governors or the councils* granted patents and this was thus not a "miniature version" of the English patent grants. See Bracha, *supra* note 7, at 98 n.224 (citing sources on colonial assemblies).

193. *Livingston*, 9 Johns. at 583–84. Section 7 of the Statute of Monopolies indeed exempts all grants by Parliament. See Nachbar, *supra* note 173, at 1352 (listing statute's provisions and exemptions). Kent's view that Parliament, versus the king, still had *carte blanche* to grant monopolies also has some historic support. In the decades following the passage of the Statute of Monopolies, Parliament permitted exclusive trade privileges to guilds and corporations to continue and also exercised its own powers to grant monopolies. See *id.* at 1358–63, 1360 n.224.

194. *Livingston*, 9 Johns. at 572.

195. *Id.* at 573.

196. In his response to Thomas Jefferson's request to constitutionally ban all government-sanctioned monopolies, Madison famously stated:

Is there not also infinitely less danger of this abuse in our governments than in most others? Monopolies are sacrifices of the many to the few. Where the power is in the few it is natural for them to sacrifice the many to their own partialities and corruptions. *Where the power, as with us, is in the many not in the few, the danger can not be very great that the few will be thus favored.*

Letter from James Madison to Thomas Jefferson (Oct. 17, 1788), in 1 THE REPUBLIC OF LETTERS: THE CORRESPONDENCE BETWEEN THOMAS JEFFERSON AND JAMES MADISON, 1776–1826, at 566 (James Morton Smith ed. 1995) (emphasis added).

Kent might have asked, should state patents be treated with any less deference, or any more skepticism, than the new patents granted by Congress—especially when each state patent was granted by an elected legislature with the specific goal of benefiting the public, and under the same procedures the states and colonies had been using for two hundred years? So long as the states did not directly interfere with Congress’s responsibility to “secure” inventors’ “exclusive rights” in their inventions, Kent saw no reason why states should not possess concurrent powers to grant patents on whatever terms and conditions they saw fit in order to promote economic development and technological innovation in their own territories.

C. STATE PATENT LAWS IN THE WAKE OF *GIBBONS*

Given how controversial the idea of state patents appears today, we might assume Kent’s position would prove untenable even in his own time and that courts would have quickly overruled it.¹⁹⁷ But the record supports the opposite. All of Kent’s fellow justices on the New York Supreme Court concurred with his decision.¹⁹⁸ His co-justice, Justice Smith Thompson, who would eventually serve on the U.S. Supreme Court alongside Justice Story and Chief Justice Marshall, wrote his own concurring opinion reflecting all of the same points.¹⁹⁹ The State Senate reinforced the holding by a unanimous vote of thirty to none.²⁰⁰ But the real test of concurrent state patent power came when, following the deaths of Livingston and Fulton, their licensee, Aaron Ogden, sued a competitor named Thomas Gibbons, who was running two steamboats between locations in New York and New Jersey, thereby violating Ogden’s exclusive right to navigate steamboats in New York. After Kent granted Ogden’s motion for an injunction, Gibbons appealed the case all the way to the Supreme Court, arguing that Ogden’s New York rights were preempted by Congress’s power to grant patents under the IP Clause and its power to regulate interstate commerce under the Commerce Clause.²⁰¹

197. Walterscheid takes this stance, at least with regard to state patents covering formerly patented inventions, concluding that Kent’s “statement that the Constitution did not take away from the states the power to enlarge the patent privilege, for example, by extending the term of a federal patent within their jurisdiction, was of doubtful validity in 1812 and most certainly would be deemed invalid today.” WALTERSCHEID, *THE NATURE OF THE INTELLECTUAL PROPERTY CLAUSE*, *supra* note 18, at 442.

198. Justice Spencer did not give an opinion as he was related to some of the parties. *Livingston*, 9 Johns. at 563.

199. *Id.* at 563–66, 567 (Thompson, J., concurring). On Thompson’s service as a justice alongside Marshall, see Kutler, *supra* note 39, at 55–56.

200. See COX, *supra* note 21, at 60.

201. *Gibbons v. Ogden*, 22 U.S. 1, 1–3 (1824); see also JOHNSON, *supra* note 21, at 55–57; Williams, *supra* note 21, at 1408.

In defending the steamboat monopoly, Ogden's attorneys reiterated Justice Kent's argument that states possessed concurrent power to grant their own patents in a variety of cases, particularly for costly inventions like the steamboat that had been "given to the public" but were not yet working in a particular state because "great expense must be incurred to put [them] into use" ²⁰² However, although Chief Justice Marshall struck down the New York law and ended Ogden's monopoly, he decided the case on narrow grounds, finding that because Gibbons possessed a federal coasting license authorizing him to navigate steamboats in U.S. waters, New York's law, which prevented Gibbons from doing so, directly conflicted with federal law and was therefore preempted by the Supremacy Clause. ²⁰³ Marshall also suggested in dicta that the Commerce Clause generated an implicit constitutional prohibition on state laws that interfered with interstate commerce. ²⁰⁴ But Marshall provided no guidance whatsoever on whether New York's law conflicted with Congress's power under the IP Clause and on whether or not states retained power to grant their own patents. Instead, he dispensed with the issue in a single sentence, concluding that "[a]s [preemption by the federal coasting license] decides the cause, it is unnecessary to enter in an examination of that part of the Constitution which empowers Congress to promote the progress of science and the useful arts." ²⁰⁵

Marshall's choice to abstain may simply have been a strategy for securing the votes of his fellow justices. ²⁰⁶ He also may have been wary of placing the

202. In his oral argument in *Gibbons*, Thomas Oakley, arguing on behalf of Livingston and Fulton, indicated that states retained the power to grant patents in a variety of cases, including where an invention had been "given to the public, and great expense must be incurred to put it into use . . ." *Gibbons*, 22 U.S. at 48 (emphasis added).

203. *See id.* at 176–79 (noting that the power of patenting is "exclusively vested in Congress," not in the States); *see also* Williams, *supra* note 21, at 1416–20, 1446 (outlining and analyzing Marshall's holding).

204. Marshall stated in dicta, though did not hold, that some aspects of commerce regulation are exclusive in the federal government. Marshall conceded that the "acknowledged power of a State to regulate its police, its domestic trade, and to govern its own citizens, may enable it to legislate on [the subject of regulation of commerce] to a considerable extent." *Gibbons*, 22 U.S. at 208. But he was particularly struck by Daniel Webster's argument that if state commercial regulations were allowed to "perform the same operation on the same thing" as federal commercial regulations, this might disrupt the federal scheme for fluid interstate commerce and require preemption even where there was no specific conflict. *See id.* at 209.

205. *Id.* at 221; *see also id.* at 239 (Johnson, J., concurring) ("I have not touched upon the right of the states to grant patents.").

206. *See* Williams, *supra* note 21, at 1422 (noting the common explanation that Marshall was trying to accommodate the views of the other justices). For another interpretation, *see*

Supreme Court at the center of a battle between state and federal authority to grant patents.²⁰⁷ However, I assert that the simpler explanation is that Marshall, a friend and fellow Federalist, *agreed* with Justice Kent's position that, generally speaking, states could continue granting patents alongside Congress. Marshall's dicta presaging modern "Dormant Commerce Clause" doctrine suggests that he might preempt a state patent that interfered with interstate commerce, as New York's steamboat patent arguably did. However, Marshall did not address or even mention the Attorney General's argument that concurrent state patent powers risked undermining the U.S. patent system and so must be completely abolished under the IP Clause.²⁰⁸ The reason, I suggest, is that Marshall, like Kent and Hamilton, required more than vague fears of conflict to completely divest states of their sovereign powers to grant patents. The fact was that Gibbons did not have a

WHITE, *supra* note 27, at 579 (suggesting that Marshall's decision was the result of "discomfort" with the idea of dual sovereignty).

207. See Williams, *supra* note 21, at 1498 (arguing that Marshall abstained from basing his holding directly on the preemptive power of the Commerce Clause because he was concerned about the political implications of placing the judiciary at the forefront of policing protectionist state legislation).

208. In his oral argument in favor of striking down the monopoly under the IP Clause, Attorney General William Wirt presented three major objections to allowing states to continue offering their own patents. See *Gibbons*, 22 U.S. at 167–71. First, state patents would compete with federal patents for applicants. See *id.* at 169 ("Who would apply to the power of Congress for a patent or a copy-right, while the States held up higher privileges?"). Second, states would compete amongst *one another* to offer the most attractive patent rights in order to convince inventors to locate in their territories. *Id.* ("This concurrent legislation would degenerate into advertisements for custom. These powers would be in the market, and the highest bidder would take all."). Finally, if states could grant their own patents this would interfere with Congress's authority to decide what must remain unpatented and free for future innovators to use (i.e., part of what we would call the "public domain"). See *id.* at 171. The Attorney General states:

The law of Congress declares, that all inventors of useful improvements throughout the United States, shall be entitled to the exclusive right in their discoveries for fourteen years *only*. The law of New-York declares, that this inventor shall be entitled to the exclusive use of his discovery for thirty years, and as much longer as the State shall permit. The law of Congress, by limiting the exclusive right to fourteen years, in effect declares, that after the expiration of that time, the discovery shall be the common right of the whole people of the United States. The law of New-York declares that it shall not, after fourteen years, be the exclusive right of the people of the United States, but that it shall be the exclusive right of this inventor for thirty years, and for so much longer as she, in her sovereign will and pleasure, may permit. If this be not repugnance, direct and palpable, we must have a new vocabulary for the definition of the word.

Id. (emphasis in original). On the evolving conception of a "public domain" in this period, see *supra* note 188.

U.S. patent for his steamboats to hold up against Livingston and Fulton's state patent rights; so, coasting license and Commerce Clause preemption aside, there simply was no conflict with the federal patent power.²⁰⁹ If Gibbons had possessed a U.S. patent, we might have witnessed the first infringement suit between state and federal patentees, and our patent system might be very different today.

This hypothesis becomes still more viable when we learn that, following the Supreme Court's edict in *Gibbons* striking the steamboat monopoly, courts continued to uphold Kent's opinion in *Livingston*.²¹⁰ The U.S. Supreme Court, though it never again had the chance to rule directly on the constitutionality of an actual state patent, frequently cited *Livingston* favorably.²¹¹ The Court's most extensive application of Kent's views on concurrent patent powers occurred in *Patterson v. Kentucky*, where the Court quoted from *Livingston* and held that the IP Clause did not preempt states' powers to restrict the exclusive rights of a federal patentee to sell a patented product for purposes of health and safety.²¹² Although the Court did not address the converse issue of states' powers to grant their own exclusive rights over inventions or ideas, the Court quoted Kent's comments in *Livingston* that the federal government's power to grant patents was "fully satisfied" so long as states did not interfere with patentees' federally protected period of exclusivity in their inventions.²¹³

209. See *Gibbons*, 22 U.S. at 171–72.

210. See, e.g., *North River Steamboat Co. v. Livingston*, 3 Wheeler C.C. 483, 3 Cow. 182 (1825) (Ch. J. Savage) (one in a series of cases refusing to enforce the New York monopoly in light of *Gibbons*, but citing *Livingston* as authority and declining to overrule its holding that the New York law was constitutional); see also COX, *supra* note 21, at 175–80. Thereafter, New York courts continued to cite to *Livingston* to support the state's power to grant individuals exclusive rights to use common property and waterways within the state's jurisdiction as a consequence of the state's sovereign powers over public property "within its limits." See *Langdon v. Mayor of N.Y.*, 93 N.Y. 129, 155–56 (1883).

211. In *Smith v. Turner*, 48 U.S. 283, 371 (1849), the Court cited to *Livingston*, *Gibbons*, and William Blackstone's COMMENTARIES (discussed in note 214, *infra*) for the principle that "[the power] to promote science and the arts by copy and patent rights" is not exclusive in the federal government, but rather, "to the extent of State limits, is believed to be concurrent." The Court also often cited *Livingston* for general principles of concurrent state and federal sovereignty. See, e.g., *Houston v. Moore*, 18 U.S. 1, 8 (1820); see also *Root v. Ry. Co.*, 105 U.S. 189, 192 (1889) (citing *Livingston* for Kent's rule on when an injunction should be granted in a patent case).

212. *Patterson v. Kentucky*, 97 U.S. 501, 508–509 (1878) (holding that state can prohibit sale of patented articles in the state if they are deemed to be a hazard to public health and safety).

213. "Chancellor Kent said that 'the national power will be fully satisfied if the property created by patent be, for the given time, enjoyed and used exclusively, *so far as*, under the

The cases cited above provide no affirmative statement of the constitutional validity of state patents. However, notable legal commentators explicitly endorsed concurrent state patent powers in their legal treatises.²¹⁴ The most significant endorsement came from Kent's friend and fellow Federalist Justice Joseph Story (1779–1845), whose opinions between 1813 and 1845 provided the basic outlines for much of modern intellectual property law and doctrine.²¹⁵ Like Kent, Story believed that states must generally be able to grant exclusive rights in order to encourage investment in costly and risky enterprises.²¹⁶ And, like Kent, Story apparently extended this

laws of the several States, the property shall be deemed for toleration.” *Id.* at 508 (quoting *Livingston v. Van Ingen*, 9 Johns. 507, 582 (N.Y. 1812)).

214. In his 1803 edition of William Blackstone's *Commentaries on the Laws of England* (1765–1769), George Tucker suggested in comments about the U.S. Constitution that states had concurrent patent powers even though they were unlikely to use them given the lesser jurisdiction of state patents and the 1793 Patent Act's requirement that inventors choose between state and U.S. patents. *See* 1 ST. GEORGE TUCKER, BLACKSTONE'S COMMENTARIES: WITH NOTES OF REFERENCE TO THE CONSTITUTION AND LAWS OF THE FEDERAL GOVERNMENT OF THE UNITED STATES; AND OF THE COMMONWEALTH OF VIRGINIA APP. NOTE D 265–66 (Philadelphia, William Young Birch & Abraham Small 1803). Tucker writes:

This is another branch of federal authority, in which I presume the states may possess some degree of concurrent right within their respective territories; but as the security which the state could afford, would necessarily fall short of that which an authority co-extensive with the union may give, it is scarcely probable that the protection of the laws of any particular state will hereafter be resorted to; more especially, as the act of 2 Cong. c. 55, declares, that “where any state before it's adoption of the present form of government shall have granted an exclusive right to any invention, the party claiming that right, shall not be capable of obtaining an exclusive right under that act, but on relinquishing his right under such particular state, and of such relinquishment his obtaining an exclusive right under that act, shall be sufficient evidence.”

Id.; *see also* Pollack, *supra* note 17, at 301–02 (citing Tucker's 1803 edition of *Blackstone's Commentaries* to support her contention that “some authority exists for the proposition that the Intellectual Property Clause was intended, or originally considered, to be a non-exclusive grant to Congress”). Also note that Tucker's mention of the 1793 Patent Act's relinquishment provision, which he suggests would hinder most inventors from choosing state rather than federal patents, directly counters Walterscheid's suggestion that the provision did not implicitly extend to state patents granted after ratification. *Compare with* WALTERSCHEID, THE NATURE OF THE INTELLECTUAL PROPERTY CLAUSE, *supra* note 18, at 437–38.

215. *See* Frank D. Prager, *The Influence of Mr. Justice Story on American Patent Law*, 5 AM. J. LEGAL HIST. 254 (1961).

216. Story fought against the decision of Andrew Jackson's Democratic appointee, Chief Justice Roger Taney, to strike down Massachusetts' monopoly to the proprietors of the Charles River bridge. Like Kent, Story believed that allowing competition to develop in the market would destroy incentives to invest in public works and “arrest all public improvements, founded on private capital and enterprise” *Charles River Bridge v.*

rationale to states' powers to grant patents. In his *Commentaries on the Constitution of the United States*, Story cited to *Livingston* and to *Gibbons*, stating that “[i]t has been suggested, that [the IP Clause] power is not exclusive, but concurrent with that of the States, so always, that the acts of the latter do not contravene the acts of [C]ongress.”²¹⁷ Story allowed that concurrent patent powers remained a possibility; and then went on to conclude with confidence that, at any rate, “as the power of [C]ongress extends only to authors and inventors, a [S]tate may grant an exclusive right to the *possessor or introducer* of an art or invention, who does not claim to be an inventor, but has merely introduced it from abroad.”²¹⁸

This statement clearly indicates that Story believed states could continue to grant exclusive rights to introducers of inventions “from abroad.” Story’s use of the term “possessor” along with his citation to *Livingston* also shows his belief that states could even continue to grant exclusive rights to American citizens like Robert Livingston on technologies that were already known, if not yet in successful operation in a given territory.²¹⁹ Thus, in Story’s interpretation, states’ power to grant patents to first inventors remained an open possibility, so long as they did not “contravene” Congressional patent laws. Meanwhile, states’ power to grant patents to those *not* claiming to be first inventors, but who merely introduced an invention into the state for the first time, was an established constitutional reality based on the principles of federalism and states’ residual sovereign powers under the Tenth Amendment.

VI. JUSTIFYING CONCURRENT STATE PATENT POWERS

At this point I have laid out Justice Kent’s position that states possessed broad powers to grant patents alongside Congress, both to original inventors and to those who simply introduced a certain technology into the state. I have shown that, although they appear controversial to us, Kent’s views received widespread support throughout the nineteenth century. The question that now bears answering is *why* Justice Kent, Justice Story, and perhaps Chief Justice Marshall himself believed that the Constitution did not significantly restrict states’ powers to grant their own patents. One hypothesis, which I discuss briefly below, is that Justice Kent believed U.S. patents did not effectively protect inventors’ exclusive rights and that state

Warren Bridge 36 U.S. (11 Pet.) 420, 608 (1837) (Story, J., dissenting) (citing Kent’s opinion in the steamboat controversy); *see also* KUTLER, *supra* note 39, at 54–61.

217. STORY, COMMENTARIES, *supra* note 25, at 50.

218. *Id.* (emphasis added).

219. *See* *Livingston v. Van Ingen*, 9 Johns. 507, 583 (1812).

patents might therefore provide inventors with an important alternative or supplement to U.S. patents. But I argue that the real and lasting case for concurrent state patent powers came from a more fundamental, inherently Hamiltonian concern that U.S. patents, which are limited to universally novel inventions and are designed to be anti-regulatory incentives, did not fill the policy role of state patent laws. Thus, state patents might still provide necessary market correctives for encouraging investment in important new technology.

A. THE PERCEIVED WEAKNESS OF U.S. PATENTS IN THE EARLY NINETEENTH CENTURY

The first hypothesis for why people like Justice Kent believed state patents remained necessary is that they believed U.S. patents were of limited value and unlikely to be upheld by courts.²²⁰ Between 1793 and 1836, the period in which *Livingston* and *Gibbons* were decided, the federal patent system employed a registration system, meaning that there was no mechanism for screening patents even for the basic criteria of novelty and disclosure. Instead, the role of policing patent validity was left to courts after patents were challenged in litigation.²²¹ In his *Commentaries* (1826–1830), Kent expressed skepticism about the workability of such a system, stating that because the Secretary of State had no power to refuse a patent for want of novelty or usefulness, “a great many worthless and fraudulent patents were issued, and the value of the privilege was degraded, and in a great degree destroyed.”²²² However, in 1836 Congress reformed the Patent Act in an attempt to make U.S. patents more effective.²²³

After the 1836 reforms, patented inventions were increasingly licensed and deployed in markets across the country.²²⁴ Based on this situation, it

220. See WALTERSCHEID, TO PROMOTE THE PROGRESS OF SCIENCE AND USEFUL ARTS, *supra* note 128, at 243–44 (contending that until the Patent Act was reformed in 1832 and 1836, U.S. patents were of inconsistent quality and were not certain to be upheld by courts); Andrew P. Morriss & Craig Allen Nard, *Institutional Choice & Interest Groups in the Development of American Patent Law: 1790–1865*, 19 SUP. CT. ECON. REV. 143, 155 n.41 (2011) (citing some contemporary views that U.S. patents were of limited value); see also B. Zorina Khan, *Property Rights and Patent Litigation in Early Nineteenth-Century America*, 55 J. ECON. HIST. 58, 63 (1995) (noting that from 1800 to 1839, only 67 patents were litigated in 116 patent cases).

221. See Bracha, *Onning Ideas*, *supra* note 7, at 416–17.

222. 2 JAMES KENT, COMMENTARIES ON AMERICAN LAW *366.

223. See Patent Act of 1836, ch. 357, § 18, 5 Stat. 117, 124 (repealed 1861).

224. See Naomi R. Lamoreaux & Kenneth L. Sokoloff, *Market Trade in Patents and the Rise of a Class of Specialized Inventors in the 19th-Century United States*, 91 AM. ECON. REV. 39, 39–41 (2001) (showing that within just a few years of the reforms of 1836, inventors increasingly began assigning the rights to practice their technologies to buyers in different geographic areas in the United States). On doctrinal developments in the federal laws associated with the

could be hypothesized that the only reason people like Justice Kent thought state patents were still necessary prior to 1836 was their low estimation of the value and quality of U.S. patents.²²⁵ If this were the case, we might decide that the 1836 reforms eliminated the need for state patents once and for all.²²⁶ I reject this hypothesis. Instead, I argue in the next Section that there was a more fundamental reason Kent and other jurists believed in the need for concurrent state patent laws.

B. U.S. PATENT LAW LEFT MAJOR GAPS IN AMERICAN INNOVATION POLICY

Everyone, including Kent, recognized the importance of U.S. patents for protecting inventors' exclusive rights in their ideas on a national level.²²⁷ However, no matter how strong U.S. patents were, they did not fulfill, or even *attempt* to fulfill, many of the policy functions performed by the colonial and state patents described above. Indeed, beyond a shared general concern for promoting innovation with the enticement of a limited monopoly, U.S. patents were entirely different. When compared to state counterparts, U.S. patent law generated three major policy gaps.

1. *No Concern for the Social Utility of the Subject Matter*

The first crucial policy gap left by U.S. patent law was its indifference to the social utility of the subject matter. As discussed above, the states granted patents in individualized statutes to incent the undertaking of worthy projects with the purpose of directly improving the welfare and standard of living of local inhabitants.²²⁸ In contrast, as early as 1817, it became established U.S.

nineteenth century transformation of patents into valuable and marketable general rights, see Bracha, *Owning Ideas*, *supra* note 7, at 401–518.

225. See COX, *supra* note 21, at 15 (commenting that inventors turned to the states to protect their work once they realized that the 1793 Patent Act “allowed applicants to secure patents for inventions regardless of how similar their discoveries might be”).

226. As noted, Walterscheid appears to take the view that U.S. patents largely removed the need for state patents. See WALTERSCHEID, *THE NATURE OF THE INTELLECTUAL PROPERTY CLAUSE*, *supra* note 18, at 76–77, 436–37, 438–42; see also BUGBEE, *supra* note 7, at 102–03 (suggesting federal patents made state patents obsolete).

227. Kent wrote in his COMMENTARIES that U.S. patents would encourage “ingenious men” to create “works useful to the country and instructive to mankind” by providing them the “hope of profit, as well as by the love of fame or a sense of duty.” He also wrote that inventors “should enjoy the pecuniary profits resulting from mental as well as bodily labor.” KENT, COMMENTARIES, *supra* note 222, at *365, *474; see also Mossoff, *supra* note 44, at 982 (arguing that Kent expressed a natural rights view of IP as indicated by his choice to title this section of his Commentaries “of original acquisition, by intellectual power”).

228. See Bracha, *Owning Ideas*, *supra* note 7, at 99–101; see also Biagioli, *Patent Republic*, *supra* note 9, at 1133–36 (describing the enhanced reduction to practice and utility requirements of early state patent laws in the United States, and contrasting these to federal patent rights).

patent law doctrine that so long as an invention worked for some purpose, it need not be better than prior technology or have any social utility at all.²²⁹ The value of the invention—the amount the patentee could charge for sale or licensing and the amount he could demand for an infringement—was determined through private exchanges based upon its perceived value in the marketplace.²³⁰ In theory, the public would benefit nonetheless because inventors would have an incentive to create inventions that could eventually be sold or licensed to others.²³¹ But Congress would play no role in ensuring that inventors create technologies of social value—let alone technologies of value to a particular community.

2. *No U.S. Patents for Previously Known or Used Technology*

The second policy gap was that U.S. patents were only allowed for universally novel inventions that were actually invented by the patentee.²³² As

229. See *Lowell v. Lewis*, 15 Fed. Cas. 1018, 1019 (C.C.D. Mass. 1817) (expressing Story's view that "[a]ll the law requires is, that the invention should not be frivolous or injurious to the well-being, good policy, or sound morals of society"); see also *Bedford v. Hunt*, 3 Fed. Cas. 37, 37 (C.C.D. Mass. 1817) (J. Story) (the law "does not look to the degree of utility; it simply requires, that it shall be capable of use, and that the use is such as sound morals and policy do not discountenance."). On courts' adoption of Story's view as stating the law on patent utility, see MERGES & DUFFY, *supra* note 43, at 246–47. Under the 2001 PTO guidelines, so long as the applicant shows "any particular purpose (i.e. a 'specific utility') and that assertion would be considered credible by a person of ordinary skill in the art," the PTO examiner should not reject the application based on lack of utility. MERGES & DUFFY, *supra* note 43, at 238 (quoting U.S. Patent and Trademark Office, *Utility Examination Guidelines*, 60 Fed. Reg. 36,263 (July 14, 1995)). In modern patent law, the utility requirement has served mainly as a "timing device" to ensure that an invention is ready for patenting, not to ensure that the invention has socially beneficial functions. See Rebecca S. Eisenberg, *Analyze This: A Law and Economics Agenda for the Patent System*, 53 VAND. L. REV. 2081, 2086–87 (2000).

230. This of course assumes no transaction costs, which is rarely the case in the context of bargaining over the rights to intellectual property. See generally Robert P. Merges, Comment, *Of Property Rules, Coase, and Intellectual Property*, 94 COLUM. L. REV. 2655, 2661 (1994) ("Despite a few brave attempts to assume away the obvious, those who have considered the application of the Coase theorem to IPRs have noted the pervasiveness of transaction costs."). On the development of the market-based conception of utility, see also MERGES & DUFFY, *supra* note 43, at 212–13.

231. As Justice Story put it: "[i]f its practical utility be very limited, it will follow, that it will be of little or no profit to the inventor; and if it be trifling, it will sink into utter neglect." *Bedford v. Hunt*, 3 F. Cas. 37, 37 (C.C. Mass. 1817) (J. Story); see also *Earle v. Sawyer*, 8 F. Cas. 254, 256 (C.C. Mass. 1825.) (J. Story) (C.C.D. Mass. 1825) ("[T]he degree of positive utility is less important in the eye of the law, than some other things, though in regard to the inventor, as a measure of the value of the invention, it is of the highest importance.").

232. The Act stated that the applicant must submit a petition "setting forth, that he, she, or they, hath or have invented or discovered any useful art, manufacture, engine, machine, or device, or any improvement therein *not before known or used.*" Patent Act of 1793, ch. 11, § 1, 1

noted above, this decision was based on a belief that the Constitution's use of the term "Inventors" prohibited Congress from granting patents to "introducers" of foreign inventions from abroad, and that Congress was "tied down" to a single means for "promot[ing] the Progress of Science and useful Arts."²³³ Although Hamilton conceded on this point of law, he disagreed with the policy and strongly favored granting patents to encourage "introduction" of foreign technology in the infant nation.²³⁴ America might already possess "*knowledge* of several of the most important [European machines,]" he stated; but actually deploying "all such machines as are known in any part of Europe, [could] only require proper provision and due pains."²³⁵ Patents covering imported technology could therefore provide a much needed market incentive.²³⁶

Congress appeared to some extent to bow to this view when it reformed the Patent Act in 1836 and made changes that specifically allowed inventors to obtain U.S. patents even if they had previously patented the subject matter in a foreign country.²³⁷ Furthermore, U.S. courts interpreted the Patent Act to allow an inventor to obtain a patent upon proving that he was an original inventor, despite some prior use of the invention in a foreign country.²³⁸ This "domestic bias" continued throughout the twentieth century: under the Patent Act of 1952, there remained various ways in which copying and

Stat. 318, 318 (repealed 1836) (emphasis added). The applicant had to "swear or affirm that he does verily believe, that he is the true inventor or discoverer of the art, machine, or improvement, for which he solicits a patent . . . before any person authorized to administer oaths . . ." *Id.*; see also WALTERSCHEID, THE NATURE OF THE INTELLECTUAL PROPERTY CLAUSE, *supra* note 18, at 327–35 (describing the unique universal novelty requirement in the United States patent laws).

233. See U.S. Const. Art. I, sec. 8, cl. 8; see also Walterscheid, *Patents and Manufacturing*, *supra* note 18, at 875 (quoting James Madison's response to Tench Coxe's request for a land premium scheme). Walterscheid believes, based on this correspondence, that the limitation was "because of a concern expressed by Madison that patents of importation were unconstitutional." *Id.* at 873; see also WALTERSCHEID, TO PROMOTE THE PROGRESS OF SCIENCE AND USEFUL ARTS, *supra* note 128, at 148–56 (discussing Hamilton's proposal in his *Report on the Subject of Manufacturers* to grant patents of importation and Hamilton's concerns about the constitutionality of such patents).

234. See Walterscheid, *Patents and Manufacturing*, *supra* note 33, at 860–78.

235. HAMILTON, REPORT ON MANUFACTURERS, *supra* note 33, at 992 (emphasis added).

236. *Id.* at 1014 (recommending patents for introducers of foreign technology).

237. See Patent Act of 1836, Ch. 357, § 8, 5 Stat. 117, 120–21 (repealed 1861).

238. In 1853, the Supreme Court upheld the validity of Samuel Morse's patent over the electro-magnetic telegraph, for which a patent was issued to him in 1840 and re-issued in 1848—despite allegations of prior inventorship in Germany in England. *O'Reilly v. Morse*, 56 U.S. 62, 62 (1853); see also *id.* at 128 (Grier, J., concurring with the result though dissenting on the issue of costs) (explaining that the "policy and spirit" behind the 1836 Act was in part "to encourage the introduction of foreign inventions and discoveries . . .").

importation of pre-existent foreign inventions was allowed.²³⁹ However, for better or worse, the rules favoring importation of foreign inventions have now largely been eliminated once and for all in the recent patent law reforms, which prohibit inventors from obtaining patents over inventions that were previously published, patented, in public use, on sale, or “otherwise available to the public” *anywhere in the world* before the application filing date.²⁴⁰ As in 1790, the assumption is that, so long as information is accessible to diligent copyists and profit-hungry entrepreneurs, someone will find it and put it to use.²⁴¹

3. *No Assurance of Local Working of Inventions*

The final piece of the U.S. patent laws’ unique, entirely “hands off” approach to promoting innovation in America was the abandonment of local working requirements. As we saw, under colonial and early state patent practices, patents had been tailored to the perceived risk and cost of the project and contained working clauses to ensure local working within an appropriate time frame. If the patentee failed to set up a working technology in the jurisdiction, the legislature could simply end the arrangement or transfer the patentee’s rights to another operator to ensure successful local implementation and commercialization.²⁴² In contrast, U.S. patents did not demand, let alone facilitate, the local implementation of patented technology. This is curious in light of the fact that the Framers believed a major goal of patents was to promote the *actual influx* of new technologies into markets throughout the country.²⁴³ Indeed, when the Framers decided to give Congress the constitutional power to grant patents, they also toyed with a variety of incentives besides patent rights, such as premiums and rewards, for ensuring the actual “advancement of useful knowledge and discoveries.”²⁴⁴

239. See MERGES & DUFFY, *supra* note 43, at 495, 500–03.

240. As of September 16, 2011, a person is entitled to a U.S. patent unless “the claimed invention was patented, described in a printed publication, or in public use, on sale, or otherwise available to the public before the effective filing date of the claimed invention.” 35 U.S.C. § 102(a)(1) (2011).

241. Possibly the most extreme example of this policy is the holding that a doctoral thesis held in a library in Germany (albeit indexed, cataloged, and shelved) that discloses the principles of an invention can count as prior art against a patent and destroy an inventor’s claim of novelty under § 102. See *In re Hall*, 781 F.2d 897 (Fed. Cir. 1986) (applying Patent Act of 1952).

242. See Bracha, *Onning Ideas*, *supra* note 7, at 102, 112.

243. See Vivian J. Fong, *Are We Making Progress?: The Constitution as a Touchstone for Creating Consistent Patent Law and Policy*, 11 U. PA. J. CONST. L. 1163, 1165 (2009) (exploring historical interpretations of the word “Progress” in the IP Clause).

244. See BUGBEE, *supra* note 7, at 126 (citing Madison’s proposal). As Merges and Duffy point out, early drafts of the IP Clause “called for both exclusive rights *and* outright subsidies

And when it was drafting the Patent Act in 1790, the Senate had proposed a set penalty of \$1000 for any patent infringement.²⁴⁵ This would effectively have forced inventors to license their inventions to others for a set fee even if they were not planning to implement them themselves—just as state patent laws had previously done.

Yet, for whatever reason, premiums, working clauses, and compulsory licensing never became a part of the U.S. patent system. Instead, the Patent Act of 1793 mandated full disclosure of the information required to practice an invention, but did not mandate local practice of the invention itself.²⁴⁶ The U.S. patentee certainly had a financial incentive to market or license out his invention in order to take advantage of his exclusive rights and had to disclose his invention sufficiently for another person to practice it upon expiration of his patent.²⁴⁷ But he was not under a statutory obligation to ensure the invention would be practiced anywhere—let alone within a particular state.

This is still the case today. The U.S. Patent Act requires, at most, only a “constructive reduction” to practice.²⁴⁸ Although the ostensible goal of this system is to improve the public welfare and enhance innovation and economic development generally, these rules are designed to only indirectly benefit the community by granting patentees economic incentives to bring potentially profitable inventions to market and by mandating disclosure of those ideas to the public. At the same time, non-practicing patent owners are

for new inventions.” MERGES & DUFFY, *supra* note 43, at 7; *see also* WALTERSCHEID, TO PROMOTE THE PROGRESS OF SCIENCE AND USEFUL ARTS, *supra* note 128, at 46–47 (discussing Madison’s and Charles Pinckney’s early proposals leading to the IP Clause).

245. *See* MERGES & DUFFY, *supra* note 43, at 6–7 (describing rejection of other proposals like a compulsory licensing provision to ensure local working, and stating that “following the somewhat ‘minimalist’ view of government involvement in the economy enshrined in the Constitution, the Convention endorsed exclusive rights only.”).

246. *See* Biagioli, *supra* note 9, at 1138; *see also* Jessica Silbey, *On the Mythical Beginnings of Intellectual Property*, 15 GEO. MASON L. REV. 319, 328 n.59 (2008).

247. *See* Patent Act of 1793, ch. 11, § 3, 1 Stat. 318, 321–22 (repealed 1836) (requiring disclosure sufficient to enable a “person skilled in the art” to practice the invention); *see also* Blanchard v. Sprague, 3 F. Cas. 648, 650 (C.C. Mass. 1839) (J. Story); Grant v. Raymond, 31 U.S. (6 Pet.) 218, 247 (1832). In *Grant*, Chief Justice Marshall wrote:

The third section [of the Patent Act of 1793] requires, as preliminary to a patent, a correct specification and description of the thing discovered. This is necessary in order to give the public, after the privilege shall expire, the advantage for which the privilege is allowed, and is the foundation of the power to issue the patent.

Id.

248. *See* Jeanne C. Fromer, *The Layers of Obviousness in Patent Law*, 22 HARV. J.L. & TECH. 75, 89–90 (2009).

free to bring infringement suits against infringers, even in localities in which the technology is not otherwise practiced.²⁴⁹ The overall effect of this policy is that patentees are not required to practice their inventions in a particular locality, yet are free to prevent others from doing so.

C. JUSTIFYING CONCURRENT STATE PATENT POWER IN THE AGE OF
Laissez Faire

Thus, no matter how strong they were, U.S. patents simply did not fill the role of their state counterparts for promoting investment in useful technology and innovation. In a nation of dual sovereignty, this left open the argument that state patents might still be necessary. Justice Kent made the case most clearly in *Livingston*. For whatever reason, Congress had decided to limit its powers to protecting inventors' exclusive rights in their original inventions. Yet it was obvious, Kent stated, that many "imported improvements, no less than original inventions, ought to be encouraged by patent"²⁵⁰ The power to wield patents to ensure local implementation of new and valuable technology was "important in itself" and, if "well and judiciously exerted," might "ameliorate the condition of society, by enriching and adorning the country with useful and elegant improvements."²⁵¹ If this power did "not reside in the states," Kent concluded, "it reside[d] nowhere, and [would be] wholly extinguished. *This would be leaving the states in a condition of singular and contemptible imbecility.*"²⁵²

In his oral argument before the Supreme Court in *Gibbons*, the lawyer Thomas Oakley built on Kent's position, highlighting the variety of circumstances in which state patents could help encourage deployment of useful technology in the states. What about importers of "useful machines or inventions from abroad?"²⁵³ he asked. How could the states attract the best foreign inventors without the power to give them patents?²⁵⁴ What about technology that, while invented in the United States, was never patented?²⁵⁵ Finally, and most broadly, Oakley prodded, what about technology that had been "given to the public, *and great expense must be incurred to put it into use*?"²⁵⁶ If states lacked power to grant patents in such cases, he concluded, the most

249. See Julie S. Turner, *The Non-Manufacturing Patent Owner: Toward a Theory of Efficient Infringement*, 86 CALIF. L. REV. 179 (1998).

250. *Livingston v. Van Ingen*, 9 Johns. 507, 584 (1812).

251. *Id.* at 584–85.

252. *Id.* at 584 (emphasis added).

253. *Gibbons v. Ogden*, 22 U.S. 1, 48 (1824).

254. *Id.*

255. See *id.*

256. *Id.* (emphasis added).

important new technologies of the day might never be put into working operation in the states. And Congress's policy for promoting innovation in America would be a failure.

The obvious objection to Kent's position was that profit-hungry capitalists did not need state patents to encourage them to develop valuable technology. After all, this was the heyday of *laissez faire* economics. Commercially viable projects, it was assumed, would be undertaken without government subsidy—let alone the protections of a legal monopoly.²⁵⁷ Jurists like Kent were quite familiar with Adam Smith's *Wealth of Nations* and its aversion to government interference in the market.²⁵⁸ Kent agreed with the conventional wisdom that monopolies generally checked the "free circulation of labour" and "enhance[ed] the price of the fruits of industry" by making goods and services more expensive than the free market would have allowed.²⁵⁹ For the most part, "competition and free entry would encourage innovation and expansion far more than the monopoly franchise would."²⁶⁰ However, despite the eagerness of jurists and policymakers to call themselves believers in the "invisible hand" of the markets, they continued to think government was necessary to stimulate development of costly and risky enterprises.²⁶¹

The most influential advocate for additional government subsidy was of course Alexander Hamilton himself. Hamilton, who was otherwise a firm believer in strong U.S. patents, was skeptical that patents for original inventions alone would be sufficient to ensure America stayed competitive in developing and adopting the newest and best technology of the day.²⁶² While

257. See Hovenkamp, *The Classical Corporation in American Legal Thought*, 76 GEO. L.J. 1593, 1594–95 (1988).

258. In his *Dissertations*, Kent concurred with Smith's proposition that the duties of the sovereign should be limited to three: "the duty of protecting the society from foreign violence, the duty of protecting every member from domestic injury, by establishing an exact administration of justice, and lastly the duty of erecting and maintaining certain public works and institutions." JAMES KENT, DISSERTATIONS 19 (1795).

259. 2 JAMES KENT, COMMENTARIES ON AMERICAN LAW *271; see also ASA KINNE, KENT'S COMMENTARIES REDUCED TO QUESTIONS AND ANSWERS 91 (2d ed. 1840).

260. Hovenkamp, *Technology, Politics and Regulated Monopoly*, *supra* note 37, at 1268.

261. See FRIEDMAN, *supra* note 37, at 177; see also Joseph Dorfman, *Chancellor Kent and the Developing American Economy*, 61 COLUM. L. REV. 1290, 1315 (1961). Dorfman speculated:

[Kent] would doubtless, like Jefferson, have called himself a strong believer in *laissez-faire*. However, in his day, *laissez-faire* meant, as a general principle, little more than that its advocates opposed, in most cases, any direct interference by government with the market determination of the prices of goods or of the factors of production.

Id.

262. See HAMILTON, REPORT ON MANUFACTURERS, *supra* note 33, at 1013–14, 1033.

he certainly recognized the argument that entrepreneurs would invest in new enterprises without subsidy, he believed private forces might not always be sufficient because various barriers stood in the way: “[t]hese have relation to the strong influence of habit and the spirit of imitation—the fear of want of success in untried enterprises—the intrinsic difficulties incident to first essays towards a competition with those who have previously attained to perfection in the business to be attempted.”²⁶³ Thus, Hamilton recommended that Congress adopt a strong national policy of encouragements and subsidies to overcome investors’ fear of failure and reluctance to enter new or highly competitive markets.²⁶⁴

I contend that Hamilton’s concerns were reflected at the state and local level in the apology for concurrent state patent laws. Historically, the job of promoting local economic development and construction of infrastructure in the nineteenth century did not fall to Congress. It fell to the states.²⁶⁵ With some exceptions—such as the National Road, the central bank (until Andrew Jackson abolished it in 1832),²⁶⁶ and (after the Civil War) the railroad system²⁶⁷—state governments were almost exclusively responsible for stimulating local economies and improving the standard of living of their constituencies by constructing internal improvements like turnpikes, roads, ferries, and bridges.²⁶⁸

Sometimes the states paid for these construction projects with cash, such as New York’s highly successful venture developing the Erie Canal at a cost of \$7,000,000, financed through collection of tolls.²⁶⁹ But to avoid bankrupting local treasuries and exposing taxpayers to the risk of failure, states more often chartered private companies to take on these projects,

263. HAMILTON, REPORT ON MANUFACTURERS, *supra* note 33, at 988; *see also* CHERNOW, *supra* note 127, at 377 (quoting and summarizing the Hamiltonian barriers to private incentives to invest).

264. HAMILTON, REPORT ON MANUFACTURERS, *supra* note 33, at 1008–34. On Hamilton’s policy agenda for promoting domestic industry and manufacturing through a variety of means including prizes for inventions and direct payments to businesses, *see* WOOD, *supra* note 30, at 102.

265. Lawrence Friedman writes that “[m]ost government intervention, and government regulation, was carried on, in 1800 or 1830, by the *states* not the federal government.” FRIEDMAN, *supra* note 37, at 177 (emphasis added).

266. *See id.* at 179.

267. As Friedman notes, the railroads were initially constructed by private investors or the states, and they were not fully nationalized until 1920. *Id.* at 178; *see also* Herbert Hovenkamp, *Regulatory Conflict in the Gilded Age: Federalism and the Railroad Problem*, 97 *YALE L.J.* 1017, 1018 (1988) (addressing the conflict of sovereignty that eventually occurred over which government was the optimal regulator of the railroads).

268. FRIEDMAN, *supra* note 37, at 178.

269. *Id.* at 181.

granting them monopolies and other exclusive privileges to encourage them to invest.²⁷⁰ In his role as Chancellor of New York, Kent frequently upheld state monopoly grants to operators of public works. The theory was that, without the prospect of a period of freedom from competition and the right to charge an “exclusive toll,” no one would “expend money upon great, and expensive, and hazardous public works, as roads and bridges,” or “become bound to keep them in constant and good repair.”²⁷¹ For this reason, they were considered constitutional and strictly enforced against states under the Contracts Clause.²⁷²

The analogy to state patents for developers of technology was obvious.²⁷³ Just like patents, state monopolies on bridges and roads provided, expressly or implicitly, the exclusive right to develop a risky enterprise and charge the public tolls for its use based on the theory that investors would not invest without a guaranteed period of freedom from competitive entry.²⁷⁴ As with patents, these grants were not seen as monopolies “in the odious sense of the

270. *Id.* at 180–81; see also Peter Karsten, *Supervising the “Spoiled Children of Legislation”: Judicial Judgments Involving Quasi-Public Corporations in the 19th Century U.S.*, 41 AM. J. LEGAL HIST. 315 (1997) (describing monopoly franchises that legislators granted to corporate entities offering to provide mills, ferries, toll roads and railways to communities in need of such entities, and arguing that courts generally protected the original grantees’ monopolies from later legislative alterations and prospective competitors).

271. See *Newburgh & C. Turnpike Road Co. v. Miller*, 5 Johns. Ch. 101 (N.Y. Ch. 1821) (enjoining as a nuisance a rival bridge constructed close to a toll bridge that had been operating for over ten years); see also Dorfman, *supra* note 174, at 1293 n.16 (discussing Kent’s a dissenting opinion in *Palmer v. Mulligan*, 3 Cai. 307, 314 (N.Y. Sup. Ct. 1805)).

272. See *In re Binghamton Bridge*, 70 U.S. 51 (1865) (upholding state’s legislative guaranty to a toll bridge company, making it unlawful for competitors to erect a bridge within a certain distance of the original structure, under the Contracts Clause); *Enfield Toll Bridge Co. v. Hartford & N.H.R. Co.*, 17 Conn. 40 (1845) (upholding state’s grant of the exclusive right of erecting a bridge and taking tolls in light of the public benefit resulting from the monopoly); see also Karsten, *supra* note 270 (describing courts’ willingness to protect legislative monopolies from later competitors). *But see* *Charles River Bridge v. Warren Bridge*, 36 U.S. (11 Pet.) 420 (1837) (C.J. Taney) (holding that a legal monopoly will not necessarily be implied in a corporate charter to bridge proprietors); see also Mossoff, *supra* note 44, at 966 n.58. In 1873, the Supreme Court famously held that a state can place a monopoly privilege in a corporate charter without violating individual rights. See *Slaughter-House Cases*, 83 U.S. 36, 83 (1873) (upholding a Louisiana law that allowed the city of New Orleans to charter a single private corporation with the exclusive right to control the slaughter-house business in the city for twenty-five years); see also Hovenkamp, *The Classical Corporation in American Legal Thought*, *supra* note 257, at 1602–03 (discussing *Slaughter-House Cases*).

273. I am not the first to make this analogy. See Oskar Liivak, *Rethinking the Concept of Exclusion in Patent Law*, 98 GEO. L.J. 1643, 1645 n.13 (2010) (citing several scholars of patent law who see a resemblance between the monopoly granted to the proprietors of the Charles River Bridge and modern patents).

274. See Hovenkamp, *Technology, Politics and Regulated Monopoly*, *supra* note 39, at 1268; see also KUTLER, *supra* note 39, at 18–21.

term” because they were given out “in consideration of expenses to be incurred by the grantees, and in contemplation of a public benefit . . . to reimburse such expenses,” and were granted only “within certain limits, for a limited time”²⁷⁵ Indeed, for Justice Kent, there was no difference between states’ authority to make these grants and states’ authority to grant patents for costly inventions like the steamboat.²⁷⁶ When viewed in the Hamiltonian economic framework, this stance makes perfect sense. Inventions, regardless of whether they were new or already employed in other localities, were costly, difficult, and time consuming to develop and market to the public.²⁷⁷ Besides the cost of conception itself, there was the cost of actually building the invention, requiring expensive materials and facilities; the cost of performing the experiments and real-world testing; the difficulty of gaining access to land or public waterways; and the cost of advertising and marketing something new to the public.²⁷⁸ Obtaining funding for such an operation would have been difficult. Additional government subsidy in the form of a period of freedom from competition would certainly have been justified.²⁷⁹

D. NEW YORK’S STEAMBOAT PATENT AS A MARKET CORRECTIVE

No contemporary technology better illustrates the Hamiltonian appeal for state patents than the steamboat and the efforts of American inventors to put it into practice. The problem had nothing to do with a lack of theoretical knowledge. When New York granted its patent to Robert Livingston in 1798, Europeans and Americans had been experimenting with steamboats for

275. *Enfield Toll Bridge Co.*, 17 Conn. at 40.

276. “The capacity to grant separate and exclusive privileges appertains to every sovereign authority. It is a necessary attribute of every independent government. All our bank charters, turnpike, canal and bridge companies, ferries, markets, &c. are grants of exclusive privileges for beneficial public purposes.” *Livingston v. Van Ingen*, 9 Johns. 507, 573 (1812).

277. Major categories of patented inventions in this period included bridges, boats, boilers, distilleries, methods for making iron, nail machines, stoves, salt-making machines, thrashing machines, machines for raising water, washing machines, steam engines, and vast varieties of mostly water-operated mills—including flax mills, flour mills, saw mills, and wind mills. See WILLIAM ELLIOT, U.S. STATE DEP’T. A LIST OF PATENTS GRANTED BY THE UNITED STATES, FOR THE ENCOURAGEMENT OF ARTS AND SCIENCES: ALPHABETICALLY ARRANGED FROM 1790 TO 1820 (Washington, D.C., S. Alfred Elliot, 1823), available at <http://hdl.handle.net/2027/loc.ark:/13960/t0ns1mn0z> [hereinafter LIST OF PATENTS].

278. See GAVIN WRIGHT, HISTORICAL FOUNDATIONS OF AMERICAN TECHNOLOGY 3 (Stanford University, 2007) (explaining that American manufacturing methods, in particular, were capital-intensive, relying on special-purpose machinery that required material inputs and fuels, and on unskilled labor).

279. See HAMILTON, REPORT ON MANUFACTURERS, *supra* note 33, at 988–90.

years.²⁸⁰ John Fitch allegedly conceived of his own idea for a boat propelled by steam-powered paddle wheels in April of 1785.²⁸¹ But Fitch spent the final thirteen years of his life trying unsuccessfully to raise the capital he required to put a steamboat into wide scale operation. He relied initially on wealthy friends.²⁸² But he soon sought larger investors, eventually forming a company of shareholders.²⁸³ They demanded quick success and grew impatient with the never-ending trial runs, at least one of which was an embarrassing failure.²⁸⁴ By 1790, Fitch had finally succeeded in developing a steamboat large and fast enough to efficiently transport people or freight from Philadelphia to Burlington and to run thousands of miles without a hitch.²⁸⁵ In 1791, he obtained a U.S. patent on his design, hoping to leverage his national patent rights into the capital he needed to monetize his steamboats by charging the public tolls for their use.²⁸⁶ Nonetheless, his investors gradually lost faith and shifted their support to other inventors like James Rumsey, whom they believed would make them a faster profit.²⁸⁷

But as of the year 1798, neither Fitch nor Rumsey nor anyone else had set up a single commercial steamboat on the Hudson River. People still had to rely on slower alternatives like sloops and ferries. This was a major problem for the New York government, entrusted with the responsibility of promoting local infrastructure and economic development in the state. So the Legislature decided to take matters into its own hands and grant its own patent to someone “more skilful in [the] business,”²⁸⁸ with the money,

280. See WESTCOTT, LIFE OF JOHN FITCH, *supra* note 140, at 373–98 (describing steamboat experiments in Europe and America).

281. For Fitch’s account of his conception, see *id.* at 120–22.

282. Westcott describes Daniel Longstreth’s account of Fitch’s efforts to develop his first model steamboat in the log shop of Cobe Scout. The “machinery was made of brass, with the exception of the *paddle-wheels*, which were made of wood by [an influential Pennsylvania politician] Nathaniel B. Boileau, whist on a visit during vacation from Princeton College.” Fitch tested the model “on a small stream on Joseph Longstreth’s meadow . . . and it realized every expectation.” *Id.* at 123.

283. In the spring of 1787, in an attempt to “carry [his invention] into an experiment” and turn his fledgling steamboat into an “effective use and service[,]” Fitch formed a company, dividing his interest in his eventual profits with several shareholders. See *id.* at 181–82 (quoting shareholder agreement), 183–84 (list of shareholders); see also COX, *supra* note 21, at 8.

284. See COX, *supra* note 21, at 13.

285. See WESTCOTT, *supra* note 140, at 376.

286. See LIST OF PATENTS, *supra* note 277, at 76 (Fitch’s 1791 patent); see also COX, *supra* note 21, at 15; WESTCOTT, *supra* note 140, at 327–28.

287. See COX, *supra* note 21, at 8, 13.

288. Oakley made this point during his argument in support of states’ powers to grant patents in *Gibbons*: “[M]ay not the State determine that it is against the public interest, that steam boats should be built or navigated, unless under the direction, or with the license; of

connections, and entrepreneurial spirit required to perfect steamboats and make them ready to service the New York public: Robert Livingston.²⁸⁹ True, Livingston was no inventor. But he was a “rich, enthusiastic, liberal, influential patron,” willing to supply Robert Fulton with “the very best machinery that could be made in Europe” and prepared to undertake the monumental task of building, testing, and operating a fleet of steamboats capable of competing in an unproven new market.²⁹⁰

According to the theory behind the U.S. patent law, New York’s patent to Livingston should not have been necessary. Fitch’s U.S. patent provided him the exclusive right to make, use, or sell his steamboats and the exclusive right to extract damages from infringers for fourteen years.²⁹¹ This should have given Livingston a monopoly-profit incentive to buy or license Fitch’s patent and introduce Fitch’s steamboat into markets throughout the country.²⁹² At the least, Fitch’s patent should have served as an important “signal” to investors and increased Fitch’s chances of securing the capital he needed to fund his operations.²⁹³ So what went wrong? The most obvious possibility is a breakdown in patent licensing.²⁹⁴ We can easily imagine a hypothetical scenario in which Livingston approached Fitch for a license, but Fitch refused to accept Livingston’s terms due to an irrational pride in his work or the parties’ failure to accurately estimate the future value of Fitch’s design.²⁹⁵

an individual, who may be thought particularly skilful in that business?” *Gibbons v. Ogden*, 22 U.S. 1, 57 (1824).

289. See *Livingston v. Van Ingen*, 9 Johns. 507, 572 (1812); JOHNSON, *supra* note 21, at 27 (noting that while “neither Fitch nor Rumsey possessed the political connections or financial patronage to fund the substantial capital investment needed to complete their experiments and turn their discoveries into fully operational steamboats,” Robert Fulton achieved this combination thanks to his partnership with Livingston).

290. WESTCOTT, *supra* note 140, at 390.

291. Patent Act of 1793, ch. 11, §§ 1 (exclusive rights), 5 (damages), Stat. 318, 322 (repealed 1836).

292. See Kitch, *supra* note 71 (describing the “prospect theory” justification for patents).

293. See Long, *supra* note 1, at 653.

294. Patent licensing can be hindered by a variety of transaction costs, especially in cases where a pioneering invention is not as profitable as subsequent derivations competing in the same market. See, e.g., Robert Merges, *Intellectual Property Rights and Bargaining Breakdown: The Case of Blocking Patents*, 62 TENN. L. REV. 75, 84–89 (1994) (describing several real and hypothetical examples of bargaining breakdowns between holders of blocking patents due to factors such as erroneous valuations of the future value of the pioneer technology or irrational motives such as the inventor’s pride in her invention, leading her to believe she is entitled to higher profits than might actually be warranted).

295. This is what happened with the British inventor James Watt, who held an early steam engine patent but refused to license it to improvers. See Shavell & Van Ypersele, *supra*

But the more likely possibility is that this was a pure market failure.²⁹⁶ A commercial steamboat operation was an incredibly risky and costly investment in 1798. No one knew for sure whether steamboats could ever be made to work efficiently and safely enough to outcompete cheaper, tried-and-true alternatives like the sailboat. Maybe Livingston crunched the numbers and decided that if he was going to take the risk, he demanded a patent with a much longer term and covering a much broader range of subject matter than Fitch's narrow U.S. patent would have allowed.²⁹⁷ Unlike Congress, New York was willing to give Livingston a thirty-year patent granting exclusive rights over "all and every species of boats or water craft, which might be impelled by force of fire or steam."²⁹⁸ What is more, as a legislative act, Livingston's patent would demand an "extremely strong" presumption of validity and be subject to minimal policing by state courts.²⁹⁹

In sum, New York's patent gave Livingston something U.S. patents did not provide: a guaranteed monopoly, specially designed to compensate him for the "uncertainty and hazard" of developing and commercializing "a very expensive experiment."³⁰⁰ Moreover, unlike a U.S. patent, it also gave Livingston an *obligation* to actually get the job done; if he failed to institute steamboats in the state, his grant would be retracted.³⁰¹ It is undeniable that Livingston's exclusive rights led to several years of higher steamboat fares.³⁰² But we might nonetheless look at the results and be inclined to agree with Justice Kent's conclusion that, thanks to New York's patent, "the experiment of navigating boats by steam has been made, and crowned with triumphant

note 67, at 543; *see also* Merges, *Intellectual Property Rights and Bargaining Breakdown*, *supra* note 294, at 84–89.

296. A "market failure," where government incentives may be required as a supplement to private funding, is generally defined in the context of innovation policy as cases where markets do not ensure investment in "long-range research, especially in high-risk technologies whose calculable value to a given firm is far smaller than their eventual social value." *See* Christopher Newfield, *Avoiding Network Failure: The Case of the National Nanotechnology Initiative*, in *STATE OF INNOVATION* 282 (Fred L. Block & Matthew R. Keller eds. 2010).

297. *See* COX, *supra* note 21, at 21–22 (recounting efforts of Livingston and others to obtain broader monopoly rights than federal patents afforded).

298. *Livingston v. Van Ingen*, 9 Johns. 507, 568 (1812).

299. *Id.* at 572–73.

300. In Kent's description, the 1798 law was "a new and original grant to [Livingston], encouraging him, by the pledge of an exclusive privilege for twenty years, to engage, according to the language of the preamble of the statute, in the '*uncertainty and hazard of a very expensive experiment.*'" *Id.* at 572 (quoting New York's 1798 grant to Livingston) (emphasis added).

301. *Id.* at 509–10.

302. *See* Williams, *supra* note 21, at 1420–21 (on decline in steamboat fares following *Gibbons*); *see also* COX, *supra* note 21, at 175–80.

success [promising] to become a great public blessing, by giving astonishing facility, despatch and safety, not only to travelling, but to the internal commerce of this country.”³⁰³ Maybe a savvier and more charming inventor than Fitch could have leveraged his national patent rights to attract investors, or a better businessman than Livingston could have obtained such a significant lead in the market that no monopoly would have been required to appropriate returns at all.³⁰⁴ But we will never really know, just as we do not know today whether pharmaceutical companies would bring life-saving drugs to market without the prospect of a limited period of charging monopoly prices.³⁰⁵ The fact remains that until the New York legislature got involved, no one had managed to do it.

VII. CONCLUSION

After reading this history of support for state patent laws, the rational question to ask is: could a state grant a patent today? The Supreme Court’s current statutory preemption doctrine, limiting states’ ability to grant “patent-like” rights, would seem to preclude the possibility of states granting exclusive rights for inventions that do not meet federal criteria for patentability, or for importers or “mere possessors” of new technology.³⁰⁶ However, as we have seen, this doctrine bears little resemblance to the principles articulated by Alexander Hamilton in the *Federalist No. 32*, requiring preemption only if concurrent state patent powers would create an

303. *Livingston*, 9 Johns. at 585.

304. On patents as signals to investors, see Long, *supra* note 1, at 653. According to this oft-cited empirical study, in most industries U.S. patents are not viewed as an effective incentive compared with other market advantages like lead-time and secrecy. Richard C. Levin, Alvin K. Klevorick, Richard R. Nelson, Sidney G. Winter, Richard Gilbert & Zvi Griliches, *Appropriating the Returns from Industrial Research and Development*, 3 Brookings Papers on Econ. Activity 783 (1987).

305. See A. Samuel Oddi, *Plagues, Pandemics, and Patents: Legality and Morality*, 51 IDEA 1, 12–14 (2011) (“Pharmaceutical enterprises developing drugs for the prevention or treatment of diseases rely heavily on the patent system. The development of such drugs is extremely costly, as is the governmental approval process. Without strong patent protection, the incentive to develop new drugs is considerably undercut.”); see also Jonathan M. Barnett, *Cultivating the Genetic Commons: Imperfect Patent Protection and the Network Model of Innovation*, 37 SAN DIEGO L. REV. 987, 989–90 (2000) (“Patent rights in genetic resources encourage private investors to sink funds into uncertain pharmaceutical projects that generate enormous development costs, take many years to yield a marketable product, and are often vulnerable to relatively low-cost imitators.”).

306. See *Bonito Boats, Inc. v. Thunder Craft Boats, Inc.*, 489 U.S. 141, 151–67 (1989). On Robert Livingston’s status as a “mere possessor,” see *Livingston*, 9 Johns. at 542.

“immediate constitutional repugnancy.”³⁰⁷ In this Article, I have shown that well-respected jurists in the nineteenth century did not believe such a “repugnancy” existed and that the Constitution therefore left states with broad residual authority to grant their own patents.³⁰⁸ As Kent put it in *Livingston*, there was no need to interpret Congress’s patent power more broadly than necessary to accomplish the goal “for which it was granted, which was to reward the beneficent efforts of genius and to encourage the useful arts.”³⁰⁹ So long as states respected the exclusive rights conferred by U.S. patents, Congress’s power was “fully satisfied”³¹⁰ In the modern world of patent licensing, this simply means that a state patentee would be wise to obtain a license before practicing a technology that potentially infringed a U.S. patent, or risk costly litigation in federal court and invalidation of the state patent under the Supremacy Clause.³¹¹

Due to the stark contrast between these views and modern presumptions of federal supremacy in patent law, I have taken the time to explain and justify them in contemporary context. I hypothesize that the perceived need for concurrent state patent powers was not based solely on the weakness of

307. See THE FEDERALIST NO. 32, *supra* note 110, at 157 (Alexander Hamilton). Hamilton stated with regard to the concurrent power of taxation:

The particular policy of the national and of the State systems of finance might now and then not exactly coincide, and might require reciprocal forbearances. It is not, however a mere possibility of inconvenience in the exercise of powers, but an *immediate constitutional repugnancy* that can by implication alienate and extinguish a pre-existing right of sovereignty.

Id. (emphasis added).

308. See U.S. CONST. amend. X; see also *Livingston*, 9 Johns. at 581–82 (stating that states have concurrent powers to grant patents); *Gibbons v. Ogden*, 22 U.S. 1, 171–72 (1824) (leaving Kent’s opinion on state patent powers intact); *Patterson v. Kentucky*, 97 U.S. 501, 508–09 (1878) (upholding Kent’s opinion in *Livingston* on concurrent state patent powers); HAMILTON, REPORT ON MANUFACTURERS, *supra* note 33, at 1014 (Congress lacked power to grant patents to introducers of inventions from abroad); STORY, COMMENTARIES, *supra* note 25, at 79 (states may have concurrent powers to grant patents to true inventors and certainly have independent powers to grant patents to “possessors” of technology); 1 WILLIAM BLACKSTONE, COMMENTARIES *265 (stating that patent powers are concurrent, even if states are unlikely to grant their own patents in most cases, given requirement that inventors choose between state or federal rights).

309. *Livingston*, 9 Johns. at 582.

310. *Id.* at 582–83.

311. *Id.* If the state patentee did obtain a license from any U.S. patent implicated by his or her technology, there would be no infringement, so the state patent would not be invalid in that scenario under Kent’s reading of the preemptive power of the IP Clause and the Supremacy Clause. As discussed above, we must assume the possibility that transaction costs, such as problems of valuation and inter-personal conflicts, might hinder the fluid licensing of U.S. patents to state patentees. See *Merges, Intellectual Property Rights and Bargaining Breakdown*, *supra* note 294, at 84–89.

U.S. patent rights in the early nineteenth century. Instead, the appeal for state patents was motivated by a more fundamental, inherently Hamiltonian concern that market forces, even when supplemented by U.S. patent rights—which, as we have seen, were themselves based on a uniquely *laissez-faire* approach to promoting innovation—would not provide sufficient incentive for inventors and their financial backers to develop and actually bring to market costly innovations. Therefore, state governments must retain autonomy to grant their own patents for technologies that they deemed to be of immense public value. This, I assert, is the main reason New York chose to grant a patent on the steamboat to Robert Livingston in 1798 and the main reason that Federalist jurists like Justice Kent supported New York’s decision and upheld states’ powers to grant similar exclusive rights in the future.

Based on this history, I suggest that we may wish to reexamine our assumption that states’ powers to grant their own patents are, or should be, preempted. Congress’s powers to grant federal patents (at least within the limits of the IP Clause)³¹² are not in serious doubt today.³¹³ This means that,

312. Whether Congress can rely on the Commerce Clause to do more than the IP Clause specifies is a hotly debated topic. See Paul Heald and Suzanna Sherry, *Implied Limits on the Legislative Power: The Intellectual Property Clause as an Absolute Constraint on Congress*, 2000 U. ILL. L. REV. 1119, 1183–87 (2000) (arguing that the IP Clause, when interpreted in its historical context, places absolute limits on Congress’s powers to grant patents or other exclusive rights over material otherwise in the public domain in some though not all circumstances); Jeanne Fromer, *The Intellectual Property Clause’s External Limitations*, 61 DUKE L.J. 1329, 1344 (2012) (“All in all, the IP Clause’s text and the constitutional structure volunteer a suggestive—but not ironclad—argument that the Clause’s unique construction operates externally to forbid Congress from using its other powers to promote the progress of science and useful arts beyond the means specified in the Clause.”). But see Thomas Nachbar, *Intellectual Property and Constitutional Norms*, 104 COLUM. L. REV. 272 (2004) (arguing that the limits of the IP Clause should not be read externally into other parts of the Constitution, that Congress could use the Commerce Clause to grant exclusive rights even in cases where it could not do so under the IP Clause, and that the Constitution generally offers little protection against rent-seeking).

313. The Supreme Court has held that Congress’s powers to regulate economic activity include the power to regulate the “channels of interstate commerce,” such as water travel, railroads, highways and hotels; the “instrumentalities of interstate commerce, or persons or things in interstate commerce,” such as suspects carrying contraband across state powers; and, most broadly, the power to regulate “those [economic] activities that substantially affect interstate commerce.” *United States v. Morrison*, 529 U.S. 598, 609 (2000). However, in recent years the Supreme Court has placed some limits on Congress’s power to regulate activity in the “non-economic” arena. See *United States v. Lopez*, 514 U.S. 549, 561 (1995) (striking a federal law prohibiting possession of firearms on school premises because gun toting in school zones “has nothing to do with ‘commerce’ or any sort of economic enterprise,” the law “is not an essential part of a larger regulation of economic activity” and the statute contained no “jurisdictional element that would ensure, through case-by-case

unlike in 1824, Congress probably has plenary power to preempt state patent laws entirely if it wished.³¹⁴ However, given America's long history of state patents and historic views that they should remain constitutional, it would be unwise for Congress, or federal courts, to nationalize patent law prematurely, without engaging in a thorough analysis of the benefits of state patent laws, on the one hand, and the potential for confusion, waste, and conflict, on the other.³¹⁵

The Supreme Court, although it has effectively prohibited states from granting patents, has never performed this analysis.³¹⁶ Chief Justice Marshall

inquiry, that the firearm possession in question affects interstate commerce.”); *see also Morrison*, 529 U.S. at 617–618 (invalidating the Violence Against Women Act, which provided a federal cause of action for victims of crimes of violence “motivated by gender,” on the ground that gender-based crime, though apparently a nationwide problem, involved “noneconomic, violent conduct,” and an essential aspect of states’ “police powers” is the “suppression of violent crime and vindication of its victims”; again, the statute had no jurisdictional element when it was challenged). Most recently, the Court placed limits on Congress’s power, at least under the Commerce Clause, to “regulate” inactivity by compelling individuals to buy health insurance they might not otherwise purchase or pay a fine. *See Nat’l Fed’n of Indep. Bus. v. Sebelius*, 132 S. Ct. 2566, 2587 (2012).

314. “By definition, preemption disputes involve lawmaking in an area in which both the federal government and the states have the power to legislate. It is generally accepted that if the federal government chooses to do so, it has the power to displace state law altogether in those areas.” Susan Stabile, *Preemption of Federal Law by State Law: A Task for Congress or the Courts?*, 40 VILL. L. REV. 1, 9 (1995) (citing *Garcia v. San Antonio Metro. Transit Auth.*, 469 U.S. 528, 554–56 (1985) (holding that the Commerce Clause permitted the federal government to preempt state and afford wage and hour protection to transit employees even though state also had this ability)).

315. As Susan Stabile explains:

When the preemption balance is struck incorrectly, negative consequences result. In some cases, there will be an improper interference with a state sphere of authority, preventing the state’s attainment of its goals without appropriate justification. Also, preempting state law without adequate federal justification limits the ability of states to act as innovators of change. “Experiments” conducted at the state level may lead to solutions to social problems that may later be adopted at a national level. Finally, improper preemption decisions give insufficient regard to the purposes and goals of Congress in passing federal legislation.

Stabile, *supra* note 314, at 10.

316. *See Bonito Boats, Inc. v. Thunder Craft Boats, Inc.*, 489 U.S. 141, 167 (1989) (concluding—without mentioning any potential benefits of state laws designed to induce investment in innovation by granting exclusive rights—that because Florida’s antimolding statute “substantially restricts the public’s ability to exploit an unpatented design in general circulation, raising the specter of state-created monopolies in a host of useful shapes and processes for which patent protection has been denied or is otherwise unobtainable[,] [i]t thus enters a field of regulation which the patent laws have reserved to Congress.”) (citing *Rice v. Santa Fe Elevator Corp.*, 331 U.S. 218, 230 (1947)); *see also* discussion *supra* Part I (discussing *Bonito Boats*).

struck down the steamboat patent in *Gibbons* in 1824. But Marshall provided no insights on the potential consequences of concurrent state patent laws, instead relying on less controversial, and far less complicated, reasons for preempting the monopoly.³¹⁷ The steamboat patent covered a recent innovation that was essential to the transportation of goods and people across state lines and that would transform the entire American economy.³¹⁸ Livingston and Fulton's exclusive rights to operate steamboats in New York clearly affected interstate commerce and, by modern standards, produced a flagrant violation of the Dormant Commerce Clause³¹⁹—making it unnecessary for Marshall to address whether their monopoly would otherwise have independently interfered with Congress's power to grant patents under the IP Clause.³²⁰ Thus, the costs and, more so, the benefits of concurrent state patent regimes have never been fully examined.

Whether state patent laws make any sense in the modern economy is the subject of another article, in progress, where I suggest that in some cases state patents could provide cash-strapped state governments with a cheaper policy tool for promoting innovations that are of particular importance to certain states and that are connected to state resources.³²¹ Examples include technologies related to public utilities, local infrastructure, and region-specific industries like agriculture in California or natural gas drilling in Texas. At the same time, introducing “bottom-up” experimentation into the process of designing effective patents and effective patent laws could improve the efficacy of the American patent system and innovation policy as a whole.³²² Whether or not readers will agree with any of these arguments, it is nonetheless enlightening to learn that our predecessors recognized what we seem to have forgotten: that U.S. patents were never intended as a complete

317. See *Gibbons v. Ogden*, 22 U.S. 1 (1824). As said, in his actual holding, Marshall relied on Thomas Gibbons' federal coasting license to preempt the state law. See discussion *supra* Section V.C.

318. See discussion *supra* Section IV.A.

319. Under the modern Dormant Commerce Clause, courts will strike down any state law—whether a tax, a tariff, a licensing requirement, or presumably a patent—that overtly discriminates against out-of-state entities or that “unduly burdens” interstate commerce, “impeding free private trade in the national marketplace.” *Reeves, Inc. v. Stake*, 447 U.S. 429, 437 (1980).

320. *Gibbons*, 22 U.S. at 177–78.

321. The Supreme Court itself has suggested in dicta that states may have an interest in granting patents on innovations that are of “local importance.” See *Goldstein v. California*, 412 U.S. 546, 557 (1973) (“The patents granted by the States in the 18th century show . . . a willingness on the part of the States to promote those portions of science and the arts which were of local importance.”).

322. Hrdy, *State Patents as a Solution to Underinvestment in Innovation*, *supra* note 40. On the need for bottom-up reform in patent law, see Nguyen, *supra* note 41.

replacement for their state counterparts and need not represent the exclusive means for encouraging private investment in technology and innovation.

THE BACKGROUND OF OUR BEING: INTERNET BACKGROUND CHECKS IN THE HIRING PROCESS

Alexander Reicher[†]

ABSTRACT

Many employers are searching job applicants on Google, Facebook, and any number of other search engines and social networks. For some, this search is a cause of great concern, leading the FTC and Senators Al Franken and Richard Blumenthal, among others, to investigate the issue. So-called “internet background checks” can vary greatly in their degree of thoroughness; on the one hand, a third-party agency might produce a formal “internet background report” documenting all websites consulted and evaluating every source, while on the other hand, a member of the employer’s hiring committee might simply search the candidate off the record. Both of these practices can inform the decision-makers in the hiring process, and both, ultimately, afford the employer access to information about the candidate they might not otherwise find in the rest of the candidate’s application. This Note analyzes the legal and normative issues surrounding internet background checks. After reviewing studies showing that, at minimum, a fifth to a quarter of employers use internet search engines or social networks to screen candidates at some point during the hiring process, this Note suggests a taxonomy of three different approaches to internet information gathering. It then considers how fair credit reporting and equal employment laws might apply to these three approaches. The Note then concludes with recommended best practices for employers in light of this legal analysis.

© 2013 Alexander Reicher. This Note is the first-place winner of the Berkeley Technology Law Journal’s annual student writing competition.

[†] Associate, Latham & Watkins LLP, San Francisco; J.D., University of California Berkeley School of Law (2012). I would like to thank Chris Hoofnagle for his helpful guidance while writing this Note and Chris Conley and Tom Reicher for their comments on earlier drafts. I would also like to thank my wife, Janice, for her assistance and support throughout the entire process. The views expressed in this Note are solely those of the author.

TABLE OF CONTENTS

I.	INTRODUCTION.....	116
II.	INTERNET BACKGROUND CHECKS ON PROSPECTIVE EMPLOYEES: THE CURRENT LANDSCAPE	118
III.	A TAXONOMY OF INTERNET BACKGROUND INFORMATION GATHERING	121
IV.	GATHERING INTERNET BACKGROUND INFORMATION	123
	A. THE FAIR CREDIT REPORTING ACT	124
	B. THRESHOLD DEFINITIONS FOR DETERMINING FCRA APPLICABILITY: “CONSUMER REPORTING AGENCY” AND “CONSUMER REPORT”	125
	C. THIRD-PARTY INFORMATION GATHERERS MUST ENSURE ACCURACY	132
	D. HEIGHTENED FCRA STANDARD FOR “MATTERS OF PUBLIC RECORD”	135
	E. STATE FAIR CREDIT REPORTING LAWS—CALIFORNIA’S APPROACH.....	137
V.	USING INTERNET BACKGROUND INFORMATION.....	139
	A. TITLE VII PROTECTIONS AGAINST RELIGIOUS DISCRIMINATION IN THE HIRING PROCESS.....	139
	B. ADA PROHIBITION ON PREEMPLOYMENT INQUIRIES.....	145
	C. STATE PROHIBITIONS ON LIFESTYLE DISCRIMINATION	148
VI.	RECOMMENDATION: THE LEGAL AND STRUCTURAL ADVANTAGES OF THE THIRD-PARTY APPROACH	149
VII.	CONCLUSION	152

I. INTRODUCTION

The Internet has changed the way we meet and evaluate people, and in particular, has altered the way employers evaluate prospective employees during the hiring process. Although the studies assessing the percentage of employers that use social networks and search engines to screen candidates vary widely in their conclusions, by the most conservative estimate about a fifth to a quarter of employers are searching job applicants on Google,

Facebook, or any number of other search engines and social networks.¹ Some estimate that this number is much higher—claiming that as many as 91% of employers are using social networks at some point during the hiring process.² In light of these findings, this Note is concerned with the legal and normative issues surrounding various methods of performing preemployment internet background checks.

Some consider internet background checks inherently unfair because these checks can be inaccurate—mixing information about job applicants with the same name—or because they expose the hiring individual to information relating to the applicant’s status as a member of a protected class under equal employment laws.³ In 2006, the Finnish Data Protection Ombudsman outlawed internet searches of potential employees entirely.⁴ More recently, Senators Al Franken and Richard Blumenthal wrote to the CEO of Social Intelligence Corp., one of the more prominent internet background screening services, expressing their concern that “their business practices relating to personal privacy were unfairly detrimental to prospective employees.”⁵ The Federal Trade Commission also investigated Social Intelligence for compliance with the Fair Credit Reporting Act, ultimately deciding to take no further action.⁶

State and federal laws, including fair credit reporting and equal employment statutes, regulate the traditional forms of preemployment

1. See SHRM Survey Findings: The Use of Social Networking Websites and Online Search Engines in Screening Job Candidates, SOCIETY FOR HUMAN RESOURCE MANAGEMENT (Aug. 25, 2011), <http://www.shrm.org/Research/SurveyFindings/Articles/Pages/TheUseofSocialNetworkingWebsitesandOnlineSearchEnginesinScreeningJobCandidates.aspx>.

2. *Managing Your Online Image Across Social Networks*, REPPLER (Sept. 27, 2011, 5:00AM), <http://blog.reppler.com/2011/09/27/managing-your-online-image-across-social-networks/>.

3. See *infra* Section IV.C (discussing the challenges of gathering accurate information for an internet background check); *infra* Part V (discussing the application of equal employment laws to internet background checks).

4. William McGeeveran, *Finnish Employers Cannot Google Applicants*, INFORMATION, LAW, AND THE LAW OF INFORMATION BLOG (Nov. 15, 2006), <http://blogs.law.harvard.edu/infolaw/2006/11/15/finnish-employers-cannot-google-applicants/>.

5. Press Release, Senator Richard Blumenthal, Blumenthal, Franken Call on Social Intelligence Corp to Clarify Privacy Practices (Sept. 19, 2011), *available at* <http://blumenthal.senate.gov/newsroom/press/release/blumenthal-franken-call-on-social-intelligence-corp-to-clarify-privacy-practice>.

6. Letter from Maneesha Mithal, Associate Director, Federal Trade Commission Division of Privacy and Identity Protection, to Renee Jackson, Att’y for Social Intelligence (May 9, 2011), <http://www.ftc.gov/os/closings/110509socialintelligenceletter.pdf>.

screening, including credit, criminal, and character background checks.⁷ This Note will address how these laws apply to internet background reporting, and specifically will seek to answer two fundamental questions. First, how does this mosaic of statutes apply to internet information gathering? Second, once the information about a prospective employee is gathered, how do these laws control how that information is used in the hiring process?

The discussion proceeds in five parts. Part II considers several studies assessing the prevalence of internet background checks in the hiring process. Part III explains the taxonomy of different approaches to information gathering in the context of internet background checks. Part IV then analyzes how the Fair Credit Reporting Act and one state fair credit reporting statute might apply to the information gathering process. Part V considers the application of equal employment laws to the internet background process. Finally, based upon this legal analysis, Part VI recommends several best practices for employers to follow.

II. INTERNET BACKGROUND CHECKS ON PROSPECTIVE EMPLOYEES: THE CURRENT LANDSCAPE

The term “internet background checks”⁸ refers to the general phenomenon in which employers gather information from the internet about a person—in this case, a job candidate. As discussed, *infra*, the hiring committee (or other hiring decision-making body) can acquire this information in a variety of ways—by searching the candidates themselves, by delegating the task to a special department with no decision-making authority within the organization, or by contracting with a third-party service outside of the organization. The internet background check, therefore, can vary greatly in its thoroughness; on the one hand, a third-party agency might produce a formal “internet background report” documenting all websites consulted and evaluating every source, while, on the other hand, a hiring committee member might simply “google” the candidate off the record.⁹ Both of these practices can inform the decision makers in the hiring process, and both ultimately afford the employer access to information about the candidate they might not otherwise find in the candidate’s application.

7. See *infra* Part IV (discussing state and federal fair credit reporting laws); *infra* Part V (discussing state and federal equal employment laws).

8. I suggest the use of this term instead of “internet and social media background checks,” as many have called them, since the latter phrase is tautological (social media networks, after all, are part of the Internet).

9. See *infra* Part III (proposing a taxonomy of internet background information gathering).

Moreover, both approaches to internet background checking search the same source: the Internet. It is the wide accessibility of the Internet that makes internet background checks readily accessible, in contrast to, for example, credit checks or criminal background checks, which are generally only available for specific purposes through consumer reporting agencies or the government.¹⁰

Unfortunately, the majority of studies assessing the number of employers that use internet background checks in the hiring process appear to fail, at least according to their reported survey methods, to account for the multitude of ways in which internet background information is gathered. The studies by the Society for Human Resource Management, CareerBuilder (2009 and 2012 studies), and Cross-Tab—four of the five surveys summarized in Figure 1 and Figure 2—only surveyed the internet background checking habits of individuals who worked as human resource professionals, hiring managers, or recruiters.¹¹ In other words, these studies failed to account for the more casual forms of internet background checking, such as internet searches performed by a member of a hiring committee. Only Repler, shown in Figure 1, surveyed the broader group of “individuals involved in the hiring process at their company.”¹²

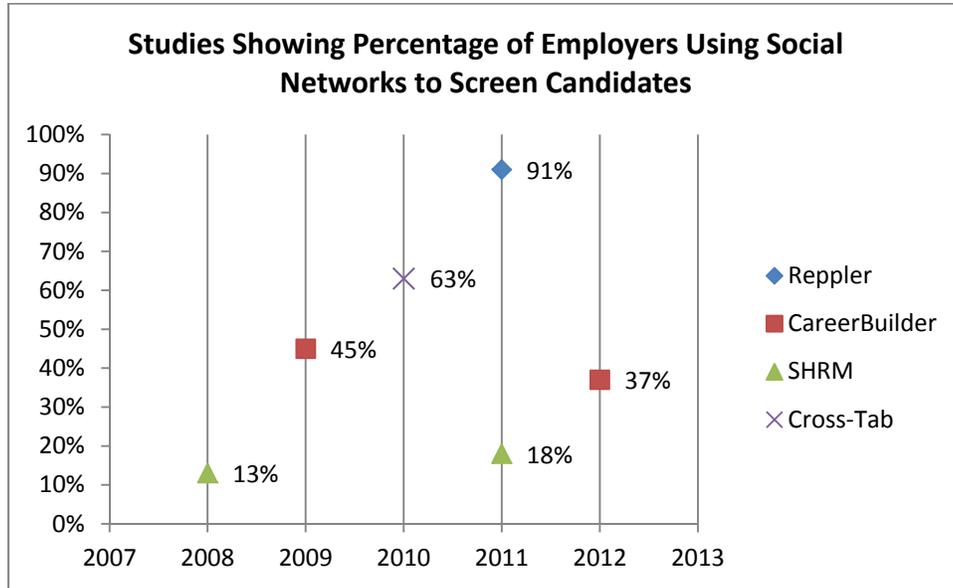
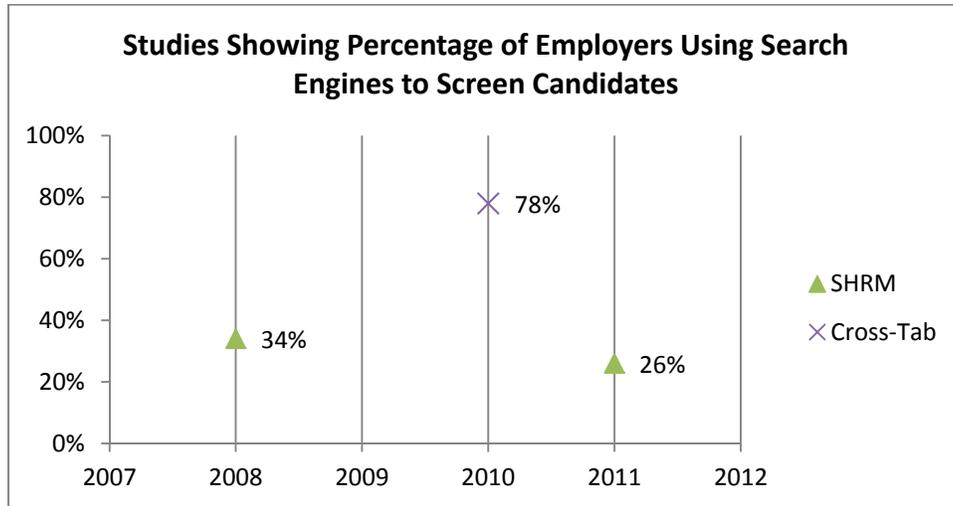
It would be premature to jump to the conclusion that Repler’s statistic—that 91% of employers use social networks to screen candidates¹³—is significantly greater than the four other studies simply because Repler used the right methods to determine the number of employers performing internet background checks. But it is safe to assume that employees, outside of traditional HR functions, are engaging in some internet background checks. When considered together, these five studies suggest that—at a minimum—a fifth to a quarter of employers are using some form of internet background checks in the hiring process.

10. *See, e.g.*, 15 U.S.C. § 1681b (2011) (limiting the permissible uses of consumer reports).

11. *Forty-Five Percent of Employers Use Social Networking Sites to Research Job Candidates, CareerBuilder Survey Finds*, CAREERBUILDER (Aug. 19, 2009), <http://www.careerbuilder.com/share/aboutus/pressreleasesdetail.aspx?id=pr519&sd=8/19/2009&ed=12/31/2009> [hereinafter CAREERBUILDER 2009]; SOCIETY FOR HUMAN RESOURCE MANAGEMENT, *supra* note 1; *Online Reputation in a Connected World*, CROSS-TAB (Jan. 2010), http://download.microsoft.com/download/C/D/2/CD233E13-A600-482F-9C97545BB4AE93B1/DPD_Online%20Reputation%20Research_overview.pdf; *Thirty-seven Percent of Companies Use Social Networks to Research Potential Job Candidates, According to New CareerBuilder Survey*, CAREERBUILDER (Apr. 18, 2012), <http://www.careerbuilder.com/share/aboutus/pressreleasesdetail.aspx?id=pr691&sd=4%2F18%2F2012&ed=4%2F18%2F2009> [hereinafter CAREERBUILDER 2012].

12. REPLER, *supra* note 2.

13. *Id.*

Figure 1¹⁴Figure 2¹⁵

14. *Id.*; CAREERBUILDER 2009, *supra* note 11; SOCIETY FOR HUMAN RESOURCE MANAGEMENT, *supra* note 1, at 2, 6; CROSS-TAB, *supra* note 11, at 8; CAREERBUILDER 2012, *supra* note 11.

15. SOCIETY FOR HUMAN RESOURCE MANAGEMENT, *supra* note 1, at 2; CROSS-TAB, *supra* note 11, at 8.

III. A TAXONOMY OF INTERNET BACKGROUND INFORMATION GATHERING

As a theoretical matter, there are three general paradigms of information gathering in the context of internet background checks, ranging in the degree of separation between the person gathering the information and the person using it in the hiring process.¹⁶ In the first approach—which I will call the Hiring Committee Approach—people involved in making the hiring decision research the candidate themselves using search engines, social networks, and other internet databases. In other words, the information gathering and information using stages are collapsed as an employee on the hiring committee simultaneously researches and evaluates whatever information he can find about the candidate. In the second approach—which I will call the Special Department Approach—an employee with no hiring decision-making power is delegated the task of preparing the internet background report on the prospective employee. Compared to the Hiring Committee Approach, this approach involves actual separation between the information gatherer and the information user. In the third approach—which I will call the Third-Party Approach—a separate consumer reporting agency prepares the internet background check and delivers a formal report to the employer.

Social Intelligence, based in Santa Barbara, CA, is one of the more prominent third-party internet background screening services.¹⁷ Founded by Max Drucker and Geoff Andrews, the company uses “a combination of automated research and manual, multi-tier analyst review” processes to gather information about prospective and current employees.¹⁸ Social Intelligence scours public internet sources, including, according to Andrews, “social networking websites (i.e., Facebook and others), professional networking websites (i.e., LinkedIn and others), blogs, wikis, video and

16. See Lester S. Rosen & Thomas Ahearn, *Managing the Risks of Using the Internet for Employment Screening Background Checks*, EMPLOYMENT SCREENING RESOURCES, <http://www.esrcheck.com/Download/> (last visited Oct. 19, 2011) (identifying independent in-house and third-party background screening as alternatives to screening by the hiring decision-maker).

17. SOCIAL INTELLIGENCE, <http://www.socialintel.com/about/> (last visited Oct. 17, 2011). While, at the time of this Note, Social Intelligence is the best known third-party internet background screening service, other background screening services have announced plans to offer social network and internet background reports. See, e.g., Molly Armbrister Share, *Tandem Select Now Part of Global Firm*, NORTHERN COLORADO BUSINESS REPORT (Sept. 23, 2011), <http://www.ncbr.com/article.asp?id=59980> (announcing that Colorado-based background screening company, Tandem Select, will begin offering “social media background checks”).

18. SOCIAL INTELLIGENCE, <http://www.socialintel.com/social-media-policy-enforcement/> (last visited Oct. 21, 2011).

picture sharing websites, etc.”¹⁹ From these sources, the company then delivers a background report to the requesting employer that highlights objectionable material, such as “racists remarks or behavior, explicit photos and video, and illegal activity,” but filters out any information that would relate to an employee’s status as a member of a protected class under equal employment laws.²⁰

To better understand this service, journalist Mat Honan of Gizmodo requested his own internet background check from Social Intelligence which, as he admits, he “flunked hard,” meaning it turned up truthful but objectionable information about him.²¹ The report included screen shots of his blogs, his public LinkedIn and Facebook profiles, an article he wrote for Wired magazine, and his personal website.²² For each of these internet sources, Social Intelligence scored them either “pass” or “negative,” and included comments such as “subject admits to use of cocaine as well as LSD,” and “subject references use of Ketamine [another recreational drug].”²³ Importantly, Social Intelligence blocked out every part of every image that might have revealed Honan’s ethnicity, including his hands in one picture because they show skin color.²⁴ The report also excluded a line on his personal web page that read “I drink too much beer.”²⁵ Honan points out some of the information that Social Intelligence missed in their report, including a tweet he allegedly meant in jest that reads: “Glad I am childless. Would not want a socialist black man telling my kids to work hard & not do drugs. Related: am so goddamn high right now.”²⁶ Honan also noted that Social Intelligence appeared to generate its search terms from the information on his resume—name, university, email and physical address—

19. Kashmir Hill, *Social Media Background Check Company Ensures That Job-Threatening Facebook Photos Are Part Of Your Application*, FORBES (June 20, 2011), <http://www.forbes.com/sites/kashmirhill/2011/06/20/now-your-embarrassingjob-threatening-facebook-photos-will-haunt-you-for-seven-years/>.

20. SOCIAL INTELLIGENCE, <http://www.socialintel.com/hiring/> (last visited Oct. 17, 2011).

21. Mat Honan, *I Flunked My Social Media Background Check. Will You?*, GIZMODO (July 7, 2011), <http://gizmodo.com/5818774/this-is-a-social-media-background-check/>.

22. *Id.*

23. *Id.*

24. *Id.*

25. *Id.*

26. @mat, TWITTER, <https://twitter.com/mat/status/3843347713/> (last visited Oct. 21, 2011).

which are pieces of information supplied by the job applicant, giving applicants a measure of control over what the service is able to find.²⁷

IV. GATHERING INTERNET BACKGROUND INFORMATION

The method of gathering internet background information about a job candidate determines the applicable federal and state laws that govern the “information gathering process,” which I define as any formal or informal act of assembling background information from an internet source. This determination is true even if all methods of gathering the information result in exactly the same background report.²⁸ For example, as a third-party internet background screening service, Social Intelligence must comply with a number of federal regulations that do not apply to the interviewer who researches prospective employees on her own.²⁹ This Part analyzes the issues that arise from simply collecting—as opposed to using—internet background information.

There is a related discussion, beyond the scope of this Note, pertaining to situations where an employer demands access to the private, password-protected areas of the prospective employee’s social media accounts. While it is unclear how often this situation actually occurs, several state legislatures have nevertheless responded by outlawing these kinds of requests.³⁰ As of the time of this publication, California, Maryland, and Illinois have already enacted legislation prohibiting this practice, and similar legislation has been introduced in nearly a dozen other states and at the federal level.³¹ Facebook

27. For example, the service was not able to find the second Facebook profile for Gizmodo’s editor-in-chief, which he registered under an alias, because he had not supplied Social Intelligence with that information. *See* Honan, *supra* note 21.

28. However, internet background reports generated by different information gathering methods will more often be different, for reasons discussed *infra*.

29. *See infra* Sections IV.C–D.

30. *See* Linda Incoe & Joseph Farrell, *California Limits Employer Access to Employee/Applicant Social Media*, GLOBAL PRIVACY & SECURITY COMPLIANCE LAW BLOG (Oct. 17, 2012), <http://www.globalprivacyblog.com/privacy/california-limits-employer-access-to-employeeapplicant-social-media/>; *see also* Doug Gross, *ACLU: Facebook password isn’t your boss’ business*, CNN (Mar. 22, 2012), <http://www.cnn.com/2012/03/22/tech/social-media/facebook-password-employers/index.html> (reporting two instances in which an employer requested a job applicant’s password).

31. *Employer Access to Social Media Usernames and Passwords*, NATIONAL CONFERENCE OF STATE LEGISLATURES (Nov. 16, 2012), <http://www.ncsl.org/issues-research/telecom/employer-access-to-social-media-passwords.aspx>; Cyrus Farivar, *Congress doesn’t “like” it when employers ask for Facebook login details*, ARS TECHNICA (May 9, 2012), <http://arstechnica.com/tech-policy/2012/05/congress-doesnt-like-it-when-employers-ask-for-facebook-login-details/>.

itself has criticized employers for asking prospective employees to share their passwords, and has suggested that they may even initiate legal action to protect their users' security.³² While this Note will not focus on the application of these new password laws, much of the following analysis of internet background reporting based on publicly-accessible information applies equally to situations where the employer or a third party lawfully obtains accesses to password-protected social network information, either by following an exception to these password laws³³ or by acting in a state without such protections. This is true because once the employer or third party has lawfully obtained access to the password-protected information, the method of internet information gathering—the Hiring Committee, Special Department, or Third-Party Approaches—remains the factor that determines the applicable federal and state fair credit reporting laws.

A. THE FAIR CREDIT REPORTING ACT

Congress enacted the Fair Credit Reporting Act (“FCRA”) in 1970 during a time of great development in “elaborate mechanism[s]” for investigating not only a consumer’s creditworthiness, but also a consumer’s “character” and “general reputation.”³⁴ The Act was Congress’s response to abuses in credit reporting.³⁵ Today, internet background checks are arguably the next development in elaborate mechanisms for investigating consumers. It would have been impossible when the FCRA was passed for anyone to contemplate the future development of internet background checks, since the Internet was not invented until the mid-1980s,³⁶ and the World Wide Web did not go live until Christmas Day 1990.³⁷ Though the broad statutory definition of “consumer report” allows the FCRA to account for this new source of “character” and “general reputation” information, the Act does not account for the Internet’s expansion of access to that information to

32. Erin Egan, *Protecting Your Passwords and Your Privacy*, FACEBOOK (Mar. 23, 2012), https://www.facebook.com/note.php?note_id=326598317390057.

33. For example, California’s enacted legislation, Assembly Bill 1844, contains an exception allowing employers to request an employee ‘to divulge personal social media reasonably believed to be relevant to an investigation of allegations of employee misconduct or employee violation of applicable laws and regulations.’ A.B. 1844, 2012 Cal. State Assemb., 2011–2012 Sess. (Cal. 2012). However, it remains to be seen how this exception, and others like it, would apply to prospective employees.

34. 15 U.S.C. § 1681(a) (2011).

35. *St. Paul Guardian Ins. Co. v. Johnson*, 884 F.2d 881, 883 (5th Cir. 1989).

36. Christopher S. Yoo, *Innovations in the Internet’s Architecture That Challenge the Status Quo*, 8 J. ON TELECOMM. & HIGH TECH. L. 79, 81 (2010).

37. Larry Greenemeier, *Remembering the Day the World Wide Web Was Born*, SCIENTIFIC AMERICAN (Mar. 12, 2009), <http://www.scientificamerican.com/article.cfm?id=day-the-web-was-born>.

practically anyone with a mouse and keyboard.³⁸ Although Congress may have understood that third-party consumer reporting agencies—the legal term of art for regulated entities under the act—would continue to develop technologies to gain access to a consumer’s character information, it may have failed to predict advancements in technology allowing *anyone* exactly the same access to that background information.

B. THRESHOLD DEFINITIONS FOR DETERMINING FCRA
APPLICABILITY: “CONSUMER REPORTING AGENCY” AND
“CONSUMER REPORT”

The majority of the provisions of the FCRA apply to “consumer reporting agencies” that produce “consumer reports.”³⁹ Thus, whether the information gatherer in the Hiring Committee, Special Department, or Third-Party Approaches falls within the definition of a “consumer reporting agency” and produces a “consumer report” is a threshold issue for determining whether the information gatherer must comply with the strict requirements of the FCRA. Though the Act defines these terms individually, both definitions include the other term and are therefore partially circular in their logic.⁴⁰ A “consumer reporting agency,” in other words, is defined as anyone who produces a “consumer report,” while a “consumer report” is defined as the report produced by a “consumer reporting agency” (and satisfying a number of other requirements, discussed *infra*).⁴¹ A consideration of the application of these two definitions to the three paradigmatic methods of internet background information gathering reveals that the FCRA applies to only the Third-Party Approach, and therefore has no bearing on employers that choose the Hiring Committee or Special Department Approaches.

Under the FCRA, a “consumer reporting agency” is defined as:

[1] any person [2] which, for monetary fees, dues, or on a cooperative nonprofit basis, [3] regularly engages in whole or in part in the practice of assembling or evaluating consumer credit information or other information on consumers

38. See 15 U.S.C. § 1681(a); see also *Hodge v. Texaco, Inc.*, 975 F.2d 1093, 1095–96 (5th Cir. 1992) (observing that though the drug-screening reports at issue in the case seemed “far from the original purposes behind the [FCRA] . . . Congress has enacted this statutory language which covers a broad range of conduct by its very terms” and therefore the reports were not categorically excluded from the Act).

39. 15 U.S.C. § 1681a(d), (f).

40. CHI CHI WU ET AL., NATIONAL CONSUMER LAW CENTER, FAIR CREDIT REPORTING 29 (7th ed. 2010) (noting the circularity of the definitions).

41. 15 U.S.C. § 1681(a), (f).

[4] for the purpose of furnishing *consumer reports*

[5] to third parties, and which uses any means or facility of interstate commerce for the purpose of preparing or furnishing consumer reports.⁴²

A “consumer report” is, in turn, defined as:

[4a] any written, oral, or other communication of any information by a consumer reporting agency

[4b] bearing on a consumer’s credit worthiness [creditworthiness], credit standing, credit capacity, character, general reputation, personal characteristics, or mode of living

[4c] which is used or expected to be used or collected in whole or in part for the purpose of serving as a factor in establishing the consumer’s eligibility for . . . [inter alia] employment purposes.⁴³

The numbering scheme suggested above endeavors to clarify the two definitions by highlighting the dependent relationship in which “consumer report” merely expands upon element [4] of the “consumer reporting agency” definition.⁴⁴ This is an attempt to resolve the circularity of these two definitions so that they may be applied to the three paradigmatic approaches to internet background information gathering. One of the consequences of the FCRA’s circularity with respect to these definitions is that *all* elements of *both* definitions must be satisfied in order to find that the FCRA applies to the agency in question. Element [1] requires that the consumer reporting agency is a “person,” but under the Act this can include “any individual, partnership, corporation, trust, estate, cooperative, association, government or governmental subdivision or agency, or other entity.”⁴⁵ Based on element [2], that person must either charge a fee or perform the service on a “cooperative nonprofit basis,” which, according to the FTC, must still serve a “commercial purpose” to fall within the definition.⁴⁶ According to element [3], the person must engage in information collection regularly, and though the FCRA does not specify what sort of regularity is required, some courts have borrowed the definition from similar statutes such as the Fair Debt

42. *Id.* § 1681a(f) (emphasis added).

43. *Id.* § 1681a(d).

44. Because the definitions are circular, the dependency could also be written the other way, where “consumer reporting agency” is a subdefinition of “consumer reports.”

45. 15 U.S.C. § 1681a(b).

46. *Porter v. Talbot Perkins Children’s Servs.*, 355 F. Supp. 174, 177 (S.D.N.Y. 1973) (quoting the 4 CCH Consumer Credit Guide ¶ 11,305); *see also* CHI CHI WU ET AL., *supra* note 40, at 57.

Collection Practices Act or the Truth in Lending Act, which generally define “regular” as not isolated or occasional.⁴⁷

As it applies to the three methods of internet background information gathering, element [5], requiring disclosure “to third parties” in order to qualify a person as a consumer reporting agency, serves as the critical factor in distinguishing the Hiring Committee and Special Department Approaches from the Third-Party Approach under this definition. The FTC’s published interpretations of the FCRA, known as the 1990 “Official Staff Commentary,” provided broad guidance to courts, forming the basis for much of the case law still in force today.⁴⁸ Even though these were merely interpretive rules, the FTC nevertheless repealed them entirely following the omnibus Dodd-Frank Wall Street Reform and Consumer Protection Act, which transferred most of the FTC’s rulemaking authority under the FCRA to the newly created Consumer Financial Protection Bureau.⁴⁹ However, because these interpretations served as the basis for many judicial opinions prior to repeal—and, moreover, because the Dodd-Frank Act preserves the FTC as the agency charged with enforcing the FCRA—the “Official Staff Commentary” remains strong persuasive authority and an indication of how the FTC might enforce the FCRA in future applications to internet background checks.⁵⁰

Interpreting the meaning of element [5] (“to third parties”) in this definition, the FTC suggested that:

[a]n agent or employee that obtains consumer reports does not become a consumer reporting agency by sharing such reports with its principal or employer in connection with the purpose for which the reports were initially obtained.⁵¹

47. CHI CHI WU ET AL., *supra* note 40, at 57–58.

48. Proposed Interpretations of the Fair Credit Reporting Act, Fed. Trade Comm’n, 16 C.F.R. § 600 (2000).

49. Dodd-Frank Wall Street Reform and Consumer Protection Act, Pub. L. No. 111-203, § 1088, 124 Stat. 1376 (2010); *see also* Statement of General Policy or Interpretation; Commentary on the Fair Credit Reporting Act, 16 C.F.R. § 600 (2011).

50. *See, e.g., 40 Years of Experience with the Fair Credit Reporting Act*, FED. TRADE COMM’N 7 (July 2011), <http://www.ftc.gov/os/2011/07/110720fcrareport.pdf> (noting that the “Official Staff Commentary” had “no binding effect, but some courts . . . accorded it weight as a policy statement of the primary agency responsible for enforcing the FCRA,” and further expressing the FTC staff’s belief that, even after their repeal, the “[i]nterpretations from the [Official Staff Commentary] on sections of the FCRA that have not been amended . . . are timely, accurate, and helpful.”).

51. Proposed Interpretations of the Fair Credit Reporting Act, Fed. Trade Comm’n, 16 C.F.R. §§ 600, 604(f), item 8 (2000), *repealed by* Statement of General Policy or Interpretation; Commentary on the Fair Credit Reporting Act, 16 C.F.R. § 600 (2011).

This interpretation of element [5] speaks directly to the distinction between the Third-Party Approach and the employer-based (Hiring Committee and Special Department) approaches to information gathering, indicating that the latter two fall outside of the definition of third party under the FCRA.

In *Menefee v. City of Country Club Hills*, the District Court for the Northern District of Illinois confirmed this interpretation of “third party” in the context of consumer reporting agencies.⁵² There, the City of Country Club Hills collected the plaintiff’s criminal and credit history in response to his application for employment to the City.⁵³ This information was later subpoenaed by a party in an unrelated lawsuit, and the City of Country Club Hills produced the credit and criminal background information.⁵⁴ Menefee sued the City of Country Club Hills, alleging, *inter alia*, that the City falls under the FCRA’s definition of “consumer reporting agency,” and therefore failed to comply with certain requirements under the Act.⁵⁵ The court held that the City was not a consumer reporting agency because originally the background information had not been collected “for the purpose of furnishing consumer reports to third parties.”⁵⁶ In other words, the fact that the City itself collected the credit and criminal background information in conjunction with their evaluation of Menefee did not make the City a consumer reporting agency because they did not collect the information for the purpose of eventually turning the information over in the subpoena. Moreover, the court held that this single instance of providing the plaintiff’s information to a third party did not mean that the City was “regularly engaged in assembling credit information” (element [3]), and thus did not make the City a consumer reporting agency.⁵⁷

In contrast, when the information gatherer is separate from the entity using the background information to make a hiring decision, the outside information gatherer is considered a “third party” for purposes of the FCRA. For example, the district court for the District of Connecticut held in *Adams v. National Engineering Service Corp.* that a staffing agency performing a prospective applicant background check for an employer had provided that report “to third parties” within the meaning of the Act.⁵⁸ In that case, the

52. *Menefee v. City of Country Club Hills*, No. 08 C 2948, 2008 WL 4696146, at *3 (N.D. Ill. Oct. 23, 2008).

53. *Id.* at *1.

54. *Id.*

55. *Id.* at *3.

56. *Id.*

57. *Id.*

58. *Adams v. Nat’l Eng’g Serv. Corp.*, 620 F. Supp. 2d 319, 328 (D. Conn. 2009).

employer, Northeast Utilities, contracted with a staffing agency to aid in filling temporary positions.⁵⁹ The staffing agency, in turn, used another agency to perform the background investigation.⁶⁰ The background agency forwarded the report erroneously describing the plaintiff's criminal history to the staffing agency, which in turn forwarded the report to Northeast Utilities.⁶¹ The court held that the staffing agency—and, by implication, the background screening agency—had furnished their background information “to third parties” within the meaning of the FCRA.⁶²

Under the Hiring Committee and Special Department Approaches to gathering information about a prospective employee, the information gatherer does not furnish a consumer report to third parties as that term is used in the FCRA. Much like the city in *Menefee* that assembled its own consumer report of credit and criminal background information about the plaintiff, in the Hiring Committee Approach the information gatherer scours the Internet for information about the prospective employee for her own use as a member of the hiring committee.⁶³ Similarly, according to the FTC's interpretation of this element, and under general principles of principal-agent theory, the information gatherer in the Special Department Approach is not furnishing internet background information to a third party under the FCRA when he delivers his report to those responsible for making the hiring decision.⁶⁴ Because they do not furnish internet background consumer reports to third parties, the information gatherers in the Hiring Committee and Special Department Approaches are not “consumer reporting agencies,” do not produce “consumer reports,” and are therefore not subject to the requirements of the FCRA.⁶⁵

In contrast, the information gatherer in the Third-Party Approach will almost always satisfy element [3] of the FCRA definition. Much like the staffing and background investigation agencies in *Adams*, the third-party internet background checking agency is a separate entity that contracts with the hiring employer and forms an independent contractor relationship that

59. *Id.* at 324.

60. *Id.*

61. *Id.*

62. *Id.* at 328.

63. *See Menefee v. City of Country Club Hills*, No. 08 C 2948, 2008 WL 4696146, at *1 (N.D. Ill. Oct. 23, 2008).

64. *See Proposed Interpretations of the Fair Credit Reporting Act*, Fed. Trade Comm'n, 16 C.F.R. § 604(f), item 8 (2000), *repealed by* Statement of General Policy or Interpretation; Commentary on the Fair Credit Reporting Act, 16 C.F.R. § 600 (2011).

65. *See* 15 U.S.C. § 1681a(d), (f) (2011).

does not rise to the level of employer-employee.⁶⁶ As a result, the Third-Party Approach satisfies this element of the definition of “consumer reporting agency.”

The analysis of whether the information gatherer in the Third-Party Approach is subject to the FCRA does not end there, however, since the information gatherer must also prepare “consumer reports,” which is element [4] of the “consumer reporting agency” definition and is itself a separate entry in the definitions section of the FCRA (elements [4a–c]).⁶⁷ As the Fourth Circuit has noted, element [4a] captures “virtually any information communicated by a ‘consumer reporting agency.’”⁶⁸ This element, therefore, does very little to limit the definition of “consumer report.” Element [4b] has similarly broad application because the seven factors upon which the information may “bear” are listed disjunctively. As the D.C. Circuit has pointed out, “almost any information about consumers arguably bears on their personal characteristics or mode of living.”⁶⁹ Thus, this element excludes very few sources of information from the definition of “consumer report.”

When the consumer report is used in the employment context, element [4c] must be read in conjunction with the separate definition of “employment purposes,” which the Act defines as using a consumer report for “the purpose of evaluating a consumer for employment, promotion, reassignment or retention as an employee.”⁷⁰ At least one circuit has read the term “employee” to encompass more employment situations than the common law definition, including, for example, sole practitioners and independent contractors.⁷¹ This element of the definition turns on how the consumer report is used or how the consumer reporting agency expects it to be used. Most circuits have settled upon a common rule for this element of the definition, articulated by the Ninth Circuit in the following way: “If a consumer reporting agency provides a report based on a *reasonable expectation* that the report will be put to a use permissible under the FCRA, then that report is a ‘consumer report’ under the FCRA and the ultimate use to which

66. See *Adams v. Nat’l Eng’g Serv. Corp.*, 620 F. Supp. 2d 319, 324 (D. Conn. 2009).

67. 15 U.S.C. § 1681a(d), (f) (2011).

68. *Hoke v. Retail Credit Corp.*, 521 F.2d 1079, 1081 (4th Cir. 1975).

69. *Trans Union Corp. v. FTC*, 245 F.3d 809, 813 (D.C. Cir. 2001) (citing *Trans Union Corp. v. FTC*, 81 F.3d 228 (D.C. Cir. 1996)); see also *Hoke*, 521 F.2d 1079, 1081 (noting that this element encompasses “virtually any information communicated by a ‘consumer reporting agency’ for any one of the purposes enumerated in 1681a and 1681b, including but not limited to, ‘employment purposes’”).

70. 15 U.S.C. §1681a(h) (2011).

71. *Hoke*, 521 F.2d at 1082 n.7.

the report is actually put is irrelevant”⁷² The Act also excludes two sources of information that would otherwise fit within the above definition—so-called “experience information” and information shared among affiliates—which are not particularly relevant to the Third-Party Approach.⁷³

Following an investigation of Social Intelligence, the FTC concluded that the company falls within the definition of a “consumer reporting agency” producing “consumer reports,” and therefore must comply with the FCRA.⁷⁴ In the letter announcing the close of its investigation with “no further action . . . warranted at this time” to ensure compliance with the Act, the Commission briefly reasoned that “because [Social Intelligence] assembles or evaluates consumer report information that is furnished to third parties that use such information as a factor in establishing a consumer’s eligibility for employment,” the company is therefore a “consumer reporting agency” within the meaning of the Act.⁷⁵ The FTC’s determination that Social Intelligence falls within these FCRA definitions implies that the company satisfies all of the elements of these two definitions. As implied by the FTC’s conclusion, the information collected from the Internet and social media sources concerning a prospective employee is thus “written, oral, or other communication” (element [4a]) that has “bearing on” one of the seven factors of element [4b].⁷⁶ Furthermore, Social Intelligence has a reasonable

72. *Comeaux v. Brown & Williamson Tobacco Co.*, 915 F.2d 1264, 1274 (9th Cir. 1990) (citing *Hansen v. Morgan*, 582 F.2d 1214, 1218 (9th Cir. 1978); *Ippolito v. WNS, Inc.*, 864 F.2d 440, 449 (7th Cir. 1988), *cert. dismissed*, 490 U.S. 1061 (1989); *Heath v. Credit Bureau of Sheridan, Inc.* 618 F.2d 693, 696 (10th Cir. 1980)) (emphasis added).

73. 15 U.S.C. §1681a(d)(2) (2011); *CHI CHI WU ET AL.*, *supra* note 40, at 49–53 (citing *Am. Bankers Ass’n v. Gould*, 412 F.3d 1081 (9th Cir. 2005) (calling information falling under the § 1681a(d)(2)(A)(i) exception “experience information”). The experience-information provision exempts “any report based on the reporter’s first-hand experience of the subject.” For example, this exemption applies to a retailer’s or bank’s disclosure of its own experience with a customer. *Hodge v. Texaco, Inc.*, 975 F.2d 1093, 1096 (5th Cir. 1992).

74. *See* Mithal, *supra* note 6. Under the FCRA, a consumer reporting agency may also produce “investigative consumer reports,” to which attach even greater statutory requirements. However, “investigative consumer reports” are obtained “through personal interviews with neighbors, friends, or associates of the consumer,” which does not describe the process by which a third-party information gatherer produces an internet background report. *See* 15 U.S.C. § 1681a(e). Therefore, it is unlikely that the FCRA’s investigative consumer reporting requirements apply to Social Intelligence or other third-party internet background reporting agencies.

75. *Id.*

76. *See* 15 U.S.C. § 1681a(d).

expectation that these reports sold to employers will be used for one of the permissible purposes under the Act.⁷⁷

As the FTC letter indicates, a finding that Social Intelligence is a consumer reporting agency triggers a host of regulatory obligations under the FCRA.⁷⁸ The Act sets forth a list of “compliance procedures” for consumer reporting agencies, requiring, *inter alia*, that agencies make reasonable efforts to verify the identity and purposes of users of the consumer report,⁷⁹ ensure the accuracy of the consumer report,⁸⁰ and allow users of the consumer report to disclose the report to consumers in the case of an adverse action against them.⁸¹ Among these many obligations imposed upon the information gatherer in the Third-Party Approach, the issue of accuracy appears most challenging in light of the unique nature of the Internet.

C. THIRD-PARTY INFORMATION GATHERERS MUST ENSURE ACCURACY

Senators Blumenthal and Franken identified a number of accuracy-related challenges through the questions they asked in their letter to Social Intelligence, most of which are particularly relevant to the Internet as a source of background information.⁸² While any information gatherer using the Internet to construct a background report—whether in the Hiring Committee, Special Department, or Third-Party Approaches—would face these challenges, the FCRA only requires the information gatherer in the Third-Party Approach to meet an accuracy standard.⁸³ For these third-party internet background checking agencies, for example, how can the information gatherer differentiate among “Googlegangers”?⁸⁴ How is the information gatherer able to judge whether the internet source is credible or whether the source is “parody, defamatory, or otherwise false”?⁸⁵ How does the information gatherer deal with internet sources that are out-of-date but

77. *See id.*; *see also Comeaux*, 915 F.2d at 1273–74 (citing *Hansen v. Morgan*, 582 F.2d 1214, 1218 (9th Cir. 1978)); *Ippolito*, 864 F.2d at 449–50, *cert. dismissed*, 490 U.S. 1061 (1989); *Credit Bureau of Sheridan*, 618 F.2d at 696.

78. Mithal, *supra* note 6.

79. 15 U.S.C. § 1681e(a).

80. *Id.* § 1681e(b).

81. *Id.* § 1681e(c).

82. *See* Blumenthal, *supra* note 5.

83. *See supra* Section IV.B.

84. According to the Urban Dictionary, “Googleganger” describes other “individual[s] with the same name as you whose records and/or stories are mixed in with your own when you Google yourself.” *Googleganger*, URBAN DICTIONARY (Oct. 6, 2007), <http://www.urbandictionary.com/define.php?term=googleganger>.

85. *See* Blumenthal, *supra* note 5.

still accessible through a search engine?⁸⁶ As applied to internet background checks, it is unclear exactly how to implement the accuracy standard required by the FCRA.

The FCRA requires that “[w]henver a consumer reporting agency prepares a consumer report it shall follow *reasonable* procedures to assure *maximum possible accuracy* of the information concerning the individual about whom the report relates.”⁸⁷ Accuracy is an objective standard, and consumer reports are either accurate or inaccurate.⁸⁸ “Reasonable procedures” that ensure the required standard of accuracy are those that “a reasonably prudent person would [undertake] under the circumstances.”⁸⁹ Moreover, “[j]udging the reasonableness of an . . . agency’s procedures involves weighing the potential harm from inaccuracy against the burden of safeguarding against such inaccuracy.”⁹⁰ Whether a procedure is reasonable is generally question for the jury.⁹¹

In evaluating whether a consumer reporting agency followed reasonable procedures to ensure maximum accuracy, two federal circuit courts have arrived at opposite conclusions based on roughly similar consumer reporting agency procedures for ensuring accuracy. In *Thompson v. San Antonio Retail Merchants Association*, the Fifth Circuit upheld the lower court’s finding that the Association’s computer database requiring no “minimum points of correspondence” between a given consumer’s file and a furnisher of new credit information was not a reasonable procedure for ensuring accuracy.⁹² In this case, the plaintiff, William Douglas Thompson, III, was denied credit because his file had been mixed with that of William *Daniel* Thompson, Jr., who had a delinquent account. As a result of the mismerger, the plaintiff’s file had become a “potpourri of information” on the two men.⁹³ Using the Association’s computer system, when a furnisher of credit information input some sort of consumer identifier, the system would return a number of matches and near-matches.⁹⁴ The furnisher then had complete discretion

86. *See id.*

87. 15 U.S.C. § 1681e(B) (2011) (emphasis added).

88. *See* CHI CHI WU ET AL., *supra* note 40, at 109 (citing *Cushman v. Trans Union Corp.*, 115 F.3d 220, 225 (3d Cir. 1995); *Cahlin v. General Motors Acceptance Corp.*, 936 F.2d 1151, 1158 (11th Cir. 1991)).

89. *Philbin v. Trans Union Corp.*, 101 F.3d 957, 963 (3d Cir. 1996) (quoting *Stewart v. Credit Bureau, Inc.*, 734 F.2d 47, 51 (D.C. Cir. 1984) (per curiam)).

90. *Id.*

91. *Andrews v. TRW, Inc.*, 225 F.3d 1063, 1068 (9th Cir. 2000).

92. *Thompson v. San Antonio Retail Merch. Ass’n*, 682 F.2d 509, 513 (5th Cir. 1982).

93. *Id.* at 511.

94. *Id.*

over which consumer's file to update.⁹⁵ By not requiring a minimum number of matching identifiers, the court upheld the lower court's finding that the Association was negligent under the FCRA's standard.⁹⁶ Furthermore, the court held that the Association's "spot audits" to verify social security numbers did not cure the situation.⁹⁷

Nearly twenty years after *Thompson*, Judge Posner of the Seventh Circuit held in *Crabill v. Trans Union* that a computer program that continued to mix up the files of two brothers with similar names nevertheless satisfied the "reasonable procedures" requirement of the FCRA.⁹⁸ In this case, Jerry Crabill alleged that he was harmed when Trans Union mixed his credit information with that of his brother John, who has the same first initial and whose social security number differs by only one digit.⁹⁹ Trans Union defended its procedures on the basis that often "two files with similar though not identical identifying data may actually be referring to the same person."¹⁰⁰ Creditors, Trans Union continued, find it useful to have both files, allowing them to use their own judgment to determine whether the information refers to two people or one.¹⁰¹ The court agreed, holding that "the statutory duty to maintain reasonable procedures to avoid inaccuracy does not require a credit agency to disregard the possibility that similar files refer to the same person."¹⁰²

In the context of internet background reports, mixed files in which real applicant data is mixed with Googleganger data has a great potential to cause inaccuracies in a consumer's internet background report. It is unclear, however, exactly what the standard of "reasonable procedures to assure maximum possible accuracy" will mean as applied to a third-party information gatherer preparing an internet background report. Following the approach of the Fifth Circuit in *Thompson*, which stressed that procedures must require a minimum number of correspondence points beyond the "identifier" (i.e., name), a court would likely require that the internet source matches the consumer file based on more than just the consumer's name.¹⁰³ For social media sources, the consumer's birthday, educational background,

95. *Id.*

96. *Id.* at 513.

97. *Id.*

98. *Crabill v. Trans Union, L.L.C.*, 259 F.3d 662, 663 (7th Cir. 2001).

99. *Id.*

100. *Id.*

101. *Id.*

102. *Id.*

103. *See Thompson v. San Antonio Retail Merch. Ass'n*, 682 F.2d 509, 513 (5th Cir. 1982).

physical or email address might serve as a second factor in identifying the match, provided that the consumer has shared those pieces of information to the public. Notably, however, the consumer's profile picture would, in many cases, be irrelevant to the *Thompson* test. When the employer hires the third-party agency to perform the internet background check, it is unlikely that the employer would have the prospective employee's picture, as the employer's request for pictures from job applicants would risk the appearance of discrimination that the employer seeks to avoid in the first place by using a third-party agency.

Under the Seventh Circuit's more relaxed standard in *Crabill*, however, the information gatherer in the Third-Party Approach would have little difficulty satisfying the "reasonable procedures to assure maximum possible accuracy" standard of the FCRA. Like the credit report in *Crabill* that contained a hodgepodge of credit information about two consumers, an internet background report that contains a mix of true applicant information with Googleganger internet hits would require the employer to use its own judgment to determine whether the report contains false data.¹⁰⁴ According to the court's holding, it would be reasonable for the third-party information gatherer to include information sources with similar—but not identical—identifying information, because the accuracy standard does not require the agency "to disregard the possibility that similar files refer to the same person."¹⁰⁵ Under the Seventh Circuit's approach, therefore, the Third-Party internet background reporting agency would have little difficulty demonstrating that they had followed reasonable procedures.

D. HEIGHTENED FCRA STANDARD FOR "MATTERS OF PUBLIC RECORD"

On top of the normal requirements the FCRA imposes on the third-party information gatherer, the Act includes additional requirements for consumer reports that contain "public record information [used] for employment purposes."¹⁰⁶ Section 1681k of the FCRA provides that:

A consumer reporting agency which furnishes a consumer report for employment purposes and which for that purpose compiles and reports items of information on consumers which are *matters of public record* and are likely to have an adverse effect upon a consumer's ability to obtain employment shall—

104. See *Crabill v. Trans Union, L.L.C.*, 259 F.3d 662, 663 (7th Cir. 2001).

105. *Id.*

106. 15 U.S.C. §1681k (2011).

(1) at the time such public record information is reported to the user of such consumer report, notify the consumer of the fact that public record information is being reported by the consumer reporting agency, together with the name and address of the person to whom such information is being reported; or

(2) maintain strict procedures designed to insure that whenever public record information which is likely to have an adverse effect on a consumer's ability to obtain employment is reported it is complete and up to date. For purposes of this paragraph, items of public record relating to arrests, indictments, convictions, suits, tax liens, and outstanding judgments shall be considered up to date if the current public record status of the item at the time of the report is reported.¹⁰⁷

The critical issue with respect to the application of this section of the FCRA is how to define "matters public record," which the Act does not define directly. However, subsection two of this provision enumerates a nonexhaustive list of information sources that, by implication, qualify as matters of public record, including records relating to "arrests, indictments, convictions, suits, tax liens, and outstanding judgments."¹⁰⁸ The Internet provides free and public access to records falling into most of these categories. For example, websites such as GOOGLE SCHOLAR and JUSTIA (not to mention courts' own websites) provide free access to state and federal legal opinions and dockets that contain information relating to nearly all of the enumerated categories of public records.¹⁰⁹ A court might also expand the meaning of "matters of public record" beyond the types of information included in subsection two of this provision to include other internet sources providing access to government records.¹¹⁰

In the event that the internet background report includes any information relating to matters of public record, the third-party information gatherer must comply with one of the two heightened requirements described in subsections one and two of this provision if the information is "likely to have an adverse effect on the consumer's ability to obtain

107. *Id.* § 1681k(a) (emphasis added).

108. *Id.* § 1681k(a)(2); *see also* CHI CHI WU ET AL., *supra* note 40, at 153.

109. *See* Google Scholar, GOOGLE, <http://scholar.google.com/> (last visited Oct. 28, 2011); *see also* JUSTIA, <http://www.justia.com/> (last visited Oct. 28, 2011).

110. Based on the plain meaning of the term "public record," however, a court would likely restrict any additional information sources to additional *government* sources. *See* BLACK'S LAW DICTIONARY 1387 (9th ed. 2009) (defining "public record" as "[a] record that a governmental unit is required by law to keep, such as land deeds kept at a county courthouse. Public records are generally open to view by the public.>").

employment.”¹¹¹ These subsections require that the information gatherer either notify the job applicant of the fact that this public record information is being reported, or maintain “strict procedures” to ensure that the information is “complete and up to date.”¹¹² Exactly what constitute “strict procedures” remains unclear, although courts have noted that the difference between “strict” and “reasonable procedures” (as are required for consumer reports generally) is “clearly not without significance.”¹¹³

The FCRA, therefore, imposes a number of requirements on third-party information gatherers that do not apply to the Hiring Committee and Special Department Approaches. Although the Act’s precise definitions of the terms “reasonable procedures to assure maximum possible accuracy” (required for all consumer reports) and “strict procedures” (required for matters of public record) remain uncertain, an agency would have to point to *some* procedure that satisfies these standards, in contrast to the employer-based approaches.¹¹⁴

E. STATE FAIR CREDIT REPORTING LAWS—CALIFORNIA’S APPROACH

While the FCRA expressly preempts many state law causes of action, including defamation, invasion of privacy, and negligence claims arising out of consumer reports, some state fair credit reporting laws impose additional requirements on consumer reporting beyond the reach of the preemption clause.¹¹⁵ While a survey of state fair credit reporting laws is beyond the scope of this Note, California’s Investigative Consumer Reporting Agencies Act illustrates how state law can expand FCRA-like requirements to information gatherers in the Hiring Committee and Special Department Approaches.

California Civil Code section 1786.53(a) adds procedural and notice requirements to information gatherers in the Hiring Committee and Special Department Approaches. The law requires:

Any person who collects, assembles, evaluates, compiles, reports, transmits, transfers, or communicates information on a consumer’s character, general reputation, personnel characteristics, mode of

111. 15 U.S.C. §1681k(a). The FCRA also does not define “adverse effect,” but does define “adverse action” to mean, in the employment context, “a denial of employment or any other decision for employment purposes that adversely affects any current or prospective employee.” *Id.* § 1681a(k)(1)(B)(ii).

112. *Id.* § 1681k(a).

113. *See Dalton v. Capital Associated Indus., Inc.*, 254 F.3d 409, 417 (4th Cir. 2001) (quoting *Equifax v. FTC*, 678 F.2d 1047, 1049 n.4 (11th Cir. 1982)).

114. *See supra* Section IV.B.

115. 15 U.S.C. § 1681h(e).

living, for employment purposes, *which are matters of public record*, and does not use the services of an investigative consumer reporting agency, shall provide that information to the consumer pursuant to subdivision (b).¹¹⁶

Subdivision (b) requires the information gatherer to provide a copy of the public record to the applicant within seven days, provide a checkbox on the job application to opt-out of receiving the public record, allows the information gatherer to temporarily withhold the information if the consumer is under investigation, and requires the employer to provide a copy of the public record if any adverse action is taken, regardless of whether the consumer opted out.¹¹⁷ Notably, section 1786.53(a) only applies to information collection from “public record[s],” which, here, the statute defines as “records documenting an arrest, indictment, conviction, civil judicial action, tax lien, or outstanding judgment.”¹¹⁸ Contrast this with the FCRA’s definition of “consumer report,” which includes essentially any form of information bearing on one of seven broad factors, including, for example, the consumer’s character or general reputation.¹¹⁹

In *Moran v. Murtaugh, Miller, Meyer & Nelson, LLP*, a California court of appeal, confronted with a case involving the termination of an employee upon the discovery of his criminal background, held that under California Civil Code section 1786.53 the employer must provide the employee with a copy of the public record upon which they based their termination decision.¹²⁰ Although this case involved the use of internet background reports in an employee termination, rather than in a hiring, section 1786.53 applies equally when the consumer report is used for many “employment purposes,” including “employment, promotion, reassignment, or retention.”¹²¹ The employer, a law firm, terminated the plaintiff, a paralegal, after an associate anonymously placed printouts of judicial opinions in which the plaintiff was convicted of several felonies, including grand theft and second-degree burglary, on the chairs of two of the firm’s partners.¹²² According to the taxonomy of information gatherers described in Part II, *supra*, the associate in this case, who used a “computerized legal database” to research the plaintiff, is closely analogous to the Special Department

116. CAL. CIV. CODE § 1786.53(a) (West 2011) (emphasis added).

117. *Id.* § 1786.53(b).

118. *Id.* § 1786.53(a)(3).

119. 15 U.S.C. § 1681a(d).

120. *Moran v. Murtaugh, Miller, Meyer & Nelson, LLP*, 24 Cal. Rptr. 3d 275, 284 (Cal. Ct. App. 2005).

121. CAL. CIV. CODE § 1786.2(f) (West 2011).

122. *Moran*, 24 Cal. Rptr. 3d at 277.

Approach in the hiring process, since here the associate, an employee of the same firm, was not likely the one responsible for making the decision to terminate the plaintiff.¹²³ However, nothing in section 1786.53 suggests that it would not apply equally to the information gatherer in the Hiring Committee Approach. *Moran*, therefore, illustrates how California's section 1786.53 can extend the reach of fair credit reporting requirements to non-third-party information gatherers, and thus beyond the reach of the FCRA.

V. USING INTERNET BACKGROUND INFORMATION

Once an employer has acquired the job applicant's internet background information through one of the three paradigmatic approaches to information gathering, a separate set of federal and state laws governs how an employer may use that information to make an employment decision. Equal employment laws largely focus on what information may or may not serve as a basis for making a hiring decision.¹²⁴ With the notable exception of the Americans with Disabilities Act ("ADA"), these laws generally do not interact with the information gathering process. In other words, laws such as Title VII of the Civil Rights Act of 1964 do not mandate one information gathering approach over the other. This Part focuses on the application of several illustrative equal employment laws—Title VII, the ADA, and several state laws—to internet background checks, and demonstrates how the Third-Party Approach has certain *structural* advantages over the employer-based approaches, even when the law imposes no requirements on how the information is acquired.

A. TITLE VII PROTECTIONS AGAINST RELIGIOUS DISCRIMINATION IN THE HIRING PROCESS

In the most highly publicized lawsuit to date alleging employment discrimination based on the employer's internet background check, C. Martin Gaskell, an astronomy professor, alleged that the University of Kentucky violated Title VII of the Civil Rights Act of 1964 when it decided not to hire him because of articles and lecture notes he posted on his personal website espousing "creationist" views.¹²⁵ Title VII prohibits employers from refusing to hire an applicant "because of" a number of protected categories, including

123. *See id.*

124. *See, e.g.*, 42 U.S.C. § 2000e-2(a) (2011).

125. *See Gaskell v. Univ. of Kentucky*, No. 09-244-KSF, 2010 U.S. Dist. LEXIS 124572 (E.D. Ky. Nov. 23, 2010); *see also* Mark Oppenheimer, *Astronomer Sues the University of Kentucky, Claiming His Faith Cost Him a Job*, N.Y. TIMES, Dec. 18, 2010, <http://www.nytimes.com/2010/12/19/us/19kentucky.html?pagewanted=print>.

“[an] individual’s . . . religion.”¹²⁶ Born out of John F. Kennedy’s speech to the nation on June 11, 1963, the Civil Rights Act included a number of broad prohibitions on discrimination.¹²⁷ Under Title VII of the Act, it is unlawful for an employer:

(1) to fail or refuse to hire or to discharge any individual, or otherwise to discriminate against any individual with respect to his compensation, terms, conditions, or privileges of employment, because of such individual’s race, color, religion, sex, or national origin.

(2) to limit, segregate, or classify his employees or applicants for employment in any way which would deprive or tend to deprive any individual of employment opportunities or otherwise adversely affect his status as an employee, because of such individual’s race, color, religion, sex, or national origin.¹²⁸

The Act defines religion as “all aspects of religious observance and practice, as well as belief, unless an employer demonstrates that he is unable to reasonably accommodate to an employee’s or prospective employee’s religious observance or practice without undue hardship on the conduct of the employer’s business.”¹²⁹ Under this section, a plaintiff may assert a claim for discrimination in the hiring process by showing either direct or indirect evidence.¹³⁰ The Sixth Circuit described this distinction in the following way:

[D]irect evidence is that evidence which, if believed, requires the conclusion that unlawful discrimination was at least a motivating factor in the employer’s actions Such evidence does not require a factfinder to draw any inferences in order to conclude that the challenged employment action was motivated at least in part by prejudice against members of the protected group.¹³¹

Beyond obvious examples of direct discrimination, which include remarks such as “I won’t hire you because you’re a woman,” courts have also characterized evidence that “reflect[s] a propensity by the decision-maker to

126. 42 U.S.C. § 2000e-2(a).

127. John F. Kennedy, President of the United States, “Radio and Television Report to the American People on Civil Rights” (June 11, 1963), *available at* <http://www.jfklibrary.org/Research/Ready-Reference/JFK-Speeches/Radio-and-Television-Report-to-the-American-People-on-Civil-Rights-June-11-1963.aspx>.

128. 42 U.S.C. § 2000e-2(a).

129. *Id.* § 2000e(j).

130. *Tepper v. Potter*, 505 F.3d 508, 515 (6th Cir. 2007).

131. *Id.* at 516 (internal quotations omitted).

evaluate employees based on illegal criteria” as direct evidence for purposes of a Title VII claim.¹³²

Absent direct evidence, a plaintiff can proceed by showing that indirect evidence satisfies the framework established by the Supreme Court in *McDonnell Douglas Corp. v. Green*.¹³³ To make a prima facie case under *McDonnell Douglas*, the plaintiff must first establish that the challenged employment action was either intentionally discriminatory or had a discriminatory effect.¹³⁴ A plaintiff can establish a prima facie case if he can show that “(1) he is a member of a protected class; (2) he was qualified for his position; (3) he experienced an adverse employment action; and (4) similarly situated individuals outside his protected class were treated more favorably, or other circumstances surrounding the adverse employment action give rise to an inference of discrimination.”¹³⁵ If the plaintiff can establish a prima facie case, the burden then shifts to the employer to offer a legitimate and nondiscriminatory reason for employment action.¹³⁶ The plaintiff may then rebut the employer’s justification with evidence that it is a pretext for the sort of discrimination prohibited by Title VII.¹³⁷

The *Gaskell* case illustrates the risk of Title VII liability in the context of the Hiring Committee Approach to internet background checks.¹³⁸ The University of Kentucky (“UK”) sought a founding director to oversee a new astronomical observatory.¹³⁹ Of the twelve people who applied for the position, there was no dispute that Gaskell was the most qualified; as one member of UK’s search committee put it, “Martin Gaskell is clearly the most experienced”; and another remarked that Gaskell “has already done everything we could possibly want the observatory director to do.”¹⁴⁰

During the hiring process, however, one of the committee members conducted an internet search and found Gaskell’s personal website, where the plaintiff had posted an article entitled “Modern Astronomy, the Bible, and Creation.”¹⁴¹ Members of the search committee became concerned that

132. *Gaskell v. Univ. of Kentucky*, No. 09-244-KSF, 2010 U.S. Dist. LEXIS 124572, at *7 (E.D. Ky. Nov. 23, 2010) (citing *Shager v. Upjohn Co.*, 913 F.2d 398, 402 (7th Cir. 1990); *Robinson v. PPG Indus., Inc.*, 23 F.3d 1159, 1164–65 & nn.2–3 (7th Cir. 1993)).

133. *McDonnell Douglas Corp. v. Green*, 411 U.S. 792 (1973).

134. *Id.* at 802.

135. *Peterson v. Hewlett-Packard Co.*, 358 F.3d 599, 603 (9th Cir. 2004).

136. *McDonnell Douglas Corp.*, 411 U.S. at 802.

137. *Id.* at 804.

138. *Gaskell v. Univ. of Kentucky*, No. 09-244-KSF, 2010 U.S. Dist. LEXIS 124572, at *1–5 (E.D. Ky. Nov. 23, 2010).

139. *Id.* at *1.

140. *Id.* at *3.

141. *Id.* at *4.

statements in his article “blended religious thought with scientific theory,” and, moreover, that Gaskell’s current university web profile linked directly to his personal web page, which might indicate his intent to use this position to promote his own religious beliefs.¹⁴² The search committee later obtained Gaskell’s lecture notes from a talk he had delivered at UK, which members of the UK’s biology department said contained “creationist” views.¹⁴³

During the review process, one search committee member, in an email entitled “The Gaskell Affair,” expressed his concern that Gaskell would be denied the job “because of his religious beliefs” and that “no objective observer could possibly believe” that the decision was made on any other grounds.¹⁴⁴ After several committee members offered other reasons why Gaskell should not be hired, the search committee ultimately recommended Timothy Knauer, a former student and employee of UK’s Department of Physics & Astronomy.¹⁴⁵ In response, another UK professor, not a member the search committee but involved in the interview process, filed a complaint with the UK Equal Employment Office.¹⁴⁶ Gaskell filed a complaint with the EEOC, which issued a Notice of Right-to-Sue letter, and then filed the complaint with the Kentucky district court alleging that UK based its decision not to hire him on his religion and religious beliefs in violation of Title VII.¹⁴⁷

Deciding cross motions for summary judgment, the court denied both motions, holding first, with respect to UK’s motion, that Gaskell’s allegations of direct evidence of discrimination raised “a triable issue of fact as to whether his religious beliefs were a substantial motivating factor in UK’s decision not to hire him”; and second, with respect to Gaskell’s motion, that UK had produced “more than a scintilla of evidence to support its argument that religion was not a motivating factor in its decision.”¹⁴⁸ The university argued that under the *McDonnell Douglas* three-part framework for indirect evidence of discrimination, Gaskell had failed to show that UK’s proffered reasons for not hiring him were a pretext.¹⁴⁹ UK argued, alternatively, that even if Gaskell’s religion played a role in their decision, the university could not reasonably accommodate his beliefs in conjunction with a position that required public outreach for fear that he would use his affiliation to promote

142. *Id.*

143. *Id.* at *5.

144. *Id.*

145. *Id.* at *6.

146. *Id.*

147. *Id.*

148. *Id.* at *10–11.

149. *Id.* at *7.

his beliefs.¹⁵⁰ Under the so-called “safe harbor” provision of the definition of “religion” under Title VII, an employer may demonstrate that it would cause “undue hardship” to accommodate a prospective employee’s “religious observance or practice.”¹⁵¹ The court rejected that the safe harbor provision applied in this case, and denied summary judgment on UK’s claim that Gaskell failed to show that their reasons were a pretext.¹⁵²

In his cross motion for summary judgment, Gaskell, following a “mixed motives” theory, argued, *inter alia*, that UK was at least partly motivated not to hire him because of his religion.¹⁵³ Congress amended Title VII in 1991 to “eliminate the employer’s ability escape liability in mixed-motives cases by proving that it would have made the same decision in the absence of the discriminatory motivation.”¹⁵⁴ Thus, following § 2000e-2(m), a plaintiff can establish an “unlawful employment practice” when one of Title VII’s protected categories “was a motivating factor for any employment practice, even though other factors motivated the practice.”¹⁵⁵ Gaskell presented direct evidence—including evidence that members of the search committee had searched his personal web page—to support his mixed motives claim.¹⁵⁶ However, the court held that UK had presented enough evidence in response to survive summary judgment.¹⁵⁷ Soon after, the parties reached a settlement awarding Gaskell \$125,000 with no admission of liability.¹⁵⁸

Gaskell illustrates the risk of Title VII liability associated with the Hiring Committee Approach to information gathering. Here, Sally Shafer, a member of UK’s search committee, performed the internet background search herself, leading her to discover information regarding Gaskell’s religious views that she likely would not have known based on the applicant’s interviews and other materials.¹⁵⁹ Upon discovery of Gaskell’s allegedly creationist beliefs, Shafer emailed links to his personal webpage—which contained Gaskell’s article and lecture notes espousing the religious views that caused her concern—to two of her colleagues on the search committee,

150. *Id.*

151. 42 U.S.C. § 2000e(j) (2011).

152. *Gaskell*, 2010 U.S. Dist. LEXIS 124572, at *9.

153. *Id.* at *10.

154. *Id.* (quoting *Hill v. Lockheed Martin Logistics Mgmt., Inc.*, 354 F.3d 277, 284 (4th Cir. 2004)).

155. 42 U.S.C. § 2000e-2(m).

156. *Gaskell*, 2010 U.S. Dist. LEXIS 124572, at *10.

157. *Id.* at *11.

158. *Kentucky: Discrimination Suit Is Settled at University*, N.Y. TIMES, Jan. 19, 2011, http://www.nytimes.com/2011/01/19/education/19brfs-DISCRIMINATI_BRF.html.

159. *See Gaskell*, 2010 U.S. Dist. LEXIS 124572, at *4.

and they were eventually shared with the entire group.¹⁶⁰ One of the consequences of the Hiring Committee Approach to information gathering was that Shafer, as both an information gatherer and a hiring decision maker, was exposed to information about the applicant that, if the decision was not to hire, had great potential to lead to allegations of employment discrimination.

Contrast the Hiring Committee Approach with an approach that separates the information gatherer from the information user. If UK had followed the Special Department Approach, a member of the university's staff, playing no role in the hiring decision, would have conducted the internet background check independently and presented his findings to the search committee. Because of the separation between the information gatherer and the decision makers, there would have been an opportunity to redact information about Gaskell's religious beliefs. Although nothing in Title VII would compel the information gatherer to remove references to Gaskell's religion, because the Act only declares it unlawful to *use* that information as the basis for a hiring decision, the university might adopt a policy of expunging information from the internet background report that could lead to claims of employment discrimination.¹⁶¹ This type of policy would limit the UK search committee's liability under Title VII, but would not be mandatory under current law. Thus, the Special Department Approach only guarantees, by virtue of its structure, that the search committee would not *automatically* learn information relating to a protected aspect of an employee's identity, and any policies to remove protected class information would be entirely voluntary.

As a theoretical matter, Title VII similarly would not compel a third-party internet background reporting agency to redact references to Gaskell's religious beliefs, because the burden ultimately rests on the employer to base its decision not to hire an applicant on lawful grounds.¹⁶² However, as a practical matter, filtering out information related to the applicant's status as a member of a protected class is *a core service that third-party internet background reporting agencies are providing*. Consider, for example, Social Intelligence's sales pitch from the front page of their website: "Federal and state protected class information is redacted from the reports we provide. Employers are only exposed to information that is job relevant and may legally be considered in

160. *Id.* at *9.

161. *See* 42 U.S.C. § 2000e-2(a) (2011).

162. *See* 15 U.S.C. § 1681e(b) (2011).

the hiring process.”¹⁶³ In the same way that Social Intelligence blocked the GIZMODO journalist’s face and hands from the pictures included in his report to hide his ethnicity, a third-party internet background reporting agency would seek to remove the references to Gaskell’s religion from his personal website had it prepared a report for UK’s search committee.¹⁶⁴ Thus, this approach would have limited UK’s liability under Title VII and many other federal and state equal employment laws, especially because third-party internet background reporting agencies are well-positioned to develop an expertise in this area of employment law in a way that employers might not be.

B. ADA PROHIBITION ON PREEMPLOYMENT INQUIRIES

In contrast to Title VII, the Americans with Disabilities Act of 1990 (“ADA”) prohibits certain types of questions an employer may ask a candidate during the hiring process, and therefore governs both information use and, to a certain extent, the manner in which information is gathered. Congress passed the ADA for the purpose of “provid[ing] clear, strong, consistent, enforceable standards addressing discrimination against individuals with disabilities.”¹⁶⁵ One of these standards includes a prohibition on making preemployment inquiries about an applicant’s disability. Unlike Title VII, which does not prohibit an employer from *asking* about information relating to the job applicant’s status as a member of a protected class—it only controls how the employer *uses* the information—the ADA prohibits these types of inquiries relating to an applicant’s disability.

The relevant section of the ADA reads: “[A] covered entity [employer] shall not . . . make inquiries of a job applicant as to whether such applicant is an individual with a disability or as to the nature or severity of such disability.”¹⁶⁶ Congress prohibited such inquiries to ensure that “[i]ndividuals with disabilities [have] a fair opportunity to be judged on their qualifications, ‘to get past that initial barrier’ where an employment judgment might be unfairly made based on disabilities rather than abilities.”¹⁶⁷ There are three exceptions to this general rule against preemployment inquiries, which the Ninth Circuit summarized in the following way:

163. *Executive Employment Screening*, SOCIAL INTELLIGENCE, <http://www.socialintel.com/hiring> (last visited Oct. 17, 2011).

164. See Honan, *supra* note 21.

165. 42 U.S.C. § 12101(b)(2).

166. *Id.* §§ 12111(2), 12112(d)(2)(A).

167. *Harris v. Harris & Hart, Inc.*, 26 F.3d 838, 841 (9th Cir. 1999) (citing 135 Cong. Rec. § 10768 (daily ed. Sept. 7, 1989) (statement of Sen. Harkin)).

- (1) when pre-employment inquiries relate to the ability of an applicant to perform job-related functions;
- (2) when an employer could reasonably believe that an applicant's known disability will interfere with the performance of a job related function; and
- (3) when an applicant requests reasonable accommodation for the application process or for the job.¹⁶⁸

Even with these exceptions, a preemployment inquiry into to a job-related function should be “narrowly tailored and it should not be phrased in terms of [the] disability.”¹⁶⁹ One circuit has even held that non-disabled job applicants may bring suit under this provision of the ADA for making improper inquiries.¹⁷⁰

In the brick-and-mortar world, as opposed to the world of the Internet, courts have applied the ADA's prohibition on preemployment inquiries regarding an individual's disability to questions arising in the context of face-to-face job interviews and job applications. For example, in *Equal Employment Opportunity Commission v. Wal-Mart Stores, Inc.*, the Tenth Circuit upheld the district court's finding that a Wal-Mart manager's interview question—“What current or past medical problems might limit your ability to do a job?”—was a prohibited preemployment inquiry under the ADA because it “did not concern [the plaintiff's] ability to perform specific job-related functions.”¹⁷¹ However, if an applicant volunteers information, courts have held that an employer may then discuss the applicant's disability. In *Cole v. Staff Temps*, the Supreme Court of Iowa ruled that a discussion of an applicant's medical restrictions in which she volunteered information about her physical restrictions did not fall within this provision of prohibited inquiries under the ADA.¹⁷²

In the context of internet background reports, it is possible that the ADA's prohibition on preemployment inquiries could be extended to

168. *Id.* at 841–42 (9th Cir. 1999) (citing 42 U.S.C. § 12112(d)(3), (4); 29 C.F.R. § 1630.14(1)); *see also* Equal Employment Opportunity Comm'n, Enforcement Guidance: Preemployment Disability-Related Inquiries and Medical Examination Under the Americans with Disabilities Act of 1990, EEOC Notice 915.002 (outlining the EEOC's position on permitted and prohibited preemployment inquiries under the ADA).

169. *Cole v. Staff Temps*, 554 N.W.2d 699, 706–07 (Iowa 1996) (citing 45A Am. Jur. 2d Job Discrimination § 564 at 528).

170. *See* *Griffin v. Steeltek, Inc.*, 160 F.3d 591, 595 (10th Cir. 1998). *But see* *Armstrong v. Turner Indus., Inc.*, 141 F.3d 554, 562 (5th Cir. 1998).

171. *E.E.O.C. v. Wal-Mart Stores, Inc.*, No. 98-2122, 1999 U.S. App. LEXIS 33144, at *12 (10th Cir. Dec. 21, 1999) (unpublished opinion).

172. *Cole*, 554 N.W.2d at 707.

prohibit certain types of internet searches. For example, like the prohibited interview question in *Wal-Mart* asking about the employee's past medical problems, an internet search for the employee's name plus the word "disability" might constitute a prohibited preemployment inquiry.¹⁷³ Such a search query might return links to the applicant's posts on internet message boards discussing the disability, links to any information about the applicant's advocacy work for people with that disability, or even links to the applicant's prior litigation concerning disability discrimination. The District Court for the Eastern District of New York recently heard a case in which the plaintiff claimed disability discrimination based on his allegation that the employer performed an internet background check and found out about the plaintiff's string of previous lawsuits involving his disability.¹⁷⁴ The court dismissed the claim as baseless because the plaintiff did not plead any facts to support his theory (and warned the plaintiff that further abuse of the legal system would result in sanctions). Therefore, it remains an open question whether an internet background check would have qualified as a preemployment inquiry had there been evidence that the employer searched the plaintiff.¹⁷⁵ It also remains uncertain what kind of search query would be considered an "inquiry" under the ADA; would the employer need to include the word "disability" or some other magic words, or would simply searching the applicant's name be enough? Would it require clicking a search result that seemed to indicate it would reveal something about the applicant's disability?

The employer-based methods of internet background information gathering—the Hiring Committee and Special Department Approaches—expose employers to the risks associated with this uncertainty in a way that the Third-Party Approach may not. The ADA defines the term "covered entities," to which the prohibition on preemployment inquiries applies, to mean "employer, employment agency, labor organization, or joint labor-management committee."¹⁷⁶ According to the plain meaning of this definition, there is little doubt that the information gatherer in the Hiring Committee and Special Department Approaches would be considered agents of the "employer" and therefore "covered entit[ies]" subject to the prohibition on preemployment inquiries. However, it is less clear whether the information gatherer in the Third-Party Approach would fall into any of the categories of "covered entity." The closest category under the definition

173. See *E.E.O.C.*, 1999 U.S. App. LEXIS 33144, at *12.

174. *Digianni v. Am. Int'l Grp.*, No. 08CV4355, 201 U.S. Dist. LEXIS 34015, at *9 (E.D.N.Y. Mar. 31, 2010).

175. *Id.* at *13–14.

176. 42 U.S.C. § 1211(2) (2011).

would be “employment agency,” but that incorrectly implies that the third-party internet background reporting agency plays a role in matching employees to employers. Thus, the third-party information gatherer has a strong claim to being exempt from the prohibition on preemployment inquiries under the ADA, making this a less risky alternative for internet background checks from the perspective of ADA liability. Moreover, as a practical matter and as discussed in the Title VII context,¹⁷⁷ the third-party information gatherer would sanitize the applicant’s internet background report of any reference to his disability, since third-party agencies typically eliminate *all* categories of information that would be unlawful bases upon which to make an employment decision.

C. STATE PROHIBITIONS ON LIFESTYLE DISCRIMINATION

In addition to these two major federal statutes, and the panoply of other federal equal employment laws, states have enacted a number of laws banning employment discrimination in many forms, including prohibitions on so-called “lifestyle discrimination” that are particularly relevant to internet background reporting. From the employer’s perspective, some off-duty conduct can affect the employer’s financial interests. As Professor Sugarman points out, the employer may justify a decision not to hire an employee because of his lifestyle “on the ground that the consequences of the off-duty behavior in some way spill over to the workplace, affecting the employer’s legitimate interests.”¹⁷⁸ Employers are, therefore, motivated to avoid employees with behaviors that could cost them in terms of decreases in individual’s productivity, the creation of interpersonal tensions, tarnish to an organization’s reputation, or increased premiums for healthcare or other benefits.¹⁷⁹ Off-duty activities that can clash with an employer’s interests include the employee’s social/sexual relationships (e.g., having an extramarital relationship), civil/political activities (e.g., speaking out at public hearings), dangerous leisure activities (e.g., hang gliding or skydiving), moonlighting (e.g., working a second job as a “centerfold” model), illegal acts (e.g., crimes committed outside of work), and daily habits (e.g., drinking and smoking).¹⁸⁰ In response to the possible employer motivation to discriminate against an employee’s daily habits, approximately 50% of states have

177. *See supra* Section V.A.

178. Stephen D. Sugarman, “Lifestyle” Discrimination in Employment, 24 BERKELEY J. EMP. LAB. L. 377, 379 (2003).

179. *See id.* at 383.

180. *Id.* at 384–95.

prohibited discrimination based on an employee's status as a smoker.¹⁸¹ Illinois, Missouri, and Wisconsin have extended this to include protection from discrimination based on the lawful consumption of alcohol.¹⁸²

State prohibitions on lifestyle discrimination are yet another area where the employer-based methods of internet background information gathering expose employers to the risk of a discrimination lawsuit, especially because the Internet (and social networks, in particular) are so full of records documenting a prospective employee's lifestyle. For example, Facebook pictures of an applicant posing with a beer or smoking a cigarette are direct evidence of the applicant's lifestyle choices and, at the same time, are illegal bases upon which to make an employment decision in the states with prohibitions on this sort of lifestyle discrimination. Moreover, because of the heterogeneity of state prohibitions on lifestyle discrimination, third-party background screening services might be better positioned to develop the expertise necessary to comply with the state-to-state variations.

VI. RECOMMENDATION: THE LEGAL AND STRUCTURAL ADVANTAGES OF THE THIRD-PARTY APPROACH

Thus far, this Note has addressed the varied legal and structural consequences of the employer's choice of internet background information search method under fair credit reporting and equal employment laws. As discussed in Part III, *supra*, the FCRA only regulates the information gathering process of the Third-Party Approach, leaving the Hiring Committee and Special Department Approaches subject only to additional state credit reporting regulations.¹⁸³ This means, among other things, that only the information gatherer in the Third-Party Approach must maintain procedures that ensure the accuracy of the internet background reports it produces. Moreover, as examined in Part V, *supra*, although state and federal equal employment laws generally do not require any single approach to information gathering, these three approaches have different structural consequences that vary the risk of employment discrimination. Because the information gatherer in the Hiring Committee Approach is the same person who ultimately uses the internet background information, that person is doomed to discover information relating to inappropriate bases upon which

181. See 1–9 Employment Screening § 9.07 n.1, *Tobacco Usage and Other Off-Duty Activities*, MB (2011) (citing twenty-one different states with laws prohibiting employers from refusing to hire an applicant because he smokes).

182. 820 ILL. COMP. STAT. § 55/5 (2011); MO. REV. STAT. § 290.145 (2011); WIS. STAT. ANN. § 111.321 (2011).

183. See, e.g., CAL. CIV. CODE § 1786.53(a) (West 2011).

to make an employment decision, simply by virtue of that approach's structure. Though the structure of the Hiring Committee Approach, without more, does not violate most equal employment laws,¹⁸⁴ the Hiring Committee's automatic exposure to the protected class information creates a basis upon which a plaintiff could allege employment discrimination. In light of these conclusions drawn from the legal analysis of information gathering and information use, this Part will suggest recommended practices under the current law.

Both employers and prospective employees benefit from a system that produces accurate internet background reports. For the employer, an inaccurate internet background report increases search costs in the hiring process by leading the employer to decline applicants who would otherwise be fit for the job. This limits the search pool, making eligible candidates artificially scarce, and correspondingly raises the overall cost of finding a suitable substitute applicant. For the prospective employee shut out of the job, an inaccurate internet background report subjects him to an increase in *his* search costs associated with finding a suitable replacement position. Moreover, the job applicant could be rejected by all employers that use information gathering methods that produce an inaccurate internet background report, which could occur, for example, when unsophisticated information gatherers attribute offensive Googleganger internet sources to the applicant himself.

Because of the high value of accuracy for both employers and prospective employees in the hiring process, employers should adopt the FCRA standard of "reasonable procedures to assure maximum possible accuracy" regardless of their approach to internet information gathering.¹⁸⁵ Furthermore, despite the uncertainty resulting from the circuit split on how to interpret this accuracy standard, employers should implement the Fifth Circuit's approach by requiring "minimum . . . points of correspondence" between the job applicant's file and the internet source.¹⁸⁶ In practice, this approach would require "two factor" matching, meaning that the applicant's name *plus* one other identifying category (birthday, educational background, physical location, email address, etc.) must match the internet source before that source can be considered part of the applicant's internet background report. Two-factor matching would necessarily decrease the amount of Googleganger data ending up in an applicant's internet background report

184. *See supra* Section V.B.

185. *See* 15 U.S.C. § 1681e(b) (2011).

186. *See* Thompson v. San Antonio Retail Merch. Ass'n, 682 F.2d 509, 513 (5th Cir. 1982).

because of the decreased probability that two factors will match a false source of internet information. The result would be increased accuracy in the information gathering process that would redound to the benefit of both employers and prospective applicants.

Although increasing internet background reporting accuracy appears to have positive externalities, accuracy, by itself, does little to uphold the core function of equal employment laws to prevent employment discrimination. After all, a fully accurate internet background report might nevertheless include information relating to the job applicant's status as a member of a protected class, which might improperly influence the hiring decision maker. From the job applicant's perspective, an internet background report that reveals information relating to his protected class status that the applicant deliberately did not disclose in his other materials might raise concerns about the fairness of the process (and perhaps about an invasion of privacy). From the employer's perspective, an internet background report that raises questions about the impartiality of the hiring process only increases the risk of a lawsuit, which can be expensive in terms of costs and bad publicity, as the University of Kentucky discovered in *Gaskell*.¹⁸⁷

In light of these concerns about impartiality resulting from internet background reports that contain protected class information, the information gatherer should be a separate person from the information user, and the information gatherer should redact all protected class information from the report. Both the Special Department and Third-Party Approaches to information gathering could adopt these policies, because structurally the information gatherer is always a different person from the information user in these approaches. As long as the barrier between employer's staff member conducting the internet information gathering and the hiring committee remains impermeable, the Special Department Approach could, in theory, provide the same structural benefits from separation as the Third-Party Approach in many circumstances.

However, in practice the Third-Party Approach better avoids the risk of employment discrimination for at least two reasons. First, as discussed in Section V.B, *supra*, the ADA prohibits the employer from making preemployment inquiries about the job applicant's disabilities, even if the applicant is not disabled.¹⁸⁸ Courts could reasonably extend this prohibition to include internet searches about the applicant that would likely reveal information about the applicant's disability status. Although it remains

187. *See supra* Section V.A.

188. 42 U.S.C. §§ 12111(2) & 12112(d)(2)(A) (2011).

unclear what kind of internet search would constitute a prohibited preemployment inquiry, it is possible that simply searching the applicant's name and then clicking through links that suggest information about the applicant's disability could suffice. After all, it would be easy to circumvent a rule against searching the applicant's name *plus* some magic word (such as "disability"), since undoubtedly the results of this narrower search (name plus disability) would be buried throughout the broader search for only the applicant's name. Thus, in the Special Department Approach, even if the insulated information gatherer eventually redacted information relating to the applicant's disability, this approach might still expose the employer to liability for violating this provision. However, because the ADA's prohibition does not likely apply to third-party information gatherers (because they fall outside of the definition of "covered entities" under the Act), the Third-Party Approach avoids this risk.¹⁸⁹

Second, the Third-Party Approach better avoids the risk of employment discrimination because specialized internet background reporting agencies can develop an expertise in this area of law. The mosaic of federal and state equal employment laws—including state prohibitions of lifestyle discrimination—is complicated to navigate, and the application of these laws to internet background reporting can be uncertain, as this Note has explored. Given the two-layered complexity consisting of the laws themselves and their application to internet background reporting, the third-party information gatherer is in a better position to develop the expertise necessary to avoid unintentional violations. Of course, this reason for preferring the Third-Party Approach over the Special Department Approach is case-specific; a sophisticated employer with the resources to develop a Special Department with the necessary legal expertise might produce internet background reports that fully comply with all state and federal equal employment laws. Nevertheless, during this early period in the application of these laws to internet background reporting, and in light of the ADA's prohibition on preemployment inquiries that applies to employers but likely not to third-party information gatherers, the Third-Party Approach appears most prudent for employers.

VII. CONCLUSION

This Note analyzed the legal and normative issues surrounding internet background checks. After concluding that, at a minimum, a fifth to a quarter of employers use internet search engines or social networks to screen

189. *See id.* § 1211(2).

candidates at some point during the hiring process, this Note suggested a taxonomy of three different approaches to internet information gathering. It then considered how fair credit and equal employment laws might apply to these three approaches. Based on this analysis, this Note concluded that the major federal regulatory requirements of the FCRA apply only to the Third-Party Approach to internet background information gathering, though state fair credit reporting laws may reach the Hiring Committee and Special Department Approaches. Furthermore, although equal employment laws generally do not mandate any particular information gathering approach, the ADA prohibits an employer's preemployment inquiries regarding disabilities, which suggests that the employer-based approaches expose the employer to greater liability than the Third-Party Approach. This Note then concluded with recommended best practices for employers in light of this legal analysis, ultimately suggesting that the Third-Party Approach both helps ensure the accuracy of the internet background report and, at the same time, reduces the risk of discrimination by virtue of its structure. The importance of internet background reporting can only increase. As our life on the Internet becomes a true parallel to our life in the real world, our internet background report will become the full "background of our being," to borrow a phrase,¹⁹⁰ documenting everything.

190. *Cf.* RALPH WALDO EMERSON, SELF-RELIANCE: THE OVER-SOUL AND OTHER ESSAYS 57 (Coyote Canyon Press Ed., 2010).

THE LAW OF THE ZEBRA

Andrea M. Matwyshyn[†]

ABSTRACT

At the dawn of internet law, scholars and judges debated whether a “law of the horse”—a set of specific laws addressing technology problems—was ever needed. Time has demonstrated that in some cases, the answer is yes. However, some courts are confused about the appropriate trajectory of contract law in technology contexts. Today, a technology-centric analysis threatens to subvert traditional contract law and the future of entrepreneurship, and circuit splits have emerged in what might be called an undesirable “law of the zebra.” Do contracts that involve technology always require exceptional contract rules? In particular, does the use of a computer to breach a contract make the breach inherently worse in law? Using the Computer Fraud and Abuse Act (“CFAA”) as a case study, this Article introduces a paradigm of “restrained technology exceptionalism” in contract law, a paradigm predicated on a return to traditional contract law principles and contractual supremacy over technology exceptionalist legal approaches. Applying the restrained technology exceptionalism paradigm to a circuit split concerning the CFAA, this Article then introduces a “privity” model bridging contract law with CFAA analysis, which inverts the traditional legal framing of the CFAA and contract relationship. It argues that where contract formation has occurred, the contract controls the relationship and only contract remedies are appropriate; both criminal and civil CFAA analysis becomes inapposite and should not be considered. Thus, the operative legal question is not whether contract breach revokes authorization in CFAA context. Rather, it is whether contract analysis is possible, thereby rendering the CFAA unnecessary and redundant in contractual contexts.

© 2013 Andrea M. Matwyshyn.

[†] Andrea M. Matwyshyn is an Assistant Professor of Legal Studies and Business Ethics in the Wharton School at the University of Pennsylvania and an affiliate scholar of the Center for Internet and Society at Stanford Law School. She can be reached at amatwysh@wharton.upenn.edu and wishes to thank the faculty of the Oxford Internet Institute and Notre Dame Law School, where she was a visitor during the writing of this Article. She also wishes to thank Colleen M. Baker, Tricia Bellia, Christina J. DeVries, Margaret F. Brinig, Ian Brown, Erik Goldman, Cathy Kaveny, Orin Kerr, Michael Kirsch, Randy Kozel, Greg Lastowka, Miluska Linares, Brian Martin, Jennifer Mason McAward, Mark McKenna, Jennifer Mueller, John Nagle, Paul Ohm, Frank Pasquale, Jeffrey A. Pojanowski, Marty Redish, John H. Robinson, Stephen F. Smith, David Schwartz, Marcia Tiersky, Julian Velasco, and Kevin Werbach for their helpful commentary and critiques of this Article.

TABLE OF CONTENTS

I.	INTRODUCTION	157
II.	HORSES AND ZEBRAS: TOWARD RESTRAINED TECHNOLOGY EXCEPTIONALISM IN CONTRACT	160
A.	CONTRACTS + TECHNOLOGY = DOCTRINAL CONFUSION	161
1.	<i>Confusion and Circuit Splits: bOrked Doctrine and Balkanization</i>	161
2.	<i>Hack3r Cat Says “I can haz contract.”</i>	165
B.	THE PARADIGM OF “RESTRAINED TECHNOLOGY EXCEPTIONALISM”	168
1.	<i>Principle 1—Freedom to Contract: Equalizing Bargaining Power and Making Consent Meaningful in Formation</i>	170
a)	Gaming in Presentation and Content	171
b)	Technology Exceptionalism in Formation Is Appropriate	174
2.	<i>Principle 2—Freedom from Contract in Breach: Reasonable Enforcement in Line with Contractual Liberty and Rights of Exit</i>	175
a)	Duties of Good Faith in Performance and Enforcement	175
b)	Ability to Exit and “Digital Peonage” Concerns.....	176
3.	<i>Principle 3—Damages: Preserving Primacy of Traditional Contract Remedies</i>	178
a)	Damages Primacy	178
b)	The Unique Concerns of Information Harms.....	180
4.	<i>Principle 4—Supplementing But Not Supplanting Contract</i>	181
III.	HORSES AND CLAPPING COCONUTS: THE PROBLEMATIC CASE OF WEAPONIZED BREACH AND COMPUTER INTRUSION UNDER THE CFAA	182
A.	WEAPONIZING BREACH: THE CFAA AND A TALE OF FOUR “CONTRACT HACKERS”.....	184
1.	<i>Consumer Users of Technology as “Contract Hackers”: Confusing Minor Breach with Black Hat Hacking</i>	187
2.	<i>“Disloyal” Employees or Business Partners as “Contract Hackers”: Confusing Intellectual Property Harms with Black Hat Hacking</i>	188
a)	The Alleged Thieves and Vandals	189
b)	The Alleged Slackers	191
3.	<i>Entrepreneurs as “Contract Hackers”: Confusing Innovation with Black Hat Hacking</i>	193
a)	Application Builders.....	193

b)	Data Aggregators.....	194
4.	<i>Security Researchers as “Contract Hackers”: Confusing Code Auditing and White Hat Hacking with Black Hat Hacking.....</i>	196
B.	THE RISKS OF WEAPONIZING BREACH WITH THE CFAA	198
1.	<i>Contract Doctrine and Judicial Activism</i>	198
2.	<i>How Weaponized Breach Disrupts Contract Theory</i>	201
3.	<i>Private Ordering.....</i>	205
4.	<i>Innovation and Entrepreneurship Policy</i>	206
5.	<i>Competing Legal Regimes.....</i>	207
C.	WHY COURTS MAY HAVE BECOME CONFUSED	208
1.	<i>Essentialism.....</i>	209
2.	<i>Confirmation Bias.....</i>	210
IV.	APPLYING RESTRAINED EXCEPTIONALISM: A PRIVACY MODEL OF AUTHORIZED ACCESS	211
A.	THE “PRIVITY” MODEL: CONTRACT TERMS AND NORMS OF ACCESS	215
1.	<i>The Privity Model.....</i>	216
B.	HANDLING THE FOUR CONTRACT “HACKERS”.....	219
1.	<i>Users</i>	219
2.	<i>Employees and Business Partners</i>	220
3.	<i>Entrepreneurs</i>	221
4.	<i>Security Researchers.....</i>	223
C.	CRAFTING GOOD NORMS IN INFORMATION SECURITY.....	223
V.	CONCLUSION	225

I. INTRODUCTION

In a seminal debate in the early days of internet law, Professor Lawrence Lessig and Judge Frank Easterbrook exchanged legal barbs over the question of whether the Internet deserves its own regulation. Judge Easterbrook referred to the suggestion of internet-specific approaches as tantamount in irrationality to suggesting a “law of the horse.”¹ Professor Lessig, in turn, responded by identifying some spaces where existing legal frameworks fall short and suggesting what this “law of the horse” might look like.² Central to Lessig’s argument was the primacy of private ordering and contract law.³

1. Easterbrook viewed crafting internet-specific laws to be an unnecessary enterprise, asserting that existing legal regimes can adequately accommodate any harms that arise in internet spaces. Frank H. Easterbrook, *Cyberspace and the Law of the Horse*, 1996 U. CHI. LEGAL F. 207, 208 (1996).

2. Lawrence Lessig, *The Law of the Horse: What Cyberlaw Might Teach*, 113 HARV. L. REV. 501, 502 (1999).

3. *Id.*

Today, almost fifteen years after that seminal debate, there is no doubt as to whether this “law of the horse” exists—it does, without question. In fact, since the Lessig–Easterbrook exchanges, Congress has passed various technology-focused statutory regimes.⁴

However, what has happened in the shadow of the “law of the horse” debates is perhaps partially what concerned Easterbrook at the time of his writing, at least in the context of contract law. Today, courts are derailing traditional contract law approaches with an overzealous focus on the role of technology in disputes. Instead of asking whether a technology-specific “law of the horse” should be crafted to fill gaps in existing law in technology contexts, courts now ask whether technology-specific approaches should usurp the traditional space of contract law. We have arrived at something more exotic and unexpected than the “law of the horse.” We have entered the strange and strained realm of what might be dubbed the “law of the zebra.”

An aphorism states “when you hear hoofbeats, think horses, not zebras.”⁵ When we turn to contract law inquiries in technology spaces, courts increasingly look for a metaphorical zebra, rather than a horse or nothing at all. Instead of using contract law in its traditional forms to resolve disputes, and supplementing it with technology-exceptionalist approaches only where true novelty exists, some courts now reach aggressively for technology exceptionalist approaches as a first cut.⁶ These courts now seem to ignore the existence of a contract governing the exchange and instead focus on the presence of a computer.⁷ At times this approach leads to intrajurisdictional conflict. A particular court may, on the one hand, ignore the novel challenges technology presents to contract formation.⁸ But, on the other hand, that court may be quick to point to allegedly unique characteristics of the same technology when analyzing a question of breach.⁹ As a result, contract law is becoming progressively more balkanized on technology issues, and, unsurprisingly, circuit splits have emerged. This Article calls for a revitalization of contract discourse around technology questions, and it

4. *See, e.g.*, Controlling the Assault of Non-Solicited Pornography And Marketing Act, 15 U.S.C. §§ 7701–7713 (2011) [hereinafter CAN-SPAM Act]; Digital Millennium Copyright Act, Pub. L. No. 105-304, 112 Stat. 2860 (1998) (codified as amended at 17 U.S.C. § 512 (2011)).

5. *World Proverbs*, SPECIAL DICTIONARY, <http://www.special-dictionary.com/proverbs/keywords/zebra/> (last visited Feb. 28, 2012).

6. *See infra* Section II.A.1.

7. *See, e.g.*, Int'l Airport Ctrs., L.L.C. v. Citrin, 440 F.3d 418, 420 (7th Cir. 2006).

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

9. *See Citrin*, 440 F.3d at 420.

cautions against crafting a “law of the zebra”—an undesirable body of law where technology exceptionalism triumphs over traditional legal paradigms.

Using judicial interpretation of the Computer Fraud and Abuse Act (“CFAA”) as a case study of the looming “law of the zebra,” this Article breaks with existing scholarship and advocates a decidedly contractarian analysis: a paradigm of “restrained technology exceptionalism.” It also offers one operationalization of this paradigm: a “privity” model to address the circuit split with respect to contract breach and CFAA interpretation. Analysis of CFAA issues in the legal literature to date has primarily come from a criminal or property perspective. This Article adopts a different approach and examines the CFAA through the lens of contract law. It argues that if a contract exists between the parties, a contract inquiry must precede any CFAA analysis. Thus, the goal of this Article is not only to ask what is wrong with the CFAA, in particular. Instead, it asks a broader question: what has gone awry with contract law in technology contexts? Why have we wrongly permitted frameworks such as the CFAA to subvert traditional contract law with overreaching technology exceptionalism?

Part II introduces the current state of balkanization in contract caselaw around technology. Courts sometimes try to duck basic contract queries in technology cases; alternatively, they haphazardly apply technology exceptionalist analysis without a consistent paradigm. For example, courts currently tend to avoid technology exceptionalist analysis in questions of formation, but are sometimes quick to apply an exceptionalist analysis in breach. In fact, the opposite is more appropriate: while contract formation questions frequently warrant technology exceptionalism, a technology exceptionalist analysis is usually misguided in questions of breach. Next, Part II introduces one possible solution to these doctrinal challenges: a paradigm of “restrained technology exceptionalism” in contractual interpretation. This approach turns on traditional contract law concerns: freedom to contract, freedom from contract, damages as the primary remedy for harm, and preservation of private ordering in technology contexts.

Applying the restrained technology exceptionalist paradigm to the case study of the CFAA-contract law circuit split, Part III then asks a superficially simple breach question: should a breach of contract relating to a computer or network automatically provide the basis for a criminal charge of computer intrusion under the CFAA? Introducing four types of “contract hacker” cases, this Part argues that many courts are wrongly viewing contract breaches that involve computers as somehow fundamentally “worse” than or different from other contract breaches. Without any justification as to why contract damages alone are an inadequate remedy, these courts overzealously sanction defendants with CFAA penalties in addition to contract remedies.

Yet, in almost all cases, contract damages sufficiently address the information harms at issue. Hence, we witness the arrival of a “law of the zebra.” Part III further argues that this type of “weaponized breach” analysis is highly undesirable as a matter of contract doctrine, contract theory, private ordering, and innovation and entrepreneurship policy. Finally, applying paradigms from developmental and social psychology theory, Part III postulates that courts’ tendency to inconsistently exceptionalize breach in technology contexts is likely rooted in essentialism and confirmation bias rather than principle-driven contract law analysis.

Part IV offers a concrete operationalization of a model embodying the paradigm of restrained technology exceptionalism. Applying the principles articulated in Part III, Part IV proposes a “privity” model of contract and the CFAA. The privity model is predicated on a notion of contractual dominance in technology contexts; it provides a means for handling the four “contract hackers” introduced in Part III, while staying true to contract law first principles. When the alleged computer intruder and the information holder stand in contractual privity with each other, and when their agreement directly or indirectly contemplates the information in question, CFAA analysis is inappropriate. The court should only conduct a contract breach analysis. Part V concludes.

II. HORSES AND ZEBRAS: TOWARD RESTRAINED TECHNOLOGY EXCEPTIONALISM IN CONTRACT

In medical contexts, a “zebra” is a slang term used to refer to a surprising and rare diagnosis.¹⁰ As this Part will explain, the way that some courts have addressed technology contracts is indeed both surprising and rare. In instances where applications of traditional contract law will suffice, courts nevertheless increasingly reach for technology-specific paradigms to trump traditional contract law analysis. Courts are hearing and expecting proverbial “zebras” where none actually exist.

In technology contexts, contract law has always been the most prevalent law. End user license agreements have existed almost as long as software, and since the early days of the Internet, terms of use have governed virtual spaces. Yet, in a strange inversion, technology exceptionalism now threatens doctrinal coherence in contract law.

Although it may have been possible in the past to ignore these doctrinal disagreements, such an approach is no longer sustainable. Circuit splits are

10. *See About Us*, ZEBRAMEDICINE, <http://www.zebramedicine.net/about-us.html>. (last visited Feb. 28, 2012).

emerging in ways that threaten the core of contract law and the future of the technology marketplace. To date, however, contract scholarship has offered few suggestions for doctrinal paradigms marrying contract law and technology questions successfully. This Part offers one possible solution: by recognizing four fundamental tenets of contract law challenged by technology, we can begin to craft a paradigm of “restrained technology exceptionalism” to guide future contract caselaw in technology contexts.

A. CONTRACTS + TECHNOLOGY = DOCTRINAL CONFUSION

Courts are struggling to understand the implications of technology for contract law. Should courts ignore contracts in some technology contexts in favor of other legal paradigms? Are contracts involving technology inherently special? Do they warrant their own exceptionalist contract law paradigms? Courts disagree on the answers to these questions,¹¹ and this disagreement now detrimentally affects everyday users of technology.

1. *Confusion and Circuit Splits: b0rked Doctrine and Balkanization*

Courts seem uncomfortable with contract law in technology contexts. Even the most contract-friendly of circuits now sometimes seem to contort their analysis to avoid conducting technology contract inquiries in technology contexts. For example, the Seventh Circuit—a circuit that has historically tended toward strict enforcement of contract terms even at the expense of equitable concerns¹²—sometimes appears to inexplicably duck contract analysis when technology is involved.¹³ For example, *International Airport Center L.L.C. v. Citrin*, a somewhat recent case, should have involved a relatively simple question of breach: an employee failed to perform in accordance with his employment agreement and damaged information on his laptop in the process.¹⁴ While a traditional breach analysis and a damages remedy would have sufficiently compensated the plaintiff for the harms

11. Andrea M. Matwyshyn, *Technology, Commerce, Development, Identity*, 8 MINN. J.L. SCI. & TECH. 515, 520, 523–24 (2007) (discussing how courts and scholars struggle generally to redress technology harms, as opposed to real-space harms).

12. For example, the unconscionability doctrine in the Seventh Circuit is particularly weak, essentially rejecting the role of substantive unconscionability in the analysis. *See, e.g.,* Nw. Nat’l Ins. Co. v. Donovan, 916 F.2d 372, 377 (7th Cir. 1990) (Posner, J.) (calling unconscionability an “umbrella term” for other traditional grounds to invalidate contracts).

13. On at least one occasion, the Seventh Circuit opted in favor of a convoluted agency analysis and invoked technology-specific legislation in lieu of a straightforward contract breach analysis. *See Citrin*, 440 F.3d at 420–21.

14. *Id.* at 420.

incurred,¹⁵ the court instead resorted to a convoluted agency and computer intrusion analysis, ignoring the contract questions entirely.¹⁶

Meanwhile, when courts do undertake contract analysis in technology contexts, the results seem haphazard and without a consistent paradigm. The Seventh Circuit and the Ninth Circuits in particular have adopted sometimes internally inconsistent positions with respect to technology exceptionalist analysis in contract. These courts have also decided cases with arguably similar facts in significantly different ways, causing a circuit split. On the point of contract formation, the Seventh Circuit has aggressively adopted a pro-drafter, non-exceptionalist position. Following the reasoning of *ProCD v. Zeidenberg*, the Seventh Circuit deems digital contexts and physical space contexts to be essentially equivalent for contract formation purposes: it appears to believe that no unique consumer protection concerns exist,¹⁷

15. Using the facts of *Citrin*, an expectation damages or a disgorgement-based contract remedy could have been crafted to compensate the employer. For example, the court could have required Citrin to pay out/return a portion of salary proportional to work he destroyed or assessed damages according to the market rate of hiring an employee to recreate the work that Citrin destroyed plus the value of any lost revenue because of the destroyed data. Further, the conduct at issue might provide a basis for an injunction under some employment agreements. The use of computer intrusion law to address this type of situation is, thus, superfluous and duplicative.

16. *Id.* at 420–21.

17. *ProCD, Inc. v. Zeidenberg*, 86 F.3d 1447, 1449, 1451 (7th Cir. 1996). In *ProCD v. Zeidenberg*, the Seventh Circuit assessed whether a plaintiff who had compiled databases from phone directories had a basis for remedy in contract for the breach of a defendant who allegedly violated the contractual restrictions on database reuse. *Id.* at 1449. The defendant argued that the terms were not binding, in part, due to the presentation of the agreement, which involved additional terms not readable by a consumer until the box was opened. *Id.* at 1450–51. The Court held that terms as drafted are binding on a consumer even if the consumer has not had opportunity to read all of the terms before she begins interacting with the product. *Id.* at 1449, 1451. The *ProCD* decision, which many would argue is fatally flawed, is predicated on outdated notions of technology, such as the ability to return software. See, e.g., Wendy J. Gordon, *Intellectual Property as Price Discrimination: Implications for Contract*, 73 CHI.-KENT L. REV. 1367, 1378–86 (1998) (criticizing ProCD for, among other things, comparing price discrimination to monopoly power rather than to a regime that permits free copying); Mark A. Lemley, *Beyond Preemption: The Law and Policy of Intellectual Property Licensing*, 87 CALIF. L. REV. 111, 147–51 (1999) (arguing that ProCD did not consider other relevant preemption doctrines, and that the type of contracts that are enforceable under ProCD can create “rights against the world” that might be contrary to public policy); Yochai Benkler, *Free As the Air to Common Use: First Amendment Constraints on Enclosure of the Public Domain*, 74 N.Y.U. L. REV. 354, 432 (1999) (“The practical effect of the decision to enforce mass market information licenses is that more uses of information will be prohibited to more people.”); Julie E. Cohen, *Copyright and the Perfect Curve*, 53 VAND. L. REV. 1799, 1812 (2000) (criticizing the price discrimination model because it skews incentives for creation in favor of works that produce large private gains at the expense of works intended primarily to benefit the public); see also Randal C. Picker, *Easterbrook on Copyright*, 77 U. CHI. L. REV. 1165,

stating “[s]hrinkwrap licenses are enforceable unless their terms are objectionable on grounds applicable to contracts in general.”¹⁸

However, on the point of technology breach, reframing the Seventh Circuit’s CFAA analysis in *International Airport Center LLC v. Citrin*¹⁹ as a question of contract law reveals that the court adopted a de facto technology exceptionalist analysis. The court asserted that the mere *intention to breach a contract terminates contractually granted access*²⁰ to technology: “authorization to access the laptop terminated when, having already . . . decided to quit [his employer] in violation of his employment contract, he resolved to destroy files.”²¹

The Ninth Circuit, on the other hand, has split with the Seventh Circuit and adopted a more consumer-friendly exceptionalist posture with respect to contract formation questions. In *Douglas v. United States District Court*,²² the Ninth Circuit found that modifications to an agreement that were posted on a website were not binding on a consumer, since “a party [would not] know when to check the website for possible changes to the contract terms without being notified that the contract has been changed” or the consumer “would have had to check the contract every day for possible changes. Without notice, an examination would be fairly cumbersome” because the consumer

1178 (2010) (“*ProCD* is the opinion that the copyright casebooks love to hate.”). Yet *ProCD* continues to be considered good law and has shaped numerous subsequent decisions. *See, e.g., Canal+ Image UK Ltd. v. Lutvak*, 773 F. Supp. 2d 419 (S.D.N.Y. 2011); *Appliance Zone, LLC v. NexTag, Inc.*, No. 4:09-cv-0089-SEB-WGH, 2009 WL 5200572, 93 U.S.P.Q.2d 1540 (S.D. Ind. Dec. 22, 2009); *Health Grades, Inc. v. Robert Wood Johnson University Hosp., Inc.*, 634 F. Supp. 2d 1226 (D. Colo. June 19, 2009); *Illinois Wholesale Cash Register, Inc. v. PCG Trading, LLC*, No. 08 C 363, 2008 WL 4924817, 67 UCC Rep.Serv.2d 291 (N.D. Ill. Nov. 13, 2008); *Propet USA, Inc. v. Shugart*, No. C06-0186-MAT, 2007 WL 1306540 (W.D. Wash. May 3, 2007).

18. *Id.* at 1449. *But see* *Specht v. Netscape Commc’ns Corp.*, 306 F.3d 17, 35 (2d Cir. 2002) (holding that internet customers did not manifest assent to an arbitration clause in a click-through agreement when the clause appeared on a scroll-down screen that did not provide “reasonably conspicuous notice of the existence of contract terms” prior to downloading defendant’s software).

19. *Int’l Airport Ctrs., L.L.C. v. Citrin*, 440 F.3d 418, 420 (7th Cir. 2006).

20. It is unlikely that the court would have made this argument about a company car, for example.

21. *Citrin*, 440 F.3d at 420 (holding that an employee was liable under the CFAA and the employee’s “authorization to access the [company] laptop terminated when . . . [the employee] resolved to destroy files that incriminated himself and other files that were also the property of his employer, in violation of the duty of loyalty that agency law imposes on an employee”).

22. *Douglas v. U.S. Dist. Court for the Cent. Dist. of Cal.*, 495 F.3d 1062 (9th Cir. 2007).

“would have had to compare every word of the posted contract with his existing contract in order to detect whether it had changed.”²³

But, on the point of technology breach analysis, the Ninth Circuit has waffled as to whether it should adopt an exceptionalist position on breach. For example, in *LVRC Holdings v. Brekka*²⁴ the Ninth Circuit arguably conducted a traditional contract law breach analysis in lieu of a CFAA-focused analysis, stating that “[t]here is no dispute that Brekka was given permission to use LVRC’s computer and that he accessed documents or information to which he was entitled by virtue of his employment with LVRC.”²⁵ Yet in *United States v. Nosal* the court initially conducted a CFAA-driven analysis in lieu of merely a traditional contract breach analysis. The court exceptionalized the technology circumstances,²⁶ asserting “[a]lthough we are mindful of the concerns raised by defense counsel regarding the criminalization of violations of an employer’s computer use policy, we are persuaded that [there are adequate protections] against criminal prosecution [of] those employees whose only violation of employer policy is the use of a company computer for personal—but innocuous—reasons.”²⁷ Other circuits are similarly leaning toward a technology exceptionalist analysis in contract law.²⁸ Further, as the next sections will argue, the deference that many courts pay to technology-specific legislative approaches in lieu of asserting the primacy of a traditional contract analysis is unwarranted, undesirable, and derailing the future of contract law.

23. *Id.* at 1066 n.1 (emphasis omitted).

24. *LVRC Holdings LLC v. Brekka*, 581 F.3d 1127 (9th Cir. 2009) (holding that because the defendant was authorized under his employment agreement to access documents on his employer’s computer and e-mail them to himself, he did not violate the CFAA).

25. *Brekka*, 581 F.3d at 1135. Brekka “would have no reason to know that making personal use of the company computer in breach of a state law fiduciary duty to an employer would constitute a criminal violation of the CFAA.” *Id.*

26. *United States v. Nosal (Nosal I)*, 642 F.3d 781, 782–89 (9th Cir. 2011) (holding that despite having authorization to access employer’s files, defendant employee “exceeded authorized access” and therefore violated the CFAA by exceeding the employer’s use restrictions), *rev’d en banc*, 676 F.3d 854 (9th Cir. 2012).

27. *Nosal I*, 642 F.3d at 782. En banc, however, the Court rejected the panel’s analysis and found that the employee’s breach of the employer’s use policy does not provide adequate basis for a charge of computer intrusion under the CFAA, which depends on computer access. *United States v. Nosal (Nosal II)*, 676 F.3d 854, 863 (9th Cir. 2012) (en banc).

28. *See United States v. John*, 597 F.3d 263, 271–72 (5th Cir. 2010) (holding that an employee’s conviction under the CFAA for exceeding authorized access did not constitute a miscarriage of justice where the employee accessed confidential customer information in violation of her employer’s computer use restrictions and used that information to commit fraud).

It is not entirely clear why contract issues that involve questions of technology prove disconcerting and confusing to courts.²⁹ It is clear, however, that the result of this confusion is a slow but steady balkanization of contract law. While courts interpret contracts involving goods relatively consistently—even contracts involving noncompetition agreements have evolved to embody some relatively consistent principles in interpretation—contracts involving technology contexts are causing divergence from rather than convergence around shared principles.

We have now arrived at the untenable point where this judicial confusion and contract law balkanization may begin to paralyze consumer behavior in technology contexts. Using a hypothetical about a consumer and her cat, we can start to frame the extent of the problems that this doctrinal uncertainty causes in everyday contracting. In brief, the average consumers can no longer predictably enter into contracts in technology contexts.

2. *Hack3r Cat Says “I can haz contract.”*

Imagine that you have gifted your grandmother an iPad for her birthday. Your Grandma, a woman with a heart of gold, has two great loves in life: her grandchildren and her cat, Dr. Whiskers. As you assist Grandma in unwrapping her shiny new iPad, you explain to her that if she learns to use the iPad and creates a Facebook account, she will be able to see many pictures of her grandchildren and read about what they are doing on a daily basis. Perhaps even more thrillingly, the Internet has many pictures of cats doing funny things,³⁰ and there are even cat games available for the iPad which Dr. Whiskers might enjoy playing.³¹ Grandma is sold.

You pull up the Apple app store. You quickly click “I Agree” on the Apple end user license agreement without reading it and hand the iPad to Grandma to guide her through the registration process. She tells you that she is registering Dr. Whiskers as the primary user of the device; she suspects he will be using the device as much as she will.³²

Next, you and Grandma log into the Facebook website together. As you start to help Grandma create a Facebook account, the very long terms and

29. Part II, *infra*, will nevertheless offer a hypothesis regarding the drivers of this dynamic.

30. See, e.g., Nadia Heninger, Telex and Ethan Zuckerman’s “Cute Cat Theory” of Internet Censorship, FREEDOM TO TINKER (July 22, 2011), <https://freedom-to-tinker.com/blog/nadiah/telex-and-ethan-zuckermans-cute-cat-theory-internet-censorship>.

31. See, e.g., jashmenn, *iPad Game for Cats: The World’s Greatest Video Game (for cats, not humans)*, YOUTUBE (Dec. 15, 2010), <http://www.youtube.com/watch?v=XK2dwTVi-aQ>.

32. Showing off your contract law knowledge, you tell Grandma that since Dr. Whiskers is nineteen years old, he is over the age of contractual consent in all jurisdictions.

conditions of use appear. Grandma asks you what all these terms mean. She tells you they are very small in size and that she cannot possibly read such small font.³³ Grandma also notes that there are various other documents linked up from this particular contract on Facebook; she asks you what they are and whether she is going to be bound by all of those terms as well. She is concerned that it may not be safe for her to participate in Facebook. You tell her not to worry; even Chief Justice Roberts just clicks yes on these types of terms of use without reading them.³⁴ Besides, you tell Grandma, companies like Facebook unilaterally change their privacy policies and terms of use so frequently that it is functionally impossible to keep up, and the terms are nonnegotiable anyway. Grandma clicks “yes I agree” on the Facebook terms of use, and she tells you that Dr. Whiskers should be registered as the primary user for this account too. You help Grandma type in Dr. Whiskers’ information in the boxes on the Facebook interface, and she uploads a particularly fetching photo of him.³⁵

What you may not realize is that you may have just exposed Grandma to possible criminal prosecution for the felony of computer intrusion. As a consequence of your gift of the iPad and your tutelage in creating a Facebook account for Dr. Whiskers, you may have just led Grandma astray into a life of crime as a “hacker,” according to some prosecutors.³⁶ You also may have just participated in a criminal conspiracy with Grandma to commit computer intrusion.

How could Grandma possibly be deemed a felon for her conduct? The legal argument goes as follows: When Grandma clicked “yes” on the terms of use (that she could not read), she, in theory, agreed to all terms in the agreement, including a provision that stipulated that all users must sign up for accounts in their real names. As Dr. Whiskers is not the (human) user connected to the account, by typing in his information rather than her own, Grandma has already breached a term of the contract. According to some courts, in that magic moment when Grandma breached the agreement, Facebook revoked her authorization to use the website, even in the absence

33. You show Grandma how to make things bigger on the iPad, but she tells you it is too hard for her to make that gesture with her arthritis constantly acting up.

34. Mike Masnick, *Supreme Court Chief Justice Admits He Doesn't Read Online EULAs Or Other 'Fine Print,'* TECHDIRT (Oct. 22, 2010 9:48 AM), <http://www.techdirt.com/articles/20101021/02145811519/supreme-court-chief-justice-admits-he-doesn-t-read-online-eulas-or-other-fine-print.shtml>.

35. You also help Grandma post the first status update, which says “Dr. Whiskers is purrfect in every way,” and you help Grandma friend you and the rest of her grandchildren.

36. *See* Indictment, United States v. Drew (C.D. Cal. Feb. 2008), *available at* <http://www.citmedialaw.org/sites/citmedialaw.org/files/2008-05-15-Drew%20Indictment.pdf>.

of notice to her of her breach. From that moment she entered Dr. Whiskers' information, she was no longer authorized to use the website.³⁷ Since she continued to use the website following her breach, her actions may be construed by some courts as "unauthorized access" constituting computer intrusion.

Although this hypothetical may seem far-fetched, a variant of the question presented in this hypothetical—whether a mere breach of contract can provide the basis for a criminal conviction for computer intrusion—has already caused a real-life circuit split in the courts.³⁸ Further, a case similar to the facts of the hypothetical, one where a consumer violated the terms of use of a social networking website and was prosecuted as a "hacker," resulted in a jury convicting the defendant in question on criminal computer intrusion charges.³⁹

Perhaps even more troubling than Grandma's possible "hacker" prosecution is the fact that the same courts that might have been quick to convict Grandma—arguing that technology contexts present unique harms with respect to breach—might instead refuse to acknowledge that Grandma faced any novel consumer protection concerns when she was entering into the contracts. In other words, the same court that, on the one hand, argues

37. *Id.*

38. *Compare* Int'l Airport Ctrs., L.L.C. v. Citrin, 440 F.3d 418 (7th Cir. 2006) (finding civil liability under the CFAA using an exceptionalist approach), *United States v. Rodriguez*, 628 F.3d 1258 (11th Cir. 2010) (following *Citrin* to find CFAA liability), *WEC Carolina Energy Solutions v. Miller*, 687 F.3d 199 (4th Cir. 2012) (same), and *United States v. John*, 597 F.3d 263 (5th Cir. 2010) (same), *with* *LVRC Holdings LLC v. Brekka*, 581 F.3d 1127 (9th Cir. 2009) (declining to follow *Citrin*), and *Nosal II*, 676 F.3d 854 (9th Cir. 2012) (en banc) (same).

39. *See* *United States v. Drew*, 259 F.R.D. 449, 452–53 (C.D. Cal. 2009) (setting aside jury's verdict finding the defendant guilty of a misdemeanor CFAA violation for "accessing a computer involved in interstate or foreign communications without authorization or in excess of authorization to obtain information," based on user's violation of the MySpace Terms of Service). *Drew* highlights the complexity and confusion that pervades questions of contracts in technology contexts. The District Court ultimately decided to grant defendant's motion to set aside the guilty verdict under the "void for vagueness" doctrine because the judge was troubled by the outcome:

Treating a violation of a website's terms of service, without more, to be sufficient to constitute [a CFAA violation] would result in transforming Section 1030(a)(2)(C) into an overwhelmingly overbroad enactment that would convert a multitude of otherwise innocent Internet users into misdemeanant criminals. . . . [And would afford] too much discretion to the police and too little notice to citizens who wish to use the Internet.

Id. at 466–68 (citing *City of Chi. v. Morales*, 527 U.S. 41, 64 (1999)). However, despite setting aside the verdict, the judge seemed to indicate that a mere breach of contract might indeed support a prosecution for computer intrusion. *See id.* at 452–53.

that contracts involving technology contexts are inherently special in breach may, on the other hand, be completely unwilling to adopt the same technology exceptionalist stance in contract formation questions.

As the discussion in Section II.A.1, *supra*, illustrated, case law on both formation and breach questions lacks a consistent theoretical paradigm to determine whether technology contexts merit a special analysis. The next section will highlight four traditional contract law principles that offer a starting point for determining whether technology exceptionalist analysis is warranted in a particular contract case: freedom to contract, freedom from contract, the primacy of damages,⁴⁰ and preserving private ordering.⁴¹ This paradigm of “restrained technology exceptionalism” can assist courts in crafting a minimally disruptive and more predictable approach to contract and technology questions.

B. THE PARADIGM OF “RESTRAINED TECHNOLOGY EXCEPTIONALISM”

As Section I.A, *supra*, explained, courts are increasingly adopting inconsistent—sometimes even internally inconsistent—approaches to analyzing the role of technology in contract law. However, as our society becomes more technology-reliant, progressively greater numbers of contracts will eventually include a technology component. This slippery slope eventually leads us into a world where substantially all contracts will implicate technology in some manner. As such, finding doctrinal coherence and a theoretically logical approach to questions of technology exceptionalism in contract law becomes essential. Without doctrinal consistency, we risk inverting the traditional relationship between contracts and technology.

Current doctrinal tensions in contract law implicating technology exceptionalism have weakened traditional contract analysis at a core level. This Section argues that these tensions are leading contract law dangerously astray, losing touch with foundational principles and crafting the aberrational and undesirable “law of the zebra” described in the Introduction. By identifying strategic contract behaviors in technology contexts, we can then map them onto traditional contract law first principles. In this way, we can

40. Primacy of damages refers to the idea that the primary remedy for a contract breach should be monetary compensation, regardless of which measure of contract damages is used. For a discussion of contract damages as the primary remedy for contract breach, see *Introductory Note*, RESTATEMENT (SECOND) OF CONTRACTS, 16 (1981); CORBIN ON CONTRACTS § 55.3 (Joseph M. Perillo ed., 2005).

41. “Private ordering” refers to the ability of individuals to enter into mutually consensual economic relationships. For a discussion of private ordering, see, e.g., Tehila Sagy, *What’s So Private About Private Ordering?*, 45 LAW & SOC’Y REV. 923 (2011).

arrive at an emergent yet minimally disruptive contract approach to technology.

Four foundational principles dictate that courts should adopt a type of “restrained technology exceptionalism” that privileges contract law analysis over exceptionalized approaches to technology. Viewing technology as “special” in contract law analysis is only appropriate when it furthers and buttresses the foundational principles of traditional contract law: freedom to contract, freedom from contract, damages primacy in breach, and private ordering. In other instances, technology exceptionalism in contract law is unjustified and must be avoided. Further, technology exceptionalist legal regimes should be used to supplement, not supplant, traditional contract analysis. Stated another way, the law of the horse may be necessary sometimes; the law of the zebra is never necessary. While new technology continues to meaningfully change our world, it does not necessarily need to dramatically change our contract doctrine; both overcorrecting and undercorrecting for technology in contract disputes will be equally destructive in the long term. Viewing technology as special is appropriate only when it furthers foundational principles of contract law and helps to maintain the doctrinal status quo in light of innovation.

This approach, which determines the appropriateness of technology exceptionalism in contract law by seeking to maintain the status quo of human relations crafted by traditional contract principles—and not by looking at a particular technology itself—might be termed a paradigm of “restrained technology exceptionalism.” Particular technologies should not be determinative of contract law; however, tweaks in contract law may be necessary when the changing behaviors of contracting parties and their jockeying for power with respect to each other (which may involve use of a particular technology) threaten traditional contract law principles. Thus, the concern is actually a human and relational one,⁴² not a technological one.⁴³

Four existing principals of contract law will serve to guide courts well in avoiding overzealous (or underzealous) technology exceptionalism. First, courts should seek to preserve freedom of contract; they should be conscious of new shifts in bargaining power and obstacles to meaningful consent that are byproducts of particular technology. Second, courts must preserve freedom from contract. They must remain vigilant in situations where

42. See Stewart Macaulay, *Relational Contracts Floating on a Sea of Custom? Thoughts about the Ideas of Ian MacNeil and Lisa Bernstein*, 94 NW. U. L. REV. 775, 782 (2000) (explaining that a relational approach avoids the oversimplification of contract law).

43. This distinction will be elaborated upon using the context of contract breach and computer intrusion in Part II, *infra*.

technology creates incentives and opportunities for one party to violate duties of good faith in performance and enforcement. This gaming creates novel types of contractual instability and threatens contractual liberty with respect to enforcement, exit, and redress. Further, courts should seek to preserve traditional contract remedies and the primacy of monetary damages as a remedy for contract harms. Finally, as a corollary fourth principle, courts should seek to preserve private ordering, supplementing but never supplanting contract's role in the innovation ecosystem.

1. *Principle 1—Freedom to Contract: Equalizing Bargaining Power and Making Consent Meaningful in Formation*

As I have demonstrated empirically in other work, contract drafting norms in technology contexts are a moving target.⁴⁴ They have shifted over time,⁴⁵ and contracts have become progressively more aggressive in the rights they reserve to the drafter.⁴⁶ Similarly, as I have argued elsewhere,⁴⁷ as technology evolves, new types of possible harms arrive, as well as new interfaces and devices for contract formation.⁴⁸ However, it is impractical to wait for contract law to fight out every incarnation of the technology contract battles with each particular product or interface. The law will never be able to anticipate the evolution of technology. Yet, simultaneously, technology evolves less efficiently when its creators cannot reasonably anticipate legal outcomes—a seeming conundrum.

However, hope is not lost: what law can anticipate reasonably successfully is strategic human behavior. In particular, we know from contract law scholarship that contract drafters will tend to act in their own self-interest, and they are likely to leverage new technologies and the contracts that accompany them to maximize their own benefit.⁴⁹ Second, we

44. See Andrea M. Matwyshyn, *Mutually Assured Protection: Toward Development of Relational Internet Data Security and Privacy Contracting Norms*, in SECURING PRIVACY IN THE INTERNET AGE 73–75 (Anupam Chander et al. eds., 2008).

45. *Id.* at 75.

46. *Id.*

47. See Andrea M. Matwyshyn, *Technoconsent(t)sus*, 85 WASH. U. L. REV. 529, 538–40 (2007) (explaining that the recent use of hacker coding tactics in digital rights management technology has led to increased vulnerability of computer systems worldwide).

48. See Andrea M. Matwyshyn, *Hidden Engines of Destruction: The Reasonable Expectation of Code Safety and the Duty to Warn in Digital Products*, 62 FLA. L. REV. 109, 125 (2009) (explaining that gaps in technology regulation have necessitated looking to other areas of law for guidance).

49. See Melvin A. Eisenberg, *Why There Is No Law of Relational Contracts*, 94 NW. U. L. REV. 805, 808 (2000). Eisenberg states:

[C]lassical contract law was based on a rational-actor model of psychology, under which actors who make decisions in the face of

know that when possible, drafters employ unnegotiated form contracts, which will be likely to disadvantage the party who lacks bargaining power.⁵⁰ Third, we know that aggressive contract enforcement in an adversarial system, if judicially unchecked, is likely to continue to push the limits of the law in favor of the better-resourced litigant in perpetuity.

A foundational principle that has been prominent in contract law relates to fairness in formation: equalizing bargaining power between differently sophisticated parties and ensuring that both parties' consent is meaningful.⁵¹ As I have argued elsewhere, significant concerns exist with current contract law in the context of formation of agreements involving technology.⁵² Two types of new strategic formation behaviors now exist, which take advantage of this doctrinal inadequacy. First, the drafter of the agreement tries to game the presentation of the agreement for competitive advantage over the other party.⁵³ Second, the drafter tries to game the content of the agreement to preserve maximum flexibility and leverage in enforcement, such as limiting the other party's ability to meaningfully consent.⁵⁴ In these scenarios, technology exceptionalist approaches to contract formation are warranted.

a) Gaming in Presentation and Content

The use of digital media for contract presentation introduces new opportunities for strategic gaming. In the technology contexts, the drafter gains additional leverage in contract formation. For example, digital contracts frequently incorporate various other documents by reference—if the user is on a mobile device, reviewing these incorporated documents simultaneously

uncertainty rationally maximize their subjective expected utility, with all future benefits and costs discounted to present value. In particular, the rules of classical contract law were implicitly based on the assumptions that actors are fully knowledgeable, know the law, and act rationally to further their economic self-interest.

Id.

50. See, e.g., Mark R. Patterson, *Standardization of Standard-Form Contracts: Competition and Contract Implications*, 52 WM. & MARY L. REV. 327, 331–33 (2010) (analyzing standard form contracts with a focus on the balance of bargaining power between the parties).

51. See 8 WILLISTON ON CONTRACTS § 18:9 (4th ed. 2012).

52. See, e.g., Matwyshyn, *Technoconsen(t)sus*, *supra* note 47, at 533, 550–55 (explaining that contract doctrine surrounding digital agreement formation focuses only on procedural unconscionability and ignores substantive unconscionability entirely); Matwyshyn, *supra* note 44, at 73–75 (explaining the progression of terms of use in a draconian direction and explaining formation challenges therein); Woodrow Hartzog, *Website Design as Contract*, 60 AM. U. L. REV. 1635, 1635–39 (2011) (explaining how website presentation can be used to obscure consumer ability to meaningfully understand end user license agreements).

53. Matwyshyn, *Technoconsen(t)sus*, *supra* note 47, at 551–52.

54. *Id.* at 554–56.

becomes functionally difficult if not impossible.⁵⁵ Many users may not even understand the concept of incorporation by reference or realize that they are in essence agreeing not only to the agreement that is visible to them on the screen before them, but also to a series of additional agreements that exist elsewhere on the Internet.

The hypothetical with Grandma and Dr. Whiskers demonstrates some of the common problems that users experience when attempting to read and understand contracts in technology contexts. Fonts used in these contracts tend to be very small on a screen, and the presentation of contracts can appear in different sizes on different screens.⁵⁶ Particularly in the context of mobile devices, the ability of an average user to comfortably read these agreements is dubious. Even when the agreement initially appears in a pop-up, links to the agreement may be hidden in obscure locations on the user interface, limiting the user's ability to navigate to and subsequently review terms after initial "consent."⁵⁷

Apart from gaming in the presentation of contracts, technology also allows for gaming with respect to content and meaningful contractual consent. Many contracts involving technology are either entirely unnegotiable, such as terms of use, or functionally unnegotiable, such as a workplace technology permissible use contract or employee handbook. Take-it-or-leave-it transactions have always been a concern of contract law.⁵⁸ However, in physical space, consumers are capable of crossing out and modifying provisions with which they do not agree, and these customizations generally control.⁵⁹ In digital spaces this ability to cross out and modify form contracts does not exist for the consumer.

55. See, e.g., *Terms of Service*, FACEBOOK (Dec. 11, 2012), www.facebook.com/legal/terms (incorporating other legal terms, such as its "Data Use Policy" and "Promotions Guidelines," into its primary "Terms of Use" agreement through hyperlinks, making it difficult to navigate to those terms on mobile devices).

56. See, e.g., *Bug 743799: Strange User Interface Variation*, BUGZILLA (Oct. 27, 2012, 11:02 AM), https://bugzilla.redhat.com/show_bug.cgi?id=743799 (explaining interface bug); Alex Heath, *One Inch Makes All The Difference: Why Apple Thinks The iPad mini's Display Is In A Whole Other League*, CULT OF MAC (Oct. 25, 2012), <http://www.cultofmac.com/198279/one-inch-makes-all-the-difference-why-apple-thinks-the-ipad-minis-display-is-in-a-whole-other-league/>.

57. See, e.g., SQUARE, <http://www.square.com/> (rollover required to make legal terms link plainly visible).

58. See Yuval Feldman & Doron Teichman, *Are All Contractual Obligations Created Equal*, 100 GEO. L.J. 5, 8, 18, 25, 44 (2011).

59. See, e.g., Gail Hillebrand, *The Uniform Commercial Code Drafting Process: Will Articles 2, 2B, and 9 Be Fair to Consumers?*, 75 WASH. U. L.Q. 69, 78 (1997).

Further, substantively, these technology contracts may be written using esoteric specialized technology or legal terminology unfamiliar to most consumers.⁶⁰ These contracts also may be drafted in an intentionally ambiguous manner in order to reserve maximum flexibility for the drafter; because these agreements are frequently unnegotiated and unnegotiable, the terms almost always remain in their original ambiguous form, governing the relationship in a one-sided manner for the benefit of the drafter.⁶¹

Finally, many form technology agreements allow for the drafter—the website operator or application author—to amend the agreements in his sole discretion. The user may not even be aware that terms have changed. When these dynamics are coupled with the substantial length of the average user agreements—length that has been increasing over time⁶²—and their poor readability, concerns about unfair surprise and oppression exist.⁶³

When one party capitalizes on known deficiencies of information for strategic benefit, courts in physical space contexts often view this advantage negatively, but these norms have not yet transferred into contract caselaw involving technology contexts. For example, where one party is known to lack the capacity required to understand the agreement, courts have historically set aside these agreements in consumer contexts.⁶⁴ When the drafter leverages, capitalizes on, or exacerbates the user's deficiencies of knowledge—such as technology knowledge—courts should be equally

60. See, e.g., Florencia Marotta-Wurgler, *Will Increased Disclosure Help? Evaluating the Recommendations of the ALI's "Principles of the Law of Software Contracts,"* 78 U. CHI. L. REV. 165, 166 (2011).

61. The contract interpretation principle of *contra proferentem* is driven by the premise that an ambiguity should be construed against the drafter to avoid the possibility of gaming by drafters. See 11 WILLISTON ON CONTRACTS § 32:12 (4th ed. 2012). For a discussion of strategic vagueness and ambiguity, see B. Douglas Bernheim & Michael D. Whinston, *Incomplete Contracts and Strategic Ambiguity*, 88 AM. ECON. REV. 902 (1998).

62. See Marotta-Wurgler, *supra* note 60, at 168, 177–78, 181. Compare, e.g., Legal Information & Notices, iTUNES (May 13, 2004), available at <http://web.archive.org/web/20040627044803/http://www.apple.com/legal/default.html>, with Legal Information & Notices, iTUNES (Nov. 20, 2009), <http://www.apple.com/legal/terms/site.html> (last visited Feb. 26, 2013) (demonstrating changes in the “Accounts, Passwords and Security” section which emphasizes the burden on the consumer in addition to changes in the “Governing Law; Dispute Resolution” section. Also, the following sections were deleted: “Trademark Information”; “Copyright Information”; “Rights and Permissions”; “Piracy Prevention”; “Software Piracy”; “Software Asset Management”; “Internet Piracy”; “Apple’s Unsolicited Idea Submission Policy”; “Terms of Idea Submission”; “Product Feedback”; “Software and Documentation Information”; and “Legal Contracts.”).

63. This language intentionally mirrors the language of unconscionability, a traditional concern of contract law. See 8 WILLISTON ON CONTRACTS § 18:9 (4th ed. 2012).

64. See Wendy Chung Rossiter, *No Protection for the Elderly: The Inadequacy of the Capacity Doctrine in Avoiding Unfair Contracts Involving Seniors*, 78 OR. L. REV. 807, 807–09 (1999).

vigilant. Yet, it is quixotic to expect that drafters will act in a manner contrary to their self-interest when the digital space grants them an inherent advantage. Indeed, one may argue that such an expectation would perhaps even be inconsistent with current contract law, which lacks a duty of good faith in negotiation in the United States.⁶⁵ This deficit of a duty to negotiate in good faith makes it even more necessary to compensate for new procedural disadvantages one bargaining party faces due to technology.

b) Technology Exceptionalism in Formation Is Appropriate

The formation obstacles described above reflect drafters' use of the digital medium to gain a strategic advantage. Although some of the concerns noted above already existed in the context of standard form contracts, for example, technology alters and exacerbates them to a meaningful degree, almost always to the detriment of the weaker party. As such, a technology exceptionalist approach to formation benefits the weaker party: it rebalances the interests of the parties toward a more level playing field.

I have argued elsewhere that the concerns with respect to unfair surprise and oppression in technology contracting are so severe that they warrant a new construction of meaningful consent hinged on actual understandings of real users that evolve across time,⁶⁶ and that state contract law should expressly incorporate a series of implied promises to compensate for the lack of a negotiability of technology contracts.⁶⁷ In the sections that follow, this Article argues the opposite with respect to contract breach. The strategic contract gaming behaviors in breach contexts reflect a different dynamic than the gaming behaviors in formation contexts. Whereas a technology exceptionalist analysis in formation is appropriate to protect the weaker party and prevent unfair surprise and oppression,⁶⁸ in breach it is not. Applying an exceptionalist analysis to breach exacerbates rather than remedies existing power imbalances between the parties; exceptionalism in breach harms the weaker party and additionally skews the playing field in favor of the more powerful party—usually the drafter. Staying true to traditional contract principles in technology breach contexts instead requires treating technology breach and other contractual breach as equivalent.

65. See, e.g., John Klein & Carla Bachechi, *Precontractual Liability and the Duty of Good Faith Negotiation in International Transactions*, 17 HOUS. J. INT'L L. 1, 16 (1994) (discussing the absence of a duty of good faith in negotiations in the United States).

66. Matwyshyn, *Technoconsen(t)sus*, *supra* note 47, at 532, 560–62.

67. Andrea M. Matwyshyn, *Privacy, the Hacker Way*, 86 S. CAL. L. REV. (forthcoming 2013).

68. Matwyshyn, *Technoconsen(t)sus*, *supra* note 47, at 533–34, 559.

2. *Principle 2—Freedom from Contract in Breach: Reasonable Enforcement in Line with Contractual Liberty and Rights of Exit*

The corollary principle in contract law to the freedom *to* contract has always been freedom *from* contract. In other words, while the law should support parties' liberty to enter into deals—even deals with unwise business terms⁶⁹—the law should also support parties' liberty to exit contractual relationships in a reasonable manner.

a) Duties of Good Faith in Performance and Enforcement

A duty of good faith in performance and enforcement exists under U.S. law.⁷⁰ However, the ability to perform this duty assumes that a party is capable of understanding her contractual obligations. Currently in many contracting situations that involve technology, neither side is clear on the extent of their obligations to the other. In other words, in light of the formation challenges presented above, breach becomes significantly more likely: parties to a contract who do not understand their obligations may be more likely to fail to perform or to perform obligations incorrectly. Similarly, parties who do not understand the obligations of the other side of the contract are less likely to correctly understand when breach by the other side has occurred. Because of the unnegotiated nature of many technology contracts, drafters may insert intentionally vague language into the agreement

69. So long as some form of consideration exists, courts do not evaluate the adequacy of consideration when ruling on contract enforceability. *See, e.g.*, *In re Xonics Photochemical, Inc.*, 841 F.2d 198, 201–02 (7th Cir. 1988) (holding that a now-bankrupt company's guarantee of an affiliate's \$15–20 million loan when its own net assets totaled \$1.7 million was enforceable because the company voluntarily assented to the loan guarantee and derived some indirect benefit from the agreement, even if minimal); *Hoffa v. Fitzsimmons*, 673 F.2d 1345, 1359–60 (D.C. Cir. 1982) (holding that a union pension agreement was not unenforceable just because the intended consideration had already been conveyed prior to execution of the contract, since the agreement conferred other ancillary legal and practical benefits upon the beneficiary); *Cleveland-Cliffs Iron Co. v. Chi. & N. W. Transp. Co.*, 581 F. Supp. 1144, 1150 (W.D. Mich. 1984) (noting that unless the consideration is “so grossly inadequate as to shock the conscience, the general rule is that courts will not inquire into the adequacy of the consideration of a contract”) (internal citation and quotation omitted); *In re Estate of Duncan v. Kinsolving*, 70 P.3d 1260, 1265 (N.M. 2003) (“Absent a showing of fraud, inadequacy of consideration is not sufficient to void a contract.”) (internal quotation omitted); *Horace Mann Ins. Co. v. Gov't Emps. Ins. Co.*, 344 S.E.2d 906, 908 (Va. 1986) (“[P]arties to a contract are at liberty to determine their own valuations, and courts generally will not inquire into the adequacy of consideration.”); *Buckingham v. Wray*, 366 N.W.2d 753, 756 (Neb. 1985); *Osborne v. Locke Steel Chain Co.*, 218 A.2d 526, 530 (Conn. 1966) (“The doctrine of consideration does not require or imply an equal exchange between the contracting parties. . . . The courts do not unmake bargains unwisely made.”).

70. *See* Richard E. Speidel, *The “Duty” of Good Faith in Contract Performance and Enforcement*, 46 J. LEGAL EDUC. 537, 539 (1996).

with a goal of preserving more flexibility in enforcement, as explained above. Meanwhile, although ambiguities have traditionally been construed against the drafter in physical space contexts,⁷¹ this norm does not appear to have transferred itself into digital spaces as yet.

These types of contractual comprehension deficits harm the market stability and commercial trust that duties of good faith intended to foster.⁷² As the hypothetical with Grandma and Dr. Whiskers highlights, an average consumer simply may not understand how to conform performance to her duties of good faith. Yet a mere breach of contract in technology contexts now brings a risk of criminal prosecution and uncertain civil consequences—a misguided technology exceptionalism which will be discussed at length in Part III, *infra*.

b) Ability to Exit and “Digital Peonage” Concerns

It has long been a hallmark of contract law that parties can exit their contractual relationships, even if that exit is accompanied by a breach and damages. Further, courts are loathe to require even specific performance of contractual obligations and do so extremely rarely in the context of service agreements.⁷³ To do otherwise would potentially be a form of forced labor under penalty of law or “peonage.”⁷⁴ In technology contexts, we might label this concern as a concern over “digital peonage.”⁷⁵

71. 11 WILLISTON ON CONTRACTS § 32:12 (4th ed. 2012).

72. For a discussion of the role of contractual misunderstandings and relational trust in commercial exchange, see, e.g., Ian R. Macneil, *Relational Contract Theory: Challenges and Queries*, 94 NW. U. L. REV. 877 (2000).

73. See, e.g., RESTATEMENT (SECOND) OF CONTRACTS § 359(1) (2012) (noting that specific performance is appropriate only when money damages would be inadequate “to protect the expectation interest of the injured party”); Nathan B. Oman, *Specific Performance and the Thirteenth Amendment*, 93 MINN. L. REV. 2020, 2022–23 (2009) (recalling how courts and scholars frequently deny or argue against specific performance as a remedy for breach of service agreements on the grounds that doing so would violate the Thirteenth Amendment, which prohibits involuntary servitude).

74. Peonage, a system that allowed “the holding of any person to service or labor,” was abolished by the Anti-Peonage Act of 1867:

[A]ll acts, laws, resolutions, orders, regulations, or usages of any Territory or State, which have heretofore established, maintained, or enforced, or by virtue of which any attempt shall hereafter be made to establish, maintain, or enforce, directly or indirectly, the voluntary or involuntary service or labor of any persons as peons, in liquidation of any debt or obligation, or otherwise, are declared null and void.

42 U.S.C. § 1994 (2011).

75. As the Idaho Supreme Court explained in the context of physical space:

The Supreme Court has reminded us that policy concerns are implicated whenever a contract limits the right to exit a contractual relationship.⁷⁶ In particular, in the Peonage Cases, the U.S. Supreme Court invalidated laws that criminalized breach of employment contracts.⁷⁷ As the Peonage Cases attest, threatened criminal prosecution for breach of contract, in particular, chills exit: employees would be worried to leave their employment for fear of losing their liberty in the process. Under the English master and servant acts, a laborer could be criminally punished for breach of contract, but criminal punishment of this kind did not exist in the United States, said the Court.⁷⁸

These concerns with respect to freedom to exit contractual relationships also pervade current law surrounding noncompetition and nonsolicitation contracts, for example.⁷⁹ Courts and legislatures have placed express restrictions on the ability of employers to limit employees' exit and future work contracts.⁸⁰ The reason for this robust law around noncompetition and nonsolicitation contracts involves, first, an acknowledgement of the power imbalance inherent in an employer and employee relationship, and, second,

Life in the competitive commercial world has at least equal capacity to bestow ruin as benefit, and it is presumed that those who enter this world do so willingly, accepting the risk of encountering the former as part of the cost of achieving the latter. Absent clear evidence to the contrary we will not presume that the parties to a contract such as the one before us meant to insure each other's emotional tranquility.

Hatfield v. Max Rouse & Sons Nw., 606 P.2d 944, 952 (Idaho 1980). These same dynamics are at issue in the context of digital peonage inquiries—employers seek to obtain emotional tranquility by sanctioning employees unexpectedly heavily for exit.

76. See *Pollock v. Williams*, 322 U.S. 4, 18 (1944). In *Pollock*, the U.S. Supreme Court stated:

When the master can compel and the laborer cannot escape the obligation to go on [working], there is no power below to redress and no incentive above to relieve a harsh overlordship or unwholesome conditions of work. . . . Whatever of social value there may be, and of course it is great, in enforcing contracts and collection of debts, Congress has put it beyond debate that no indebtedness warrants a suspension of the right to be free from compulsory service.

Id. at 18.

77. *Id.* at 25.

78. *Id.* at 18.

79. See Ken Matheny & Marion Crain, *Disloyal Workers and the "Un-American" Labor Law*, 82 N.C. L. REV. 1705, 1746 (2004) ("[M]any courts remain troubled by the fact that noncompetes contravene a basic precept of capitalism—free competition in the market.").

80. See CAL. BUS. & PROF. CODE § 16600 (West 1987 & Supp. 2003) (rendering noncompetition agreements void).

the undesirability as a matter of social policy in allowing private parties to restrict mobility of talent.⁸¹

However, precisely these types of peonage concerns arise in case law involving technology breach: “digital peonage.” When technology questions emerge in breach analysis, a second set of concerns arise: technology exceptionalism in breach opens the door to unpredictable new forms of draconian punishment for exit—even criminal penalties—that violate traditional contract principles. The ambiguities in technology contracting contexts, particularly the threat of possible criminal prosecution described in Part III, *infra*, may cause parties to view themselves as unable to exit a contractual relationship.

As such, this situation is distinctly different from the dynamics of the desirable technology exceptionalism in questions of formation. Exceptionalism in breach results in granting additional rights to the drafter, rights that may not even be expressly articulated in the contract and perhaps not reasonably foreseeable to the disempowered party.

3. *Principle 3—Damages: Preserving Primacy of Traditional Contract Remedies*

As the previous Section explained, exceptionalizing technology breach skews the dynamics of contract performance and enforcement. Logically, these disruptions also have ripple effects in the context of damages: technology exceptionalism in breach also disrupts the traditional contract principle of monetary damages as the primary recourse for contract harms. Arguing in favor of technology exceptionalism in breach, courts have been willing to provide plaintiffs a method to leverage already existing regimes of punishment from outside contract law, in lieu of preserving the primacy of damages.

a) *Damages Primacy*

The primary goal of contract damages, generally speaking, is compensating harmed parties for actual economic losses, not to punish the breacher for a “bad” act of breach.⁸² For example, courts generally do not

81. Christina L. Wu, Comment, *Noncompete Agreements in California: Should California Courts Uphold Choice of Law Provisions Specifying Another State’s Law?*, 51 UCLA L. REV. 593, 608–10 (2003).

82. RESTATEMENT (SECOND) OF CONTRACTS, § 355 (1981) (“Punitive damages are not recoverable for a breach of contract unless the conduct constituting the breach is also a tort for which punitive damages are recoverable.”).

enforce punitive damages clauses⁸³ and, as a general rule, we do not criminalize breaches of contract.⁸⁴ As Farnsworth has stated, a “stipulated sum [that] is significantly larger than the amount required to compensate the injured party for its loss . . . would allow the parties to depart from the fundamental principle that the law’s goal on breach of contract is not to deter breach by compelling the promisor to perform, but rather to redress breach by compensating the promisee [sic]” and therefore is prohibited “when a court characterizes such a provision as a penalty.”⁸⁵ Attaching a draconian contract penalty such as specific performance, gratuitously high damages, or incarceration to a technology exceptionalist analysis runs opposite to this guiding contract principle.

An exceptionalized technology breach regime becomes disconnected from the goal of providing the benefit of the bargain to the wronged party; it instead focuses on whether breach implicates technology and on punishing the breacher. Even adopting a more morality-driven stance, such as that of the *Third Restatement of Restitution*, a traditional damages approach focuses on actual financial harms, not on the deterrence of future breach, retribution, incapacitation, or rehabilitation of breachers.⁸⁶ Though criminal law may adopt these goals, contract law generally does not. Even in the context of the most aggressive traditional damages calculation, expectation damages,⁸⁷ Professors Fuller and Perdue argue in favor of protecting reliance and facilitating business contracts in a system where the economy and the legal institutions are intertwined.⁸⁸ Thus, an exceptionalized breach analysis that allows remedies other than damages or, in the most severe situations, an

83. See Charles R. Calleros, *Punitive Damages, Liquidated Damages, and Clauses Penale in Contract Actions: A Comparative Analysis of the American Common Law and the French Civil Code*, 32 BROOK. J. INT’L L. 67, 69 (2006).

84. See *supra* notes 76–82 and accompanying text. Considering this general rule against criminalization of breaches of contract, the approach currently used by a number of courts in interpreting the CFAA, which purports to do precisely that, is all the more problematic. This problem will be discussed in detail in Part III, *infra*.

85. E. ALLAN FARNSWORTH, FARNSWORTH ON CONTRACTS § 12.18 (3d ed. 2004).

86. RESTATEMENT (THIRD) OF RESTITUTION & UNJUST ENRICHMENT § 1 et seq. (2011).

87. For a discussion of the definition of expectation damages and, in particular, their relationship to contractual misunderstandings, see Daniel P. O’Gorman, *Expectations Damages, the Objective Theory of Contracts, and the “Hairy Hand” Case: A Proposed Modification to the Effect of Two Classical Contract Law Axioms in Cases Involving Contractual Misunderstandings*, 99 KY. L.J. 327 (2011).

88. L.L. Fuller & William R. Perdue, Jr., *The Reliance Interest in Contract Damages: 2*, 46 YALE L.J. 373 (1937).

injunction, unnecessarily violates traditional contract principles.⁸⁹ It exceeds even the most aggressive traditional contract remedies without a compelling justification.

b) The Unique Concerns of Information Harms

But, playing devil's advocate, perhaps technology-connected harms are so severe that current calculations of damages are incapable of encompassing the extent of information harms. Perhaps a more aggressive set of remedies is needed for information harms.⁹⁰

Although this argument may hold superficial appeal, scholars have posited that even extraordinary situations of information breach can be addressed successfully using a damages regime modeled from traditional contract principles of damages. As Melvin Eisenberg has convincingly argued, disgorgement damages in particular offer a potent remedy for information-based contract harms.⁹¹ Disgorgement is also appropriate where losses would be very difficult to establish with sufficient certainty for an expectation damages calculation. Eisenberg asserts that a disgorgement calculation has worked well in a line of cases concerning damages in breaches of noncompete agreements or agreements to give the plaintiff exclusive territorial rights.⁹² Courts have successfully awarded damages based on the defendant's profits, explicitly protected the disgorgement interest, or used a disgorgement measure as a surrogate for the expectation measure.⁹³ Theoretically, this calculation protects the integrity of contract law, argues Eisenberg: "unless disgorgement is awarded in such cases, a promisor could subvert the right to specific performance simply by completing an irreversible breach before the promisee can get to court."⁹⁴ In other words, disgorgement offers a damages construction that approximates specific

89. A second contract theory debate that becomes permanently altered in an exceptionalized technology contract breach environment relates to damages and the duty by the nonbreaching party to mitigate losses. For a discussion of the duty to mitigate damages in contract, see, e.g., 25 C.J.S. *Damages* § 46 (2013).

90. For example, although specific performance is rarely granted, this remedy does exist for some types of extreme contract harm situations. See FARNSWORTH, *CONTRACTS* § 12.4 (3d ed. 1999) ("[A]lthough the injured party can always claim damages for breach of contract, that party's right to specific relief as an alternative is much more limited.").

91. This calculation of damages is appropriate in cases in which the nonbreaching party would have been awarded specific performance if the nonbreaching party had been able to bring suit before the breacher's wrongful action. See Melvin A. Eisenberg, *The Disgorgement Interest in Contract Law*, 105 MICH. L. REV. 559, 584 (2006).

92. *Id.* at 588.

93. *Id.* at 578–81.

94. *Id.* at 584.

performance better than an expectation measure in the case where, for example, information is stolen or a system is used to harm others.

For example, in *Snepp v. United States*, a former CIA employee breached an employment agreement requiring that he “not . . . publish . . . any information or material relating to the Agency, its activities or intelligence activities generally, either during or after the term of [his] employment . . . without specific prior approval by the Agency.”⁹⁵ Snepp had left the CIA and published a book about CIA activities in South Vietnam that happened during the term of his employment, failing to submit the material for approval.⁹⁶ The Government sued Snepp for disgorgement of his profits from the book, and the Supreme Court agreed with the argument that disgorgement was warranted.⁹⁷ The Court applied a disgorgement-based calculation of damages, stating that despite the Government being harmed by Snepp’s book, proving the extent of that harm would be difficult, and the Government could not pursue other remedies “without losing the benefit of the bargain it seeks to enforce.”⁹⁸ Ergo, argues Eisenberg, a disgorgement measure of damages made the government whole for this commercial harm.⁹⁹ The same rationale should apply to any rogue insider situation where harm, including digitally-caused harm, occurs.

4. Principle 4—*Supplementing But Not Supplanting Contract*

To date, contract law has provided the dominant structures for organization of technology contexts; the reason for this supremacy rests in contract law’s flexibility. It offers nimble structures that allow for customization of obligations and self-help, as well as default shared norms of permissible conduct.¹⁰⁰ In this way, contract simultaneously fosters innovation and entrepreneurship without letting entrepreneurs cause unfettered damage to others. However, contract law is by no means the only law that impacts technology use and innovation. As the reach of existing technology exceptionalist statutes expand, and Congress drafts a growing number of statutes, technology exceptionalism threatens to unbalance the relationship between contract law and technology law. Although technology exceptionalist statutes address real harms in Congress’s estimation, such

95. *Snepp v. United States*, 444 U.S. 507, 508 (1980).

96. *Id.* at 507.

97. *Id.* at 508, 516.

98. *Id.* at 514.

99. Eisenberg, *supra* note 91, at 589.

100. See I. Trotter Hardy, *The Proper Legal Regime for “Cyberspace,”* 55 U. PITT. L. REV. 993, 1019–25 (1994) (advocating self-help, custom, and contract to regulate cyberspace).

statutes should buttress rather than override the fundamental principles of contract law.

Contract law is an evolutionary framework. For example, when emergent circumstances warrant shifts in consumer protection posture, tweaks to contract law frequently offer a more successful and less disruptive solution than a new technology exceptionalist statutory regime.¹⁰¹ As such, new technology exceptionalist approaches should begin where the structure afforded by contract law ends, or they should work in tandem with contract law without usurping its role. Future innovation and user adoption will be better served through the preservation of private ordering and contract remedies than through aggressive technology-exceptionalist legislative approaches.

As Part III, *infra*, will demonstrate through the example of computer intrusion and contract breach, when shifting contract drafting and enforcement norms are coupled with a failure to maintain contractual supremacy over technology-exceptionalist legal regimes, the negative consequences become amplified. In the context of the circuit split on contract breach/CFAA analysis, the failure to maintain contractual supremacy “weaponizes” contract breach. By weaponizing breach, courts inappropriately repurpose contract law from its traditional role as a means for private ordering. Contract law is instead corrupted into a mere springboard to leverage potentially draconian remedies available under other legal regimes—particularly intellectual property law and technology law. The next Part elaborates on this undesirable dynamic of weaponized breach and how CFAA analysis threatens to usurp the traditional role of contract remedies for technology-related contract breach.

III. HORSES AND CLAPPING COCONUTS: THE PROBLEMATIC CASE OF WEAPONIZED BREACH AND COMPUTER INTRUSION UNDER THE CFAA

“King Arthur: We have ridden the length and breadth of the land in search of knights who will join me in my court at Camelot. I must speak with your lord and master.

1st soldier: What? Ridden on a horse?

King Arthur: Yes!

1st soldier: You’re using coconuts!

101. See Matwyshyn, *Technoconsent(t)us*, *supra* note 47, at 557–59 (discussing the benefits of tweaking contract law to address consumer consent to security-invasive technologies).

King Arthur: What?

1st soldier: You've got two empty halves of coconut and you're bangin' 'em together."

– *Monty Python and the Holy Grail*¹⁰²

In an iconic scene from the movie *Monty Python and The Holy Grail*, a soldier with a keen interest in birds hears the sound of hoofbeats, only to realize that instead of a horse arriving, it is merely a person clapping coconuts together.¹⁰³ Breach in technology contexts is the contract law equivalent of a person clapping coconuts—from the noise, a hearer may believe the sound to arise from a horse or a zebra, but the source is actually neither. The previous section introduced the argument that technology exceptionalism was overwhelming traditional legal paradigms in contract, crafting an unnecessarily exotic legal regime that derails traditional contract law—a “law of the zebra.” A “law of the zebra” approach for contract breach is unwarranted, and even a “law of the horse” exceptionalized approach is unnecessary.

This Part elaborates on the importance of adopting a paradigm of restrained technology exceptionalism for contract breach analysis through an example of “the law of the zebra”: the circuit split on the relationship of breach and computer intrusion under the Computer Fraud and Abuse Act (“CFAA”).¹⁰⁴ As the previous sections explained, contract breaches have traditionally—very intentionally—not been viewed as crimes. Yet this particular legal query between the relationship of contract breach and

102. *Memorable quotes for Monty Python and the Holy Grail*, IMDB, <http://www.imdb.com/title/tt0071853/quotes> (last visited Jan. 16, 2013).

103. MONTY PYTHON AND THE HOLY GRAIL (Python (Monty) Pictures Ltd. 1974).

104. The CFAA is the primary computer intrusion or “hacking” statute in the United States. It hinges on the idea of whether “authorized access” has occurred or whether the access in question is unauthorized or exceeds the scope of authorization. It prohibits obtaining:

- (A) information contained in a financial record of a financial institution, or of a card issuer . . . or contained in a file of a consumer reporting agency on a consumer . . . ;
- (B) information from any department or agency of the United States; or
- (C) information from any protected computer

18 U.S.C. § 1030(a)(2) (2011). The CFAA defines the term “financial institution” to include a range of financial institutions, such as banks, credit unions, and broker-dealers. *Id.* at § 1030(e)(4). The CFAA is also violated when a person uses or sells passwords to access machines, uses a computer with an intent to extort money or anything of value, or transmits communications threatening damage to a protected computer though interstate or foreign commerce. *Id.* § 1030(a)(6), (a)(7).

computer intrusion has split the Seventh and Ninth Circuits in the context of CFAA analysis.¹⁰⁵ Although contract breaches involving technology may on their surface appear to warrant a technology exceptionalist analysis, the technology aspects of the breach are ultimately irrelevant—the mere use of a computer does not alter the capacity of contract law to redress the harm.¹⁰⁶

This Part introduces four very different types of “contract hackers” impacted by this weaponized breach dynamic, and argues that weaponizing breach is highly undesirable as a matter of contract doctrine, theory, private ordering and innovation, and entrepreneurship policy. Applying paradigms from developmental and social psychology, this Part then postulates that courts’ misguided tendency to weaponize breach in technology-related cases is likely rooted in essentialism and confirmation bias rather than traditional contract law analysis.

A. WEAPONIZING BREACH: THE CFAA AND A TALE OF FOUR
“CONTRACT HACKERS”

“[E]ven without [a] voluntary release there are perhaps no contracts or engagements . . . of which one can venture to say that there ought to be no liberty whatever of retraction.”

– John Stuart Mill¹⁰⁷

This Section asks a superficially simple breach question: should a breach of contract relating to a computer or network automatically provide the basis for a criminal charge of computer intrusion under the CFAA?¹⁰⁸ The CFAA has a checkered past, with several amendments that significantly expanded

105. *See infra* notes 134–43.

106. Similarly, detection of a breach may be even more likely because of the extent of digital monitoring used in an average workplace.

107. JOHN STUART MILL, ON LIBERTY 199 (Boston: Ticknor & Fields 1863).

108. Proponents of a technology exceptionalist analysis would argue, among other things, that information loss has become a more pressing problem than in the past, that stealing information is easier and more prevalent because of technology interconnection, and that lost information is harder to detect. While each of these points may be descriptively correct, none of them explains why current contract law is inadequate to address technology-related breach cases. Similarly, none of these arguments provide a justification for disrupting traditional contract law around breach as the most effective means of addressing these challenges. Further, advocates of a “law of the zebra” fail to demonstrate the presence of new categories of harm outside of those that contract law already contemplates. Contract law around breach and damages can already effectively address the harms arising from breaches involving technology. While we may need a legislative correction to buttress—but not override—traditional concerns in contract formation in technology contexts, traditional contract law around breach should not be mutated into a construct violative of its own foundational principles. Criminalizing breaches of contract would do precisely that.

the scope of the Act.¹⁰⁹ Intended primarily as a criminal statute to address third party “hackers,”¹¹⁰ it was passed originally in 1984, perhaps partially in response to a cyberwar-themed movie.¹¹¹ The statute has been amended on multiple occasions since its passage and has been expanded to include civil matters.¹¹² Despite numerous amendments to the statutory framework created by the CFAA, it continues to be beset with problems. Indeed, two circuit splits currently exist with respect to its interpretation.¹¹³ The first split is outside the scope of this Article.¹¹⁴ The second split, which is partially the subject of this article, relates essentially to the meaning of “authorized access” under the statute and the relationship of the statute to contract law.

109. For a discussion of the history behind unauthorized access statutes, including the CFAA, see Orin S. Kerr, *Vagueness Challenges to the Computer Fraud and Abuse Act*, 94 MINN. L. REV. 1561, 1563–71 (2010).

110. *See id.* (providing a history of the CFAA).

111. The movie *WarGames* was released in 1983, before the passage of the CFAA in 1986. *WAR GAMES* (United Artists 1983) (depicting a teenage hacker accidentally compromising a weapons system at the Pentagon, believing it to be a computer game); *see also In re America Online*, 168 F. Supp. 2d 1359, 1374 (S.D. Fla. 2001) (discussing the legislative history, noting that the CFAA has expanded beyond federal and financial systems). The *America Online* court quoted the Senate Report:

As computers continue to proliferate in business and homes, and new forms of computer crimes emerge, Congress must remain vigilant to ensure that the Computer Fraud and Abuse statute is up-to-date and provides law enforcement with the necessary framework to fight computer crime.

Id. (quoting S. Report, 104-357, at 5 (1996)).

112. An interesting parallel in the expansion to the CFAA for civil matters can be found in another troubled statute, RICO. Like the CFAA, RICO is simultaneously criminal and civil and has also caused circuit splits. Some of the dynamics that this section highlights in the context of contract and CFAA enforcement have predecessors in the RICO caselaw, which may hold lessons for the future of the CFAA. *See, e.g.,* Randy D. Gordon, *Clarity and Confusion: RICO's Recent Trips to the United States Supreme Court*, 85 TUL. L. REV. 677 (2011) (discussing Racketeering Influenced and Corrupt Organizations Act (“RICO”) and the challenges in application raised by its dual civil and criminal nature).

113. *Id.*

114. The first split relates to whether a plaintiff must plead both “damage” and “loss” for a recovery under the CFAA. The First Circuit has held that the CFAA provides for recovery of economic “loss” even in the absence of any physical “damage” to the computer system or its information. *EF Cultural Travel BV v. Explorica, Inc.*, 274 F.3d 577, 585 (“As we move into increasingly electronic world, the instances of physical damage will likely be fewer while the value to the victim or what has been stolen and the victim’s costs in shoring up its security features undoubtedly will loom ever-larger.”). However, in *Garelli Wong & Assocs. v. Nichols*, a district court in the Seventh Circuit read the statute to hold that a plaintiff under the CFAA must plead both “damage” and “loss.” 551 F. Supp. 2d 704, 708 (N.D. Ill. 2008). This split is outside the scope of this Article.

The CFAA creates civil and criminal penalties for unauthorized access or access that exceeds authorization to computer data.¹¹⁵ In general, two categories of alleged intruders in systems are possible—first, the group of alleged intruders with no preexisting contractual relationship to the network or computer owner,¹¹⁶ and second, the group of alleged intruders where a preexisting relationship exists.¹¹⁷ In the first situation, where the alleged intruder has no contractual relationship to the owner of the system, CFAA analysis is somewhat straightforward: no authorization exists, and determining liability becomes primarily a fact-driven question as to whether the alleged intruder has indeed violated the boundaries of the network or accessed the machine in question. These unaffiliated intruders include the malicious third parties who attack networks, and they are those “hackers” the CFAA originally sought to address.¹¹⁸

The second group of alleged computer intruders might be termed “contract hackers”—those alleged intruders who have a preexisting contractual relationship to the system owner. A “contract hacker” analysis under the CFAA inevitably starts with some variant of a traditional contract law inquiry—for example, an employment or nondisclosure or noncompetition agreement, an end user license agreement, or terms of use.¹¹⁹ The Parts that follow argue that if courts engage in a technology exceptionalist analysis of contract breach in CFAA cases, they inappropriately “weaponize” breach and cause permanent damage to the future contract law as a whole.

Contract breach weaponized with penalties under the CFAA overreaches not only by applying criminal liability for the defendants, but also by

115. 18 U.S.C. § 1030 (2011).

116. David M. Hafele, *Three Different Shades of Ethical Hacking*, SANS INSTITUTE (Feb. 23, 2004), http://www.sans.org/reading_room/whitepapers/hackers/shades-ethical-hacking-black-white-gray_1390.

117. Often this preexisting relationship takes the form of an employment relationship accompanied by a confidentiality agreement or permitted computer use policy. See Colette Thomason, Case Summary, *United States v. Nosal: Separating Violations of Employers' Computer-Use Policies from Criminal Computer Hacking Invasions*, 43 GOLDEN GATE U. L. REV. 163, 164 (2013).

118. These “hackers” include the malicious third parties who attack networks and whom are referred to colloquially in the information security community as “black hats.” Though the dynamics of this CFAA case law with respect to unaffiliated third-party “hackers” should also be revisited and clarified, these inquiries are outside the scope of this Article. The “hacker” collective Anonymous in particular has triggered a need to craft lines between computer intrusion and permissible methods of digital protest. See, e.g., Sean Captain, *The Real Role of Anonymous at Occupy Wall Street*, FAST COMPANY, (Oct. 17, 2011), <http://www.fastcompany.com/1788397/the-real-role-of-anonymous-at-occupy-wall-street>.

119. See *Nosal II*, 676 F.3d 854, 860–62 (9th Cir. 2012) (en banc).

imposing duplicative financial remedies in civil claims, and thereby threatening to derail contract law. Cases involving four different types of “contract hackers” illustrate the dangers of contract breach weaponized with the CFAA, including (1) consumer users of technology, (2) “disloyal” employees or business partners, (3) entrepreneurs, and (4) security researchers. These cases demonstrate a pressing need to resolve the tension between contract breach and computer intrusion law.

1. *Consumer Users of Technology as “Contract Hackers”: Confusing Minor Breach with Black Hat Hacking*

We again return to the Grandma and Dr. Whiskers hypothetical. Does a consumer user who breaches an end user license agreement or terms of use agreement on a website lose authorization and, therefore, create the basis for a charge of criminal intrusion under the CFAA? At least one jury has problematically concluded that the answer is yes.

In *United States v. Drew*, an adult created a fictitious profile on the social networking website Myspace and used this account to communicate with a teenager, who committed suicide after one such communication.¹²⁰ In doing so, the jury found, the defendant completed a CAPTCHA and consented to (and then violated) Myspace’s Terms of Service.¹²¹ The jury convicted the defendant of a misdemeanor under the CFAA on the basis of her contract breach.¹²²

The court, however, ultimately set aside the conviction, asserting that “[t]he pivotal issue herein is whether basing a CFAA misdemeanor violation as per 18 U.S.C. § 1030(a)(2)(C) and (c)(2)(A) upon the conscious violation of a website’s terms of service runs afoul of the void-for-vagueness doctrine.”¹²³ The court concluded that it did run afoul “primarily because of the absence of minimal guidelines to govern law enforcement, but also because of actual notice deficiencies.”¹²⁴ However, it is troubling that the same judge who set aside the verdict in the *Drew* case, nevertheless, appeared to agree with the jury’s CFAA-weaponized breach analysis. The court in *Drew* explained that “an intentional breach of the [MySpace Terms of Service] can potentially constitute accessing the MySpace computer/server without authorization and/or in excess of authorization under the statute,” apparently suggesting that a mere breach of contract can indeed provide the

120. *United States v. Drew*, 259 F.R.D. 449, 452 (C.D. Cal. 2009).

121. *Id.* at 451.

122. *Id.*

123. *Id.* at 464.

124. *Id.*

basis for a criminal prosecution under the CFAA.¹²⁵ Further, a judge's use of JNOV as a procedural mechanism is a rare event.¹²⁶ Other courts may not be willing to set aside a jury verdict, and since we know that at least one jury has already convicted based on an exceptionalized technology contract breach argument,¹²⁷ the issue is likely to surface again in another court.

As consumers increasingly rely on technology in their daily lives, this type of breach will occur with regularity. As Part I, *supra*, explained, unnegotiable, unilaterally amendable technology contracts with vague terms empower the drafter to assert breach by the other party with ease. We may all soon become "contract hackers" at the mercy of the drafters of the contracts that accompany the technologies we use in daily life. Although the Senate Judiciary Committee appears somewhat concerned about this eventuality,¹²⁸ and a bill was introduced in the Senate to exempt this user-hacker scenario from the CFAA,¹²⁹ Congress has not so amended the CFAA as of this writing.

2. "Disloyal" Employees or Business Partners as "Contract Hackers":
Confusing Intellectual Property Harms with Black Hat Hacking

Does an employee or business partner who has been initially granted access to a network through a contract lose such authorization if she accesses digital information for undesirable purposes in the opinion of her employer? By breaching her contract while accessing a network does she automatically become a "hacker"? The circuits are split in response.¹³⁰ These employee "contract hacker" cases roughly fall into two categories—what we might label cases dealing with alleged "thieves and vandals" and cases dealing with alleged "slackers."

125. *Id.* at 461.

126. See, e.g., Dick Thornburgh, *The Dangers of Over-Criminalization and the Need for Real Reform: The Dilemma of Artificial Entities and Artificial Crimes*, 44 AM. CRIM. L. REV. 1279, 1284 (2007) ("[Judge Chin] took the rare step of overriding a jury's guilty verdict and granting a motion for judgment of acquittal . . . because the government failed to prove fraudulent or deceptive conduct . . .").

127. See *Drew*, 259 F.R.D. at 453.

128. See Harley Geiger, *Senate Judiciary Committee Passes Three Data Security Bills*, CTR. FOR DEMOCRACY & TECH. (Sept. 23, 2011), <https://www.cdt.org/blogs/harley-geiger/239-senate-judiciary-committee-passes-three-data-security-bills>.

129. Cyber Crime Protection Security Act, S. 2111, 112th Cong. (as introduced, Feb. 15, 2012), available at <http://thomas.loc.gov/cgi-bin/query/z?c112:S.2111>.

130. For a discussion of the split, see, e.g., Robert C. Kain, *Federal Computer Fraud and Abuse Act: Employee Hacking Legal in California and Virginia, but Illegal in Miami, Dallas, Chicago, and Boston*, 87 FLA. BAR. J. 36 (2013).

a) The Alleged Thieves and Vandals

In *International Airport Centers, LLC v. Citrin*, Citrin, an employee who quit to start his own business, deleted data before turning in his company-issued laptop, including data that may potentially have demonstrated a breach of his noncompetition agreement.¹³¹ He also installed a program to write over deleted files in order to prevent their recovery.¹³² The employer sued, alleging violations of CFAA, violations of his employee's duty of loyalty, and breach of his employment contract.¹³³ The Seventh Circuit found a violation of the CFAA, using an inherently technology-exceptionalized analysis discussed in Part II, that "Citrin's breach of his duty of loyalty terminated . . . any rights he might have claimed as [the company's] agent . . . and with it his authority to access the laptop."¹³⁴

Other courts have disagreed.¹³⁵ In *LVRC Holdings LLC v. Brekka*, the Ninth Circuit expressly declined to follow *Citrin* on the definition of access deemed "without authorization" under the CFAA.¹³⁶ There, Brekka had been employed by LVRC, but maintained two consulting businesses with LVRC's

131. 440 F.3d 418, 419 (7th Cir. 2006).

132. *Citrin*, 440 F.3d at 419.

133. *Id.* at 418–19.

134. *Id.* at 420–21; *see also* United States v. Rodriguez, 628 F.3d 1258 (11th Cir. 2010) (upholding the conviction a worker for the Social Security Administration accessed the personal records of friends and acquaintances under 18 U.S.C. § 1030(a)(2)(B), which applies to government computers); EF Cultural Travel BV v. Explorica, Inc. 274 F.3d 577, 578–80, 585 (1st Cir. 2001) (finding a basis for CFAA liability where an employee's automated search queries violated the scope of his employment agreement); Shurgard Storage Ctrs., Inc. v. Safeguard Self Storage, Inc., 119 F. Supp. 2d 1121, 1124 (W.D. Wash. 2000) (concluding that once the employee acted on interests adverse to those of his employer by disseminating trade secret information via e-mail, he acted "without authorization" and could therefore be held liable under the CFAA). *But see* *Nosal II*, 676 F.3d 854, 855–56, 864 (9th Cir. 2012) (en banc) (finding that a group of employees' transfer of corporate information to a former employee did not provide a basis for a CFAA claim); *LVRC Holdings LLC v. Brekka*, 581 F.3d 1127, 1129–30, 1137 (9th Cir. 2007) (concluding an employee's transfer of corporate information to his personal email account did not trigger a CFAA violation).

135. A number of courts in other circuits have declined to follow *Citrin*. *See, e.g.*, *Orbit One Commc'ns, Inc. v. Numerex Corp.*, 692 F. Supp. 2d 373 (S.D.N.Y. 2010); *ReMedPar, Inc. v. AllParts Medical, LLC*, 683 F. Supp. 2d 605 (M.D. Tenn. 2010); *US Bioservices Corp. v. Lugo*, 595 F. Supp. 2d 1189, 1194 (D. Kan. 2009) ("The court agrees that the CFAA cannot be read to encompass (and criminalize) frauds that happen to involve the use of a computer someplace during the course of its commission."); *Diamond Power Intern., Inc. v. Davidson*, 540 F. Supp. 2d 1322, 1342–43 (N.D. Ga. 2007) (finding that employee's alleged misappropriation of employer information using employer's computer network did not provide a basis for a CFAA claim); *B & B Microscopes v. Armogida*, 532 F. Supp. 2d 744, 758 (W.D. Pa. 2007) (holding that accessing and deleting employer's files did not provide a basis for an employer's CFAA claim).

136. 581 F.3d 1127, 1134 (9th Cir. 2009).

knowledge.¹³⁷ While an LVRC employee, Brekka had e-mailed some LVRC documents to a personal email account, including a financial statement and marketing budget.¹³⁸ After Brekka terminated his employment, LVRC discovered that Brekka had accessed the company website using his log-in.¹³⁹ LVRC sued on the basis of both the emailed documents and the access after terminating employment.¹⁴⁰ But LVRC and Brekka had no written employment agreement, and LVRC had no published policy prohibiting employees from emailing LVRC documents to their personal computers.¹⁴¹ The Ninth Circuit ruled that criminal statutes must be interpreted to require that defendants have ample notice regarding what actions are criminal, and that Brekka had not agreed to keep emailed documents confidential or to return or destroy them at the conclusion of his employment.¹⁴² As the Ninth Circuit articulated in *Brekka*:

For purposes of the CFAA, when an employer authorizes an employee to use a company computer subject to certain limitations, the employee remains authorized to use the computer even if the employee violates those limitations . . . A person uses a computer “without authorization” under [section 1030(a)(4) only] when the person has not received the permission to use the computer for any purpose (such as when a hacker accesses someone’s computer without any permission), or when the employer has rescinded permission to access the computer and the defendant uses the computer anyway.¹⁴³

These two cases suggest that employers, trade secret owners, and copyright owners may view the CFAA as providing a “bonus” cause of

137. *Id.* at 1129.

138. *Id.*

139. *Id.*

140. *Id.*

141. *Id.*

142. *Id.*

143. *Id.* at 1133–35. In *United States v. Nosal*, a Ninth Circuit panel held that a former employee could be indicted under the CFAA for exceeding authorized access when he violated the employer’s computer access policies and persuaded three employees of an executive search firm to help him start a competing business. *Nosal I*, 642 F.3d 781, 784–89 (9th Cir. 2011). The three employees sent the defendant source lists, names and contact information from the company’s database. *Id.* at 783. The Ninth Circuit panel reversed a district court holding that the employees, who had permission to access work computers, did not exceed their authorized access for purposes of the CFAA. However, the Ninth Circuit granted a rehearing en banc, reversing the panel to bring the holding in line with *Brekka*. *Nosal II*, 676 F.3d 854, 863 (9th Cir. 2012) (en banc).

action when traditional contract breach claims are more appropriate.¹⁴⁴ Under this dynamic, when an employee extracts data through digital means rather than, for example, by copying the same information by hand, a court might apply additional damages simply because a computer was used for the copying. This type of exceptionalized analysis, apart from weaponizing breach and providing greater than appropriate remedies, also privileges one technology over another illogically: a handwritten copy of secret information becomes somehow less evil than an emailed copy of the same information. Meanwhile, as Professor Kerr has argued, because the CFAA's scope of covered technologies may now apply very broadly, the question of whether using a copier to copy the same information (instead of emailing it) now falls under the CFAA is unclear.¹⁴⁵ As an result, the damage analysis might no longer turn on the value of the information taken, but rather on the means of its copying.

b) The Alleged Slackers

Perhaps one the most obvious attempts at employer abuse of the CFAA and weaponized breach involves a case where an employer alleged that an employee who spent excessive (in the employer's opinion) amounts of work time on Facebook, violating the technology use policy of the workplace and, through this breach, the CFAA.¹⁴⁶ In *Lee v. PMSI, Inc.*, a Florida court rejected this weaponized breach analysis proposed by the employer.¹⁴⁷ The court asserted that:

Both the letter and the spirit of the CFAA convey that the statute is not intended to cover an employee who uses the internet instead of working The definition of "loss" contemplates damage to a system or data, rather than a lack of productivity Because the only information Lee allegedly accessed was on the personal

144. *See also* *Four Seasons Hotels & Resorts B.V. v. Consorcio Barr, S.A.*, 267 F. Supp. 2d 1268 (S.D. Fla. 2003). There, Four Seasons litigated to protect its database of customer preferences from use by a former business coventurer, with whom a contract existed. *Id.* at 1271–72. The court found that the Four Season's detailed customer profiles qualified as trade secrets under Uniform Trade Secrets Act: the information had economic value, was not generally known or readily available by others, and was the subject of reasonable efforts by licensor to preserve its secrecy. *Id.* at 1326. The court valued the data at \$2,090,000, doubled it as a penalty, and then added still more damages under the Computer Fraud and Abuse Act because of the means of acquisition. *Id.* at 1327; *see also* *Salestraq America, LLC v. Zyskowski*, 635 F. Supp. 2d 1178 (D. Nev. 2009) (considering the intersection of trade secret and computer intrusion law).

145. *See* Kerr, *supra* note 109, at 1562.

146. *Lee v. PMSI, Inc.*, No. 8:10-cv-2904-T-23TBM, 2011 WL 1742028, at *1 (M.D. Fla. May 6, 2011).

147. *Id.*

websites, not PMSI's computer system, Lee never "obtained or alter[ed] information in the computer." Lee accessed her facebook [sic], personal email, and news websites but did not access any information that she was "not entitled so to obtain or alter."¹⁴⁸

Florida courts have also held that checking personal email, even when it may contain information that breaches employer agreements with a defendant, does not constitute a violation of the CFAA.¹⁴⁹ However, employers in other jurisdictions will undoubtedly attempt to weaponize contract claims with the CFAA to reach alleged "slackers" and other courts, unlike this Florida court, may find the argument credible.

As I argue in Part III, *infra*, contract law and other bodies of law already embody adequate vehicles for providing remedies in almost all instances, even in complicated cases of improperly used information by "hacker" insiders. Because courts in general do not award information misuse damages under CFAA recoveries, employers' ability to threaten criminal prosecution for a contract breach in technology contexts creates a new type of punishment for disloyalty—the "digital peonage" concern discussed at Section II.B.2, *supra*. Digital peonage presents potentially problematic constitutional concerns. As the Court explained in *Pollack v. Williams*, "[w]hatever of social value there may be, and of course it is great, in enforcing contracts and collection of debts, Congress has put it beyond debate that no indebtedness warrants a suspension of the right to be free from compulsory service."¹⁵⁰ Particularly when the Supreme Court has already indicated an unwillingness to enforce contract regimes where employees' fear that merely leaving employment may give rise a criminal charge, this reasoning is likely to be expanded to the question of employee exit and the risk of computer intrusion charges. As explained in Part I, *supra*, such a regime would be highly subject to the possibility of employer abuse.¹⁵¹

148. *Id.* at *1–3.

149. *See* *Clarity Servs. v. Barney*, 698 F. Supp. 2d 1309, 1313–14, 1316 (M.D. Fla. 2010) (rejecting company's claim under 18 U.S.C. § 1030 where an employee solicited and read an email from a customer on the employee's company email account after resigning from the company, and deleted information from his company laptop before returning the laptop to his former employer, holding that the defendant did not lack authorization to access the information or exceed his authorization).

150. *Pollock v. Williams*, 322 U.S. 4, 18 (1944).

151. *See* Kathleen Kim, *The Coercion of Trafficked Workers*, 96 IOWA L. REV. 409 (2011) (discussing modern peonage contexts).

3. *Entrepreneurs as “Contract Hackers”: Confusing Innovation with Black Hat Hacking*

Arguably the most complex category of “contract hackers” for purposes of a CFAA analysis are those who agree to a contract, but whose activities later exceed the boundaries of authorized conduct when they apply their own code or specialized knowledge. At least two types of commercial cases potentially implicate these behaviors: (1) entrepreneurs who author applications performing services on behalf of users, and (2) competitors or entrepreneurs who use code to aggregate information.

a) Application Builders

The first category of these entrepreneur “contract hacker” cases involve computer applications that have been assigned express authorization from a user—a user who has entered into a contract with each party—to perform services on behalf of that user. A recent California case, *Facebook v. Power Ventures*, raises the specter of this type of CFAA case in the future.¹⁵² In that case, Facebook sued a company providing a content curation application.¹⁵³ The application allowed Facebook users to access their own messages, friends lists, and other content from their profiles on various social networking websites inside a single application.¹⁵⁴ Facebook argued that by offering these services,¹⁵⁵ Power Ventures violated Facebook’s terms of use and, therefore, the CFAA and California computer crime law.¹⁵⁶

This weaponized breach argument is problematic. It can be argued that a Facebook user simply assigns her access rights to Power Ventures when using the aggregation application. As a matter of basic contract, it is not clear that Power Ventures stands in privity with Facebook; if Power Ventures is not in privity with Facebook, a breach of a contract by Power Ventures cannot exist. Thus, the weaponized breach argument must fail immediately purely on the most basic contract formation grounds. However, assuming that Power Ventures lacks privity, then the party who is in privity with Facebook and may have breached the terms is the average consumer user who authorized the Power Ventures application. In other words, we return

152. 844 F. Supp. 2d 1025, 1027 (N.D. Cal. 2012).

153. *Id.*

154. *Id.*

155. *Id.* at 1027–28. Power Ventures also sent out advertising emails that appeared to be originated by Facebook email accounts. *Id.* Consequently, Facebook also argued that Power Ventures violated the CAN-SPAM Act. The court agreed, and granted summary judgment. *Id.* at 1030.

156. *Id.* at 1036–38.

right back to the deeply troubling Grandma and Dr. Whiskers hypothetical and its implications discussed earlier in this section.

However, even assuming Power Ventures is in privity with Facebook, a contract breach argument against an application builder like Power Ventures is weak at best. Conducting a traditional breach analysis, the existence of any economic harm to Facebook due to this type of intermediation will be difficult to prove. In fact, intermediation and applications that facilitate user content creation may be examples of desirable technology entrepreneurship, not commercially harmful conduct. Here again the appeal of a weaponized breach analysis for plaintiffs becomes apparent: in a breach analysis weaponized with the CFAA, the extent of economic damages becomes basically irrelevant for criminal application of the CFAA. Consequently, the potency of a possible CFAA charge is clear, even when no recourse would be appropriate for the breach under contract law.¹⁵⁷ Weaponized breach offers a vehicle for existing market players to limit entrepreneurial efforts that build on their business models.

b) Data Aggregators

The second group of entrepreneur “contact hacker” cases involves data aggregators seeking to technologically capture and use information. A number of cases to date have involved competitor aggregators using automated programs such as spiders to capture information from a company’s website for their own business interests in this manner.

Recently, in *United States v. Lowson*, a New Jersey court allowed a criminal prosecution arising out of a weaponized breach claim to continue to trial.¹⁵⁸ The defendant ticket aggregators and resellers were accused of taking various steps to defeat Ticketmaster’s code-based security measures, including CAPTCHAs and encryption and implementing “hacks” and using backdoors to enable automated programs to purchase tickets.¹⁵⁹ The defendants also

157. In *Power Ventures*, this question of whether any contract damages existed was not reached. On summary judgment the court chose not to analyze the question of whether a contract existed, instead focusing on the changes in IP addresses made by Power Ventures as acts of circumvention of technological limitations on access. *Id.*

158. Indictment at 1, *United States v. Lowson*, No. 2:10-cr-00114-KSH (D.N.J. Feb. 23, 2010), available at http://www.wired.com/images_blogs/threatlevel/2010/03/wiseguys-indictment-filed.pdf.

159. CAPTCHAs or HIPs are typing tests used to ascertain that a particular user is a human rather than a bot. See *Ticketmaster L.L.C. v. RMG Technologies, Inc.*, 507 F. Supp. 2d 1096, 1112 (C.D. Cal. 2007) (“Here, CAPTCHA both *controls access* to a protected work because a user cannot proceed to copyright protected webpages without solving CAPTCHA, and *protects rights* of a copyright owner because, by preventing automated access to the ticket purchase webpage, CAPTCHA prevents users from copying those pages.”).

allegedly disregarded cease-and-desist letters and hired programmers to defeat security restrictions.¹⁶⁰ The prosecution's argument hinged at least in part on a weaponized breach argument: the indictment stated that the defendants' behavior was an act of computer intrusion in part because "[ticket v]endors using CAPTCHA technology on their websites routinely added Terms of Service that expressly stated that users were not permitted to access a CAPTCHA-protected website using automated software."¹⁶¹ By including this argument, the prosecution implied that the basis for unauthorized access under the CFAA arose out of a breach of the Terms of Service.¹⁶² Problematically, although the matter ultimately ended in a plea bargain, the court initially denied the motion to dismiss, apparently indicating that the court afforded some traction to the argument that a criminal violation of the CFAA can be predicated on a breach of contract.¹⁶³

The use of this weaponized breach line of argument in these data aggregator entrepreneur cases is a relatively new development. Looking back to earlier case law dealing with data aggregators, such as *Register.com v. Verio*, we see allegations of both contract breach and violations under the CFAA.¹⁶⁴ In *Verio*, however, the court identified those claims as separate rather than combining them under a single cause of action.¹⁶⁵ Even Ticketmaster, one of the vendors most interested in the prosecution of the defendants in *Lowson*, had not connected breach of contract claims with CFAA claims in previous rounds of litigation against data aggregators.¹⁶⁶ For example, in *Ticketmaster v. Tickets.com*, Ticketmaster had attempted to pursue claims under contract,

160. United States v. Lowson, No. 2:10-cr-00114-KSH, at *12 (D.N.J. Oct. 12, 2010) (ruling on motion to dismiss), available at http://www.wired.com/images_blogs/threatlevel/2010/10/Wiseguys_Ruling-on-Motion-to-Dismiss.pdf.

161. *Id.* at 10.

162. Indictment at 1, United States v. Lowson, No. 2:10-cr-00114-KSH (D.N.J. Feb. 23, 2010), available at http://www.wired.com/images_blogs/threatlevel/2010/03/wiseguys-indictment-filed.pdf.

163. United States v. Lowson, No. 2:10-cr-00114-KSH, at *8 (D.N.J. Oct. 12, 2010) (ruling on motion to dismiss), available at http://www.wired.com/images_blogs/threatlevel/2010/10/Wiseguys_Ruling-on-Motion-to-Dismiss.pdf.

164. *Register.com, Inc. v. Verio, Inc.*, 356 F.3d 393, 425, 439 (2d Cir. 2004) (separately assessing the merits of a breach of contract claim and a CFAA claim in connection with an alleged violation of terms of use).

165. *Id.*

166. *See, e.g.*, Defendant's Motion to Dismiss, *Ticketmaster Corp. v. Tickets.com, Inc.*, No. CV99-7654HLHVBKX, 2000 WL 525390 (C.D. Cal. Mar. 27, 2000); Defendant's Motion for Summary Judgment, *Ticketmaster Corp. v. Tickets.com, Inc.*, No. CV99-7654HLHVBKX, 2003 WL 21406289 (C.D. Cal. Mar. 7, 2003).

copyright, and trespass to chattels as independent bases of recourse.¹⁶⁷ Only the contract law claim survived summary judgment.¹⁶⁸ As such, it is perhaps unsurprising that in this next iteration of commercial wrangling between established internet companies and data aggregators, this new and particularly aggressive strategy has emerged—an attempt to buttress contract claims with a second harsher regime through breach weaponized with the CFAA.

4. *Security Researchers as “Contract Hackers”: Confusing Code Auditing and White Hat Hacking with Black Hat Hacking*

Perhaps the most problematic group of “contract hackers” are information security researchers who may violate terms in an end user license agreement in the course of their research into the behaviors of websites or other technology-related products. These researchers hold a vital consumer protection role in the ecosystem of information security vulnerabilities that cannot easily be filled any other way.¹⁶⁹ Although determining who qualifies as a security vulnerability researcher, the permissible scope of information security research, and proper disclosure processes are a matter of debate,¹⁷⁰ what has become clear is that some researchers’ work in the security space, such as the work of computer academics, is essential as a code audit mechanism.¹⁷¹ Knowing that code is reviewed by third parties after release for security vulnerabilities and privacy invasive conduct keeps companies honest, and it prevents severe information security harms from spreading throughout the digital ecosystem.¹⁷²

As I have explained in other work, sometimes code can behave in ways that damage consumers and enterprise information security, but the authors of the code do not always feel obligated to prevent these harms or disclose the risk.¹⁷³ Third-party information security researchers frequently uncover these risks and alert the public.¹⁷⁴ For example, researchers play a vital role in identifying vulnerabilities related to digital rights management (“DRM”) tactics that companies sometimes use in the name of intellectual property

167. *Ticketmaster Corp v, Tickets.com, Inc.* No. CV99-7654 HLHVBKX, 2003 WL 21406289, at *1 (C.D. Cal. Mar. 7. 2003).

168. *Id.* at *2–3, *6.

169. For a discussion of the role of information security researchers in the software ecosystem, see Andrea M. Matwyshyn, *Hacking Speech*, 107 NW. L. REV. 795 (2013).

170. For a discussion of diverging norms in information security vulnerability disclosure, see *id.* at 825–27.

171. *Id.*

172. *Id.*

173. See Matwyshyn, *supra* note 47, at 438–41.

174. See Bruce Schneier, *Sony’s DRM Rootkit: the Real Story*, SCHNEIER ON SECURITY (Nov. 17, 2005), http://www.schneier.com/blog/archives/2005/11/sonys_drm_rootk.html.

protection, tactics which are sometimes functionally similar or identical to the tactics used by black-hat hackers.¹⁷⁵ The key difference is that, ostensibly, the DRM is placed on the user's hard drive with contractual consent from the user through the product's end user license agreement. In practice, however, the agreement may not explain the manner in which the code behaves. Without a third party-technology expert analysis of the behavior of the code, the marketplace will never know how this code functions—functionality that may cause millions of dollars of information security damage.¹⁷⁶

Each time a security researcher analyzes a product or decompiles it to track behaviors for the benefit of all users of the product, the researcher now potentially risks criminal prosecution as a result of an alleged contract breach in the course of a code audit. Numerous examples exist with respect to security researchers—including Ivy League academics—being threatened with litigation and even criminal prosecution in connection with their research into information security vulnerabilities.¹⁷⁷ If weaponized breach analysis persists, society imposes on information security researchers a very high legal cost for engaging in socially beneficial research.

Meanwhile, in a less traditional research space, sometimes one company's information security researchers discover vulnerabilities in another company's products. In this scenario, a company that becomes annoyed with a competitor's disclosing their vulnerabilities may attempt to legally sanction the individual researchers working for the competitor, as well as the competitor itself, with CFAA-weaponized breach claims. There have already

175. *Id.* Digital rights management technologies are code-based methods of limiting a user's ability to interact with digital content.

176. See Deirdre K. Mulligan, *The Magnificence of the Disaster: Reconstructing the Sony MBG Rootkit Incident*, 22 BERKELEY TECH. L.J. 1157, 1158 (2007) (discussing the Sony DRM rootkit incident).

177. By way of example, let us take the case of Professor Edward Felten at Princeton University, formerly the chief technologist at the Federal Trade Commission. *FTC Names Edward W. Felten as Agency's Chief Technologist; Eileen Harrington as Executive Director*, FED. TRADE COMM'N (Nov. 4, 2010), <http://www.ftc.gov/opa/2010/11/cted.shtm>. Professor Felten, his colleagues, and his graduate students in his lab at Princeton frequently audit code in order to assess the behaviors for consumer protection reasons. See CENTER FOR INFORMATION TECHNOLOGY POLICY AT PRINCETON UNIVERSITY, <https://citp.princeton.edu/research/> (last visited Feb. 28, 2012). Their research, which undoubtedly breaches end user license agreements and terms of use at least occasionally, now carries with it a risk of criminal prosecution under the CFAA. Yet it is indisputable that their research fulfills a critical role in the information ecosystem. Professor Felten himself was already once threatened with criminal prosecution for his research under the Digital Millennium Copyright Act in connection with publishing a paper exposing flaws in a DRM technology. See Letter from Matthew J. Oppenheim, Esq., RIAA Counsel, to Professor Edward Felten (Apr. 9, 2001), available at <http://cryptome.org/sdmi-attack.htm>.

been squabbles between major technology companies over precisely these dynamics,¹⁷⁸ and major information security conferences have been disrupted with speakers being sued over allegedly violating the CFAA via a breach of contract.¹⁷⁹ Most such disputes end in settlement. Currently, however, the legal questions remain unresolved and the dynamics remain volatile.

B. THE RISKS OF WEAPONIZING BREACH WITH THE CFAA

As the previous Section's discussion illustrated, contract breach weaponized with the CFAA implicates at least four different types of "contract hackers" and strategic litigant interests. As a result, a blanket weaponized breach approach further balkanizes traditional contract law in a particularly destructive manner. First, weaponized breach derails contract doctrine through judicial activism. Second, it eviscerates the possibility of a contract theory discourse regarding the nature, morality, and appropriate consequences of breach. Third, it disrupts the private ordering that has been the main regulator of technology-mediated spaces to date. Fourth, weaponized breach harms entrepreneurship and innovation policy. Finally, it stunts development of other, perhaps more suitable legal regimes to address information harms. The remainder of this Part considers each of these negative consequences in turn.

1. *Contract Doctrine and Judicial Activism*

Weaponizing breach with a CFAA analysis threatens the future of contract doctrine. The CFAA offers no guidance on the relationship between contract and computer intrusion, and Congress has never directly addressed their interplay.¹⁸⁰ The paradigm that existed in the minds of legislators at the

178. See, e.g., Graham Cluley, *Tavis Ormandy – are you pleased with yourself? Website exploits Microsoft zero-day*, SOPHOS NAKED SECURITY (June 15, 2010), <http://nakedsecurity.sophos.com/2010/06/15/tavis-ormandy-pleased-website-exploits-microsoft-zero-day/> (describing tension between Microsoft Corporation and a Google employee's chosen method of vulnerability disclosure). Vulnerabilities may be found by one company in another's product when, for example, a researcher is trying to customize or secure their own products to interoperate with the vulnerable product.

179. See, e.g., Jennifer Granick, *An Insider's View of CiscoGate*, WIRED.COM, Aug. 5, 2005, <http://www.wired.com/science/discoveries/news/2005/08/68435?currentPage=2> (describing the events around Michael Lynn's disclosure of a vulnerability in Cisco IOS software).

180. In 1986, Congress deleted the part of the statute that prohibited those with authorization from using the system for unauthorized purposes and substituted the phrase "exceeds authorized access." See *Int'l Ass'n of Machinists and Aerospace Workers v. Werner-Masuda*, 390 F. Supp. 2d 479, 499 n.12 (D. Md. 2005) (quoting S. Rep. No. 99-432, at 9 (1986), reprinted in 1986 U.S.C.C.A.N. 2479, 2486). The *Werner-Masuda* court explained the change as follows: "By enacting this amendment, and providing an express definition for exceeds authorized access, the intent was to eliminate coverage for authorized access that

time of passage of the CFAA was most likely that of the “hacker”—a third-party outsider or “black hat” hacker. It is unlikely that entrepreneurs, information security researchers and, in particular, the Grandma and Dr. Whiskers of the world were within the mental model of “hackers” legislators used as the basis for the CFAA.¹⁸¹ As such, judicial restraint in interpretation of breach and the CFAA is warranted rather than building a common law of the zebra. Until Congress speaks directly on this matter, judicial restraint is appropriate. In the words of a lower court, even assuming that “[w]hat the Government is seeking to do is to punish conduct that reasonable people might agree deserves the sanctions of the criminal law,” the wiser course was to leave it to Congress to prescribe crime and establish penalties.¹⁸²

Sound policy, as well as history, supports our consistent deference to Congress when major technological innovations alter the market . . . Congress has the institutional authority and the institutional ability to accommodate fully the varied permutations of competing interests that are inevitably implicated by such new technology.¹⁸³

Contract law is predominantly a state law construct, and without clear evidence of Congressional intent to the contrary, courts should not interfere with state-law contract claims by weaponizing them under the federal CFAA. In *Morrison v. National Australia Bank*, the Supreme Court cautioned against “judicial speculation-made-law divining what Congress would have wanted if it had thought of the situation before the court.”¹⁸⁴ Yet, that is precisely what judicial activism in expanding the reach of the CFAA to mere breaches of contract does. As the Court has elsewhere noted, “the historic police powers

aims at purposes to which such authorization does not extend, thereby removing from the sweep of the statute one of the murkier grounds of liability, under which a [person’s] access to computerized data might be legitimate in some circumstances, but criminal in other (not clearly distinguishable) circumstances that might be held to exceed his authorization.” *Id.* at 499 n.12 (quoting S. Rep. No. 99-432, at 21, reprinted in 1986 U.S.C.C.A.N. at 2494–95) (internal quotations omitted).

181. Recent Senate Judiciary Committee activity indicates that some members of Congress disagree with the idea of extending the CFAA to cover consumers who breach contracts. See Geiger, *Senate Judiciary Committee Passes Three Data Security Bills*, *supra* note 128.

182. *United States v. LaMacchia*, 871 F. Supp. 535, 544 (D. Mass. 1994) (citing *Dowling v. United States*, 473 U.S. 207, 225 (1985)).

183. *LaMacchia*, 871 F. Supp. at 544 (citing *Sony Corp. of Am. v. Universal City Studios, Inc.*, 464 U.S. 417, 431 (1984)).

184. 130 S. Ct. 2869, 2881 (2010).

of the States [are] not to be superseded by [a] [f]ederal Act unless that was the clear and manifest purpose of Congress.”¹⁸⁵

One of the strategic benefits a weaponized breach claim may offer a litigant is a basis for removing a run-of-the-mill state contract law dispute to federal court.¹⁸⁶ While this may be desirable for some litigants, flooding the federal courts with contract claims has systemic consequences. Contract claims belong in state courts as a general rule. If increasing numbers of contracts involve some aspect of technology, and if the weaponized breach litigation trend continues, soon the federal courts may find themselves overwhelmed with litigants asserting weaponized breach claims.¹⁸⁷

Furthermore, in the context of entrepreneur “contract hackers,” another reason why the CFAA-weaponized contract breach claims have gained popularity among litigants may be because they are the only claims that “stick” in technology contexts that offer more aggressive penalties.¹⁸⁸ However, this type of overreaching by leveraging contract to latch onto federal law remedies is unnecessary and misguided. Other regimes already exist to provide remedy for these harms. State trade secret law, for example, already protects confidential information of various types that is frequently not otherwise protectable under other areas of intellectual property law.¹⁸⁹ However, if trade secret does not cover the alleged harm at issue, then perhaps only a contract law remedy, or no remedy at all, is appropriate.

185. *Wyeth v. Levine*, 129 S. Ct. 1187, 1194–95 (2009) (quoting *Rice v. Santa Fe Elevator Corp.*, 331 U.S. 218, 230 (1947)).

186. *See, e.g.*, Justin R. Long, *Against Certification*, 78 GEO. WASH. L. REV. 114, 119 (2009) (“Plaintiffs can choose where to file, and defendants can choose whether to remove, based on strategic considerations of how the lower state and federal courts have treated the open question of state law in the case.”).

187. When faced with a similar tension in the context of tort law and federal actions under Section 1983, the Supreme Court clearly indicated that only state law remedies apply. *Paul v. Davis*, 424 U.S. 693, 700–01 (1976) (considering whether a privacy interest was violated by a flyer posted by an officer alleging Davis was an “active shoplifter” and presented cause of action under Section 1983). As the Court in *Davis* asserted, “Congress should not be understood to have attempted ‘to make all torts of state officials federal crimes.’ It brought within (the criminal provision) only specified acts ‘under color’ of law and then only those acts which deprived a person of some right secured by the Constitution or laws of the United States.” *Id.* at 700 (citing *Screws v. United States*, 325 U.S. 91, 109 (1945)).

188. *See* discussion *supra* Section III.A.3.

189. For example, in some states, trade secret law may protect a client list, while neither copyright nor patent law would provide protection for a client list. For a discussion of trade secret law and client lists, see Erini R. Svokos, *What About the Client? Trade Secret Law and Fiduciary Duty Law as Applied to Law Firm Client Lists*, 24 GEO. J. LEGAL ETHICS 937, 941 (2011).

But, one might argue, perhaps penalizing the mere use of technology as part of a commercial transaction in breach is a socially desirable state of affairs.¹⁹⁰ I argue it is not. Punishing breach more harshly in digital contexts in essence creates a type of technology contracting “tax” that is anathema to Congress’s asserted intent of crafting parity for digital and offline contracts. Specifically, in the E-Sign Act, Congress has articulated an express goal of legal equivalence between technology-mediated contract and physical, offline contracts.¹⁹¹ Thus, an exceptionalized breach analysis runs contrary to Congress’s explicit intent. Congress’ goal in the E-Sign Act might also be described as seeking to extend contract law into digital spaces, but only in a minimally disruptive manner, and certainly not in a manner that supplants existing contract doctrine.¹⁹²

2. *How Weaponized Breach Disrupts Contract Theory*

Just as a weaponized breach analysis disrupts the relationship of contract doctrine and may run afoul of Congressional intent, it also problematizes contract theory and scholarship. Neither a traditional efficient breach/“bad man” approach to contract analysis nor an autonomy theory analysis can co-exist with weaponized breach. In 1897 Supreme Court Justice Oliver Wendell Holmes articulated a “strict liability” approach to contract breach: “The duty to keep a contract at common law means a prediction that you must pay damages if you do not keep it, and nothing else.”¹⁹³ A traditional law-and-

190. For example, perhaps one can argue that we are inevitably moving toward a surveillance society where human relationships and legal responsibility need to be reconfigured to acknowledge this new reality. Perhaps traditional legal balances between liberty interests and criminal responsibility and civil liability should be disregarded.

191. Electronic Signatures in Global and National Commerce Act (“E-SIGN”) Pub. L. No. 106-229, 114 Stat. 464 (2000) (codified at 15 U.S.C. §§ 7001–7006, 7021, 7031 (2000)). Although Congress has usually been substantially behind the pace of technology innovation with respect to making law, one notable example of forward-thinking occurred—the E-Sign Act. In the context of legislating rules for technology-mediated contract formation, federal and state digital signature legislation established parity for physical space and virtual space technology-mediated signatures at a relatively early point in the mainstreaming and commercialization of the Internet.

192. See generally Stephen E. Friedman, *Protecting Consumers from Arbitration Provisions in Cyberspace, the Federal Arbitration Act and E-Sign Notwithstanding*, 57 CATH. U. L. REV. 377 (2008) (providing discussion of the E-Sign Act); Stephen E. Blythe, *Digital Signatures Law of the United Nations, European Union, United Kingdom and United States: Promotion of Growth in e-Commerce with Enhanced Security*, 11 RICH. J.L. & TECH. 6 (2005) (comparing digital signature legislation in the European Union, United States, and United Kingdom).

193. O.W. Holmes, *The Path of the Law*, 10 HARV. L. REV. 457, 462 (1897). As more recently expressed by Justice Scalia: “Virtually every contract operates, not as a guarantee of particular future conduct, but as an assumption of liability in the event of

economics notion of efficient breach, as well as this Holmesian “bad man” notion—i.e. one which views breach as a viable option provided one pays damages arising therefrom—view a contractual promise being no more than an option to breach and pay damages.¹⁹⁴ An efficient breach approach does not consider compensated breach to be wrongful. In fact, if it is efficient, it may even be desirable.¹⁹⁵ One such debate focuses on whether the act of breaching a contract is even itself a legal wrong or merely a choice that a party in privity makes with respect to performing or paying damages. Professor Shiffrin cites Justice Holmes and argues that breach of contract is “not a legal wrong” because the law views every contract as an agreement in the alternative—to perform or to pay an amount of money equal to the value of performance.¹⁹⁶

A weaponized breach analysis also does not comport with an autonomy theory analysis of contract, which focuses on the bilateral structure of contractual liability and expectation damages.¹⁹⁷ As Professor Oman argues:

[A]utonomy itself places limits on the sort of remedies that the law can impose. The basic intuition behind this argument is that specific performance represents a greater intrusion into personal freedom than do money damages, and so long as damages compensate the promisee for her loss, we ought to choose the remedy that intrudes on liberty the least.¹⁹⁸

Similarly, Charles Fried, a defender of an autonomy theory of contract, writes:

If I make a promise to you, I should do as I promise; and if I fail to keep my promise, it is fair that I should be made to hand over the equivalent of the promised performance. In contract doctrine this proposition appears as the expectation measure of damages for breach. The expectation standard gives the victim of a breach no

nonperformance” *United States v. Winstar Corp.*, 518 U.S. 839, 919 (1996) (Scalia, J., concurring).

194. See Barry E. Adler, *Efficient Breach Theory Through the Looking Glass*, 83 N.Y.U. L. REV. 1679 (2008) (arguing that a willful efficient breach must simply be priced accurately).

195. *Id.*

196. Thus, for Professor Shiffrin “contract diverges from promise” in that the “contents of the legal obligations and the legal significance of their breach do not correspond to the moral obligations and the moral significance of their breach.” Seana Valentine Shiffrin, *The Divergence of Contract and Promise*, 120 HARV. L. REV. 708, 709 (2007).

197. See Gregory Klass, *Promise Etc.*, 45 SUFFOLK U. L. REV. 695, 696–97 (2012) (discussing autonomy theory).

198. Nathan B. Oman, *The Failure of Economic Interpretations of the Law of Contract Damages*, 64 WASH. & LEE L. REV. 829, 869 (2007).

more or less than he would have had had there been no breach—in other words, he gets the benefit of his bargain.¹⁹⁹

In an environment where choosing not to perform a contractual promise can lead to liberty deprivation, no true “option” to breach exists. Becoming a felon can never be the efficient or autonomous outcome in a contractual relationship. Hence, the contract theory discourse dies, despite a long-standing debate in both the courts and the legal academy about the nature of breach and its connection to morality.²⁰⁰ Regardless of whether one subscribes to an “efficient breach” or autonomy-driven analysis of contract law, the theory behind providing redress for contract “wrongs” is driven by financial redress on an individual level—not righting a wrong against society as a whole like in criminal law. In the most generous calculations, contract damages provide the benefit of the anticipated bargain to the victim of the breach or, in limited cases, contract law seeks to disgorge ill-gotten gains as a consequence of the breach.²⁰¹ Even punitive damages are a disfavored remedy, and liquidated damages provisions are generally enforced only when the stipulated sums are connected to actual economic loss arising from breach, not driven by a desire to punish the breacher.²⁰²

The possibility of criminal consequences of a mere breach does not, therefore, generally appear in contract discourse. Even if we assume that breach is immoral, contract theory discourse does not equate “immorality”

199. CHARLES FRIED, *CONTRACT AS PROMISE: A THEORY OF CONTRACTUAL OBLIGATION* 17–19 (1981).

200. See generally Steven Shavell, *Is Breach of Contract Immoral?*, 56 *EMORY L.J.* 439, 439–57 (2006) (discussing the contract theory debate over breach and morality).

201. See *RESTATEMENT (SECOND) OF CONTRACTS* § 355 cmt. a (1981) (“The purpose[] of awarding contract damages is to compensate the injured party . . . For this reason, courts in contract cases do not award damages to punish the party in breach or to serve as an example to others unless the conduct constituting the breach is also a tort for which punitive damages are recoverable.”).

202. See 22 *AM. JUR. 2d Damages* § 574 (2012).

Punitive damages are generally not available under a contract theory . . . Specifically, courts generally hold as a general rule that punitive damages are not available as a remedy for breach of contract without an underlying tort . . . Courts have explained that punitive damages are not ordinarily recoverable in actions for breach of contract because: (1) the damages for breach of contract are generally limited to the pecuniary loss sustained; and (2) the purpose of punitive damages is not to remedy private wrongs but to vindicate public rights . . . Nonetheless, it has also been held that a contract provision immunizing a party from liability for punitive damages is substantively unconscionable as violating public policy.

Id. (citations omitted).

with criminal penalties like weaponized breach analysis does.²⁰³ Weaponized breach analysis means that, because of the stick of criminal law, one side automatically loses—the side of permissible exit. Thus, with only one logical outcome possible because of the fear of criminal sanctions, the contract theory debates are obliterated.

Further, the weaponization of breach destroys one of the core doctrinal distinctions in contract law—the distinction between a material and a minor breach.²⁰⁴ When breach is weaponized, even a minor breach can ostensibly provide the basis for a criminal penalty under the CFAA or a duplicative financial penalty under both civil CFAA remedies and traditional contract remedies. A weaponized breach analysis is counter to the contract default rule of preserving contractual relationships whenever possible.²⁰⁵ When a party's breach is material, the nonbreaching party has the option of receiving damages and terminating the contractual relationship. However, in an instance of a minor breach, courts prefer to preserve the existence of the relationship and simply award damages to compensate for actual losses suffered by the nonbreaching party.²⁰⁶ As such, it can be said that the dominant lens courts apply is closer to a relational notion of contract than a moralistic one that views breaches as irretrievable wrongs.²⁰⁷ But, again, in a

203. See DAVID HUME, A TREATISE OF HUMAN NATURE 523 (L.A. Selby-Bigge ed., 2d ed. 1978) (describing promise-keeping as a social good, saying that “a sentiment of morals concurs with interest, and becomes a new obligation upon mankind.”). As Fried explains:

There exists a convention that defines the practice of promising and its entailments. This convention provides a way that a person may create expectations in others. By virtue of the basic Kantian principles of trust and respect, it is wrong to invoke that convention in order to make a promise, and then to break it.

FRIED, *supra* note 199, at 17. *But see* Richard Craswell, *Contract Law, Default Rules, and the Philosophy of Promising*, 88 MICH. L. REV. 489, 489 (1989).

[A]nalyzes such as Fried's have little or no relevance to those parts of contract law that govern the proper remedies for breach, the conditions under which the promisor is excused from her duty to perform, or the additional obligations . . . imputed to the promisor as an implicit part of her promise.

Id.; see also Steven Shavell, *Why Breach of Contract May Not Be Immoral Given the Incompleteness of Contracts*, 107 MICH. L. REV. 1569, 1569 (2009) (“[Although it is a] widely held view that breach of contract is immoral . . . breach may often be seen as moral, once one appreciates that contracts are incompletely detailed agreements and that breach may be committed in problematic contingencies that were not explicitly addressed by the governing contracts.”).

204. 23 WILLISTON ON CONTRACTS § 63:3 (4th ed.)

205. See, e.g., RESTATEMENT (SECOND) OF CONTRACTS § 237 (1981).

206. *Id.*

207. See Macaulay, *Relational Contracts*, *supra* note 42, at 793–94 (discussing relational contracts).

weaponized breach analysis, particularly one marked with the possibility of prison, this distinction disappears. Every breach becomes the basis for a possible corollary criminal charge. Thus, with weaponized breach a type of permanent specific performance mandate suddenly pertains to all contractual obligations no matter how big or small and no matter how efficient it might be for all parties involved to breach obligations: each time a party debates breaching an agreement, that party essentially accepts the risk of a potential criminal prosecution.

3. *Private Ordering*

The law has generally adopted a restrained approach to technology regulation because of the rapid evolution of internet business models, traditionally deferring to business partners' privately ordered arrangements through contract as defining the relationship.²⁰⁸ Weaponized breach analysis that triggers an immediate right to pursue remedies under the CFAA eliminates this deference. It limits the ability of parties to set their own deal terms through private ordering: an enforcement remedy through the CFAA essentially eviscerates the negotiated contractual deal as the definitive set of rules governing the exchange. The person controlling the technology always holds the ability to inflict criminal punishment for breach on the other party.²⁰⁹

In this manner, weaponized breach analysis creates a new technology-driven imbalance in the relationship that the parties cannot contract around: the specter of criminal punishment exists *regardless of the deal terms and even when the two parties are similarly situated*, except for one side managing the network that is implicated by the contract. In a legal regime of weaponized breach, the balance of power between otherwise equally situated parties becomes disrupted if one controls the network and has possible recourse under the CFAA while the other does not. For example, weaponized breach destroys the efficacy of indemnification provisions and limitations of liability as the primary monetary terms of recourse and risk management. Regardless of how the parties negotiated out liability terms and indemnification obligations,

208. See Henry H. Perritt, Jr., *The Internet is Changing the Public International Legal System*, 88 KY. L.J. 885, 921–31 (1999) (discussing private ordering and internet regulation).

209. See, e.g., Edward A. Morse, *Private Ordering In Light Of The Law: Achieving Consumer Protection Through Payment Card Security Measures*, 10 DEPAUL BUS. & COM. L.J. 213, 215 (2012); Robert M. Yeh, Note, *The Public Paid for the Invention: Who Owns It?*, 27 BERKELEY TECH. L.J. 453 (2012); D. Gordon Smith, *Private Ordering With Shareholder Bylaws*, 80 FORDHAM L. REV. 125, 127 (2011) (discussing private ordering and various contracts).

the party that controls the technology suddenly holds a trump card—the ability to weaponize any breach.²¹⁰

4. *Innovation and Entrepreneurship Policy*

The ability of entrepreneurs to experiment with innovative technology business models is threatened when each contractual transgression may result in a criminal sentence. Entrepreneurs face not only the possibility of their own imprisonment under weaponized breach, but further challenges when raising capital, as it creates perverse incentives for venture capitalists to invest in the most novel digital enterprises when those enterprises carry a higher likelihood of contract breach and founder felony convictions. Further, the approach of weaponized breach privileges corporate entities over human ones: a lone inventor who breaches an agreement may find himself facing computer intrusion charges and potentially prison, but a corporation that breaches an agreement faces no incarceration risk under the CFAA.²¹¹ In this way, legally sophisticated entrepreneurs may avoid prison, but the smallest startups where founders have not correctly selected a corporate form to insulate themselves are vulnerable. This state of affairs is highly undesirable as a matter of innovation policy.

Further, weaponizing breach chills innovation policy debates about technology conduct in gray legal areas. For example, norms in the technology community around the appropriate protocols for white-hat, well-intentioned hacking are in flux. Many companies such as Google and Mozilla expressly encourage breaking of their products and offer monetary rewards²¹² to researchers who find and submit problems with corporate code. However,

210. Further, this type of a dynamic creates disincentives for the party operating a network or a database to practice good network security and data management practices. As will be discussed later, where the ability to harshly punish through weaponized breach exists, the desire to take precautions against breach diminishes. *See* discussion *infra* Section II.B.4. The failure to take these information security precautions has negative effects outside the business relationship and simultaneously devalues corporate information assets. *See, e.g.,* Andrea M. Matwyshyn, *Imagining the Intangible*, 34 DEL. J. CORP. L. 965, 976–80 (2009) (discussing the devaluation of intangible assets through information security breaches).

211. In the absence of a basis for personal responsibility of the officers and directors of a company, a business entity cannot physically be incarcerated because it is inanimate, even if its officers can be. In other words, only individuals participating in criminal activity may be incarcerated. For a discussion of corporate criminal responsibility, see 18B AM. JUR. 2D *Corporations* § 1640 (2012).

212. Google recently offered a one million-dollar bug bounty for vulnerabilities in Chrome. *See* Kim Zetter, *Google Offers \$1 Million in Hacker Bounties for Exploits Against Chrome*, WIRED.COM (Feb. 28, 2012), <http://www.wired.com/threatlevel/2012/02/google-1-million-dollar-hack-contest>; *see also* *Bug Bounty Program*, MOZILLA (Mar. 3, 2013), <http://www.mozilla.org/security/bug-bounty.html>.

other companies view this type of vulnerability research as an intrusion and attack on the integrity of their code.²¹³ As such, it is likely that a company will attempt to use the CFAA as a weapon against a security researcher on the basis of a contract breach argument.²¹⁴ Weaponizing breach in this manner short-circuits the technology policy discussion around white hat hacking and security vulnerability disclosure. Such chilling of this debate is highly undesirable; courts should seek to preserve, not destroy, this ongoing social policy and technology business conversation.

Finally, concerns regarding children's experimentation with technology warrant review. It is clearly in our national interest to encourage children's interest in technology: the next Bill Gates or Steve Jobs is likely in grade school today. Yet, as generations of children grow up as "digital natives"²¹⁵ using technology from their very early years, a world of weaponized breach may result in inquisitive six-year olds charged with CFAA offenses for registering their pets as social network users.²¹⁶ While I argue elsewhere that a strong form of the capacity doctrine with respect to children's right to disavow their agreements in most cases should be expressly extended to digital spaces, courts do not always adopt this approach.²¹⁷ A weaponized breach regime leaves children vulnerable when they are tinkering with code.

5. *Competing Legal Regimes*

Finally, the weaponized breach approach is appealing to litigants despite its flaws. It is an expedient path to imposing high costs on a breaching party, and litigants are utilizing it with increasing frequency.²¹⁸ Because of the

213. For example, Apple recently banned a security researcher from its application store after he demonstrated a vulnerability in iOS. See Alex Heath, *Apple Kicks Security Researcher Out Of The App Store After iOS Exploit Demonstration*, CULT OF MAC (Nov. 7, 2011), <http://www.cultofmac.com/128577/apple-kicks-security-researcher-out-of-app-store-and-developer-program-after-ios-vulnerability-demonstration/>.

214. Companies frequently adopt aggressive postures with security researchers. See Robert MacMillan, *Black Hat: ISS researcher quits job to detail Cisco flaws*, IDGNS, INFOWORLD.COM (July 27, 2005), <http://www.infolworld.com/d/security-central/black-hat-iss-researcher-quits-job-detail-cisco-flaws-088>.

215. See JOHN PALFREY AND URS GASSER, BORN DIGITAL (2008) (discussing how today's children are growing up as digital natives).

216. See *Nosal II*, 676 F.3d 854, 861 (9th Cir. 2012) (en banc) (noting that the government's asserted statutory construction would criminalize the conduct of "vast numbers of teens and pre-teens" under the CFAA).

217. Andrea Matwyshyn, *Generation C: Childhood, Technology and the Future of Identity*, 87 NOTRE DAME L. REV. 1979, 1983-89 (2012).

218. A query of the Westlaw ALLFEDS database with the search term "CFAA /20 contract" yields an upward trajectory in volume of cases. Between 1996 and 2005, thirteen cases fit these search criteria, but between 2006 and 2011, eighty-five cases fit these criteria. See ALLFEDS database, Westlaw, <http://www.westlaw.com> (Nov. 7, 2012).

convenience and potency of reaching for a weaponized breach CFAA allegation, plaintiffs will consequently be unlikely to invest in litigating more novel and legally appropriate measures of harm. In this way, weaponized breach potentially stunts the development of other bodies of law that could more successfully address exceptional harms that result from the involvement of technology.

For example, tort and other bodies of law may be better suited than weaponized breach to address financial harms that arise in a manner tangential to but not contemplated by contracts. As Professor Bellia argues:

[A]n approach recognizing a system owner's right to set the conditions of access, so long as she provides adequate notice of those conditions (through actual notice or adopting a system configuration that makes restrictions plain to the user), provides a better baseline for access to network resources than a pure liability rule or one requiring strong technical measures to trigger injunctive relief. In addition, where necessary to compensate for the inadequacies of this sort of property-rule protection, we must look to technology-displacing rules rather than pure liability or strong code-based approaches to achieve the appropriate level of access.²¹⁹

Through crafting a restrained, contract-based approach to digital harms in “contract hacker” cases, disputes which are better addressed through regimes other than contract and criminal law will become more clearly visible for legislators and judges. Yet the law of the horse cannot successfully develop in the shadow of the law of the zebra.

C. WHY COURTS MAY HAVE BECOME CONFUSED

Given the parade of horrors articulated in the sections above, it seems incomprehensible that any court would sanction a weaponized breach analysis. But a number of courts have indeed sent contract doctrine down this path of the law of the zebra. Although the reasons that courts select this destructive choice are not entirely clear, several theories exist.

Professor Ohm argues that a myth of a “superuser” with extraordinary technology skills emerged in policy circles in other legal contexts.²²⁰ Ohm argues that this mythology leads legislators to engage in overzealous regulation to stop the mythical force of this dangerous being.²²¹ Relying on the theory of moral panic of sociologist Stanley Cohen, Ohm argues that the

219. Patricia L. Bellia, *Defending Cyberproperty*, 79 N.Y.U. L. REV. 2164, 2272 (2004).

220. Paul Ohm, *The Myth of the Superuser: Fear, Risk, and Harm Online*, 41 U.C. DAVIS L. REV. 1327, 1338 (2008).

221. *Id.* at 1396–97.

superuser or “hacker” is a type of folk devil.²²² Ohm asserts that “[t]he trope arises in every single branch of Internet law, including intellectual property, computer crime, information privacy, information security, Internet governance, telecommunications, innovation policy, First Amendment law, and jurisdiction.”²²³

Ohm does not expressly reference contract law in his list of fields affected by this approach, but it is a logical addition. What Ohm calls a folk devil and mythology, I will reframe here as a problem of psychology.²²⁴ Psychology has long studied the heuristics that humans use to classify our existence, especially when we are confronted with novelty. Ascribing negative characteristics to objects or skills we cannot fully understand or fear reflects two psychological phenomena: essentialism and confirmation bias.

1. *Essentialism*

In developmental psychology, “essentialism” is a form of early cognitive bias describing the tendency to search for hidden, non-obvious features of things.²²⁵ A type of misguided technology essentialism is now visible in judicial contract analysis. Without recognizing their error, people frequently attempt to cope with new situations by trying to generalize existing knowledge to new categories and construct causal explanations in order to make sense of new information.²²⁶ They look for familiarity through the isolation of key shared characteristics across various people or objects. Each person or object of this “kind” then is ascribed a list of characteristics by default. This type of processing, known as “cognitive essentialism,”²²⁷ can give rise to incorrect generalizations, despite its utility in assisting humans in coping and learning. Sometimes this classification taxonomy results in a

222. *Id.* at 1337–38.

223. *Id.* at 1338.

224. I also must respectfully differ with Prof. Ohm when he states: “Computer experts rarely assess a risk of online harm as anything but, ‘significant,’ and they almost never compare different categories of harm for relative risk.” Paul Ohm, *The Myth of the Superuser, Part Three, The Failure of Expertise*, VOLOKH CONSPIRACY (Apr. 11, 2007, 1:09 PM), http://volokh.com/archives/archive_2007_04_08-2007_04_14.shtml#1176311368. As I have argued elsewhere, skilled information security professionals today apply scales of risk to determine the severity of various attacks. Many attacks are routinely found to be unlikely, low-priority and, ergo, not a significant risk. See Matwyshyn, *Hidden Engines*, *supra* note 48, at 139–40; see also Ryan Hurst, *The Contribution of a Security Practitioner*, UNMITIGATED RISK (May 27, 2010), <http://unmitigatedrisk.com/?p=6> (analyzing these dynamics from the perspective of experienced IT professionals).

225. SUSAN GELMAN, *THE ESSENTIAL CHILD: ORIGINS OF ESSENTIALISM IN EVERYDAY THOUGHT* 6–13 (2003).

226. *Id.*

227. *Id.*

classification fortuitously consistent with that of the external world, but sometimes the taxonomy is incoherent in context of the external world.²²⁸

A similar dynamic can occur when courts and legislatures encounter seemingly novel legal questions; cognitive biases inevitably creep into legal analysis because law is made by humans for humans. Courts sometimes erroneously fixate on computer use in their analysis of questions of breach. However, this computer use is merely a façade of essentialism and is substantively irrelevant. With increasing frequency, the digital breacher is being essentialized through his association with a computer. Immediately mislabeled “hacking” simply because the breach involves a computer, courts erroneously allow plaintiffs and prosecutors to reach for the CFAA. As discussed in Section III.B.2, *supra*, the various types of breaches of these exceptionalized “contract hackers” actually have little in common conceptually; hence, the one salient characteristic that appears to link them is the involvement of a computer.²²⁹

In addition to the essentialism visible in contract breach analysis, a similar dynamic exists with courts’ and Congress’s perceptions not only of “hackers” but of technology expertise in general—an “essentialism” of computer “experts” or “hackers.” In the recent debates over SOPA, for example, legislators frequently referenced, in arguably disrespectful fashion, “bring[ing] in some nerds” to explain the technology aspects on digital piracy.²³⁰ This act of “othering” technology experts in a derisive fashion is consistent with the harbingers of essentialist analysis in contract law performed by courts.

2. Confirmation Bias

Framing these observations another way, human decision-making also frequently reflects a “confirmation bias”—the tendency that many people hold to confirm their pre-existing beliefs.²³¹ Particularly when the topic is an emotionally-charged or threatening issue, confirmation bias is a common occurrence.²³² In essence, confirmation bias demonstrates a form of limited information processing: in lieu of processing information in a scientific way, when a person demonstrates confirmation bias he seek out information that

228. *Id.*

229. *See* discussion *supra* Section III.B.2.

230. *See, e.g.*, Kriss Kritto, *Congress Speak: Nerd*, THE HILL, Dec. 26, 2011, *available at* <http://thehill.com/capital-living/congress-speak/201345-congress-speak-nerd>.

231. *See, e.g.*, DANIEL HAHNEMAN, THINKING, FAST AND SLOW 62 (2011) (discussing confirmation bias).

232. *Id.*

tends to confirm his existing biases and beliefs about a given problem.²³³ Therefore, if a judge or jury holds fears about computers and their destructive potential, this finder of fact may be more susceptible to confirmation bias and fixate on the role of the computer in a breach.

Perhaps because contract disputes involving technology contexts seem novel and more threatening than other contract scenarios, judges may have unconsciously given in to essentialism and confirmation bias. Instead of preserving contractual supremacy and viewing technology exceptionalist legislation as merely a type of “gap filler,” judges have overreacted to the novelty of technology in contract. This overreaction should be doctrinally corrected; the law of the zebra should be avoided.

The next Part presents an operationalization of using the restrained technology exceptionalism paradigm for correcting the relationship between contracts, breach, and the CFAA, and resolving the circuit split on this issue.

IV. APPLYING RESTRAINED EXCEPTIONALISM: A PRIVACY MODEL OF AUTHORIZED ACCESS

Although legal literature discusses various aspects of the CFAA, contract law questions connected with this statute have thus far been analyzed primarily as a corollary to an argument based on CFAA analysis or rooted in another area of law. To the extent the civil applications of the CFAA²³⁴ and

233. See R.S. Nickerson, *Confirmation Bias: A Ubiquitous Phenomenon in Many Guises*, 2 REV. GEN. PSYCHOL. 175, 177 (1998); P.C. Wason, *On the Failure to Eliminate Hypotheses in a Conceptual Task*, 12 Q. J. EXPERIMENTAL PSYCHOL. 129, 132 (1960).

234. The EFF discusses the CFAA as follows:

The CFAA is primarily a criminal statute. However, in 1994 a civil suit provision was added that provides a private cause of action if a violation causes loss or damage, as those terms are defined in the statute. 18 U.S.C. § 1030(g) (2006). To state a civil claim for violation of the CFAA, (1) a plaintiff must allege damage or loss; (2) caused by a violation of one of the substantive provisions set forth in § 1030(a); and (3) conduct involving one of the factors in § 1030(c)(4)(A)(i)(I)–(V). *Id.* An action under this section must be brought within two years of the date the act is complained or the date of the discovery of the damage. No action may be brought under this subsection for the negligent design or manufacture of computer hardware, computer software, or firmware. *Id.* In 2008, Congress amended the CFAA through the Identity Theft Enforcement and Restitution Act, Pub. Law 110-326, 122 Stat. 3560. This amendment enhanced a number of aspects of the CFAA. Most notably, the 2008 amendment eliminated the need for Plaintiff’s loss to be greater than \$5,000 and made it a felony for a user to cause damage to ten or more computers. *Id.* Thus, while the previous \$5,000 threshold has been

breach have been analyzed, such discussions have been few,²³⁵ sometimes in the context of the prospective tort of cybertrespass.²³⁶

Arguing through the lens of property law, Professor Bellia argues that contract breach should not provide a basis for criminal CFAA prosecution. She states that “as a matter of statutory interpretation, coverage of unauthorized ‘access’ should be limited to activities that breach some technical limitation on access.”²³⁷ Professor Galbraith similarly argues against contract breach serving as a basis for criminal CFAA prosecution, particularly in the case of intellectual property harms on publicly accessible websites.²³⁸ Meanwhile, Professor Madison presents a thought-provoking analysis, asking whether confusion over metaphors of the internet as place or thing have led us to confusion regarding contract and computer intrusion.²³⁹

One of the most thorough and recent examinations of the subject of contract, breach, and CFAA criminal computer intrusion is that of Professor Kerr, who similarly notes that a breach of contract should not constitute an act of unauthorized access for purposes of CFAA criminal prosecution.²⁴⁰ He argues that violations of code-based restrictions and contract-based restrictions usually divide into two relatively easily distinguishable, discrete categories.²⁴¹ For Kerr, violations of code-based restrictions and whether the user has tricked the computer are the dispositive inquiries with respect to CFAA criminal analysis. He asserts that “[r]egulation by contract offers a significantly weaker form of regulation than regulation by code. Regulation

eliminated, a plaintiff still needs to show that they suffered damage or loss.

Computer Fraud and Abuse Act (CFAA), INTERNET LAW TREATISE, [https://ilt.eff.org/index.php/Computer_Fraud_and_Abuse_Act_\(CFAA\)](https://ilt.eff.org/index.php/Computer_Fraud_and_Abuse_Act_(CFAA)) (last visited Mar. 14, 2013).

235. See Bellia, *supra* note 219.

236. Peter Winn has also argued in favor of revitalizing cybertrespass and crafting a standard triggered by reasonableness of conduct. See Peter A. Winn, *The Guilty Eye: Unauthorized Access, Trespass and Privacy*, 62 BUS. LAW. 1395, 1428 (2007).

237. Bellia, *supra* note 219, at 2262. Professor Bellia’s approach successfully conceptually addresses many of the CFAA questions raised in my argument above, albeit through a more property-focused lens. One area where this property-based approach as well as Professor Kerr’s approach arguably fall short is addressing hybrid contract-code restrictions on access such as CAPTCHAs, which seem to be increasing in use. The privacy approach presented in this Article is capable of addressing hybrid code-contract scenarios more effectively.

238. Christine D. Galbraith, *Access Denied: Improper Use of the Computer Fraud and Abuse Act to Control Information on Publicly Accessible Internet Websites*, 63 MD. L. REV. 320, 324 (2004).

239. Michael J. Madison, *The Narratives of Cyberspace Law (Or, Learning from Casablanca)*, 27 COLUM. J.L. & ARTS 249 (2004).

240. Orin Kerr, *Cybercrime’s Scope: Interpreting “Access” and “Authorization” in Computer Misuse Statutes*, 78 N.Y.U. L. REV. 1596, 1600 (2003).

241. *Id.*

by code enforces limits on privileges by actually blocking the user from performing the proscribed act, at least absent circumvention.”²⁴² Professor Kerr’s analysis is clearly appropriate for the first category of intruder; strictly third-party criminal context of accessing a “protected computer” where no contractual relationship of any kind is formed.²⁴³

However, Kerr’s approach is somewhat less satisfying in current commercial contexts. Many commercial contexts today present a blended set of code-based and contract-based restrictions: for example, CAPTCHAs or human interactive proofs provide a case of a hybrid code-based and contract-based restriction.²⁴⁴ Is employing a bot that acts like a human to solve CAPTCHAs a violation of a code-based restriction for Kerr? Is circumventing an (easily-avoided) end user license agreement text copy-protection measures also an act of computer intrusion? If a consumer accesses content on the *New York Times* website and avoids a technical restriction in a manner requiring no technical skill—perhaps his cat stands on the delete key and deletes the last piece of an article’s URL²⁴⁵—is he properly classified a “hacker” for CFAA purposes, despite already being bound by the *New York Times* terms of use contract?²⁴⁶ The dynamics of code and contract are becoming increasingly interwoven, and the definition of a “code-based”

242. *Id.*

243. 18 U.S.C. §1030(a)(2)(A)–(C) (2011). A “protected computer” is one that is exclusively for the use of the U.S. government or a financial institution, or if not exclusively for such use, when a computer is used by or for the U.S. government or a financial institution and the conduct constituting the offense affects that use; or is used in interstate or foreign commerce or communication. *Id.* §1030(e)(2). A person who obtains anything of value by accessing a protected computer, knowingly, without authorization and with the intent to defraud, violates the CFAA unless (a) the only thing of value that is obtained is the use of the computer itself, and (b) the use is valued at less than \$5,000 during a one-year period. *Id.* §1030(a)(4).

244. A CAPTCHA (Completely Automated Public Turing test to tell Computers and Humans Apart) is a code-based test to determine whether a particular user is a human or code, such as a bot. CAPTCHAs usually accompany a requirement that the user agree to terms of use of a website. For a discussion of CAPTCHAs, see *Telling Humans and Computers Apart Automatically*, GOOGLE, <http://www.google.com/recaptcha/captcha> (last visited Nov. 18, 2012).

245. See, e.g., Lauren Indvik, *How to hack the New York Times Paywall... With Your Delete Key*, MASHABLE, (Mar. 28, 2012), <http://mashable.com/2011/03/28/how-to-bypass-new-york-times-paywall/> (describing how merely deleting the last portion of the New York Times URL would break the website’s copy protection on articles at the time of this author’s writing).

246. I submit that none of these qualify as triggering the CFAA. Each of these acts can be analyzed to have occurred within the context of a contractual relationship. As such, the privity model deems CFAA analysis inappropriate.

restriction has become an increasingly slippery slope. As the discussion of the four types of contract “hackers” demonstrates, gray areas abound.

I strongly agree in principle with Professor Kerr, Professor Bellia, and Professor Galbraith’s conclusion that a breach of contract does not provide the basis for a criminal CFAA prosecution (or any other technology exceptionalist prosecution for that matter), but I submit that this question should be framed differently. The mere existence of an enforceable contract between the parties that covers the information at issue is dispositive. The existence of this contract trumps the applicability of the CFAA. Therefore, framing a contract inquiry with the CFAA as a starting point is overly deferential to the CFAA: this is the law of the zebra. Further, this framing captures only half of the problematic cases: I submit that a breach of contract also does not provide the basis for any civil CFAA action for the same reason as it does not support criminal actions. Contract law and other already existing legal regimes can continue to effectively address the types of harms implicated by the four “contract hackers” described in Section III.A, *supra*.

In the privity model that follows, I fundamentally invert Kerr’s argument. Contract should be considered a superior and stronger form of regulation to that of code and technology exceptionalist approaches such as the CFAA: CFAA analysis is only appropriate provided there is no contract analysis possible. While the behaviors of code can cast doubt on the meaning of a contract term, as I describe in the next Part, if a contract has been validly formed, contractual agreements hold a superior position to CFAA analysis. They override the need for it, both civilly and criminally.

The privity model that follows thus takes into account hybrid contract-code restrictions. I argue that as a matter of contract law, when ambiguities arise in technology breach contexts and both contract and code breaches may exist, a contract-based analysis must control and ambiguities must be construed against the drafter of the contract at issue. In other words, I am arguing for the blanket supremacy of contract law over computer intrusion analysis where a contract between the parties giving access to the information in question was properly formed. Any other analysis results in an exceptionalized and undesirable construction of contract breach in technology contexts that threatens the future of contract law.²⁴⁷ This “privity” model is explained below.

247. Additionally, breach weaponized with the CFAA presents the constitutional void-for-vagueness concerns that Professor Kerr addresses elsewhere. *See* Kerr, *supra* note 109, at 1562.

A. THE “PRIVITY” MODEL: CONTRACT TERMS AND NORMS OF ACCESS

In this Part, I propose a “privity”²⁴⁸ model of understanding the relationship of contract and the CFAA. This privity model provides an operationalization of the paradigm of restrained technology exceptionalism introduced in Part II, *supra*. Embodying these principles, the privity model avoids the pitfalls of a weaponized breach analysis. It expressly divorces contract law analysis from CFAA analysis and frames contract as the superior regime. It does not, however, eliminate any other remedies that already exist in law apart from the CFAA: if the conduct in question is otherwise criminal or tortious, those penalties remain.

As the name of the model implies, the privity model is an analytical framework that gives supremacy to contract law analysis over CFAA analysis. As courts have recognized in various contexts, including even with respect to copyright, a contract inquiry is not preempted as an initial matter over other areas of law.²⁴⁹ But for very limited circumstances,²⁵⁰ contract law controls when an agreement exists between the parties: as *ProCD, Inc. v. Zeidenberg* suggests,²⁵¹ where a contract between the parties exists, contract law is usually not preempted.²⁵²

248. Privity refers to the contract law concept of the existence of a bilateral contractual relationship that gives the parties certain rights with respect to each other. For a discussion of privity, see, e.g., J.W. Neyers, *Explaining the Principled Exception to Privity of Contract*, 52 MCGILL L.J. 757, 760–63 (2007). Much like the concept of privity, the model focuses on the existence of a contractual relationship as the focal point of analysis.

249. See generally *ProCD, Inc. v. Zeidenberg*, 86 F.3d 1447, 1454–55 (7th Cir. 1996).

250. In *Rano v. Sipa Press, Inc.*, the Ninth Circuit held that copyright preempted state law relating to the termination at will of a license with an indefinite duration because “California law and federal law are in direct conflict, federal law must control.” *Rano v. Sipa Press, Inc.*, 987 F.2d 580, 585 (9th Cir. 1993). Assignability of a licensee’s rights would provide another preemption basis because under federal law such rights cannot be assigned in a nonexclusive license without the consent of the licensor. See *In re CFLC, Inc.*, 89 F.3d 673 (9th Cir. 1996); cf. *Chamberlain v. Cocola Assocs.*, 958 F.2d 282 (9th Cir. 1992) (applying California statute regarding transfer of tangible object in case of transfer of intangible rights to use object).

251. *ProCD, Inc. v. Zeidenberg* was the first appellate ruling dealing with the enforceability of shrinkwrap licenses and that the contract restrictions it placed on the use of a noncopyrightable database were not preempted by copyright law. *ProCD*, 86 F.3d at 1454–55; see also *Davidson & Assocs. v. Jung*, 422 F.3d 630 (8th Cir. 2005) (holding that license not preempted by fair use); *Altera Corp. v. Clear Logic, Inc.*, 424 F.3d 1079, 1080 (9th Cir. 2005) (copyright does not preempt contract enforcement, citing *Zeidenberg*); *DaimlerChrysler Servs. N. Am., LLC v. Summit Nat., Inc.*, 144 Fed. App’x. 542, 545 (6th Cir. 2005) (holding that copyright defenses irrelevant to contract enforcement); *Bowers v. Baystate Techs., Inc.*, 320 F.3d 1317, 1325 (Fed. Cir. 2003) (holding that license not preempted by fair use).

252. See Guy A. Rub, *Winter, 2011 Symposium: the Licensing of Intellectual Property: Contracting Around Copyright: The Uneasy Case For Unbundling Of Rights In Creative Works*, 78 U. CHI. L. REV. 257, 257 (2011).

By framing the analysis from a contract law perspective, the privity model limits the reach of the CFAA and enables it to act as the basis for tort-like and equitable remedies (and criminal prosecution) only in cases of truly novel “hacking”²⁵³ harms that happen outside the realm of privately ordered contractual relationships. In other words, the privity model converts the CFAA from being an instrumentality of the law of the zebra into a law of the horse approach—an approach limited to those circumstances where legal gaps cannot be successfully filled by traditional law.

1. *The Privity Model*

A privity model of contract breach and CFAA analysis can be summed up in one sentence: In the case where a contract was properly formed between the parties and a breach occurs, the use of a computer in connection with the breach is irrelevant. The CFAA is not implicated.

As demonstrated by Figure 1, *infra*, the privity model arises from a simple distinction in contract law: that breach can only arise when a contract has been properly formed.²⁵⁴ As such, a privity model starts with a formation inquiry: has an oral or written contract been formed between the parties? Are the parties in privity with each other? Courts should define this act of contract formation broadly. Just as contracts can be formed in both oral and written form in physical space contexts, so too the mere extension of technology credentials, such as logins and passwords that allow access to a portion of a network, constitutes the formation of at least an oral agreement granting access to the extent of the credentials’ authorization inside the system.

If the terms expressly give access to the allegedly violated information resource, the matter is a regular contract breach. Traditional contract breach analysis applies, and no claim under the CFAA is appropriate. If the terms ambiguously discuss access to the information, the contract is construed against the drafter, and a contract law claim for breach with traditional remedies is the appropriate recourse. No CFAA claim is appropriate.

253. I use the word hacking here colloquially. A “hacker” in the mind of an average consumer, as reflected by media depictions, refers to someone who possesses specialized computer skills and uses them to circumvent code-based restrictions on access to networks. Reasonable people will differ about the desirability and criminality of certain types of *Bowers v. Baystate Techs* “hacking” conduct, such as white hat hacking of a database to determine the existence of a dangerous vulnerability in the code. That question is beyond the scope of this Article.

254. *See Fontanella v. Marcucci*, 877 A.2d 828, 834–35 (Conn. App. Ct. 2005) (“To recover for breach of contract, the plaintiffs must establish the formation of an agreement, performance by one party, breach of the agreement by the other party, and damages.”).

Construing any vagueness against the drafter is an analysis consistent with a traditional contract law approach. It is also the more equitable position: the drafter is in a better position to choose the desired scope of access with respect to information security default settings, both contractually and in practice. Similarly, the drafter is in the superior position to terminate the rights of access in both of these granting vehicles.²⁵⁵

If the access that permitted the alleged information theft was not expressly contemplated in the contract, however, the information security access settings of the network, profile, or information at issue are incorporated as implied terms of the contract. The contract is then construed against the drafter, and a traditional breach analysis is appropriate. CFAA remedies are again inapposite. An aggressive damages calculation for information harms can be constructed through coupling expectation damages with an injunction and disgorgement damages for extreme cases. This type of an aggressive damages regime would enable traditional contract remedies to sufficiently compensate damages plaintiffs for information harms in the vast majority, if not all, information theft instances arising out of a contractual relationship.

If, however, the access that occurred was permitted neither by the express language of the contract nor by the information security settings of the information owner at the time, then the particular act may fall outside the scope of the contractual relationship. In such case, a CFAA claim may be appropriate. However, in this case, the CFAA claim arises *as an independent claim that is unrelated to the contract. The basis for a CFAA claim does not arise out of the contract breach.* Even in a situation where a contractual relationship existed and was terminated, a CFAA claim is appropriate only if termination was done correctly. Access terminates through a notification of termination of the agreement itself in accordance with the terms stipulated in the agreement and, simultaneously, the reality of access must be terminated using information security controls. A failure to either provide notice or to terminate access in fact can be rightfully construed as a failure to terminate by the drafter and construed against him. If termination of access was incomplete or incorrect, no CFAA claim is appropriate—the existing contract was still in effect and a breach claim is appropriate.

The privity model in this way encompasses a greater scope of activities and recourse than Professor Kerr's model, which focuses only on code-based

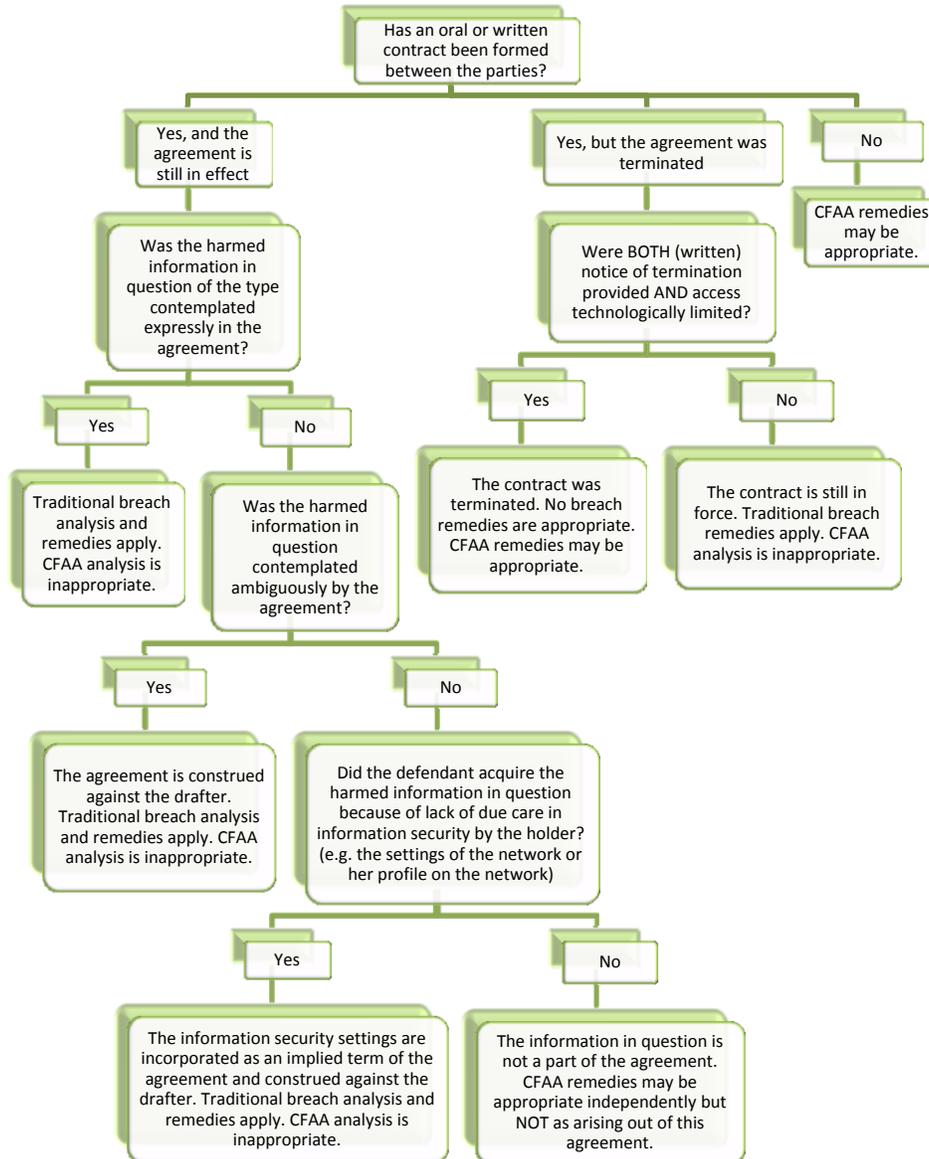
255. See, e.g., David Horton, *Flipping the Script: Contra Proferentem and Standard Form Contracts*, 80 U. COLO. L. REV. 431 (2009) (discussing the contract law rule of construction of terms against the drafter).

regulation and criminal CFAA application. The privity model also extends the reach of contract law to address hybrid contract-code situations such as a human interactive proofs, an approach in line with the suggestion made by Professor Bellia regarding the merits of contemplating hybrid law and code notice-based approaches.²⁵⁶ Most importantly, however, the privity model minimizes the extent to which technology is treated as “special.” It also places an affirmative burden of care in information security on the information holder, limiting the holder’s ability to weaponize contractual relationships. This approach is consistent with the duty to mitigate damages in contract law, as well as and basic principles of due care in information security. It also ensures that both the civil and criminal penalties of the CFAA are used only in extraordinary circumstances: to address truly novel harms arising out of technology, not as a means of contracting parties inflicting transaction costs on each other.

The privity model will next be applied to each of the four “contract hacker” situations in the following Section, demonstrating its effectiveness at handling the vast majority of the breach questions that arise in a technology contracts.

256. *See* Bellia, *supra* note 219, at 2272–73 (arguing that “the literature has thus far neglected the complex relationship between law and technical measures in this context—in particular, the possibility that too-weak legal protection will induce greater reliance on too-strong technical measures, whether or not the law in fact backs those measures.”).

Figure 1: The “Privity” Model of Contract and CFAA Remedies



B. HANDLING THE FOUR “CONTRACT HACKERS”

The privity model offers contract law tools to address the harms caused by each of the four “contract hackers.”

1. Users

With respect to Grandma, Dr. Whiskers, and end users in general, it is highly unlikely that any basis for a CFAA claim would exist under the privity

model. For example, if a user clicks on a portion of a site that is improperly coded and finds herself funneled into an interface which should be protected with restricted access, actual access capabilities fall within the scope of use permitted by the end user license agreement/terms of use. The privity model would thus incorporate the reality of this access as an implied term of the user agreement.

Grandma's use of Dr. Whiskers' information rather than her own on the website may indeed be a breach of contract. However, such a breach of contract would likely carry with it a very small remedy in damages, if anything, because actual damages would be difficult to prove on the part of the website owner. As such, the distinction of minor versus material breaches becomes intrinsically important again in contractual relations, and no CFAA civil or criminal claim would exist.

2. *Employees and Business Partners*

In the context of employee and business partner "contract hackers," the existence of the contractual relationship controls under the privity model if accessed information is misused. As one court correctly analyzed the question of employee "contract hackers," a distinction must be drawn between the existence of a contract that grants authorization to access and the bad use of information in a manner which breaches the contract. In *International Association of Machinists and Aerospace Workers v. Werner-Masuda*, the defendant, a union officer, was charged with exceeding her authorization to use a computer when she violated the terms of a terms-of-use agreement which granted her access to a membership list.²⁵⁷ The court rejected the application of § 1030 of the CFAA, holding that even if the defendant breached a contract, that breach of a promise did not mean her access to that information was unauthorized or criminal. In the words of the court:

[Even if] Werner-Masuda may have breached the Registration Agreement by *using* the information obtained for purposes contrary to the policies established by the [union] Constitution, it does not follow, as a matter of law, that she was not authorized to access the information, or that she did so in excess of her authorization in violation of the . . . CFAA . . . the gravamen of [the plaintiff's] complaint is not so much that Werner-Masuda improperly accessed the information . . . but rather what she did with the information once she obtained it. . . . Nor do [the] terms [of the CFAA]

257. *Int'l Ass'n of Machinists and Aerospace Workers v. Werner-Masuda*, 390 F. Supp. 2d 479, 495–96 (D. Md. 2005).

proscribe authorized access for unauthorized or illegitimate purposes.²⁵⁸

This distinction explained by the court preserves the role of contract breach in technology contracting. When a rogue insider such as an employee or a business partner chooses to inflict commercial harm that includes a technology-based operationalization—meaning involving a computer—this harm comes usually in the form of copied information that is published or shared with a competitor (potentially eviscerating trade secret protection), destruction of information assets inside the company, or using access to the corporate network to inflict technology harms to third parties. In each of these situations, both confidentiality agreements and separate agreements related to the business relationship should exist between the parties. Provided these agreements are well drafted, they will offer the basis for a full recovery for losses. With respect to the first and second scenario of damaged intellectual property assets, in addition to a suit in contract, the harmed party has an additional option to sue the breacher (and any recipient of information) for theft of trade secrets, and any party receiving information for possible tortious interference with contract. In the case of use of the corporate network to inflict technology pain on innocent third parties, breach of contract will cover any actual losses suffered by the company. Meanwhile, the third-party victims of the technology harms retain recourse against the individual under the CFAA—no contract controls that relationship and the parties do not stand in privity. As such, regardless of how technology is used in the course of the breach, a combination of a restitution and a disgorgement calculation of damages in contract, particularly if coupled with a trade secret remedy, can make a harmed party whole in the instance of a “contract hacker” insider.

3. *Entrepreneurs*

A similar analysis holds when an application developer or data aggregator is sued by an information holder or content owner. With respect to the

258. *Id.* at 499 (citations omitted); *see also* Shamrock Foods v. Gast, 535 F. Supp. 2d 962, 967 (D. Ariz. 2008) (holding the CFAA did not apply when an employee emailed himself files for the benefit of a rival company in violation of the defendant’s confidentiality agreement); Diamond Power Int’l, Inc. v. Davidson, 540 F. Supp. 2d 1322, 1343 (N.D. Ga. 2007) (rejecting a CFAA claim against an employee who violated an employment agreement by using his access to his employer’s computer system to steal data for a competitor and holding that “a violation for accessing ‘without authorization’ occurs only where initial access is not permitted,” while explaining further that a violation for ‘exceeding authorized access’ occurs where initial access is permitted but the access of certain information is not permitted).

“contract hacker” entrepreneurs, the question again turns on whether a contract was formed. If a contract exists between the particular entrepreneur and the information holder who alleges that the entrepreneur has wrongly accessed information, then the privity model applies. If, however, contract formation has not occurred between the entrepreneur and the information owner, then the analysis may indeed implicate the CFAA.

In the case where the entrepreneur is acting through an assignment of access right from a consumer user,²⁵⁹ the technological reality of the assignability of the credential controls. In other words, the contract rights of the user to access the website can be assigned to the entrepreneur by the consumer user. Although, again, this may be a contract breach on the part of the user and/or the entrepreneur, the access is through a contract right. If the information owner chooses to use technological means to restrict the access of the particular entrepreneur or all such entrepreneurs, this blocking is the prerogative of the information owner. Similarly, the information owner may seek to terminate its agreement with all users relying on the disfavored products.

The privity model also offers guidance in data aggregation cases. The analysis yet again turns on whether a contract was formed. The inquiry may be framed by a court in a manner which analyzes whether the code of the data aggregator behaved in a way consistent with the behaviors that a human user would demonstrate. If the code of the data aggregator behaved in a humanlike way—simply stepping into the shoes of an authorized human user—and a contract was indeed formed between the data aggregator and the information holder, then a contract law analysis applies. If no contract was formed or if the contractual relationship was appropriately terminated, then CFAA analysis may be appropriate.

It is in these cases where the potency of the remedy of an injunction in contract becomes apparent. Let us presume that the data aggregator, posing as a human user, formed an agreement with a ticketing website, and it sends queries in a manner that are human-like but in exceedingly high volume. The ticketing website may then notify the data aggregator that the access contract between the data aggregator and the website is terminated and that no additional access is appropriate. However, if the ticketing website creates

259. For example, the program Vtok is an application that facilitates communication through Google talk, text, and video on iPhones. Because no official Google chat application is available on iPhone, users rely on third-party applications such as Vtok to access their accounts. In order to access a user's Google account, Vtok requires the user's login credentials. The access has been expressly authorized by the user, but the user's access is intermediated by Vtok. See VTOK, <http://www.vtokapp.com> (last visited Mar. 15, 2013).

agreements with its users through the standard method of a terms of service contract on the homepage, i.e. forming a contract with anyone who cares to visit the website, each time the data aggregator revisits that homepage, a new contract is formed and access is again granted. Hence, if the ticketing website wishes to prevent the data aggregator from accessing the site and entering into additional contracts with the site, seeking injunctive relief from a court may be the appropriate remedy. A breach of contract, however, should not provide a basis for a CFAA claim simply because the ticketing website is technologically incapable of preventing the data aggregator from posing as a human and accessing the site.

4. *Security Researchers*

With respect to information security researchers, provided that an agreement has been formed through an end user license agreement, any acts of reverse engineering, decompiling, and any other analysis, even if these acts constitute a breach of the terms of the agreement, would not provide any basis for a CFAA civil suit or criminal prosecution. Although the security researcher may still be open to claims under the Digital Millennium Copyright Act and other intellectual property regimes, those are outside the scope of the privity model. Notably, none of these four contract hacker scenarios impacts the use of the CFAA in its traditional sphere of application—criminally prosecuting third party unaffiliated hackers who do not sit in contractual privity with an information holder or content owner.

C. CRAFTING GOOD NORMS IN INFORMATION SECURITY

Because the privity model provides an operationalization of restrained technology exceptionalism, it creates a line between contract law and other areas of law that has been sometimes blurred in recent case law. It avoids the list of deleterious consequences to contract law that were set forth in Part II and it resolves the fact-specific problems of the four contract hackers introduced in that Part. Further, the privity model allows for other areas of law to develop in this space around information regulation.

In particular, because the privity model incorporates the realities of information security control and construes them as an implied term of the contract, it creates a strong incentive for good information security behaviors on the part of information holders. As I have argued elsewhere, the incentives for good information security behaviors are not always apparent to information holders.²⁶⁰ As rampant information security breaches

260. Andrea M. Matwyshyn, *Material Vulnerabilities: Data Privacy, Corporate Information Security and Securities Regulation*, 3 BERKELEY BUS. L.J. 129 n.174 (2005).

demonstrate, information management success and information security practices vary dramatically across entities, and widespread deficits exist.²⁶¹ One of the most basic principles of information security is the principle of least privilege—the idea that access to information should be granted in as stingy a manner as possible, with the fewest people possible having unfettered access and with others having the least amount of access necessary to fulfill responsibilities.²⁶² The privity model embodies this idea of least privilege and blends the technological and contractual access limitations. This approach is aligned with the spirit of the CFAA, which, as Professor Bellia notes, is a statute that “does in fact strike a balance that heavily favors technical measures as a predicate for legal protection.”²⁶³

By creating rights of recourse for information holders that are contingent on exercise of care in their information handling, both the information holders and the rest of the information ecosystem benefit. I have explained elsewhere that neglecting security of key information assets can contribute to meaningful devaluation of corporate assets, a devaluation that potentially can be construed as a breach of fiduciary duty.²⁶⁴ Further, the underlying approach of strong information security practices is fundamentally consistent with the duty to mitigate damages that already exists in contract law.²⁶⁵ Even if a party is blameless in a breach, that party nevertheless has a duty in contract law to mitigate the damages resulting from that breach.²⁶⁶ When an information holder chooses to exercise poor information control practices, the information holder runs afoul of the spirit of this duty to mitigate the damages arising from a breach. Similarly, trade secret law relies on the holder of the allegedly secret information to maintain reasonable precautions of secrecy.²⁶⁷ As such, the privity model penalizes lack of due care in information handling by the information holder in a manner consistent with both contract and trade secret law.

261. See, e.g., DATALOSS DB, <http://www.datalossdb.org> (last visited Feb. 18, 2012) (listing information security breaches); PRIVACY RIGHTS CLEARINGHOUSE, <http://www.privacyrights.org> (last visited Feb. 18, 2012) (same).

262. See DEPARTMENT OF DEFENSE, TRUSTED COMPUTER SYSTEM EVALUATION CRITERIA 109 (1985), available at <http://src.nist.gov/publications/history/dod85.pdf>.

263. Bellia, *supra* note 219, at 2271.

264. Matwyshyn, *Imagining the Intangible*, *supra* note 210, at 990.

265. See RESTATEMENT (SECOND) OF CONTRACTS § 350 (1981).

266. *Id.*

267. Unif. Trade Secrets Act, 14 U.L.A. 433 (1990).

V. CONCLUSION

This Article warns of the “law of the zebra,” an inherently disruptive paradigm seeping into contract law that prioritizes technology driven analysis over traditional contract law. In place of the law of the zebra, this Article argues in favor of restoring contract law to its traditional trajectory. It seeks to craft a minimally disruptive, predictable approach to analyzing contract law questions in technology contexts by offering the paradigm of “restrained technology exceptionalism”—a paradigm based on traditional contract law principles. In particular, through one operationalization of this paradigm in a “privity model,” this Article offers a suggested roadmap for judicial interpretation of the CFAA contract hacker cases that are currently causing a circuit split in the federal courts.

EXACTITUDE IN DEFINING RIGHTS: RADIO SPECTRUM AND THE “HARMFUL INTERFERENCE” CONUNDRUM

Thomas W. Hazlett[†] and Sarah Oh^{‡‡}

ABSTRACT

In the century since the Radio Act of 1912 initiated U.S. spectrum allocation rules, a precise definition of “harmful interference”—the control of which forms the rationale for regulation—has eluded policymakers. In one sense, that result is unsurprising; rights are always defined incompletely. In another sense, however, the regulatory system is dysfunctional, severely limiting the productive use of spectrum while locked down in years-long border disputes. These disagreements have, in turn, triggered calls to develop brighter lines and fuller engineering specifications of harmful interference. However, this emphasis on exact definitions is misguided. Spectrum use rights generate more robust market development when they feature technically fuzzy borders but are awarded in economically efficient bundles. The key ingredients are (a) exclusive, flexible rights; (b) frequency borders set via standardized edge emission limits; (c) large bundles of complementary rights that limit fragmentation; and (d) fluid secondary trading that allows mergers to end border disputes by eliminating borders. Regulators should focus less on delineating precise interference contours, and instead expeditiously distribute standard bandwidth rights to economically responsible agents, taking care to avoid undue fragmentation (and tragedy of the anti-commons). Many episodes illustrate these lessons, including those involving reallocation of the broadcast TV band, the emergence of HD radio, the Nextel/public safety “spectrum swap,” and the ongoing WCS/SDARS dispute. Each instance reveals that economic incentives, not engineering complexity, drive—or block—productive coordination of radio spectrum use.

© 2013 Thomas W. Hazlett and Sarah Oh. All Rights Reserved.

[†] Professor of Law and Economics, George Mason University School of Law; FCC Chief Economist, 1991–1992. This Article follows from a presentation on *Creating Efficient Spectrum Property, Towards an Economic Definition of “Harmful Interference” in Radio Transmissions*, Presentation at the Conference on Spectrum Markets: Challenges Ahead, Kellogg School of Management, Northwestern University (June 2–3, 2011). The author thanks Brent Skorup for excellent research assistance.

^{‡‡} Research Fellow, Information Economy Project. Ph.D. Student in Economics, George Mason University; J.D. George Mason University; B.S. Stanford University, Management Science & Engineering.

TABLE OF CONTENTS

I.	INTRODUCTION	229
A.	THE CRISIS IN U.S. SPECTRUM POLICY.....	229
B.	EXACTITUDE IN SPECTRUM USE RIGHTS BOUNDARIES AT THE FCC.....	231
C.	THE FCC'S APPROACH TO SPECTRUM USE RIGHTS DEFINITIONS.....	232
D.	LIBERAL LICENSES THAT REMEDY TRAGEDIES OF THE ANTI- COMMONS.....	236
E.	DEFINING HARMFUL INTERFERENCE AS IF ECONOMICS MATTERS.....	239
II.	EXACTITUDE	240
III.	THE TROUBLE WITH EXACTITUDE: EXAMPLES	245
A.	TV BAND REALLOCATION: PROBLEMS WITH RIGID RIGHTS.....	245
B.	HD RADIO: BUNDLING RIGHTS.....	247
C.	LICENSED PCS VERSUS UNLICENSED PCS: ADMINISTRATIVE RULES VERSUS MARKET MECHANISMS.....	251
D.	WCS/SATELLITE RADIO: THE BENEFITS OF MERGER.....	254
E.	NEXTEL/PUBLIC SAFETY RADIO: ELIMINATING BORDERS.....	256
F.	INTERFERENCE TEMPERATURE: A FAILED ATTEMPT AT GREATER CLARITY.....	269
IV.	BEYOND EXACTITUDE: ENABLING OPTIMAL COMBINATIONS	277
A.	BASIC STRATEGY.....	277
B.	LICENSED AND UNLICENSED DEPLOYMENTS IN AN EFFICIENT MIXED USE REGIME.....	282
C.	RULES TO ENABLE MARKETS.....	288
D.	OBSERVED EFFICIENCIES IN EMERGING SPECTRUM MARKETS.....	292
E.	MOBILE OPERATORS AS SPECTRUM OWNERS AND NEIGHBORS.....	301
F.	THE COSTLY AND UNPRODUCTIVE SEARCH FOR ADDITIONAL SPECIFICITY.....	305
G.	THE AM RADIO CRITIQUE.....	314
H.	LIABILITY RULES VERSUS PROPERTY RULES.....	320
I.	OPTIMAL SPECTRUM RIGHTS DEFINED.....	323
J.	AUCTIONS AS INTERFERENCE CONTROL MECHANISMS.....	324
K.	OVERLAYS.....	327
V.	CONCLUSION	335

I. INTRODUCTION

“Regardless of how or to whom particular rights are assigned, ensuring that all rights are *clearly delineated* is important to avoiding disputes, and provides a clear common framework from which spectrum users can negotiate alternative arrangements.”

—Federal Communications Commission (2002)¹

“Commenters . . . almost uniformly cited the FCC’s interference rules as the prime example of rules that are not *clearly defined*. A common refrain was that the FCC rules speak of the right to be protected from “harmful interference,” but this term is not defined in technical terms”

—Federal Communications Commission (2002)²

A. THE CRISIS IN U.S. SPECTRUM POLICY

There is officially a crisis in U.S. spectrum policy. Congressionally chartered studies,³ top U.S. policymakers,⁴ and the Presidential Administration,⁵ citing the need for additional wireless bandwidth for economic growth, have pushed for dramatic improvements in the process whereby the Federal Communications Commission (“FCC”) allocates the radio spectrum. The emergence of wireless 3G and 4G data networks, as well as a slew of popular handsets and applications after the 2007 introduction of

1. FCC, SPECTRUM POLICY TASK FORCE REPORT, ET DOCKET NO. 02-135, 18 (Nov. 2002) [hereinafter SPTFR 2002] (emphasis added), *available at* http://hraunfoss.fcc.gov/edocs_public/attachmatch/DOC-228542A1.pdf.

2. *Id.* (emphasis added).

3. The National Broadband Plan, issued by the FCC, was mandated in the American Recovery and Reconstruction Act of 2009 (also known as “the stimulus bill”). The FCC issued its report in March 2010; Chapter 5, “Spectrum,” deals with the issues discussed in this Article. FCC, CONNECTING AMERICA: THE NATIONAL BROADBAND PLAN, CH. 5: SPECTRUM 73 (2010) [hereinafter NBP 2010], *available at* <http://www.broadband.gov/download-plan/>.

4. Lawrence H. Summers, Technological Opportunities, Job Creation, and Economic Growth, Remarks at the New America Foundation on the President’s Spectrum Initiative (June 28, 2010), <http://www.whitehouse.gov/administration/eop/nec/speeches/technological-opportunities-job-creation-economic-growth>; Julius Genachowski, Chairman, Federal Communications Commission, Prepared Remarks for the 2011 International Consumer Electronics Show (Jan. 7, 2011), http://hraunfoss.fcc.gov/edocs_public/attachmatch/DOC-303984A1.pdf.

5. White House, Office of the Press Secretary, Presidential Memorandum: Unleashing the Wireless Broadband Revolution (June 28, 2010), <http://www.whitehouse.gov/the-press-office/presidential-memorandum-unleashing-wireless-broadband-revolution>.

the iPhone, triggered a “mobile data tsunami” that has prompted policy makers to respond to the demands of the market.⁶

But institutional stasis hampers this effort at virtually every step. In March 2010, the National Broadband Plan set a goal of making available another 300 MHz of spectrum (a bit more than one-half of that currently available to mobile carriers) by 2015.⁷ Yet today, every one of the proceedings comprising that plan is either behind schedule or has been abandoned.⁸ Scholars are unshaken in their assessment that “[t]he FCC’s traditional system for managing the radio spectrum is a paradigm of economic inefficiency.”⁹

Prominent communications experts advance the following diagnosis of the underlying problem: Regulators have failed in the century of radio regulation since the Radio Act of 1912¹⁰ to precisely define radio “interference.”¹¹ Such conflicts have been of central importance to regulators, who are charged with creating rules for wireless operations that avoid the “tragedy of the commons.”¹² Arguably, the delineation of border

6. See Thomas W. Hazlett, Roberto E. Muñoz & Diego B. Avanzini, *What Really Matters in Spectrum Allocation Design*, 10 NW. J. TECH. & INTELL. PROP. 93, 95 (2012).

7. See NBP 2010, *supra* note 3, at 75.

8. This assessment is per FCC Commissioner Ajit Pai. See John Eggerton, *Pai Concerned About Where FCC Is Headed on Big Issues*, BROADCASTING & CABLE (July 18, 2012, 10:59 AM), http://www.broadcastingcable.com/article/487370-Pai_Concerned_About_Where_FCC_Is_Headed_on_Big_Issues_.php.

9. Dale Hatfield & Phil Weiser, *Toward Property Rights in Spectrum: The Difficult Policy Choices Ahead*, CATO INST. POLY ANALYSIS NO. 575, at 4 (Aug. 17, 2006).

10. The Radio Act instructed the Secretary of Commerce & Labor to issue licenses to those wishing to transmit radio signals, and in doing so to “minimize interference.” This authority remained with the Department of Commerce when the Department of Labor was split off in 1913. Radio Act of 1912, CH. 287, § 4, P.L. NO. 264, 37 STAT. 302 (1912) (“That for the purpose of preventing or minimizing interference . . . said private and commercial stations shall be subject to the regulations of this section. These regulations shall be enforced by the Secretary of Commerce and Labor”); An Act to Create a Department of Labor, P.L. NO. 62-426, § 1, 37 STAT. 736 (1913).

11. 47 C.F.R. § 2.1 (2010) (“Harmful Interference[.] [i]nterference which endangers the functioning of a radionavigation service or of other safety services or seriously degrades, obstructs, or repeatedly interrupts a radiocommunication service operating in accordance with [the ITU] Radio Regulations.”); see also *id.* (“Accepted Interference[.] [i]nterference at a higher level than defined as permissible interference and which has been agreed upon between two or more administrations without prejudice to other administrations.”) (citation omitted).

12. The “tragedy of the commons” relates to a situation in which the distribution of ownership rights preempts resource optimization. The classic formulation, by biologist Garrett Hardin, focuses on the “over-grazing” problem in a pasture that is, effectively, owned by no one. A better description of the resulting inefficiency would have been, “tragedy of open access.” See Garrett Hardin, *Tragedy of the Commons*, 162 SCIENCE 1243

conditions (demarcation of where one set of wireless users' rights ends and another begins) has become increasingly important. Not only have wireless services become far more widespread, and spectrum deployments much more economically important, but key wireless licenses—specifically, those authorizing mobile voice and data networks—have been dramatically liberalized in recent decades. The combination of significantly increasing wireless traffic and delegation of spectrum use choices to decentralized private actors leads many to insist that regulators must get far more serious about specifying the precise boundaries delimiting use rights.

B. EXACTITUDE IN SPECTRUM USE RIGHTS BOUNDARIES AT THE FCC

Theoretically, the FCC endorses exactitude: “Regardless of how or to whom particular rights are assigned, ensuring that all rights are clearly delineated is important to avoiding disputes.”¹³ Yet competitive economic forces point in exactly the reverse direction. Where particular rights are assigned to responsible economic agents in packages that avoid excessive fragmentation, “clearly delineated” rights are not much needed. Rather, approximately defined spectrum rights can be sufficient to launch markets. Self-interested actors then help to govern the delicate edges, between which—or through—intense airwave conflicts occur. These agents are rewarded by the net value of the traffic they generate, and are incentivized to discover how to efficiently mitigate mutually exclusive claims. Bureaucrats crafting use rights so as to avoid harmful interference have neither access to such information, nor the means to acquire it. As economist Thomas Sowell laments: “Why the transfer of decisions from those with personal experience and a stake in the outcome to those with neither can be expected to lead to better decisions is a question seldom asked, much less answered.”¹⁴

In radio spectrum the question of how to precisely define spectrum rights is often raised, but then buried beneath an underlying assumption that the distribution of all additional rights should be delayed until greater

(1968); Thráinn Eggertsson, *Open Access Versus Common Property*, in PROPERTY RIGHTS: COOPERATION, CONFLICT AND LAW 75 (Terry L. Anderson & Fred S. McChesney eds., 2002). The term “anticommons” is also used to describe situations where complementary rights are too widely distributed, preventing the emergence of efficient economic activity. The problem, however, is simply a reverse angle on the initial formulation by Hardin. See Michael Heller, *The Tragedy of the Anticommons: Property in the Transition from Marx to Markets*, 111 HARV. L. REV. 621 (1998); Lee Anne Fennell, *Common Interest Tragedies*, 98 NW. L. REV. 907 (2004).

13. SPTFR 2002, *supra* note 1, at 17–18.

14. THOMAS SOWELL, *INTELLECTUALS AND SOCIETY* 19 (2011).

exactitude can be specified.¹⁵ Even leading champions of further liberalization—sometimes identified as a shift to “private property spectrum”¹⁶—caution that “inexact” borders, without articulation of the diminishing returns of exactitude, will stifle the growth of wireless networks.¹⁷ The claim is made as a categorical proposition without considering the alternative costs of delaying market development. In fact, certainty is never achieved in drafting property rights or contracts, criminal laws, or liability rules. The quest to refine spectrum use contours is socially costly.¹⁸ The relevant comparison, therefore, is not between approximate use rights and a theoretically ideal set of precise, fully-specified spectrum use rights covering all dimensions over which frequencies may be usefully deployed (including those yet to be discovered). Rather, the relevant comparison is between continuing administrative control and decentralized choices made by profit-seeking operators who possess incomplete spectrum rights.

C. THE FCC’S APPROACH TO SPECTRUM USE RIGHTS DEFINITIONS

As shown in the FCC comments above, policymakers claim the following paradox: (a) harmful interference, defining the contours of spectrum use rights, must be clearly and fully detailed for wireless markets to work efficiently; (b) this feat has never been achieved; (c) mobile markets are

15. Hatfield and Weiser make this argument for proceeding slowly and carefully, despite their admission that “[t]he FCC’s traditional system for managing the radio spectrum is a paradigm of economic inefficiency.” Hatfield & Weiser, *supra* note 9, at 4.

16. MICHAEL HELLER, *THE GRIDLOCK ECONOMY* 89 (2008) (The evolution of private-property spectrum has come via wireless “licenses that look a lot like ordinary private property. Bands are exclusively assigned. Licensees may manage this spectrum with a measure of autonomy. Flexible uses are allowed. License holders may change technologies, services and business models to maximize their own profits—much as ordinary merchants shift inventory in bricks-and-mortar stores.”).

17. Philip J. Weiser & Dale Hatfield, *Spectrum Policy Reform and the Next Frontier of Property Rights*, 15 *GEO. MASON L. REV.* 549, 550–51 (2008) (“Commentators have recognized the need for reform . . . [but] a poorly designed property rights regime for spectrum might even be worse than the legacy model of spectrum regulation.”).

18. For more on property and the costs and gains of border definition, see Thomas W. Merrill, *Property as Modularity*, 125 *HARV. L. REV. F.* 151 (2012) (“Exclusion does more than minimize information costs for third parties. Even more importantly, it gives the owner residual managerial authority and residual accessionary rights—the rights to sow and to reap what has been sown.”), a response to frequent co-author, Henry E. Smith, *Property as the Law of Things*, 125 *HARV. L. REV.* 1691 (2012) [hereinafter Smith, *Property as the Law of Things*]. Merrill and Smith are leading voices on definitional efficiency in property law. See generally Thomas W. Merrill & Henry E. Smith, *Optimal Standardization in the Law of Property: The Numerus Clausus Principle*, 110 *YALE L.J.* 1 (2000); Thomas W. Merrill & Henry E. Smith, *The Property/Contract Interface*, 101 *COLUM. L. REV.* 773 (2001).

nevertheless robust: “U.S. consumers continue to reap significant benefits—including low prices, new technologies, improved service quality, and choice among providers—from competition in the [mobile] marketplace”¹⁹

This botched syllogism reveals critical truths. The latter two points are factually correct—(b) border contours in spectrum allocations are inexact, and (c) large-scale networks, created by private capital, host intense economic activity that generates extremely high social value. This latter contribution to economic welfare, moreover, is intrinsically related to the manner in which the underlying frequency space (“mobile spectrum”) is defined, with flexible use authorized for an exclusive licensee. As the FCC has seen fit to note, no other regulatory model that it employs could plausibly achieve such results.²⁰

To wait six years for a superior specification of spectrum use rights eliminates 25.4% of the total social value of a project, assuming a 5% social discount rate.²¹ Six years is, in the FCC’s estimate, a lower bound on how long spectrum allocations take.²² And there is no guarantee that such delays improve the delineation rights much (if at all). Indeed, the stronger inference is that, by avoiding templates in favor of customizing each allocation case by case, regulators who are dependent on data from key participants in the legal/lobbying process may obscure the path to simple, predictable harmful interference rules that could streamline allocations and put more spectrum in far more productive use.

19. In the Matter of Implementation of § 6002(b) of the Omnibus Budget Reconciliation Act of 1993 Annual Report and Analysis of Competitive Market Conditions With Respect to Commercial Mobile Services, Thirteenth Report, 24 FCC Rcd. 6185, 6189 ¶ 1 (2009), *available at* http://hraunfoss.fcc.gov/edocs_public/attachmatch/DA-09-54A1.pdf.

20. *Wireless Bureau Chief Daniel Phythyon Hails Success of Market-Based Spectrum Policies*, FCC (Sept. 11, 1997), http://transition.fcc.gov/Bureaus/Wireless/News_Releases/1997/nrw17037.html (The Wireless Telecommunications Bureau, through spectrum auctions, “has successfully moved wireless services towards a flexible, competitively neutral regulatory scheme that has resulted in an unprecedented growth in the number of new service offerings and providers Permitting CMRS providers to offer fixed wireless services on a co-primary basis with mobile services has afforded them flexibility to meet consumer demand for fixed services, mobile services, and combinations of the two. A number of PCS licensees and other CMRS providers are developing fixed as well as mobile service offerings.”). *See generally* In the Matter of Implementation of Section 6002(b) of the Omnibus Budget Reconciliation Act of 1993 Annual Report and Analysis of Competitive Market Conditions With Respect to Commercial Mobile Services, Fifteenth Report, 26 FCC Rcd. 9664 (2011) [hereinafter Fifteenth Annual Competition Report], *available at* http://hraunfoss.fcc.gov/edocs_public/attachmatch/FCC-11-103A1.pdf.

21. The net present value of a six-year delay at a 5% discount rate is calculated $1 - 1/(1.05^6)$.

22. The FCC’s National Broadband Plan found that such allocations take six to thirteen years. NBP 2010, *supra* note 3, at 79.

Incomplete rights, when assigned with sufficient scope to yield effective spectrum ownership rights, power cellular markets.²³ These markets enable rapid economic growth in the United States²⁴ and internationally, particularly in developing economies.²⁵ Notably, such markets coordinate continual spectrum redeployments. Under traditional regulation, new services, network architectures, business models, and technologies are subject to regulatory approval. In this process, government agencies consider whether proposed changes are in the public interest and, in particular, whether the change would result in harmful interference with other wireless activities. Typically, such deliberations encompass competitive concerns—meaning that competitors widely object to granting rivals new authority to change service offerings. The FCC itself has become an attractive nuisance, allowing spectrum use decisions to serve anti-competitive firm strategies. Leading FCC officials themselves have bemoaned the fact, citing the term “Forever Captured by Corporations” as the operational reality.²⁶

A better system is one in which private wireless operators, rather than regulators, make decisions about spectrum use. Under the flexible uses permitted to Commercial Mobile Radio Services (“CMRS”) licensees,²⁷ technical considerations have been delegated to wireless operators themselves. Within the bandwidth allocated to their licenses, carriers coordinate complex activities that involve extensive spectrum sharing. In emission spillovers with bandwidth allocated to other carriers’ licenses, operators routinely negotiate win-win agreements to deal with radio

23. This Article uses the terms “cellular markets,” “mobile markets,” “wireless markets,” “CMRS” (commercial mobile radio services, an FCC acronym), “PCS” (personal communications service, an FCC acronym), “3G,” “4G,” and “3G/4G,” interchangeably, except as otherwise noted. “G” in the preceding abbreviations stands for “Generation.”

24. At least \$200 billion in U.S. consumer surplus is generated annually by the use of mobile voice services. Rapidly evolving mobile data services (via smartphones, tablets, netbooks, notebooks and M2M—“machine to machine”—devices) are on top of this sum. Hazlett, Muñoz & Avanzini, *supra* note 6, at 100 (noting that \$211.8 billion in U.S. consumer surplus was generated in 2009 (in 2008 dollars)). It is estimated that, in 2011, mobile services added about \$146 billion to GDP, accounting (directly and indirectly) for nearly four million jobs. ROGER ENTNER, *THE WIRELESS INDUSTRY: THE ESSENTIAL ENGINE OF U.S. ECONOMIC GROWTH* 4, 14 (2012), <http://reconanalytics.com/wp-content/uploads/2012/04/Wireless-The-Ubiquitous-Engine-by-Recon-Analytics-1.pdf>.

25. See, e.g., Leonard Waverman et al., Presentation at the IFC Seminar: The Socio-Economic Impact of Mobile Phones in Africa (May 6, 2005), <http://info.worldbank.org/etools/BSPAN/PresentationView.asp?PID=1522&EID=741>; Sara Corbett, *Can the Cellphone Help End Global Poverty?*, N.Y. TIMES, Apr. 13, 2008, at MM34; Robert Jensen, *The Digital Divide: Information (Technology), Market Performance, and Welfare in the South Indian Fisheries Sector*, 122 Q. J. OF ECON. 879 (2007).

26. Ken Auletta, *The News Rush*, NEW YORKER, Mar. 18, 1996, at 42.

27. 47 C.F.R. § 20.3 (2011); see generally *infra* note 28.

conflicts.²⁸ And, more importantly, the vast majority of potential conflicts are remedied by the aggregation of spectrum rights: U.S. mobile operators hold well over 50,000 CMRS licenses, creating national and regional networks that largely eliminate border disputes by eliminating borders themselves.²⁹

CMRS licenses define radio spectrum rights far more liberally than do traditional FCC licenses, which narrowly specify what wireless users can do. Historically, traditional licenses have been called “radio station authorizations,” and have been analogous to *use permits* rather than *spectrum licenses*.³⁰ That is because the underlying resource, radio spectrum, was under the control of regulators, not licensees. The operator could only transmit (or receive) wireless signals as permitted. Regulators determined where transmissions could take place, what frequency and emitted power they would use, with which technologies they would operate, and what services would be supplied. These regulators even decided which business models (e.g., subscription versus ad-supported, common carrier versus private carrier, etc.) were deployed.³¹ The development of mobile phone networks saw a relaxation of such rules, which transitioned from in personam use rights (specified by the FCC) to in rem spectrum ownership rights. The latter

28. An executive for Verizon Wireless describes the manner in which disputes among mobile networks are resolved:

Under current rules, licensees negotiate to extend rights into each others’ licensed spectrum on a daily basis. These . . . involve hundreds of individual negotiations between companies’ engineers who are tasked with the day-to-day operations of the network [T]here is no benefit to seeking regulatory redress.

Charla Rath, *Defining Radio Rights: Theory and Practice*, in J. Pierre de Vries & Kaleb A. Sich, *The Unfinished Radio Revolution: Eight Perspectives on Wireless Interference Symposium*, 9 J. TELECOM. & HIGH-TECH. L. 501, 529 (2011). It is of interest that Rath attributes the satisfactory process of dispute resolution to the fact that “the rights of both licensees are clear.” *Id.* In fact, the rights are *clear enough* such that Verizon Wireless and other carriers may create and operate large networks, but the border definitions are themselves fairly cryptic and are defined in four dimensions—frequency, power at the edge, geography, and time. Various spectrum experts recommend defining borders in more than four dimensions, as discussed *infra* note 56 and accompanying text.

29. Thomas W. Hazlett, *Is Federal Preemption Efficient in Cellular Phone Regulation?*, 56 FED. COMM. L.J. 155, 193 (2003) [hereinafter Hazlett, *Federal Preemption*]. In 2003, the count for U.S. CMRS licenses was 51,597. *Id.* Since that time, 1,087 were issued in the 2006 auction of Advanced Wireless Service (AWS) licenses and another 1,090 in the 2008 auction of 700 MHz licenses. *Auctions Summary*, FCC, http://wireless.fcc.gov/auctions/default.htm?job=auctions_all.

30. Thomas W. Hazlett, *The Wireless Craze, the Unlimited Bandwidth Myth, the Spectrum Auction Faux Pas, and the Punchline to Ronald Coase’s ‘Big Joke’: An Essay on Airwave Allocation Policy*, 14 HARV. J.L. & TECH. 335 (2001) [hereinafter Hazlett, *Wireless Craze*].

31. *Id.*

cede greater scope to rights holders to determine resource uses.³² This creates net benefits when regulatory rules are relatively inefficient, even as it increases the direct costs of defining rights: “in rem rights conserve on information when it is cheaper to define the resource itself and appoint a single manager (the owner) who has the discretion to choose among multiple permitted uses.”³³

D. LIBERAL LICENSES THAT REMEDY TRAGEDIES OF THE ANTI-COMMONS

One major reason for the policy change in cellular communications was the greatly increased complexity of the network architectures³⁴ over previous wireless services such as broadcasting or point-to-point microwave relays. Modern U.S. mobile networks construct 50,000 or more base stations, distribute hundreds of millions of mobile handsets to subscribers using matching technologies, build high-capacity (fixed) links connecting wireless calls to multiple voice and data networks, and supply platforms for the operation of additional millions of applications and machine-to-machine devices.³⁵ The coordination requirements are large. Competitive forces push efficiency and innovation in these busy ecosystems, discovering trade-offs unknown to disinterested observers and illustrating the information cost savings enabled by delegating resource decisions to profit-seeking licensees.³⁶

Years prior to the deregulatory actions yielding “flexible-use” spectrum rights, scholars argued on theoretical grounds for such a policy. In papers

32. Merrill & Smith, *The Property/Contract Interface*, *supra* note 18, at 778.

33. *Id.*

34. This theory is elaborated upon in Thomas W. Hazlett, *Assigning Property Rights to Radio Spectrum Users: Why Did FCC License Auctions Take 67 Years?*, 41 J.L. & ECON. 529 (1998). One of the major outcomes of the liberalized approach to spectrum regulation was the introduction of license auctions in the United States and at least thirty other countries—a trend coincident with the rise of cellular networks. *Id.*

35. Thomas W. Hazlett, David J. Teece & Leonard Waverman, *Walled Garden Rivalry: The Creation of Mobile Network Ecosystems*, George Mason Law & Economics Research Paper No. 11-50 (Nov. 22, 2011) (paper presented at CITI, Columbia University (Oct. 14, 2011)).

36. See Julius Genachowski, Chairman, FCC, Prepared Remarks for the GSMA World Congress, Barcelona (Feb. 27, 2012), <http://www.fcc.gov/document/chairman-genachowskis-remarks-gsma-mobile-world-congress> (“490 million smartphones were sold worldwide in 2011, exceeding the number of PCs sold over the same period.”); Julius Genachowski, Chairman, FCC, Prepared Remarks for the Broadband Acceleration Conference, Washington D.C. (Feb. 9, 2011), <http://www.fcc.gov/document/prepared-remarks-chairman-julius-genachowski-federal-communications-commission-broadband-ac> (“It has also been estimated that removing red tape and expediting approval processes could unleash \$11.5 billion in new broadband infrastructure investment over two years.”).

written in 1962³⁷ and 1969,³⁸ such researchers considered, *tabula rasa*, how spectrum ownership rights could be defined with respect to three variables: time (T), geographic area (A), and frequency (S)—a “TAS package.”³⁹ Even as a theoretical exercise, the components of what has proven successful in CMRS licenses were presented, including the specification of maximum field strength at geographic and frequency boundaries.⁴⁰ Today, scholars still use these bundles to shape appropriate rights packages, even in sophisticated discussions of how to achieve efficient economic outcomes in wireless markets.⁴¹ As developed over the past several decades, the FCC has crafted CMRS (including cellular and PCS) license rights to authorize delivery of wireless services of any type, technology, or business model, limited by emission boundaries.⁴² These rights have formed the resource rights used to generate far more consumer welfare in the wireless sector than any competing method and, more importantly, to significantly out-perform any alternative rights regime yet considered.

That borders are not costless to define or police, however, raises another important consideration. In defining and distributing spectrum rights, methods that permit license aggregation serve a valuable function. In integrating ownership, borders are reduced and border disputes eliminated—“since there is no externality if ownership is unified . . . [t]o merge or not to

37. RONALD COASE, WILLIAM H. MECKLING & JORA MINASIAN, PROBLEMS OF RADIO FREQUENCY ALLOCATION (1995), <http://www.rand.org/pubs/drafts/DRU1219.html>. This monograph, published in 1995, was written in May 1963. *See also* Thomas W. Merrill & Henry S. Smith, *Making Coasean Property More Coasean*, 54 J.L. & ECON. S77, S85 n.6 (2011) (describing the delay in public release).

38. Arthur S. De Vany, Ross D. Eckert, Charles J. Meyers, Donald J. O’Hara & Richard C. Scott, *A Property System for Market Allocation of the Electromagnetic Spectrum: A Legal-Economic-Engineering Study*, 21 STAN. L. REV. 1499 (1969).

39. *Id.* at 1501.

40. *Id.* at 1513–17.

41. *See, e.g.*, Timothy K. Forde & Linda E. Doyle, *A Combinatorial Clock Auction for OFDMA-based Cognitive Wireless Networks*, in PROCEEDINGS OF 3RD INTERNATIONAL SYMPOSIUM ON WIRELESS PERVASIVE COMPUTING 329, 329–33 (May 7–9, 2008), http://ieeexplore.ieee.org/xpl/freeabs_all.jsp?arnumber=4556224.

42. As FCC experts Evan Kwerel and John Williams have written: “[W]e suggest setting objective limits on some of the principal factors that cause interference (*e.g.*, transmitter power at boundaries) and allowing licensees to deploy unilaterally, and control actual interference, within those limits The values used in the current PCS rules should be appropriate in most instances.” Evan Kwerel & John Williams, *A Proposal for a Rapid Transition to Market Allocation of Spectrum* 44–45 (FCC, OPP Working Paper No. 38, Nov. 15, 2002), available at <http://wireless.fcc.gov/auctions/conferences/combin2003/papers/masterevanjohn.pdf>. The passage cites to the FCC’s PCS boundary rules: 47 C.F.R. §§ 24.236, 24.238 (2001).

merge ownership interests is the question.”⁴³ It is appropriate here to cite the Demsetzian critique of Coase’s analysis. Coase assumed, as have many others, that market structure is exogenous. In instances where small numbers of parties were involved in a spillover problem—for example, a doctor and a confectioner occupying adjacent quarters⁴⁴—negotiations to maximize social welfare were seen as likely to occur. Where, however, large numbers of parties were affected by a given activity—for example, many farmers absorbing pollution from a railroad⁴⁵—Coase readily conceded that government regulation might be preferred. With legal rights widely dispersed, transaction costs stymie the “correction” of externalities; spillovers that produce less value than the productivity they destroy will continue because the expense of fixing them (bargaining between parties) outweighs the benefits that could be gained.⁴⁶

Demsetz observes, however, that *given* fragmented ownership rights, the efficient solution is achieved by *not remedying* the externality. “A decision that something is not worth taking into account is not, because of this, a source of inefficiency.”⁴⁷ Hence, no “market failure” exists. The more efficient distribution of legal rights might improve social outcomes. That is to say, markets are decentralized either because there are relatively important advantages that accrue from decentralization (for example, the costs of managing larger businesses are avoided, a situation where diseconomies of scale obtain), or because legislators, courts, or regulators have chosen to create rights that impose uneconomic barriers to aggregation. This

43. Harold Demsetz, *Ownership and the Externality Problem*, in PROPERTY RIGHTS: COOPERATION, CONFLICT, AND LAW 282, 287 (Terry L. Anderson and Fred S. McChesney eds., Princeton University Press 2003).

44. Ronald H. Coase, *The Problem of Social Cost*, 3 J.L. & ECON. 1, 8–10 (1960) (citing *Sturges v. Bridgman*, 11 Ch.D. 852 (1879)).

45. *Id.* at 29–34 n.41 (citing 31 HALSBURY, LAWS OF ENGLAND 474–75 (3d ed. 1960)).

46. For an account of the Eureka moment witnessed by the twenty-one Chicago economists who attended Coase’s talk on the subject of his paper on “externalities” in the living room of Aaron Director, greeted by unanimous hostility but, within two hours, producing stunning consensus, see GEORGE STIGLER, MEMOIRS OF AN UNREGULATED ECONOMIST (1982) and Steven G. Medema, *A Case of Mistaken Identity: George Stigler, “The Problem of Social Cost,” and the Coase Theorem*, 31 EURO. J.L. & ECON. 11, 22–23 (2011). The economists engaged in this intellectual event included George Stigler, Aaron Director, Reuben Kessel, Milton Friedman, Gregg Lewis, Arnold Harberger, and Martin Bailey, among others, which Stigler remarked as “one of the most exciting intellectual events of my life.” *Id.* at 24 (citing Edmund W. Kitsch, *The Fire of Truth: A Remembrance of Law and Economics at Chicago, 1932–1970*, 26 J.L. & ECON. 163 (1983)).

47. Harold Demsetz, *The Problem of Social Cost: What Problem? A Critique of the Reasoning of A.C. Pigou and R.H. Coase*, 7 REV. L. & ECON. 1, 10 (2011).

inefficiency flows from a standard anti-commons tragedy, and is fully illustrated in numerous FCC spectrum allocations.⁴⁸

Common characteristics of these anti-commons “gridlock” situations are: highly fragmented rights, in personam rights, and licenses held by non-profit organizations. In such situations, regulatory methods and restrictions block markets from rearranging rights or restructuring markets, such that self-interested economic parties can mitigate border disputes. These restrictions ensure that one or more of the parties involved in a border dispute are severely restricted in their access to financial markets. Whereas a profit-seeking firm could raise capital (debt or equity) to buy up TV licenses (for example, moving TV broadcasts to alternative frequencies or video delivery platforms to unleash the radio spectrum allocated to the stations for higher valued uses, thus paying back investors and reaping, with luck, a profitable residual), the process is thwarted when TV licenses are locked into TV broadcasting. Society is left to develop means to mimic the role of financial markets,⁴⁹ or to devise policies that expand the spectrum use rights owned by responsible economic agents.⁵⁰

E. DEFINING HARMFUL INTERFERENCE AS IF ECONOMICS MATTERS

The Demsetz critique—that because fragmented ownership rights exist, efficiency is achieved by not remedying externalities—has deep salience when considering how spectrum rights are usefully defined.⁵¹ The manner in which productive economic activity will result from the creation of rights crafted by regulators is key, and it cannot be assumed that the dispersion of rights—or,

48. See HELLER, *supra* note 16. In particular, Chapter 4 deals with anti-commons in wireless. For a discussion on the tragedy of the anti-commons, see *infra* note 93.

49. This is what the FCC’s National Broadband Plan attempted in advocating an “incentive auction.” The procedure, authorized by a congressional statute in early 2012, has the regulatory agency conducting two sets of auctions. In the first, a reverse auction, TV stations state offer prices to exit broadcasting. In the second, the spectrum made available (by the vacating stations and a relocation plan implemented by the FCC for stations that remain) in the TV Band is reallocated to flexible-use licenses, which are sold in a forward auction. See *In the Matter of Expanding the Econ. & Innovation Opportunities of Spectrum Through Incentive Auctions*, Notice of Proposed Rulemaking, 27 FCC Rcd. 12,357 (2012), available at <http://apps.fcc.gov/ecfs/document/view?id=7022026834> [hereinafter 2012 Incentive Auction NPRM]; H.R. Rep. No. 112-399 (2012) (Conf. Rep.). For an overview, see THOMAS W. HAZLETT, DAVID PORTER & VERNON SMITH, “INCENTIVE AUCTIONS”—ECONOMIC AND STRATEGIC ISSUES (2012) [hereinafter HAZLETT ET AL., INCENTIVE AUCTIONS PAPER], available at <http://www.arlingtoneconomics.com/studies/WhitePaper.pdf>.

50. See, e.g., Thomas W. Hazlett, *Unleashing the DTV Band: A Proposal for an Overlay Auction*, Comment to the FCC, NBP Public Notice No. 26 (Dec. 18, 2009) [hereinafter Hazlett, *A Proposal for an Overlay Auction*], http://mason.gmu.edu/~thazlett/pubs/NBP_PublicNotice26_DTVBand.pdf.

51. See *infra* note 60.

if preferred, market structure—is exogenous. Regulators are perfectly capable of spreading rights to a large number of parties that cannot economically coordinate, of truncating rights issued to licensees such that market participants are prevented from using the most efficient methods, and of using authority over license transfers to delay or deter the useful aggregation of rights. Focusing on what policies regulators should pursue so as to define spectrum use rights—abstracting from the use of those rights to aid coordination on the one hand, and enabling production efficiencies on the other—leads analysts to completely miss the central components of the harmful interference policy question.

This Article addresses the crisis in spectrum allocation as a property rights conundrum. In Part I, we critiqued the common reasoning that a lack of clarity in defining the term “harmful interference” accounts, in large measure, for gridlock. However, we argued that the answer to the harmful interference problem in spectrum allocation is not greater exactitude in rights delineation, but rather the delegation of exclusive, flexible use rights to responsible economic agents. In Part II of this Article, we illustrate the FCC’s position on “exactitude”—the idea that the regulator bears the burden to clearly define spectrum usage rights for parties to avoid harmful interference and best utilize radio spectrum. The FCC’s approach to spectrum use rights boundaries is responsible for endemic failure and costly administrative delay. In Part III, we provide examples of situations in which the FCC’s search for exactitude failed, and examples of situations in which exact-enough border definitions generated economic robustness. In Part IV, we propose specific ways in which spectrum use rights can optimally be defined to both avoid the harmful interference problem and promote social utilization.

II. EXACTITUDE

Despite game-changing mobile markets arising from the deployment of liberal CMRS licenses, the dominant meme among policy experts⁵² and

52. In fairness, not all spectrum policy scholars espouse this view. Counter-examples include MARTIN CAVE, SPECTRUM MANAGEMENT FOR A CONVERGING WORLD: CASE STUDY ON THE UNITED KINGDOM (2006), <http://www.itu.int/osg/spu/ni/spectrum/UK-RSM.pdf>; Stuart M. Benjamin, *Spectrum Abundance and the Choice between Private and Public Control*, 78 N.Y.U. L. REV. 2007 (2003); Gregory L. Rosston, *The Long and Winding Road: The FCC Paves the Path With Good Intentions*, 27 TELECOMM. POL’Y 501 (Aug. 2003), <http://www-siepr.stanford.edu/papers/pdf/01-08.pdf>; Gerald R. Faulhaber, *The Future of Wireless Communications: Spectrum as a Critical Resource*, 18 INFO. ECON. & POL’Y 256 (Sept. 2006); John Mayo & Scott Wallsten, *Enabling Efficient Wireless Communications: The Role of Secondary Spectrum Markets*, 22 INFO. ECON. & POL’Y 61 (2010); Hazlett, Muñoz & Avanzini, *supra* note 6.

regulatory officials⁵³ is that, going forward, superior utilization of radio spectrum depends upon clear improvement, or “exactitude,” in the technical specification of usage rights. Ellen Goodman asserts that, “[e]fficient conflict resolution requires that initial entitlements be state[d] precisely in the license (or license-free allocation),”⁵⁴ noting that we are failing to make much progress in this respect, as spectrum rights are still issued “without clear entitlements . . . [and are] lacking in the regularity and transparency that would facilitate secondary markets.”⁵⁵ Robert Matheson and Adele Morris propose a seven-dimensional format for fully defining “electrospace” units, arguing that rights must be “flexible, exhaustive, and economically efficient.”⁵⁶ Pierre De Vries and Kaleb Sieh, in a 2011 paper that seeks to create a framework for “unambiguously defining and delegating radio rights,” posit:

Clear rules are essential for radio coexistence and particularly important where conflict is more likely due to proximity in time, space or frequency; different technologies or business models; frequent change in operating parameters or operators; and multiple steps of delegation in the right to operate.⁵⁷

The concern over definitional clarity is driven by two identified concerns:

1. *Interference Disputes*. There have been quagmires in regulatory proceedings as to what constitutes harmful interference, including proceedings involving Nextel and public safety radio users, Wireless Communications Services (“WCS”) and Satellite Digital Audio Radio Services (“SDARS”) licensees, L Band LTE users (LightSquared), and Geopositioning Satellite Services (“GPS”).
2. *White Spaces*. Radio sensing technologies increasingly support spectrum sharing.⁵⁸ When a licensed service “consumes” less than all of the communications capacity of a given frequency space, it

53. A notable exception is *A Proposal for a Rapid Transition to Market Allocation of Spectrum*. Kwerel & Williams, *supra* note 42.

54. Ellen P. Goodman, *Progress Toward Rational Spectrum Rights: Are We Getting Anywhere?*, 9 J. TELECOMM. & HIGH-TECH. L. 505, 506 (2011).

55. *Id.*

56. Robert Matheson & Adele Morris, *The Technical Basis for Spectrum Rights: Policies to Enhance Market Efficiency* 40 (Brookings, 2011), available at http://www.brookings.edu/~media/research/files/papers/2011/3/03%20spectrum%20rights%20matheson%20morris/0303_spectrum_rights_matheson_morris.pdf.

57. J. PIERRE DE VRIES & KALEB A. SIEH, *THE THREE PS: INCREASING CONCURRENT OPERATION BY UNAMBIGUOUSLY DEFINING AND DELEGATING RADIO RIGHTS* (2011) (paper presented at the IEEE International Symposium on Dynamic Spectrum Access Networks), available at <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5936248>.

58. 47 C.F.R. § 15.703(l) (2011).

leaves open the possibility that if there were rules in place that better defined harmful interference, additional radios could opportunistically transmit in the unused “white spaces,” using rules for unlicensed device operation.⁵⁹

Both inferences are misleading. First, additional clarity in spectrum rights (as with property rights generally) is preferred to less clarity, all else equal. But perfect clarity (or exhaustiveness) is not achievable, and additional clarity is costly. After basic use rights (such as with a critical mass, like TAS) are established, the cost of further specification of spectrum rights rises sharply.⁶⁰ As De Vany et al. wrote in their 1969 study, “[c]omplete certainty in this regard is not necessary for the functioning of a market. Any landowner can cite many areas of uncertainty—for example, the introduction of commercial and industrial uses into a residential area may depress values—yet there is an active market in real estate.”⁶¹

Second, easily defined spectrum borders deliver enormously high social value compared to the cost—and delays—of specifying more complete use rights. When regulators succeed in delegating flexible use rights to a responsible economic agent, specifically an organization constrained by profit maximization, the problems associated with “ill-defined rights” dissipate. Such basic rights are available in the templates used to assign spectrum use rights to mobile carriers in the United States and in many other countries. As discussed below, when these rights are distributed in a manner that avoids

59. In the Matter of Unlicensed Operation in the TV Broad. Bands & Additional Spectrum for Unlicensed Devices Below 900 Mhz & in the 3 Ghz Band, Second Report and Order and Memorandum Opinion and Order, 23 FCC Rcd. 16,807 (2008). See *infra* notes 191–94 for a detailed discussion of harmful interference definitions in the white space proceedings between 2002 and 2010.

60. The idea that property rights are costly to establish and, of necessity, are subject to cost-benefit tests, dates to the landmark work by Harold Demsetz, *Toward a Theory of Property Rights*, 57 AM. ECON. REV. 347 (1967). Demsetz’s point was that private rights came to replace an “open access” environment when the net benefits of establishing and enforcing private property rights rose above zero. *Id.* The analysis seamlessly transfers to all relevant margins, not just the jump from a situation with ownership rights to one without. See Henry E. Smith, *On the Economy of Concepts in Property*, 160 U. PENN. L. REV. 2097, 2115 (2012) [hereinafter Smith, *On the Economy of Concepts in Property*] (“By setting up cheap and rough proxies like boundary crossings, property law can indirectly protect a wide range of largely unspecified interests in use, the details of which are of no particular relevance to those under a duty to respect the right”); Henry E. Smith, *Property and Property Rules*, 79 N.Y.U. L. REV. 1719, 1725 (2004) (discussing the continuum of uncertainty and risk in the information costs of boundary definition in FRANK H. KNIGHT, RISK, UNCERTAINTY AND PROFIT 19–21, 197–232 (1921)).

61. De Vany et al., *supra* note 38, at 1511.

excessive fragmentation, the ills associated with disputes over harmful interference are almost entirely avoided.

That there exist hard problems in defining private rights to control use of resources does not imply that an administrative allocation of the resource would be superior. To assume such runs smack into the “Nirvana Fallacy,” a term coined by Harold Demsetz⁶² that was also the key analytical insight of Ronald Coase’s famous 1960 essay on social cost.⁶³ It is true that property rights in land or other resources (such as subsurface minerals, oil pools, fisheries, or patents) are often very complex to define and highly costly to enforce. But the confusion is revealed when it is noted that, in all of these instances, society prefers to rely upon private ownership as the best way to allocate the resource. As Coase has opined, the idea of switching to a “Federal Land Commission” to overcome definitional problems in land rights would (a) not eliminate complexities in distinguishing the appropriate limits imposed upon conflicting uses, and (b) presumably result in far less social welfare:

No business would have any interest in economizing in the use of its land. Changes in land-use would come about only with great difficulty and would depend to a large extent on land becoming valueless in existing uses. Economic growth in the United States would be slowed by the shortage of land and the problem would no doubt call for Presidential attention. That such would be the consequences of the establishment of a Federal Land Commission is not, I think, open to serious doubt.⁶⁴

Liberal licenses cede de facto spectrum ownership, forcing licensees, through economic incentives common to markets, to rationally weigh gains against losses in choosing among myriad forms of spectrum utilization. This process incorporates the efficient control of harmful interference and far more functional band clearing efforts, such as standards migration and technology development. It prompts the creation of wireless networks via massive sunk, long-term investments, agreements for spectrum sharing with both subscribers and rival network providers, contracts altering initially assigned borders, and mergers. These productive actions result in a reliable process that delivers higher-valued outputs for consumers.

62. Harold Demsetz, *Information and Efficiency: Another Viewpoint*, 12 J.L. & ECON. 1, 1–4 (1969).

63. Coase, *supra* note 44.

64. Ronald H. Coase, *Evaluation of Public Policy Relating to Radio and Television Broadcasting: Social and Economic Issues*, 41 LAND ECON. 161, 163 (1965).

Interference disputes in the traditional spectrum allocation system are so destructive not because they deal with inherently complicated matters,⁶⁵ which are of equal technical complexity for both governments and firms, but because decision-makers in these disputes do not face strong incentives to avoid controversy and, instead, generate gains from trade.

Incomplete contracts are ubiquitous in commerce because precisely and exhaustively defined agreements covering all possible contingencies are prohibitively expensive.⁶⁶ In spectrum, marginal improvements in the delineation of rights, as well as the resolution of disputes over conflicts, may sometimes be warranted, and promising avenues have been outlined.⁶⁷ But today, there is no doubt that the most expensive path is found in delaying the distribution of “imperfect” spectrum rights for years or even decades. Sadly, we have wide and deep observational experience with which to gauge this cost. When, for instance, 30 MHz of PCS bandwidth (about 15% of what was then available to mobile carriers) was withheld from the U.S. market in 1995–2005 due to litigation over FCC rules and bankruptcy law in the so-called “PCS C Block fiasco,” the estimated loss in social welfare exceeded \$10 billion annually.⁶⁸

It has not been entirely missed that economic incentives are part of the solution. Several analysts note the success of liberal licenses⁶⁹ and the importance of “economic efficiency” in creating and awarding spectrum use rights.⁷⁰ Some have noted the difficulty in defining spectrum use in

65. Whether or not spectrum use rights have changing boundaries is less important than which entities have rights to respond. Probabilistic boundaries of other information resource uses are widespread in other intellectual property forms. See Henry E. Smith, *Intellectual Property as Property: Delineating Entitlements to Information*, 116 YALE L.J. 1742, 1822 (2007) (“In both intellectual property and property more generally, exclusion rights—as modified by governance rules—furnish, at some positive cost, modularity to the system of providing inputs and appropriating benefits from assets.”).

66. See Merrill & Smith, *The Property/Contract Interface*, *supra* note 18.

67. See *Wireless Craze*, *supra* note 30.

68. This is based on econometric estimates of the value of marginal spectrum in mobile services using cross-sectional data from international markets. Thomas W. Hazlett & Roberto E. Muñoz, *A Welfare Analysis of Spectrum Allocation Policies*, 40 RAND J. ON ECON. 424, 436 tbl.5 (2009) (showing annual welfare change in U.S. PCS C-block auctions) [hereinafter *Welfare Analysis*].

69. See, e.g., Gregory L. Rosston & Jeffrey S. Steinberg, *Market-Based Spectrum Policy to Promote the Public Interest*, 50 FED. COMM. L.J. 87 (1997); *SPTFR 2002*, *supra* note 1; Kwerel & Williams, *supra* note 42; Gerald R. Faulhaber, *supra* note 52.

70. See MATHESON & MORRIS, *supra* note 56. The debate over whether or not to provide new flexibilities or more permanent rights often convolves the efficiency and distributional effects of proposed change. Here, without taking a position on who should receive the benefit of more flexible or longer-duration rights, we merely say that certain rights are likely to support a more efficient spectrum market than uncertain rights.

unlicensed bands, as with TV Band “white spaces,” where economic incentives do not internalize the costs and benefits of “interference” for market agents.⁷¹ What these analysts generally omit is that economic efficiency resulting from rights assignments forms virtually the entire locus of relevant (correctible) policy reform. Treating “economic efficiency” as one factor among many typically results in priority confusion, sending policy analysis in pursuit of time-consuming technical definitions of spectrum boundaries and entirely missing the path to the great social welfare gains that are available.

III. THE TROUBLE WITH EXACTITUDE: EXAMPLES

Examples are both useful and a central feature of the policy literature in this area. Here we focus on those examples that critics commonly give to illustrate how and why FCC definitions of harmful interference need greater clarity, such as the WCS/Satellite Radio and Nextel/Public Safety Radio controversies, and others that are rarely mentioned. These include the problem of harmful interference in the TV Band, by consensus the most costly failure of spectrum policy, and the FCC’s attempt (from 2003–2007) to construct an “interference temperature” metric to allow for greater spectrum sharing—an admitted policy failure by the Commission that was intended to bring just the precision to spectrum use rights advocated so frequently today. Later in the Article we will describe other important examples, including those involving Low-Power FM and HD radio and the LightSquared/GPS interference dispute. Each of these episodes illustrates the economic robustness, generated not from regulatory exactitude, but from the distribution of exclusive, flexible rights awarded to economically responsible parties.

A. TV BAND REALLOCATION: PROBLEMS WITH RIGID RIGHTS

In an elaborate⁷² and influential discussion⁷³ of how spectrum use rights can and should be defined in seven dimensions—three more than PCS

71. Goodman, *supra* note 54, at 507.

72. This work derives from the pioneering attempts in the 1960s and 1970s to describe ownership rights in spectrum. See De Vany et al., *supra* note 38; COASE, MECKLING & MINASIAN, *supra* note 37, at 1 n.1; Jora Minasian, *Property Rights in Radiation: An Alternative Approach to Radio Frequency Allocation*, 18 J.L. & ECON. 221 (1975). These earlier papers were written prior to liberalization in U.S. (and other) spectrum policy, wherein licenses were created that granted de facto spectrum property rights. Among the most important policy reforms bringing such licenses—now used for cellular networks—into existence were rules delegating licensees discretion over technologies deployed, services supplied, and business models utilized. Since these market changes, some countries have fundamentally revised

license contours and, in our opinion, three too many⁷⁴—Matheson and Morris offer this conclusion on radio interference: “The FCC designs its rules regarding emission masks, transmitter separation criteria, and other licensing rules to prevent harmful interference. If all parties follow license rules, systems usually operate without interference.”⁷⁵ In fact, these FCC rules are the source of the most costly interference that exists, namely blocked spectrum that is not put to its most productive use. Under this view, there is little or no “interference” on the TV Band when, in fact, the loss to society due to the interference of TV broadcasting rules by the FCC easily exceeds \$1 trillion.⁷⁶

The television band constitutes a classic example of the tragedy of the anti-commons,⁷⁷ not because technical rules are extraordinarily difficult to construct, but because they have been so easy to construct. By imposing specific engineering and operational limits on spectrum users, conflicts have been managed to a bureaucratic fare-thee-well. Great swaths of bandwidth

their regulatory regimes to more closely mimic property regimes. Four countries where this has happened are New Zealand (1989), Guatemala (1996), El Salvador (1997), and Australia (1997). See Thomas W. Hazlett, *Property Rights and Wireless License Values*, 51 J.L. & ECON. 563, 567 (2008) [hereinafter Hazlett, *Wireless License Values*].

73. Phil Weiser and Dale Hatfield write: “The final, and most significant, development since the De Vany-led study is a 2005 paper by Robert Matheson. This paper presents the most complete analysis of property rights in radio spectrum as well as the fullest discussion of the practical challenges and limitations of actually employing such rights in the management of the resource.” Weiser & Hatfield, *supra* note 17, at 573.

74. See MATHESON & MORRIS, *supra* note 56, at 8. The four dimensions that are useful are actually just three: frequency, time, and geographic location (Matheson & Morris define geography in latitude and longitude, yielding the extra dimension). The three incremental dimensions are altitude, azimuth, and elevation angle. Liberal licenses bundle the fragments implied (or defined) by these latter criteria, ceding authority to the licensee for determining how the spaces defined in the other dimensions are best deployed. Defining more divisions than necessary likely incurs higher costs in the administrative allocation of spectrum relative to the alternative—allowing licensees to define use rights where possible. Indeed, holders of liberal licenses routinely pack in additional communications traffic by exploiting differing signal angles (say, with array antenna technology) and the altitude of such signals (as when cell towers are set high, but not too high, above adjacent terrain). *Id.* at 8. However, “[t]hat is not to say that any particular rights holder would wish to subdivide his or her rights along every dimension,” yet each additional border increases the number of possible fragments for aggregation in an exponential fashion. *Id.* at 9 (rightly declining to add polarization and/or modulation as additional dimensions to the proposed seven).

75. MATHESON & MORRIS, *supra* note 56, at 31.

76. Thomas W. Hazlett, *Tragedy TV: Rights Fragmentation and the Junk Band Problem*, 53 ARIZ. L. REV. 83, 86 (2011) [hereinafter *Tragedy TV*].

77. Lee Anne Fennell, *Commons, Anticommons, Semicommons*, in RESEARCH HANDBOOK ON THE ECONOMICS OF PROPERTY LAW (Kenneth Ayotte & Henry E. Smith, eds., Edward Elgar, 2009), available at <http://www.law.uchicago.edu/files/files/457-261.pdf>; see also *id.* at 11 (describing anticommons as “an assembly problem, nothing more and nothing less”).

have been left idle, and myriad new services have been thwarted, all to protect existing transmissions according to a plan imposed in the TV Allocation Table of 1952. The proximate cause of this non-market failure has been exact, rigid use rights limitations that rest too few utilization choices with competitive markets and too many with regulators who fail to properly weigh economic trade-offs.⁷⁸

The great technological upgrade during this six-decade policy rollout was the digital TV transition, officially initiated in 1987 and completed in June 2009 when the last full-power analog TV stations were switched off.⁷⁹ But in just thirteen years, cellular operators transitioned analog mobile phone users to a fully digital system without mandates and without massive spectrum set-asides, and away from a system that consumed most of the spectrum allocated to the service. Michael Heller asks, and answers, the key legal question: “What’s the difference between cell and television performance? Private versus anticommons ownership of the underlying spectrum—not technological limits.”⁸⁰

B. HD RADIO: BUNDLING RIGHTS

Yet, without controversy, vast vacant space was discovered in the FM band and a large number of new broadcasting rights were issued. In 2002, the

78. For a general analysis of “non-market failure,” see CHARLES WOLF, JR., *MARKETS OR GOVERNMENTS?: CHOOSING BETWEEN IMPERFECT ALTERNATIVES* (MIT Press rev. ed. 1993). “It does not follow that whenever laissez faire falls short government interference is expedient; since the inevitable drawbacks of the latter may, in any particular case, be worse than the shortcomings of private enterprise.” HENRY SIDGWICK, *PRINCIPLES OF POLITICAL ECONOMY* 414 (Macmillan 1883), cited in Charles Wolf, Jr., *A Theory of Nonmarket Failure: Framework for Implementation Analysis*, 22 *J.L. & ECON.* 107, 107 n.1 (1979); see also *id.* at 112 (citing ROBERT BACON & WALTER ELTIS, *BRITAIN’S ECONOMIC PROBLEM: TOO FEW PRODUCERS* (Macmillan 1976)).

Just as the absence of particular markets accounts for market failure, so nonmarket failures are due to the absence of nonmarket mechanisms for reconciling calculations by decision makers of their private and organizational costs and benefits with total costs and benefits. Nor, for reasons we will suggest, are prospects for invention of suitably compensatory nonmarket mechanisms to avoid nonmarket failure notably brighter than for creating suitable markets where their absence leads to market failures.

79. FCC, *Frequently Asked Questions*, DTV.GOV, <http://dtv.gov/consumercorner.html> (“In 1996, Congress authorized the distribution of an additional broadcast channel to each broadcast TV station so that they could use it for digital broadcasting while simultaneously continuing their analog broadcast channel. Later, Congress mandated June 12, 2009 (extended from February 17, 2009) as the last day for full-power television stations in the U.S. to broadcast in analog . . .”).

80. HELLER, *supra* note 16, at 96.

FCC authorized hybrid digital (“HD”) Radio.⁸¹ The new rules permitted each of the existing 8,840 full-power FM radio stations and 4,700 AM radio stations to adopt a new format that, while continuing to transmit a primary analog signal, broadcasted two new digital signals.⁸² This yielded an immediate increase in potential FM broadcasts of nearly 17,680, a 200% increase from 8,840 existing full-power FM stations. When HD receivers become more widespread, FM stations may select an all-digital broadcasting format that permits the sending of seven signals per FM channel, three full-power and four low-power.⁸³ Hence, the FCC’s 2002 authorization for HD Radio cleared a path for another 54,000 FM broadcasts.⁸⁴ In 2005, 1,272 HD stations (195 AM and 1,077 FM) were broadcasting with an in-band on-channel (“IBOC”) system and 700 FM stations had special temporary authority for multicasting with 2,000 more in the pipeline.⁸⁵ In early 2012, more than 2,100 HD radio stations were broadcasting, and some 3 million HD receivers had been sold.⁸⁶ There were more than 650 stations in the top 100 markets, and advertising sales on these stations was projected to exceed \$110 million in 2011.⁸⁷

The HD authorization sailed through the FCC without controversy.⁸⁸ That was not because the new stations do not create any interference. In

81. *See* In the Matter of Digital Audio Broadcasting Systems and their Impact on the Terrestrial Radio Broadcast Service, First Report and Order, 17 FCC Rcd. 19,990, 19,995–96 (2002) [hereinafter 2002 HD Radio R&O], *available at* http://fjallfoss.fcc.gov/edocs_public/attachmatch/FCC-02-286A1.pdf.

82. *Id.* at 19,999 (4,700 AM stations); In the Matter of Digital Audio Broadcasting Systems and their Impact on the Terrestrial Radio Broadcast Service, Second Report and Order First Order on Reconsideration and Second Further Notice of Proposed Rulemaking, 22 FCC Rcd. 10,344, 10,349 ¶ 11 n.17 (2007) [hereinafter 2007 HD Radio R&O], *available at* <http://apps.fcc.gov/ecfs/document/view?id=6520026787> (noting that in 2005, there were 10,973 commercial radio stations, 2,625 FM educational radio stations, and of the commercial stations, 6,215 were FM stations and 4,758 were AM stations).

83. In the Matter of Review of the Emergency Alert System, 20 FCC Rcd. 18,625, 18,637 (2005) (“Radio stations will eventually convert to all-digital modes of operation.”) (addressing emergency alert systems for DTV, DAB, digital cable, DBS, and SDARS).

84. FM stations have maximum ERPs of 100,000 watts (grandfathered or waived cases), 50,000 watts, or the majority with average 6,000 watts of ERP. 2002 HD Radio R&O, *supra* note 81 (6,000 watt ERP Class A FM station). Low-power FM stations are sixty times smaller at 100 watts ERP.

85. 2007 HD Radio R&O, *supra* note 82, at 10,349.

86. *What is HD Radio Broadcasting?*, IBIQUITY DIGITAL, <http://www.ibiquity.com/hd-radio> (last visited Jan. 12, 2013).

87. *HD Radio Alliance Announces 2011 Marketing Plans*, HD RADIO (Nov. 30, 2010), <http://www.hdradio.com/press-room/hd-radio-alliance-announces-2011-marketing-plans>.

88. *See* TaNoah Morgan, *Digital Radio Approved by FCC*, BALTIMORE SUN, Oct. 11, 2002, http://articles.baltimoresun.com/2002-10-11/business/0210110010_1_digital-radio-hd-radio-broadcasters (“The Federal Communications Commission gave its blessing

packing thousands of new full-power transmissions into the FM band, spillovers are technically far deeper an issue than with the insertion of tiny 100W emissions, which generated significant controversy in the Low-Power FM (LPFM) docket.⁸⁹ Much greater electromagnetic levels of “interference” for HD Radio were approved than for LPFM, with the support of the radio industry.⁹⁰

Spillovers were thus more economically managed in HD Radio than in the case of LPFM.⁹¹ This coordination of HD Radio interference was due to the form in which the spectrum rights were issued and the transaction costs surrounding the resulting economic organization.⁹² In short, the tragedy of the commons⁹³ that blocked fuller deployment of the FM band under the regulatory approach employed in LPFM evaporated when the Commission pursued an alternative path, as in the case of HD Radio. A key distinction

yesterday to a digital radio technology for local broadcasters that promises highly improved sound for consumers and new revenue streams for broadcasters.”)

89. *Cf.* 2002 HD Radio R&O, *supra* note 81, at 19,995 (“[M]inimizing interference to stations on first- and, to a lesser extent, second-adjacent channels poses the most serious analog compatibility challenge.”) (citation omitted). Some noted interference concerns for “a limited number of listeners may perceive an impact outside of the protected contour under certain conditions.” *Id.* Interested parties found a way to agree, “including all of the broadcasters that address the issue . . . that this is a reasonable tradeoff.” *Id.* (footnote omitted).

90. 2002 HD Radio R&O, *supra* note 81, at 20,002 (“According to iBiquity, the estimated costs of implementing its hybrid IBOC system range from \$30,000 to \$200,000, with an average cost of \$75,000. Conversion costs vary depending on the age and other characteristics of a station’s transmitter plant and studio equipment.”) (footnote omitted).

91. *See* In the Matter of Econ. Impact of Low-Power FM Stations on Commercial FM Radio: Report to Cong. Pursuant to Section 8 of the Local Cmty. Radio Act of 2010, Report, 27 F.C.C.R. 3 (2012) [hereinafter 2012 LPFM Report], *available at* http://transition.fcc.gov/Daily_Releases/Daily_Business/2012/db0105/DA-12-2A1.pdf. An act of Congress was required after a decade of policy deliberation. *See id.*; LOCAL COMMUNITY RADIO ACT OF 2010, § 8, PUB. L. NO. 111-371, 124 STAT. 4072 (2011), following LPFM interference definitions starting with the 2000 LPFM Order. 2012 LPFM Report, *supra*, at 7 ¶ 14 n.20 (citing In the Matter of Creation of Low Power Radio Service, Report & Order, 15 FCC Rcd. 2205 (2000)).

92. Stations have increased interference between digital and analog receivers by increasing digital power in 2010 in symmetric operation in sidebands, with requests currently pending in 2012 for asymmetric interference due to new technology. *See* In the Matter of Digital Audio Broad. Sys. & Their Impact on the Terrestrial Radio Broad. Serv., Order, 25 FCC Rcd. 1182 (2010), *available at* <http://apps.fcc.gov/ecfs/document/view?id=7020409846>.

93. For an overview on the “tragedy of the commons,” see Hardin, *supra* note 12, at 1243; Thráinn Eggertsson, *Open Access Versus Common Property*, in PROPERTY RIGHTS: COOPERATION, CONFLICT AND LAW 75 (Terry L. Anderson & Fred S. McChesney, eds., 2002); Heller, *The Tragedy of the Anticommons: Property in the Transition from Marx to Markets*, 111 HARV. L. REV. 621 (1998); Fennell, *Common Interest Tragedies*, 98 NW. L. REV. 907 (2004).

was that LPFM awarded new rights to *entrants*, and HD Radio implemented software upgrades and asymmetric sideband innovations for *incumbents*.⁹⁴ There were two forces at work to determine this bifurcated outcome, with LPFM reflecting deep structural inefficiencies in spectrum allocation and with HD Radio reflecting important efficiency considerations in the creation and distribution of property rights.

The moral is not that new entry is impossible in the absence of industry capture, or even that regulatory disputes over harmful interference have their roots in financial interest rather than in “technical” spillovers. The lesson here is that the *manner* in which spectrum rights are awarded—the legal rights yielding economic incentives—decisively determines whether “technical” opportunities by way of innovative combinations of receivers, band plans, and software will be found or lost. The FCC’s efforts to craft superior radio contours are a sideshow. When the FCC attempted to define “white spaces” to accommodate new LPFM broadcasting on the FM dial without the cooperation of full-power FM stations, it effectively found nothing. When it gave incentives to FM stations to find the iBiquity HD “white spaces,” suddenly thousands of new opportunities were discovered.

Another key aspect of the HD Radio spectrum allocation, however, is rarely noted. In permitting incumbent licensees to place three (or seven) broadcasts in the frequency space (200 KHz) where only one broadcast had previously been authorized, the FCC bundled new rights with an incumbent licensee’s existing rights.⁹⁵ By expanding rights bundles, rather than inserting new rights (and rights holders) immediately adjacent to existing broadcasters, the FCC avoided the fragmentation of new, contentious borders. There will still be borders, technically speaking, and spillovers that cross them. But with both of the rival broadcasts owned by the same licensee, the costs and benefits of emissions are internalized.

FM stations are thus able to increase their spectrum usage, balancing new stations with existing broadcasts, by upgrading to digital technology. This upgrade was mandated as part of the authorization and is a product of traditional FCC regulation, which specifies exactly how licensees use

94. Comments of iBiquity Digital Corp., In the Matter of Digital Audio Broadcasting Systems And Their Impact on the Terrestrial Radio Broadcast Service, Attachment A, at 4, MM Docket No. 99-325 (Dec. 19, 2011), *available at* <http://apps.fcc.gov/ecfs/document/view?id=7021751178> (“HD Radio software versions 4.4.x and above allow for independent, asymmetric sideband control by implementing a new peak-to-average power ratio reduction algorithm.”).

95. *See supra* note 83.

airwaves.⁹⁶ But this upgrade does not alter the underlying efficiency of the bundled nature of the new rights awarded. In fact, efficiencies of greater magnitude are observed when liberal licenses—with no technology or service mandates—are awarded.⁹⁷

Far more broadly, it is a principle in property law that *complementary resource rights* be defined together. For example, the landowner has both broad control of her real estate and the right to exclude those not granted the right to encroach. But when airplanes were invented and commercial flights began, the question arose: did each landowner have a claim to exclude them? In *United States v. Causby*, the Court found that the right to own real property did not include air routes some 30,000 feet overhead.⁹⁸ There was little complementarity between the planes (or the airspaces they used) and the land below. Moreover, tying ownership of the high sky to the land would have created significant, and likely prohibitive, transaction costs. By deleting these rights from the land ownership bundle, those transaction costs were avoided. As with HD Radio, but with the opposite result (splitting rights rather than bundling them), the transactionally efficient path was found. Such divergent results are in direct accord with Demsetz's work on the costs and benefits of rights definition.⁹⁹

C. LICENSED PCS VERSUS UNLICENSED PCS: ADMINISTRATIVE RULES VERSUS MARKET MECHANISMS

The allocation of Personal Communications Services (“PCS”), which stretched (officially) from 1989 to 1995, authorized CMRS licenses (120 MHz) and an unlicensed PCS (“U-PCS”) band (30 MHz).¹⁰⁰ The licensed bandwidth has been intensely utilized in mobile phone networks, and is responsible for a significant fraction of the more than \$200 billion in annual consumer surplus delivered by mobile services.¹⁰¹ The U-PCS allocation has

96. The HD technology was not developed by the FCC, but by iBiquity, a private company that has allied with radio broadcasters to produce HD Radio. Radio stations petitioned for the HD authorization and recommended that the FCC implement reforms that included the iBiquity-HD technology upgrade. See 2002 HD Radio R&O, *supra* note 81, at 19,992 ¶ 6 (“The NAB concurs, stating simply, ‘[i]t works; it’s ready’”) (from comments of the NAB on Feb. 19, 2002).

97. See Thomas W. Hazlett & Evan Leo, *The Case for Liberal Spectrum Licenses: A Technical and Economic Perspective*, 26 BERKELEY TECH. L.J. 1037 (2011).

98. 328 U.S. 256 (1946).

99. See *supra* note 43 and accompanying text.

100. NBP 2010, *supra* note 3, at 79 fig.5-C; see also *infra* notes 103–04.

101. See Hazlett, Muñoz & Avanzini, *supra* note 6.

been notably unsuccessful, with most of the bandwidth used lightly if at all.¹⁰² Recent FCC proceedings have either reallocated the U-PCS spectrum or reconfigured rules in an attempt to induce more device deployment.¹⁰³

A major obstacle for the entire PCS band was the presence of some 4,500 incumbent microwave licensees. These operators consumed relatively little bandwidth but were adamantly opposed to relocating (by either using new frequencies or switching operations to alternative communication modes, such as fiber or satellite links).¹⁰⁴ A years-long political standoff froze the U.S. PCS allocation even as E.U. countries were, between 1989 and 1992, allocating 2G licenses (equivalent to PCS). What broke the deadlock was a compromise in the form of “overlay” licenses.¹⁰⁵ New PCS licenses would be issued as overlays, meaning that incumbent licensees would be grandfathered.¹⁰⁶ Once the new licenses had been auctioned and assigned, these overlay licensees could then bargain with the incumbents to release their spectrum by moving to an alternative option—induced by payments from the overlay licensees. Supported by an FCC decision to impose arbitration procedures (to determine the costs of relocation, for which the

102. U-PCS operations were authorized by the FCC in 2004 with spectrum etiquette standards. In the Matter of Amendment of Part 15 of the Commission’s Rules Regarding Unlicensed Pers. Communications Serv. Devices in the 1920–1930 MHz Band, Notice of Proposed Rulemaking, 25 FCC Rcd. 5118, 5119–20 ¶ 4 (2010) [hereinafter 2004 U-PCS NPRM], available at http://hraunfoss.fcc.gov/edocs_public/attachmatch/FCC-10-77A1.pdf (where “listen-before-transmit” protocols involved channel monitoring between unlicensed devices). The Commission provided additional technical flexibility in 2004 to promote greater use of the spectrum, *id.* at 5120–21 ¶ 5, and again in the 2012 U-PCS R&O, *infra* note 103. The band suffered from underuse, precipitating requests to modify the rules to allow for “more devices to access usable channels . . . currently [] restricted from use under the existing 50 dB above thermal noise threshold, but that are actually acceptable for use.” 2004 U-PCS NPRM at 5123 ¶ 12.

103. In the Matter of Amendment of Part 15 of the Commission’s Rules Regarding Unlicensed Pers. Communications Serv. Devices in the 1920–1930 MHz Band, Report & Order, 27 FCC Rcd. 3645 (2012) [hereinafter 2012 U-PCS R&O], available at http://hraunfoss.fcc.gov/edocs_public/attachmatch/FCC-12-33A1.pdf.

104. 2004 U-PCS NPRM, *supra* note 102, at 5126–27 ¶ 18 (describing the formation of a coalition in 1993 of a Unlicensed PCS Ad Hoc Committee for 2 GHz Microwave Transition and Management (UTAM, Inc.) to manage the fixed microwave incumbent transition); *see also* In the Matter of Amendment of the Commission’s Rules to Establish New Personal Communications Services, Second Report & Order, 8 FCC Rcd. 7700, 7775–78 (1993) [hereinafter 1993 Broadband PCS Second R&O], *modified on recon.*, 9 FCC Rcd. 4957 (1994), available at http://hraunfoss.fcc.gov/edocs_public/attachmatch/FCC-94-144A1.pdf.

105. 1993 Broadband PCS Second R&O, *supra* note 104; *see* Cramton et al., *infra* note 122, at 664–65.

106. 1993 Broadband PCS Second R&O, *supra* note 104; *see* Cramton et al., *infra* note 122, at 664–65.

overlay licensee was liable) and time limits for negotiations, the set-up worked to clear licensed PCS bands—an effective approach to mitigating harmful interference in that PCS licensees would not pay moving costs were these costs to exceed the benefits of microwave relocation.¹⁰⁷

U-PCS had a more difficult time clearing 1,100 incumbents to mitigate interference. Unlike the licensed PCS bands, no entity internalized both the benefits and costs of band clearing. This created a problem anticipated by the FCC, which set up rules to deal with it in its initial U-PCS Order.¹⁰⁸ First, it mandated that devices (such as cordless phones) used in the band employ a “listen before talk” protocol.¹⁰⁹ Restricting access to vacant channels presumably prevents interference with existing transmitters. Second, a frequency coordinator, Unlicensed Transition and Management (“UTAM”), was assigned the task of monitoring conflicts and relocating incumbents.¹¹⁰ This organization was supported by a fee on U-PCS devices, originally set by the FCC at \$20 per unit sold. The purpose was to mimic what a spectrum owner would do, paying conflicting users to relocate in a situation where the benefits of clearing U-PCS spectrum would not accrue to any particular device vendor (all would potentially realize the gain) or licensee (there was none).¹¹¹

While incumbents were eventually cleared from the U-PCS band in April 2005, the process in licensed PCS spectrum (“L-PCS”) proceeded far more expeditiously.¹¹² Moreover, the use of tools used to mitigate interference

107. *See, e.g.*, Wireless Telecommunications Bureau Announces Commencement of the Voluntary Negotiation Period for 2 GHz Microwave Incumbents Operating in the Broadband PCS “C” Block, DA 96-838 (May 24, 1996), http://fallfoss.fcc.gov/edocs_public/attachmatch/DA-96-838A1.pdf.

108. In the Matter of Amendment of the Commission’s Rules to Establish New Personal Communications Services, Fourth Memorandum Opinion and Order, 10 FCC Rcd. 7955 (1995) [hereinafter 1995 Broadband PCS Fourth MO&O], *available at* http://hraunfoss.fcc.gov/edocs_public/attachmatch/FCC-95-167A1.pdf.

109. *See* 2004 U-PCS NPRM, *supra* note 102, at 5119–20 ¶ 4.

110. 1995 Broadband PCS Fourth MO&O, *supra* note 108, at 7955 ¶ 1.

111. *Id.* at 7957 ¶ 9.

112. There were severe problems in deploying some of the licensed PCS spectrum, but they were not associated with the transaction costs of band clearing. The A and B PCS licenses, allocated 30 MHz each, were auctioned in 1995 and the spectrum was productively deployed soon thereafter. The assignment of D and E licenses (10 MHz each) took place following a 1996 FCC auction, with similarly smooth deployment. However, the C (30 MHz) and F (10 MHz) licenses, auctioned in 1996 with “designated entity” (“DE”) bidding preferences extended to small businesses and rural telephone carriers met with disaster. DEs received long-term loans at Treasury bond interest rates to buy licenses, a risk shifting the FCC came to regret. The largest DEs overbid for licenses, declared bankruptcy, and were then protected by bankruptcy law from having to return their FCC licenses. Indeed, they received substantial price reductions after bankruptcy declarations. The FCC objected and

proved far more onerous in U-PCS than in L-PCS.¹¹³ While there are other important factors at work in this natural experiment (pitting licensed against unlicensed spectrum allocations), the *technical* nature of the interference faced from microwave incumbents in either case was *identical*. Administrative rules for interference control specified in 2004 and modified in 2012, in the case of U-PCS, were tested against market mechanisms deployed in the case of L-PCS through spectrum auctions and liberal licenses. By delegating broad, exclusive authority over how to use designated spectrum spaces to licensees, the FCC efficiently resolved an “interference dispute.” Hold-ups were overcome while onerous technology restrictions were avoided. The resulting productivity—at least \$40 billion annually¹¹⁴—illustrates the principle: border issues are often best dealt with not by regulating receivers or transmitters, but by issuing rights that obviate the need for such complicated regulatory tasks.

D. WCS/SATELLITE RADIO: THE BENEFITS OF MERGER

The dispute between Wireless Communications Service (“WCS”) licensees and Satellite Digital Audio Radio Service (“SDARS”) licenses over interference rules is often cited as an example of interference disputes that require administrative decree. The WCS/SDARS border has made, and continues to make, the WCS band unusable for high valued products such as mobile voice and data services.¹¹⁵ Allocated 25 MHz in the 2.3 GHz band,

took the case to the U.S. Supreme Court, where it lost. Deals were then struck between the bankrupt parties and the FCC, with the Commission ending up with a large number of the C and F block licenses. These were finally re-auctioned in 2005. See Thomas W. Hazlett & Babette E.L. Boliek, *Use of Designated Entity Preferences in Assigning Wireless Licenses*, 51 FED. COMM. L.J. 639 (1999); *FCC v. NextWave Personal Commc’ns., Inc.*, 537 U.S. 293 (2003) (holding that the FCC cannot “revoke a license held by a bankruptcy debtor upon debtor’s failure to make timely payments”); HAROLD W. FURCHTGOTT-ROTH, *A TOUGH ACT TO FOLLOW?: THE TELECOMMUNICATIONS ACT OF 1996 AND THE SEPARATION OF POWERS FAILURE* 126–27 (AEI Press, 2006).

113. See *In Re* Amendment of Commission’s Rules to Establish New Pers. Communications Services, Narrowband PCS, Second Report & Order & Second Further Notice of Proposed Rule Making, 15 FCC Rcd. 10,456, 10,463 ¶¶ 11, 13 (2000); *id.* at 10,464 ¶ 15 (declining to create additional licenses, but allowing for aggregation of licenses, and supporting larger licensing areas for narrowband PCS).

114. *Welfare Analysis*, *supra* note 68, at 433. Hazlett & Muñoz estimated a marginal social value of a 60 MHz license equal to approximately \$10 billion in 2003. *Id.* at 434. The infra-marginal PCS allocation, allocated 120 MHz, would likely be worth more per MHz (i.e., exceed \$40 billion in annual welfare gains). NBP 2010, *supra* note 3, at 85 fig.5-F.

115. The FCC adopted an order to modify rules on the contentious WCS/SDARS band to allow for LTE Mobile Broadband. See *In the Matter of* Amendment of Part 27 of the Comm’n’s Rules to Govern the Operation of Wireless Communications Services in the 2.3 Ghz Band, Order on Reconsideration, 27 FCC Rcd. 13,651 (2012) [hereinafter 2012 WCS/SDARS Order on Reconsideration], available at http://transition.fcc.gov/Daily_

WCS airwaves could potentially generate several billions annually in consumer surplus.¹¹⁶

The proximate cause of the spectrum border dispute is that satellite radio signals arrive at the subscriber's radio relatively weakly (i.e., with low power), given the long distances they travel in geosynchronous orbit around the equator. The satellite radio signals are susceptible to very low levels of radio emissions spilling from adjacent bands—and more so when those levels are raised due to the use of millions of mobile receivers such as handsets. This is because mobile use of both those WCS devices and SDARS receivers would tend to put many of the latter in close proximity to the former in unpredictable ways. Perhaps better, more technically sophisticated use rights for the respective services would solve the problem.¹¹⁷ This point is almost irrelevant without a new structure for the regulatory system, which for fifteen years has failed to adopt whatever new technical rules are advocated now,¹¹⁸ and ignores the obvious market-based solution that was rejected by the FCC: merger.

In 2005, XM Satellite Radio, an SDARS licensee, offered to pay \$198 million to buy the WCS licenses.¹¹⁹ The combination would have eliminated the WCS/SDARS border dispute by eliminating the border through economic alignment. Indeed, the most widespread and effective solution to technical definition issues is to sidestep them by allowing integration of

Releases/Daily_Business/2012/db1023/FCC-12-130A1.pdf. For descriptions of LTE and WiMax uses on the band, see *id.* at 13,660–61 ¶ 18 n.57; *id.* at 13,673–74 ¶ 51.

116. The 2.3 GHz band is close, and similar in propagation characteristics, to the PCS band (1.9 GHz) and the AWS band (1.7/2.1 GHz). Frequencies in these bands are commonly used to provide mobile voice and data services, including 3G and 4G high-speed data connections. This suggests that the WCS allocation could, under alternative regulatory rules, host traffic about as valuable as seen in cellular markets—where an incremental 30 MHz was found, in 2003, to add over \$10 billion in social value to the U.S. economy. *Welfare Analysis*, *supra* note 68.

117. Consider the implications of the 2012 order, where AT&T and Sirius XM agreed upon ground power level targets, where the FCC could not broker a resolution for many years. 2012 WCS/SDARS Order on Reconsideration, *supra* note 115, at 13,663–64 ¶¶ 22–26.

118. See 2012 WCS/SDARS Order on Reconsideration, *supra* note 115, at 13,655–56 ¶ 7 (discussing the procedural history beginning in 1997). Recall that the 30 MHz on the WCS/SDARS band was heralded for the WiMAX innovation in 2010, with the need for administrative action in 2012 to specify LTE. See In the Matter of Amendment of Part 27 of the Commissions Rules to Govern the Operation of Wireless Communications Services in the 2.3 Ghz Band, Report & Order & Second Report & Order, 25 FCC Rcd. 11,710, 11,718 ¶ 15 (2010), available at http://hraunfoss.fcc.gov/edocs_public/attachmatch/FCC-10-82A1.pdf.

119. *XM to Buy WCS Wireless for Nearly \$200 Million*, MEDIA POST (July 14, 2005), <http://www.mediapost.com/publications/article/32096/xm-to-buy-wcs-wireless-for-nearly-200-million.html>.

ownership.¹²⁰ The FCC blocked the merger between XM Radio and WCS, bending to pressure by satellite radio rivals (terrestrial radio stations) to deny the evidently pro-competitive transaction.¹²¹ Importantly, the efficient solution to this particular interference dispute was to integrate ownership of the conflicting licenses, a remedy that market forces sought to implement and that regulators blocked. That leading analyses of how to craft policies dealing with radio interference routinely ignore the importance of merger is emblematic of the confusion in the discussion.

E. NEXTEL/PUBLIC SAFETY RADIO: ELIMINATING BORDERS

The long-running dispute between Nextel and public safety radio users, such as police and fire departments, is almost universally cited in the interference literature.¹²² In brief, the problems at issue came up when little used dispatch licenses, known as Specialized Mobile Radio (“SMR”) licenses, allocated in this 800 MHz band, were relaxed to permit the use of digital services to supply mobile phone calls to the public. In going from “pizza delivery” radio¹²³ to a full-fledged competitor in cellular, the SMR licenses

120. This is a subtle but far-reaching point. A very general efficiency in the creation and distribution of property rights is to award such rights, initially, such that fragmentation—and associated issues of border definition and enforcement—does not preempt efficient resource uses. A leading goal of legal rules or regulatory policy, then, must be to allow or even enable efficient rights aggregation so that optimally sized bundles of ownership rights are not preempted by transactions costs. Importantly, legal processes that overly divide and separate complementary rights here impose such transaction costs implicitly. See HAROLD DEMSETZ, *Reinterpreting the Externality Problem*, in FROM ECONOMIC MAN TO ECONOMIC SYSTEM: ESSAYS ON HUMAN BEHAVIOR AND THE INSTITUTIONS OF CAPITALISM 112–13 (2008). (“Legal error has caused the problem, not positive transaction cost. There is no inefficiency in the way the market accommodates to the court’s mistake.”) (emphasis omitted). See also HELLER, *supra* note 16, at 104 (“Given the forced nature of the exchange, markets couldn’t price network fragments.”); *id.* at 77 (“The crucial point is that the emerging structure of drug discovery clashes more and more with old-fashioned patent law and competition policy.”).

121. *XM Satellite, WCS Wireless Abandon Merger*, FIERCE WIRELESS (May 22, 2006), <http://www.fiercewireless.com/story/xm-satellite-wcs-wireless-abandon-merger/2006-05-23> (“XM Satellite and WCS Wireless abandoned their agreement that called for XM to buy WCS and its wireless spectrum licenses because they were unable to receive regulatory approvals for the deal.”). Since then, Sirius Satellite Radio and XM Satellite have merged to form Sirius XM Radio Inc. SIRIUS and XM Complete Merger, Sirius XM Radio (July 29, 2008), <http://investor.sirius.com/ReleaseDetail.cfm?ReleaseID=324858>.

122. See, e.g., Dale Hatfield, *Radio Regulation Summit: Defining Out-of-Band Operating Rules*, Silicon Flatirons Center for Law, Technology, and Entrepreneurs, Sept. 8–9, 2009, available at <http://www.silicon-flatirons.org/documents/misc/OOBSummit/SFC%20Interference%20Summit%20-%20September%202009wo.ppt>; Peter Cramton, Evan Kwerel & John Williams, *Efficient Relocation of Spectrum Incumbents*, 41 J.L. & ECON. 647 (1998).

123. The SMR licenses had been used for dispatch calls such as those made for taxi pick-ups and pizza deliveries.

were transformed from a wireless backwater into a high-productive addition to the emerging Information Economy.¹²⁴ At the same time, the millions of mobile handsets using the SMR frequencies were generating more noise, and radio emissions were much more likely to spill over into adjacent frequencies.¹²⁵

Many public safety agencies used those adjacent bands for their radio communications. As has been well documented by the FCC, these users complained that vital transmissions were being endangered. Regulators took such complaints, which implicated life or death outcomes, seriously.¹²⁶ For many years, however, the dangerous situation continued.¹²⁷ Finally, the FCC embraced a Nextel-proposed “spectrum swap.”¹²⁸ Nextel would (a) give up some of its licenses to use bands next to fire and police bands; (b) pay for new equipment like radios and base stations so public safety users could begin using Nextel’s abandoned SMR bands; (c) receive new FCC licenses allowing them to use an altogether different frequency location (10 MHz at

124. See HELLER, *supra* note 16, at 93. For an account of entrepreneur Morgan O’Brien’s successful efforts to turn under-used spectrum into wireless gold, see JAMES B. MURRAY, JR., WIRELESS NATION: THE FRENZIED LAUNCH OF THE CELLULAR REVOLUTION IN AMERICA 251–66 (2001).

125. In the Matter of Improving Pub. Safety Communications in the 800 Mhz Band, Report & Order, Fifth Report & Order, Fourth Memorandum Opinion & Order, and Order 19 FCC Rcd. 14,969, 14,983 ¶ 21 (2004) [hereinafter 800 MHz R&O], available at http://hraunfoss.fcc.gov/edocs_public/attachmatch/FCC-04-168A1.pdf (showing a diagram of the multiparty band).

126. 800 MHz R&O, *supra* note 125, at 14,975–76 ¶ 7 (“These considerations require that we take the most effective actions, in the short-term and long-term, to promote robust and reliable public safety communications in the 800 MHz band to ensure the safety of life and property.”).

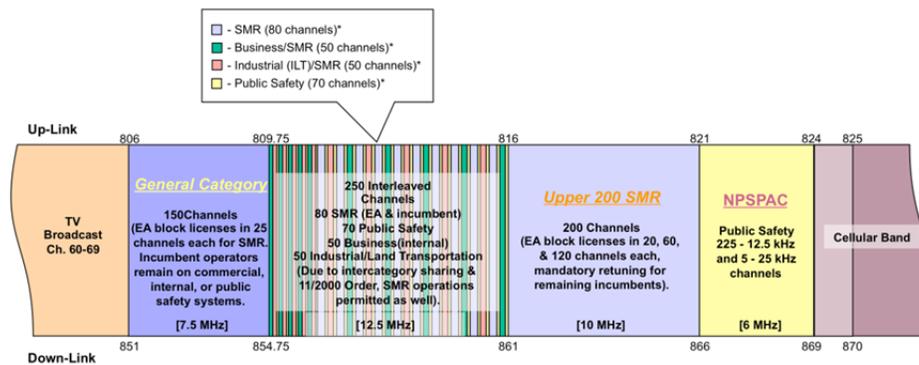
127. *Id.* at 14,976 ¶¶ 8–9 (citing the ongoing problem of interference to 800 MHz public safety communications systems, which the FCC could not resolve by either band reconfiguration or case-by-case “technical fixes”).

128. *Id.* at 14,987–88 ¶ 31; *id.* at 14,989–94 ¶ 35–42. The Consensus Plan reordered the Nextel-Public Safety border. For a detailed timeline, see *id.* at 15,002–10 ¶ 61. The Notice of Proposed Rulemaking began in 2002, in *In Re* Improving Pub. Safety Communications in 800 Mhz Band, Notice of Proposed Rule Making, 17 F.C.C.R. 4873, 4882 ¶ 16 (2002). Private solutions for “voluntary technical changes to prevent or reduce interference” of the multidimensionally inefficient borders were proposed in 2000 but gridlocked with regulatory rights of interference. 800 MHz R&O, *supra* note 125, at 14,979–80 ¶¶ 14–15 (“The Consensus Parties have proposed a band reconfiguration plan that would move ESMR systems—most notably Nextel—to the upper portion of the 800 MHz band, move all public safety and “high site” operators to the lower portion of the band, and make additional spectrum in the band available for public safety use.”); *id.* at 14,982–83 ¶ 20 (“The method of interference abatement we adopt herein leaves to the involved parties—and not the Commission—the choice of how best to ensure that their systems do not cause unacceptable interference.”).

1.9 GHz); and (d) pay \$4.8 billion for the new award, less the relocation costs of \$850 million or more, which Nextel would pay for public safety radios.¹²⁹

The common conclusions taken from this episode are (1) that border definitions are difficult, which is true; and (2) that more careful technical specification of spectrum rights is needed to deal with this problem, which is false. The lack of specificity was neither the true cause of the problem nor the answer that was eventually implemented. The underlying problem was that the FCC spectrum allocation scheme for SMR was a model of regulatory mischievousness. The band plan featured small slices of bandwidth, channels to be assigned to different licenses, and licenses to be assigned to different users. These narrow channels were “interleaved,” the opposite of the solution needed and later adopted, which featured contiguous blocks of spectrum.¹³⁰

Figure 1: FCC 800 MHz SMR/Public Safety Band: 2001¹³¹



129. 800 MHz R&O, *supra* note 128, at 14,987–88 ¶¶ 29, 31.

130. See FCC, *SMR and Cellular Frequencies Presentation*, <http://wireless.fcc.gov/releases/011121-exhibit.ppt> [hereinafter *SMR and Cellular Frequencies Presentation*]; NEXTEL COMMUNICATIONS, INC., WHITE PAPER, PROMOTING PUBLIC SAFETY COMMUNICATIONS: REALIGNING THE 800 MHz LAND MOBILE RADIO BAND TO RECTIFY COMMERCIAL MOBILE RADIO-PUBLIC SAFETY INTERFERENCE AND ALLOCATE ADDITIONAL SPECTRUM TO MEET CRITICAL PUBLIC SAFETY NEEDS 7 (2001), available at http://wireless.fcc.gov/releases/011121-whitepaper_final.pdf; Letter from Nextel Communications, Inc., to FCC (Nov. 21, 2001), available at <http://wireless.fcc.gov/releases/011121-letter.txt>. In the public safety band, interleaving occurred at the license layer. Interleaving of spectrum can also occur on a time, frequency, or geographic basis, where modern (code division) CDMA and (frequency division) OFDMA modulation standards are embedded within mobile hardware.

131. *SMR and Cellular Frequencies Presentation*, *supra* note 130 at 6; *Nextel 800 MHz Interference Plan*, DISPATCH MAGAZINE, http://www.911dispatch.com/info/800_transition/nextel_slides.html [hereinafter *Nextel Interference Plan*] (as presented in Nextel’s slides in FCC presentation).

A picture is worth a thousand words or, at least, 6.25 MHz. Notice how this modest block of spectrum was split into 250 separate channels in the original SMR band plan, depicted in Figure 1. Each channel was then assigned to one license, and licenses were distributed to different licensees. In bandwidth that would constitute less than one license elsewhere—cellular licenses are allocated 25 MHz each—literally hundreds of borders were created by the FCC rights creation process.¹³² This dictated needless costs in social coordination, which could only be undone by a process of license aggregation.

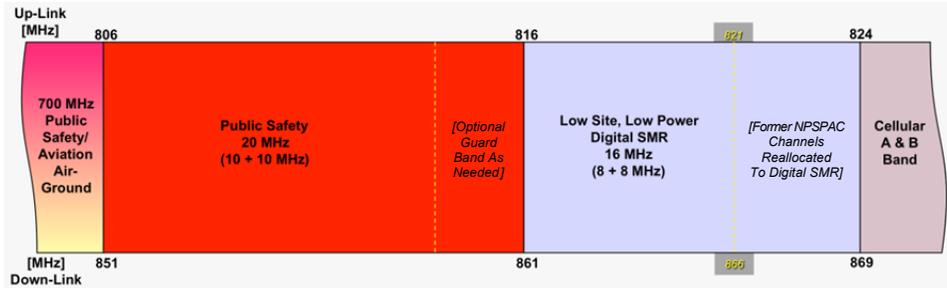
Nextel pursued license aggregation, advancing a plan for FCC spectrum reorganization and lobbying public safety interests to support it, as seen in Figure 2, *infra*. The general approach of aggregating licenses was nothing new for Nextel. Indeed, the Nextel network had been stitched together by the acquisition of over 40,000 SMR licenses purchased in secondary markets.¹³³ This strategy enabled Nextel, a national carrier, to emerge to compete with the cellular duopoly.¹³⁴ It created a resource so valuable that Nextel ultimately served 15 million subscribers and was sold to Sprint for \$35 billion in 2005.¹³⁵ However, the process of license aggregation stopped short when Nextel targeted licenses controlled by public safety organizations. These licenses were not owned by profit-seeking enterprises, so trades were problematic, if not illegal. Moreover, the balkanized nature of the rights, strewn throughout tens of thousands of local agencies, frustrated collective action.

132. Over 2,200 filings claimed different positions on interference in the 800 MHz band, with “engineering, economic, legal and policy analysis” to align for a solution that would be “technically sound, effective, and equitable to the parties.” 800 MHz R&O, *supra* note 125, at 15,002–10 ¶ 61. Stakeholders included APCO, Nextel, CTIA, Motorola, Public Safety Wireless Network (“PSWN”), National Association of Manufacturers (“NAM”), MRFAC (a FCC-certified frequency coordinator), B/ILT and cellular SMR licensees, 800 MHz Users Coalition (whose “balanced approach” claimed to technically solve interference to obviate need for band reconfiguration in the Consensus Plan), Anne Arundel County of Maryland (a public safety user), Verizon Wireless, Industry Canada, the city of Denver, and the city and county of San Diego. *Id.*

133. Hazlett, *Federal Preemption*, *supra* note 29, at 193 tbl.8.

134. Cellular licenses were awarded two per market, primarily by lotteries conducted by the FCC, in 1984–1989. *See* MURRAY, *supra* note 124.

135. John Shinal, *Sprint, Nextel Holders Approve Deal*, WALL ST. J. MARKETWATCH, July 13, 2005, http://articles.marketwatch.com/2005-07-13/news/30803396_1_nextel-share-holders-sprint-nextel-corp-antitrust-approval; Nextel Communications, Inc., 10-K, File No. 0-19656, (Dec. 31, 2004), <http://www.sec.gov/Archives/edgar/data/824169/000095013305001019/w05804e10vk.htm>; 800 MHz R&O, *supra* note 125, at 15,020 ¶ 82 (“Allocating spectrum to establish a long-term solution to the public safety interference problem and support the associated rebanding is a valid use of spectrum in the public interest.”).

Figure 2: Nextel's Spectrum Reorganization Proposal: 2001¹³⁶

De Vries and Sieh note that:

Nextel did not pursue its claims but instead offered to provide the upfront cash to fund rearrangement of the operations in the 800 MHz band in exchange for a nation-wide license to 10MHz of contiguous frequencies in a separate band. As a result, the FCC never had to decide which parties may have been at fault.¹³⁷

This passage describes how a profit-maximizing firm, interested in welfare-generating outcomes, pursues gains from trade. The transaction eliminates disputed borders, and dispenses with the thankless and expensive task of finding fault. Indeed, that fault determination might take forever, as the authors explain the situation as a conflict that emerged despite being a “paradigmatic example of where radio operators have both been operating within their rights.”¹³⁸ In other words, the fault lay with the rules, not the parties. Yet the FCC spent years in a fruitless search for liability. By terminating that source of delay, the “spectrum swap” productively remedied harmful interference.

De Vries and Sieh, rather than heralding the mutually beneficial outcome or recommending ways to smooth the path for such solutions in the future, decry the lack of technical specificity in FCC rules: “We contend that the root cause of the [conflict] was that there had not been a sufficiently clear delineation of *who* had *what* rights according to their individual license, and what actions each party could be required to take to resolve the conflict.”¹³⁹

This view ignores (a) the costs of the FCC striving to create “sufficiently clear delineation,” and (b) alternative ownership mechanisms that efficiently

136. *SMR and Cellular Frequencies Presentation*, *supra* note 130; *Nextel Interference Plan*, *supra* note 131 (graphic altered for readability).

137. DE VRIES & SIEH, *supra* note 57, at 2.

138. *Id.*

139. *Id.* at 3.

solve the coordination problem at issue. The FCC had, indeed, imposed detailed emission rights in the SMR band. These decrees were clear enough, in the De Vries and Sieh analysis, to handle the interference problem until Nextel, gaining certain regulatory waivers, began packing far more traffic—creating far more economic value—into the band. The FCC could have, at any point, adjusted boundary conditions. It failed to do so. The problem festered, and was only resolved when the FCC authorized Nextel’s suggested “spectrum swap,” which mimicked a secondary market transaction.¹⁴⁰

Hence, the FCC’s own actions and admissions reveal that the task of further rights delineation was onerous compared to the alternative of simply rearranging ownership rights using existing boundary delineations. This is not to say that there was never a problem in the conflicts between Nextel and public safety users. Quite the reverse. The actual “root” cause of the issue lay in the manner in which economic rights were defined and enforced, separate and apart from the technical specifications of border spillovers. First, the FCC allocation plan had inefficiently interspersed the use rights of rival and disparate parties, interleaving hundreds of licenses per market.¹⁴¹ Second, it had then distributed licenses to thousands of parties across the nation. Third, the licensees were largely nonprofit organizations. The effect of these regulatory choices was to create a tragedy of the anti-commons, a situation where highly complementary rights are held by widely dispersed entities, making efficient rights-aggregations a high transaction cost proposition.¹⁴²

140. *SMR and Cellular Frequencies Presentation*, *supra* note 130; *Nextel Interference Plan*, *supra* note 131.

141. *SMR and Cellular Frequencies Presentation*, *supra* note 130, at slide 6.

142. This point on positive transaction costs from inefficiently fragmented complementary rights by which to measure price and conduct exchange is not to be confused with strategic behavior of owners of inefficient property rights “[b]ut if the reader insists on confusing the two, so be it, but, then, please remember that situations in which strategic behavior is important are but a part, and a small part at that, of all situations that involve positive transaction cost.” DEMSETZ, *supra* note 120, at 117. Demsetz also discusses the two categories of ownership problems, “the content of a privately owned bundle of rights and those that relate to the identity of the owner of the bundle.” *Id.* at 95. As does Heller:

The Russian storefront experience taught me a fundamental lesson about ownership: the *content* of property rights can matter as much as the *clarity* of ownership Regulators inadvertently fragment ownership into so many uncoordinated rights that it becomes impossible to reassemble the ownership egg, even for the state Don’t forget that governments are creating property rights every day, right here, all along the innovation frontier. Corporate law, mergers and acquisitions, subprime mortgage regulations, bankruptcy law, Internet regulation—they can each create invisible tragedies of the anticommons.

Here the situation was exacerbated by the involvement of government agencies, notably difficult to remunerate for cooperating in economic bargains due to the non-economic incentives of enterprise decision-makers.¹⁴³

The Nextel spectrum swap took too long to implement and should not have been necessary. But if the solution lay in the FCC identifying, specifying, and then implementing precise spectrum use provisions along the pre-swap boundaries, it is likely that the regulatory problem would remain unresolved and, in any event, far less value would be provided to wireless users. Instead, the problems and implemented solutions should be grasped, resulting in better regulatory policy in the future. FCC spectrum allocations should studiously avoid creating anti-commons in wireless markets through the practice of interleaving. At a minimum, licenses should be distributed by auction—including rules conducive to efficient aggregation, namely combinatorial bidding¹⁴⁴—that permit markets to assemble heterogeneous assets into productive bundles. Distributing licenses to independently operated nonprofit or government organizations imposes a huge cost on future wireless solutions. Where possible, the wireless services necessary for the operations of these organizations should be purchased in the competitive marketplace, not manufactured in compartmentalized, local, government-run networks.¹⁴⁵ Public safety radios are bought from Motorola or Raytheon; public safety network services should also be purchased from Verizon, AT&T, Sprint, T-Mobile, Clearwire, or Globalstar (or some combination

HELLER, *supra* note 16, at 147–48 (emphasis in original). He continues: “[Russian] socialist law was more concerned with the identity of the owner than with the type of property or scope of rights. Property owned by the state received more protection than cooperative property Personal property held by individuals received the least regard of all.” *Id.* at 149.

143. Cramton et al., *supra* note 122, at 664–65. Specifically note the problem created when non-profit licensees are involved in spectrum reallocations. The problem, when PCS licensees sought to buy out incumbent microwave users, extended to rate-of-return regulated utilities. The problem mirrors that of non-profit organizations in that the utility, when making an ostensibly “win win” deal to switch spectrum use, is duty-bound to relinquish any payments to ratepayers (in the form of rate reductions) so as to keep the firm’s rate-of-return constant.

144. David Porter, Stephen Rassenti, Anil Roopnarine & Vernon Smith, *Combinatorial Auction Design*, 100 PROC. NAT’L ACAD. SCI. 11,153 (2003).

145. This point was discovered and argued by University of Chicago Law School student Leo Herzel in 1951. Having made the point in a comment in the law review, he was subject to mocking counter-attack from Dallas Smythe, who at one point served as FCC chief economist: “Surely it is not seriously intended that noncommercial radio users (such as the police) . . . should compete with dollar bids against the broadcast users for channel allocations.” Leo Herzel, *Facing Facts About the Broadcast Business: Rejoinder*, 20 U. CHI. L. REV. 106, 106 (1952) (citing Dallas Smythe, FCC Economist).

thereof).¹⁴⁶ Moving license awards from “beauty contests” granting rights to non-profits reluctant to abandon them towards economically motivated for-profits enterprises will do far more to solve interference.¹⁴⁷

That Nextel had to enlist the FCC to implement its trade, involving thousands of non-profit government enterprises, caused a years-long delay when compared to corresponding processes conducted with for-profit enterprises. But the solution arrived at is crucial to understand: moving use rights among organizations, not improvements in technical rules or receiver regulations, corrected the problem such that spectrum ultimately was put to supplying higher valued uses.

The delineation of interference rules has little to do with this outcome. Suppose that the FCC had, in the Nextel/public safety dispute, imposed regulations that, with full specificity and zero ambiguity, prohibited any harmful interference.¹⁴⁸ Suppose, further, that this rule would determine that Nextel’s mobile phone network and the public safety radio services were

146. Even before accounting for very large in-kind spectrum subsidies, publicly run networks are so expensive to operate that many public safety agencies contract with private vendors (using commercial spectrum) for service. See Jerry Brito, *Sending Out an S.O.S.: Public Safety Communications Interoperability as a Collective Action Problem*, 59 FED. COMM. L.J. 457, 475 (2007).

147. See Hazlett, Muñoz & Avanzini, *supra* note 6, at 105–06 (comparing subjective “beauty contest” awards of spectrum licenses before spectrum auctions; see also Hazlett, *supra* note 34).

148. The FCC explicitly avoided the predictive exercise in its 800 MHz R&O:

We also conclude we should adopt an interference protection standard in the 800 MHz band based on measured, rather than predicted signal strength. While one approach would be to define the coverage area of public safety system by a predicted signal contour, signal level prediction is an inexact science and 800 MHz radio signal propagation can be affected by multiple factors such as buildings and other obstructions, reflection of signals from nearby man-made surfaces, terrain, and foliage. Moreover, system designers frequently predict signal strengths in terms of statistical probability, e.g., the charts and algorithms used for coverage determinations predict the distance from a transmitter at which a given level of signal will be equaled or exceeded at fifty percent of the locations, fifty percent of the time. [See, e.g., 47 C.F.R. § 73.699, Figures 9, 10 and 10b)]. Thus, while signal strength predictions are useful for obtaining an overall picture of system coverage, we believe they are of limited utility in predicting the strength of an 800 MHz public safety signal in a localized and relatively small area, which is exactly the type of area in which interference may be encountered from an ESMR or cellular system. Consequently, we conclude that we need to use a basis other than distance separations or predicted signal contours in establishing the threshold determination of entitlement to interference protection.

800 MHz R&O, *supra* note 128, at 15,025 ¶ 95.

incompatible: both could not continue to operate under the FCC's interleaved channel plan without spewing harmful interference.¹⁴⁹ Suppose, too, that the net social value of Nextel's wireless services equaled \$100 billion, an estimate almost surely on the low side.¹⁵⁰ Suppose, finally, that the net social value of the public safety radio services supplied using the original (1982) FCC band plan was \$10 billion.

Four implications of this hypothetical question are important. First, the dispute would very likely have stretched on for many more years. There is no technical solution to the question of liability, and given the dispositive importance of this determination, the FCC would be choosing between two rival economic outcomes. We surmise that public safety would be the likely winner, given its incumbency and the importance of "headline risk" to rational, utility-maximizing bureaucrats. Regardless, a considerable swath of SMR spectrum would be rendered unusable under a zero-tolerance harmful interference standard and the interleaved channel allocation.¹⁵¹

Second, under the FCC's initial band allocation, at least \$10 billion, the lesser of the social contributions made by the rival SMR band services, would be lost. The simultaneous accommodation of both Nextel and public safety users was achieved not by better SMR borders but by reorganizing claims in frequency space. Imposing improved border definitions, in this situation, is

149. 800 MHz R&O, *supra* note 125, at 15,001 ¶ 101 ("In this connection, we note that almost all participants in this proceeding agree that the *status quo*—addressing interference to public safety systems on an *ad hoc* basis and reactive fashion—is no longer workable in the 800 MHz band.") (emphasis in original).

150. The 2005 price paid for Nextel was \$35 billion, yielding an estimate of the present value of producers' surplus. The ratio of consumers' to producers' surplus is commonly set at ten or above. *See, e.g.*, Hazlett, Muñoz & Avanzini, *supra* note 6, at 119–20; *Welfare Analysis*, *supra* note 68, at 425 (citing J.A. Hausman, *Valuing the Effect of Regulation on New Services in Telecommunications*, Brookings Papers on Econ. Activity, Microeconomics, Vol. 1997 (1997) (for annual consumer surplus in cellular telephone licenses exceeding annual producer surplus by ten times); Rosston, *supra* note 52 (estimating "an order of magnitude more weight to consumer surplus than to the private license values").

151. The FCC avoided proposed "zero tolerance" policies as well:

Given this fact, we believe that it would be inappropriate, as a matter of responsible spectrum management, to afford public safety systems the noise-limited coverage that some proponents have recommended. For example, were we to do so for a given public safety system in the 800 MHz band, it would not only restrict the availability of public safety spectrum in adjoining areas but also would make it virtually impossible for CMRS systems to use channels that contributed the slightest amount of noise to a public safety receiver in the far fringes of its noise-limited coverage area. Such an outcome would result in inefficient utilization of CMRS spectrum.

800 MHz R&O, *supra* note 125, at 15,024–25 ¶ 94 (citations omitted).

all about excluding the party determined to be at fault for spillovers. The clarity of rights, offered in the standard property rights context as beneficial in yielding efficiencies by internalizing effects, does not here bring with it improved modes of organization. That is because the ownership rights in public safety radio remain mired in a regulatory commons.¹⁵² Transactions to simply exchange channels, turning an interleaved band into contiguous blocks of spectrum, are blocked on three levels: determinations of the “public interest,” incentive misalignment, and the mismatch between public and private gains.

Formally, license transfers must be approved by regulators as being in the “public interest.” In fact, FCC regulators generally approve license trades that have been negotiated between the parties. But such negotiations in public safety allocations are stymied by (a) the nonprofit status of the organizations, making it difficult to reward decision-makers for discovering and executing efficient deals; and (b) the extreme fragmentation of SMR licensees, numbering over 10,000 nationwide.¹⁵³ The transaction costs thus thwart the aggregation of public safety radio licenses into efficient blocks.

Third, the regulator does not internalize the social gains from choosing the correct—i.e., low cost—party to find at fault. The difference between a \$100 billion and a \$10 billion loss is large, but FCC officials do not pay it. They are more affected by the responses from interests that, charged with protecting life and limb in partnership with local governments, may become severely unpleasant when their slices of spectrum are threatened. An FCC decision maker is aware that \$100 billion of other peoples’ money may be used in two ways: not only to purchase interference protection for certain worthy radio users, but also to ensure that said official will never have to explain to a Congressional committee why America’s police, sheriff, and fire departments are more important than the text-messaging addictions of middle school students.

Fourth, given the biases of regulators—who asymmetrically fear Type I errors (visible interference) far more than Type II errors (invisible interference: the services silenced to protect others)¹⁵⁴—the likelihood that

152. The idea is presaged in many treatments, including the externality problem endemic in government regulation. See WOLF, *supra* note 78. The most direct statement may be found in HELLER, *supra* note 16, at 26: “If the regulatory drama involves too many uncoordinated actors . . . the sheer multiplicity of players may block use of the underlying resource.”

153. See generally Brito, *supra* note 146.

154. This owes to the incentives of regulators, as briefly described in the paragraph above. The paradigm is widespread in regulatory decision-making. See, e.g., Daniel Carpenter, *The Political Economy of FDA Drug Review: Processing, Politics, and Lessons for Policy*, HEALTH

these regulators can foresee which is the \$100 billion opportunity and which is the \$10 billion service is very low. The purpose of granting spectrum use rights to firms or individuals is to enlist the energy, dynamism, and information of competitive market forces. Because entrepreneurs are rewarded for correctly valuing assets (which is to say, buying under-priced rights and selling over-priced rights), they reveal values. Government allocations supersede this process and preempt data only the market can uncover. The administrative assignment of use rights thus embeds basic inefficiencies due to its reliance on non-market estimates of value,¹⁵⁵ the essence of the critique made by Ronald Coase in 1959.¹⁵⁶ More recent research has confirmed this perspective, mocking traditional FCC spectrum allocation as “command and control”¹⁵⁷ or “Gosplan.”¹⁵⁸

Because many of the spectrum rights at issue were not held by profit-seeking enterprises able to contract in the marketplace, the establishment of clearer borderlines between Nextel and public safety radio users was predictably an inefficient solution. It would have simply frozen usage by one of the two conflicting sets of wireless services, and outlawed the other, given

AFFAIRS, Jan.–Feb. 2004, at 52, 55–57, <http://content.healthaffairs.org/cgi/reprint/23/1/52>.

155. Should someone seriously believe that FCC economists can be called on to deliver timely estimates of alternative spectrum values—on a par with the values revealed by actual market transactions—the best that can be said is that they do not understand the enormity of this task. For simplification, we here consider two specific systems in conflict and give an assumed value to each. In reality, there are limitless possibilities; Nextel and the public safety radio users could combine, disaggregate, or cooperate in myriad ways, some with a little more space for Nextel, some with a little less. Mixing and matching other spectrum resources, while trading interleaved channels (the “spectrum swap” forwarded by Nextel and eventually executed by the FCC) is one example of how this could be done. That the FCC was led (here and in innumerable other instances) to act on the solution brought to it by private parties is one further bit of evidence that the valuation trade-offs are best left to parties that internalize impacts. See generally *Feds OK Nextel Spectrum Swap*, WIRED (July 8, 2004), <http://www.wired.com/techbiz/media/news/2004/07/64142>. For more on the “informational responsibilit[ies]” of market participants and central planners, consider Smith, *On the Economy of Concepts in Property*, *supra* note 60, at 2114.

156. Coase’s analysis leaned heavily on Adam Smith’s view of the importance of decentralized decision making. It also incorporated Friedrich A. Hayek’s view that such decentralization was efficient when decision makers were incentivized to productively use specific information of “time and place.” Asset owners pursuing value maximization are subject to such incentives. See R. H. Coase, *The Federal Communications Commission*, 2 J.L. & ECON. 1, 18 (1959); see generally Friedrich A. Hayek, *The Use of Knowledge in Society*, 35 AM. ECON. REV. 519 (1945).

157. NBP 2010, *supra* note 3, at 79.

158. Gerald R. Faulhaber & David J. Farber, *Spectrum Management: Property Rights, Markets and the Commons* 6 (AEI-Brookings Joint Center for Regulatory Studies, Working Paper 02-12, Dec. 2002), <http://www.ictregulationtoolkit.org/en/Document.3629.pdf>.

the impossibility (prohibitive transaction costs) of recontracting. This is what results when the goal of the regulatory task is limited to finding fault—“Gosplan.”¹⁵⁹

A superior result was, in fact, obtained. The FCC executed a spectrum swap that left technical rules in place but rearranged the economic structure of spectrum rights.¹⁶⁰ The implication for public policy is that such transactions—routine in the market when exclusive, liberal rights control the relevant spectrum—demonstrably improve upon the performance of administrative allocation. To focus on more sophisticated rights definitions, as suggested by Matheson & Morris (2011) or De Vries & Sieh (2011), ignores the available lessons.¹⁶¹ Efforts to fine tune border definitions were, at best, costly diversions.¹⁶² The solution came in undoing the problems of interleaving, fragmentation, non-profit rights ownership, and spectrum integration by merger.¹⁶³ That FCC fiat was required to impose this mutually

159. *Faulhaber*, *supra* note 158.

160. The FCC rejected over specification and relied on economic reconfiguration:

Thus, although we have discussed herein the technical means disclosed in the record to avoid unacceptable interference—especially those that come within the definition of Enhanced Best Practices—we reject as unnecessary, the recommendations of some parties for mandatory restrictions on all ESMR and cellular systems with respect to such parameters as maximum cell ERP, combiner technology, and specific antenna pattern characteristics.

800 MHz R&O, *supra* note 125, at 15,028 ¶ 103 (footnotes omitted).

161. *See infra* note 163.

162. The FCC declined receiver standards, and interference temperatures, proposed to technically solve border interference:

The Consensus Parties proposed that full interference protection would be provided only for systems using receivers that satisfy TIA Class A specifications. Receivers not conforming to these specifications would be protected only to some higher desired signal threshold power level. Several parties supported the Consensus Parties in this regard; while others disagreed, pointing out that some of the TIA standard parameters, for example, operating temperature range of the radio are irrelevant to 800 MHz interference and therefore that the Commission should not require compliance with the entire standard but, instead, should simply adopt minimum intermodulation rejection ratios for receivers.

800 MHz R&O, *supra* note 125, at 15,026–27 ¶ 99 (footnotes omitted).

163. In fairness, MATHESON & MORRIS, *supra* note 56, and DE VRIES & SIEH, *supra* note 57, at 66, both mention the importance of avoiding fragmentation in rights issuance. Goodman also touches on the complex burden on interference rules when spectrum sharing is authorized not by liberal licenses but in unlicensed allocations:

Things take longer when no one can be held accountable for interference. One of the complications of unlicensed use, however desirable it may be, is that it's hard to assign responsibility for interference. This difficulty

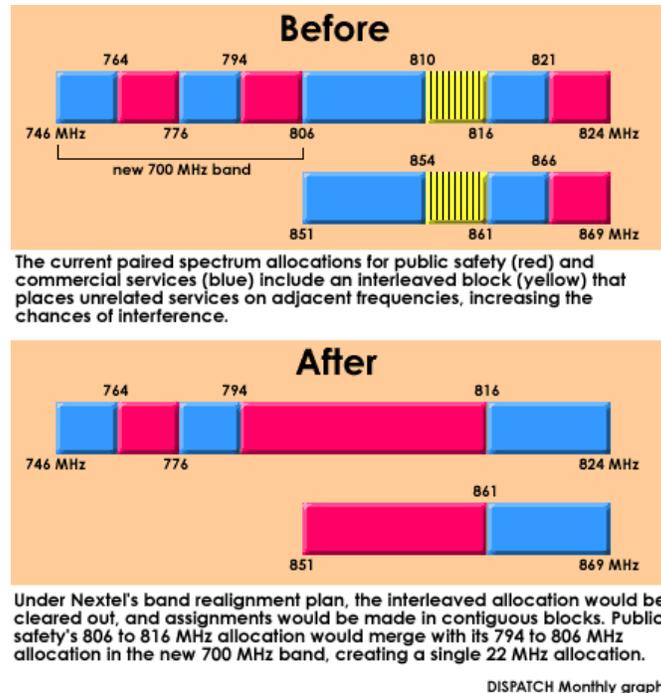
beneficial solution on nonprofit licensees is an important lesson first explained by Leo Herzel in 1951.¹⁶⁴ And, as Demsetz has more recently explained, the transactionally efficient way to solve most externality problems is to allow the market trades to eliminate borders altogether, a proxy for what finally remedied the conflicts in the SMR band (as seen in Figure 3).¹⁶⁵

buttresses the already existing tendency towards conservative allocations and is one of the reasons the White Spaces decision took so long.

Goodman, *supra* note 54, at 507. These are important inclusions, but they are discussed in each of the analyses as stand alone considerations. They are not. The reliance on well-crafted, economically efficient spectrum rights bundles is key to welfare maximization. The expenditure of time and scarce regulatory resources on technical specificity directly undermines this path. “Such [cellular/PCS] systems can substantially increase the ability of base stations to re-use frequencies without increasing interference, and illustrate the importance of including direction of propagation as a dimension of spectrum rights.” MATHESON & MORRIS, *supra* note 56, at 14. “The direction of propagation is likely to be an increasingly important dimension across which to partition rights to access spectrum.” *Id.* at 15. MIMO, dynamic sensing and spectrum reuse should allow regulators *less* involvement in detailing borders, but path-dependency of the specification exercise remains. *See* DE VRIES & SIEH, *supra* note 57, at 63–64 (“Rights not assigned are deemed to be in the public domain and can be appropriated by any operator—but only until the next license renewal point, at which time the regulator may add rights to the licenses specifying these parameters”); *see also* MATHESON & MORRIS, *supra* note 56, at 37 (“[D]ifferent technologies may be better suited to different sets of unlicensed rules, and there is no way for the FCC to ensure that its portfolio of different rules in different unlicensed bands is efficient.”).

164. Herzel, *supra* note 145.

165. *See* Demsetz, *supra* note 47.

Figure 3: Nextel/Public Safety Conflict Resolution¹⁶⁶

F. INTERFERENCE TEMPERATURE: A FAILED ATTEMPT AT GREATER CLARITY

In the spectrum policy literature, an idea that is prevalent (and correct) is that there is a great deal of unused or under-utilized spectrum: “[T]here is no real debate that our current system of spectrum rights keeps some lower valued uses on wireless frequencies at the expense of higher-valued uses.”¹⁶⁷

166. 800 MHz Interference Issue Rebanding, DISPATCH MAGAZINE, http://www.911dispatch.com/info/800_transition/index.html (last visited Mar. 1, 2013).

167. Stuart Minor Benjamin, *Roasting the Pig to Burn Down the House: A Modest Proposal*, 7 J. TELECOMM. & HIGH-TECH L. 95, 95–96 (2009); see also NBP 2010, *supra* note 3, at 78–79:

In the case of commercial spectrum, the failure to revisit historical allocations can leave spectrum handcuffed to particular use cases and outmoded services, and less valuable and less transferable to innovators who seek to use it for new services. The market for commercial, licensed spectrum does not always behave like a typical commodities market. Commercially licensed spectrum does not always move efficiently to the use valued most highly by markets and consumers. For example, a megahertz-pop may be worth a penny in one industry context and a dollar in another. Legacy “command and control” rules, high transaction costs and highly fragmented license regimes sometimes preserve outmoded band plans and prevent the aggregation (or disaggregation) of spectrum into more valuable license configurations.

But there is a distinct split as to how that process should be achieved. Economists tend to see the underutilization, as did Coase, as caused by a lack of ownership rights.¹⁶⁸ Without responsible economic agents to conserve or extend resource value, productive opportunities are squandered. The sharp contrast between spectrum allocated to liberal licenses—generating heavy investment in complementary network infrastructure, so it can yield value by being intensely shared—and restrictive licenses, such as those governing over-the-air television broadcasting, is emblematic of the efficiencies that might be realized if more licenses were liberalized.¹⁶⁹ As is, traditional licenses allocating prime spectrum typically return far less than their social opportunity cost in social welfare gains.¹⁷⁰

Lawyers and engineers, on the other hand, often see the underutilization as endemic in all licensed spectrum. Even in liberally licensed bands, full utilization is rare: almost all bands could host additional wireless services during certain parts of the average day.¹⁷¹ Given sophisticated radio technologies that identify vacant channels and avoid conflicts with competing radio emissions, society could benefit; however, the government would need to craft rules that allowed ad hoc access to almost any unused spectrum resources. This would, arguably, make much more bandwidth

Id.

168. Not only economists take this position, of course. *See* Benjamin, *supra* note 167.

169. Underuse and overuse, compared to ordinary use and optimal use, are concepts that “reorient[] policymaking from relatively simple either-or choices to the more contentious trade-offs that make up modern regulation of risk.” HELLER, *supra* note 16, at 35–37. “Today, for many observers, the property trilogy can be reduced to an opposition of private and commons property, what one scholar calls simply, ‘all or none.’ I believe a substantial cause of our cultural blindness to gridlock arises from this too simple image of property.” *Id.* at 34 (citing YORAM BARZEL, *ECONOMIC ANALYSIS OF PROPERTY RIGHTS* 71 (Cambridge University Press 1989)).

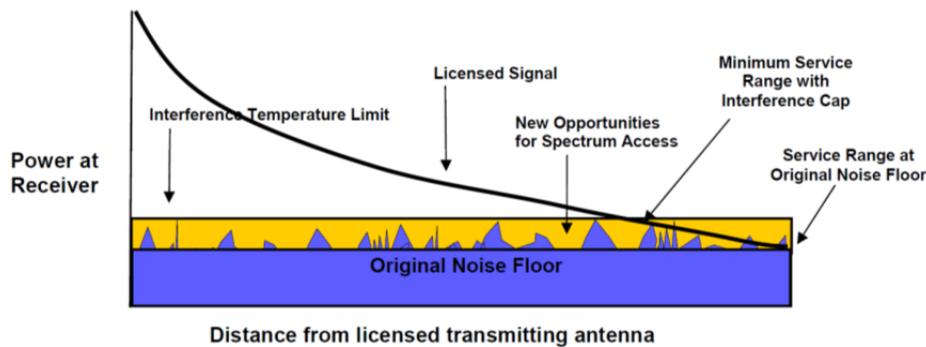
170. This is not to say that broadcast TV content is lacking in value, but that the broadcast TV distribution platform—transmissions to roof-top antennas—is obsolete. As over 90% of households receive their video content (including broadcast TV signals) via cable or satellite TV connections, and given that boosting that rate to 100% would require another \$3 billion to “complete” the grid, the forty-nine TV channels (6 MHz each, or 294 MHz all told) allocated for broadcasting have extremely modest social value. *See* Hazlett, *A Proposal for an Overlay Auction*, *supra* note 50. In comparison, the consumer surplus that would be generated by allowing all TV band frequencies to be used for mobile communications services easily exceeds \$1 trillion. *Id.*

171. Capacity utilization, whether considered optimal, average, stochastic, or probabilistic, is rarely at full capacity in any resource channel; consider fluctuations of electricity usage, vehicular traffic, seasonal factory orders, auto-scaling Amazon Web Services, emergency safety services, and military services. Spectrum users face the same economic capacity utilization choices. *See* SPTFR, *supra* note 1, at 21.

available for both existing and new, innovative wireless services, fueling economic growth.

The latter view sees the problem of white spaces as a technically defined problem that can be fixed by engineering rules that provide for optimal spectrum sharing. As seen in Figure 4 below, wireless systems using licensed spectrum (liberal or traditional regimes) tend to leave some frequency space unoccupied at least some of the time. By simply drawing rules about where new ad hoc radio transmissions can fit (in an additional noise floor, denoted in yellow in Figure 4), the theory is that extra communications capacity can be forged.

Figure 4: FCC Proposal for an “Interference Temperature”¹⁷²



This view is deeply flawed. Were such spaces being wasted, liberal licensees would have every incentive not only to know that fact, but to invest in technologies, radios, or new network architectures to exploit the waste.¹⁷³ Of course, the gains from using the vacant spaces must profitably exceed the costs of packing more traffic into the frequency space. Because all such traffic is stochastic and powerful, exhibiting both cyclical and random

172. In the Matter of Establishment of an Interference Temperature Metric to Quantify and Manage Interference and to Expand Available Unlicensed Operation in Certain Fixed, Mobile and Satellite Frequency Bands, Notice of Inquiry and Notice of Proposed Rulemaking, 18 FCC Rcd. 25,309, 25,315 (2003) [hereinafter 2003 Interference Temperature Metric NOI & NPRM], available at http://hraunfoss.fcc.gov/edocs_public/attachmatch/FCC-03-289A1.pdf.

173. That new technologies increase spectrum productivity from spectrum inputs only emphasizes the importance of liberalizing the spectrum to absorb technical specifications into larger bundles of complementary use rights. Consider the opposing forces between technical innovation and regulatory liberalization. See DE VRIES & SIEH, *supra* note 57, at 67 (“DSA aims for higher density and more transience of use than conventional static allocations . . .”). But see *id.* (“However, there are many flavors of DSA that can be classified according to distinctions, such as: coordinated vs. uncoordinated; cooperative vs. non-cooperative; homogeneous vs. heterogeneous uses; and co-primary, primary-secondary or co-secondary legal status.”).

fluctuations in demand, along with similar perturbations in propagation characteristics, networks are typically designed to feature some optimal level of slack. Indeed, quality of service (“QoS”) is highly correlated with the consistency of network performance.¹⁷⁴ However, the probabilistic aspects of radio technology that are raised as impediments to drawing private property rights in spectrum actually cut just the opposite way.

First, lines that are hard to draw for private owners are equally hard to draw, in technical terms, for limited use rights narrowly defined and then assigned by regulators.¹⁷⁵ This is the basic point Coase made in his “Social Cost” paper.¹⁷⁶ Second, the essential trade-offs of harmful interference are both complex and economic in nature. By both logic and historical observation, spectrum regulators have no comparative advantage in determining what these trade-offs are or how best to resolve them.

This was demonstrated rather powerfully in the proceeding opened by the FCC in 2003 to establish an “interference temperature.”¹⁷⁷ Driven by the technical view of radio interference,¹⁷⁸ the Commission sought to allow ad hoc access, by using unlicensed radio devices, in the low-power spaces found in the noise floor of frequency bands allocated to both liberal and traditional licenses.¹⁷⁹ In the bands allocated to traditional licenses, vast white spaces

174. See Gregory Sidak & David Teece, *Innovation Spillovers and the “Dirt Road” Fallacy: The Intellectual Bankruptcy of Banning Optional Transactions for Enhanced Delivery Over the Internet*, 6 J. COMPET. L. & ECON. 521 (2010) (discussing the economics of innovation in quality of service (“QoS”).

175. Consider the role of the regulator in renewal of limited rights:

Rights not assigned are deemed to be in the public domain and can be appropriated by any operator—but only until the next license renewal point, at which time the regulator may add rights to the licenses specifying these parameters. For example, a transmission permission may limit resulting energy at or below an altitude of 1.5 meters, but say nothing about operation at 1,000 meters; thus a surveillance drone operator a [sic] would be able to exceed the specified 1.5 meter transmit power at 1,000 meters.

DE VRIES & SIEH, *supra* note 57, at 8–9.

176. Coase, *supra* note 44. For a detailed explanation of this point, often lost in confusion over the “Coase Theorem,” see Thomas W. Hazlett, *Ronald H. Coase*, in *PIONEERS IN LAW AND ECONOMICS*, CHAPTER 1 (Lloyd R. Cohen & Joshua D. Wright, eds., Edward Elgar 2009). See generally Interview by Russ Roberts with Ronald Coase, *Coase on Externalities, the Firm, and the State of Economics*, Podcast on *EconTalk*, LIBRARY OF ECONOMICS AND LIBERTY (May 21, 2012), available at http://www.econtalk.org/archives/2012/05/coase_on_extern.html (explaining misunderstandings of the “Coase Theorem”).

177. 2003 Interference Temperature Metric NOI & NPRM, *supra* note 172, at 25,312 ¶ 8.

178. *Id.* at 25,317–18 ¶ 21.

179. *Id.* at 25,315 ¶ 16.

often exist, making some new unlicensed uses relatively uncontroversial. But the spectrum licensed to liberal licenses is not only intensely used, but mobile carriers using such frequencies are also continually engaged in the process of making the noise floor quieter, so as to allow greater traffic to generate higher revenues.¹⁸⁰ Moreover, such networks are continually cutting deals with third parties, such as wholesale mobile operator agreements, vendor equipment (handsets, tablets, modems, etc.) joint marketing contracts, and application platforms, to use spectrum allocated liberal licenses more fully and profitably.¹⁸¹

A broad misunderstanding concerning the nature of exclusive rights merits attention here. It is commonly asserted that unlicensed spectrum access categorically entails lower (or no) licensing expense. Transactions costs in using exclusive spectrum rights are incurred due to the “hassle” inherent in the “need for a license to operate the device”¹⁸² These costs, it is asserted, are avoided in unlicensed bands, saving resources and speeding innovative wireless services to market.¹⁸³ Charles L. Jackson provides a numerical example, estimating that “the spectrum occupied” by one remote wireless car key in one day “is worth one ten-thousandth of a cent” in terms of the bandwidth consumed.¹⁸⁴ The efficiency of unlicensed access is found in Jackson’s further estimate that “a consumer would need spend only about one-thousandth of a second contemplating an FCC license form before the transactions costs exceeded ten times the value of the spectrum used by the device.”¹⁸⁵

The value estimates are plausible, but the conclusion is flawed. When exclusive spectrum rights are issued to profit-maximizing enterprises, licensees then seek to avoid needless transaction costs. Hence, when cellular carriers resell spectrum rights to mobile subscribers, subscribers need not

180. See Rath, *supra* note 28, at 529 (“These are not massive, one-time negotiations between companies, but involve hundreds of individual negotiations between companies’ engineers who are tasked with the day-to-day operations of the network.”).

181. Fifteenth Annual Competition Report, *supra* note 20, at 9698 ¶ 32, 9751–52 ¶ 138; 9757–58 ¶ 154 (2011).

182. Kenneth R. Carter, Ahmed Lahjouji & Neal McNeil, *A Joint OSP-OET White Paper on Unlicensed Devices and Their Regulatory Issues* 5 (FCC, OSP Working Paper No. 39, May, 2003), available at http://hraunfoss.fcc.gov/edocs_public/attachmatch/DOC-234741A1.pdf.

183. *Id.* (“Because they are free from the delays inherent in the licensing process, unlicensed devices can frequently be designed to fill a unique need and be introduced into the marketplace rather quickly. The availability of spectrum for use by unlicensed devices has spawned a variety of new applications.”).

184. Charles L. Jackson, *The Genesis of Unlicensed Wireless Policy*, 11 INFO 2 n.4 (Aug. 2009).

185. *Id.*

obtain licenses to use that spectrum—not once, let alone every time they dial a call, send a text, or check email. Contracts are arranged by monthly billings or prepaid cards with set prices. Indeed, marketing innovations like “digital one rate,” a fixed price for a monthly “bucket of minutes,” or unlimited off-peak, on-net, or texting usage remove even more of the pricing overhead.¹⁸⁶ The cost of billing is not zero, of course, but the opportunity to bill is a feature, not a bug. Payments from spectrum users reveal the value of service and support the creation of complementary infrastructure, including wide area wireless networks that dominate value creation (including those found in TV and unlicensed bands).¹⁸⁷

All costs, including the direct costs of customer billing for use, shrink when economic incentives are properly aligned. Exclusive ownership rights tend to do that. Whatever approaches the FCC may employ to reduce licensing costs, after licensing exclusive spectrum rights with liberal use rules, can be employed by private parties. The unlicensed approach, allowing customers to buy approved FCC devices and then “plug-and-play” without further transactions, is commonly adopted in the use of licensed spectrum.¹⁸⁸ Phones purchased at 7-Eleven or Walmart sold by mobile virtual network operators (“MVNO”) such as TracFone, need no carrier contract.¹⁸⁹ Nor does Amazon Kindle, where book or movie downloads are delivered seamlessly over a carrier’s mobile network, unseen by the user, who bargains with Amazon.¹⁹⁰ It is undeniably true that the initial licensing round is costly, but the regulatory process largely imposes that cost. Moreover, parallel processes are implicated in the initial allocation of spectrum for unlicensed use. The FCC proceeding to permit unlicensed access to TV Band white

186. Fifteenth Annual Competition Report, *supra* note 20, at 9724 ¶ 81 (“bucket” of minutes); In the Matter of Implementation of Section 6002(b) of the Omnibus Budget Reconciliation Act of 1993 Annual Report & Analysis of Competitive Mkt. Conditions with Respect to Mobile Wireless, Including Commercial Mobile Servs., Eleventh Report, 21 FCC Rcd. 10,947, 10,983–94 ¶ 90 (2006) [hereinafter Eleventh Annual Competition Report], available at http://hraunfoss.fcc.gov/edocs_public/attachmatch/FCC-06-142A1.pdf (AT&T Wireless’s introduction of the “digital one rate” in 1998).

187. Hazlett & Leo, *supra* note 97.

188. Eleventh Annual Competition Report, *supra* note 186, at 10,961 ¶ 30.

189. Fifteenth Annual Competition Report, *supra* note 20, at 9731 ¶ 98; Search Result for “TracFone,” WALMART, http://www.walmart.com/search/search-ng.do?search_query=tracfone (last visited Nov. 26, 2012) (offering prepaid TracFone phones); No-Contract Phones, 7-ELEVEN, <http://www.7-eleven.com/Play/No-Contract-Phones/Default.aspx> (last visited Nov. 26, 2012) (offering a TracFone Samsung T245G).

190. Press Release, Amazon.com, Inc., Amazon Takes on the High-End—Introducing the New Kindle Fire HD Family (Sept. 6, 2012), <http://phx.corporate-ir.net/phoenix.zhtml?c=176060&p=irol-newsArticle&ID=1732546> (stating that Amazon named AT&T as a contracted provider of LTE data service).

spaces has dragged on for ten years, yet there are still no commercial deployments.¹⁹¹ Difficult choices are made about rules to control harmful interference, choices made by regulators having no stake in the matter and subject to no cost in delays.¹⁹² In point of fact, licensing is not avoided in unlicensed allocations—it is the device that is licensed, not the operator. These approvals, as seen, may take many years.

It is even more costly when a spectrum is frozen in sub-optimal use because the unlicensed adopted (or contemplated) use does not allow efficient spectrum reallocation to take place in the marketplace. This is precisely the situation in the TV Band. Thanks to both the traditional TV licensing scheme, which reserves all but a very few narrow, well-specified, seemed-like-a-good-idea-in-1952-broadcasting spectrum rights in the hands of regulators, and the decade-long, policy focus on sprinkling unlicensed access rights in the voluminous white spaces (a product of the rigid and obsolete government spectrum use rules), the U.S. economy loses perhaps \$100 billion or more in annual social welfare.¹⁹³

191. In the Matter of Additional Spectrum for Unlicensed Devices Below 900 MHz and in the 3 GHz Band, Notice of Inquiry, 17 FCC Rcd. 25,632 (2002); In the Matter of Unlicensed Operation in the TV Broadcast Bands & Additional Spectrum for Unlicensed Devices Below 900 MHz and in the 3 GHz Band, Notice of Proposed Rulemaking, 19 FCC Rcd. 10,018 (2004); In the Matter of Unlicensed Operation in the TV Broadcast Bands & Additional Spectrum for Unlicensed Devices Below 900 MHz and in the 3 GHz Band, Second Memorandum Opinion and Order, 25 FCC Rcd. 18,661 (2010) [hereinafter 2010 White Spaces Second Memorandum R&O]. “[A]ll TV bands devices to be certified by the FCC Laboratory. The Laboratory will request samples of the devices for testing to ensure that they meet all the applicable requirements.” 2010 White Spaces Second Memorandum R&O, *supra*, at 18,667 ¶ 11. Since the 2008 *SR&O*, the latest progress includes orders to certify database administrators, and a report of basic protocol specifications. Two white space devices have been tested by the FCC Laboratory as of 2011, after IEEE released 802.22 in July 2011. *See* In the Matter of Unlicensed Operation in the TV Broad. Bands & Additional Spectrum for Unlicensed Devices Below 900 Mhz & in the 3 Ghz Band, Order, 26 FCC Rcd. 554 (2011); In the Matter of Unlicensed Operation in the TV Broad. Bands & Additional Spectrum for Unlicensed Devices Below 900 Mhz & in the 3 Ghz Band, Order, 26 FCC Rcd. 10,599 (2011); Report of the White Space Database Administrator Group, DA 11-131 (Sept. 12, 2011); *Database-to-Database Synchronization Interoperability Specification*, Version 1.01; IEEE-Standards Association, IEEE 802.22™-2011 *Standard for Wireless Regional Area Networks in TV Whitespaces Completed*.

192. Goodman, *supra* note 54, at 506 (noting the added difficulty inherent in the unlicensed rules, specifically tying this to the TV band “white spaces” regulatory delay); *see also* DE VRIES & SIEH, *supra* note 57, at 56 (“The operating parameters for unlicensed devices would be revisited at regular intervals, say ten years, just as licensed parameter sets are reviewed at license renewal.”). Historically, the time of review through the FCC administrative process has taken six to thirteen years. *NBP 2010*, *supra* note 3, at 79.

193. Hazlett, *A Proposal for an Overlay Auction*, *supra* note 50, at 3 n.2. This is a back of the envelope calculation, leveraged on the proposition that if just 200 MHz of actively used

Indeed, the real transaction cost issue is the reverse of what Jackson has offered. While private rights to control spectrum can and should produce resource owners highly motivated to reduce transaction costs, a regulatory system that ambitiously attempts to preempt the market's approach must largely engage in the "command and control" system of spectrum allocation that yields such poor results today. That is, when regulators are to decide where unlicensed access is more efficient, they must then evaluate all band rules individually. Unlicensed allocations, further, are not unregulated. They require policing, as Charles Jackson, Raymond Pickholz and Dale Hatfield have explained.¹⁹⁴ Such bands, to avoid the tragedy of the commons, come equipped with coordinating rules: power limits, technology restrictions, and (sometimes) business model regulations. These rules are imposed by regulators, and can only be changed by regulators. The bottom line is that liberal licensed and unlicensed allocations are two rival means of dealing with harmful interference. Neither path is free. It is appropriate to compare the total costs and benefits of either approach in selecting a framework for choosing among them.

The FCC's attempt to establish an interference temperature underscores how elusive and expensive the unlicensed approach could be. The Commission argued that its effort to establish clear metrics would give "greater certainty" through "specifying a potentially more accurate measure of interference" to licensees regarding interference their services might suffer from the noise floor.¹⁹⁵ But this path to great clarity was worth much less than it cost—or so liberal licensees believed, as witnessed by their strong opposition to the proceeding. They already exercised de facto control of the spectrum resources in question. Borders were not explicit, but were good enough for large networks to be created and operated. As explained in detail elsewhere,¹⁹⁶ the government did not solve any market failure with its purported imposition of a technically defined border for the noise floor, but threatened to impose a tragedy of the commons on highly productive resource owners by introducing fragmented spectrum use rights that could not be reconfigured due to open access. Despite the FCC's aims to pave the way for dynamic access to licensed spectrum with rules protecting owners

CMRS spectrum in 2008 generated at least \$200 billion in voice service consumer surplus, another 294 MHz of (even better VHF-UHF) spectrum might be worth, at the margin, at least half again as much. The guesstimate here is to establish order of magnitude.

194. See Charles Jackson, Raymond Pickholtz & Dale Hatfield, *Spread Spectrum Is Good—But it Does Not Obsolete NBC. v. U.S.*, 58 FED. COMM. L.J. 245 (2006).

195. 2003 *Interference Temperature Metric NOI/NPRM*, *supra* note 172, at 25,309 ¶ 1.

196. See, e.g., Thomas W. Hazlett & Matthew L. Spitzer, *Advanced Wireless Technologies and Public Policy*, 79 S. CAL. L. REV. 595 (2006) [hereinafter *Advanced Wireless Technologies*].

against harmful interference, as measured by spectral analyzers evaluating radio emissions, the Commission conceded failure. In 2007 it abandoned the effort, stating:

Commenting parties generally argued that the interference temperature approach is not a workable concept and would result in increased interference in the frequency bands where it would be used. While there was some support in the record for adopting an interference temperature approach, no parties provided information on specific technical rules that we could adopt to implement it.¹⁹⁷

It should not go unnoticed that the enthusiasm for FCC-driven solutions to spectrum sharing is so great that commentators today lament this outcome, arguing that the Interference Temperature proceeding and Receiver Standards proceeding should be brought back to life, despite the fact that the regulators needed to define such metrics have closed the matter.¹⁹⁸

IV. BEYOND EXACTITUDE: ENABLING OPTIMAL COMBINATIONS

A. BASIC STRATEGY

There are various combinations of resources—transmission power, antenna height and directivity, frequency of transmission, method of propagation, etc.—that can be utilized to achieve a given level of (received) power at a point distant from the point of transmission. The *range* of alternative combinations is determined by technology—the state of the arts—and is an engineering problem. The “proper” combination actually to use to achieve a given goal is, however, an *economic* problem and is not (properly)] soluble solely in terms of engineering data.¹⁹⁹

197. In the Matter of Establishment of an Interference Temperature Metric to Quantify & Manage Interference & to Expand Available Unlicensed Operation in Certain Fixed, Mobile & Satellite Frequency Bands, Order, 22 FCC Rcd. 8938 (2007) (internal citations omitted).

198. DE VRIES & SIEH, *supra* note 57, at 62 (“However, it would likely not take such a step since limiting the number of receivers . . . is politically unrealistic, and attempting to specify receiver standards . . . has proven difficult in the US as a general matter”) (citing In the Matter of Interference Immunity Performance Specifications for Radio Receivers, Notice of Inquiry, 18 FCC Rcd. 6039 (2003)).

199. COASE, MECKLING & MINASIAN, *supra* note 37, at 23. For an explanation of why Coase’s RAND paper was suppressed for over three decades, see Ronald H. Coase, *Comment on Thomas W. Hazlett, Assigning Property Rights to Radio Spectrum Users: Why Did FCC License Auctions Take 67 Years?*, 41 J.L. & ECON. 577, 579 (1998) (responding to Hazlett, *Why Did FCC License Auctions Take 67 Years?*, *supra* note 34).

This profound summary of the coordination problem in spectrum allocation was offered by three economists a half-century ago.²⁰⁰ Their analysis was rendered a priori, based almost wholly on economic theory and before they could directly observe the empirical result of their proffered system for allocating radio spectrum rights via market mechanisms. They had observed the rigidities and presumed inefficiencies of the administration allocation model, however, where airwave transmissions were narrowly constrained by law. The performance of this system led them to conclude that an alternative regime vesting decentralized licensees with the flexibility to change spectrum uses according to profit-and-loss criteria would improve social coordination, producing substantially greater economic output.

This approach has been strongly supported by the data generated over the intervening decades when, in fact, liberal spectrum rights were defined and distributed by regulators in the United States and elsewhere. In mobile markets licensees have been generally afforded wide latitude to control pricing, services, business models, the deployment of wireless technologies, the location of base stations, and the mobile devices used by subscribers. Private market transactions determine, for example, how a mobile phone communicates with the base station, what power it uses, what technical format is deployed (GSM, CDMA, LTE), and how the customer pays for her subscription. These are all dimensions that were (and in many cases still are) determined by the FCC in traditional wireless licenses.²⁰¹ The outcome of this liberalization in mobile services has been an eruption in productive activity, with a complex ecosystem emerging in which technology suppliers, app developers, equipment manufacturers, investors, wireless carriers, and consumers coordinate the use of myriad inputs, including radio spectrum, in a competitive quest to achieve the “‘proper’ combination.”²⁰²

How is it known that the market outcome in mobile is superior per the liberal spectrum rights issued to enable it? It is not a trivial question. Perhaps administrative allocation would do as well to accommodate the emerging marketplace. The evidence against this view is persuasive, however. First, the regulators themselves have come to this conclusion, which might be characterized as an admission against interest. In the United States, the FCC

200. See COASE, MECKLING & MINASIAN, *supra* note 37.

201. For an overview of commercial broadcast licenses, see In the Matter of Annual Assessment of the Status of Competition in the Market for the Delivery of Video Programming, Fourteenth Report, 27 FCC Rcd. 8610, 8646 ¶¶ 176–177 (2012) [hereinafter Fourteenth Video Competition Report], available at http://hraunfoss.fcc.gov/edocs_public/attachmatch/FCC-12-81A1.pdf (describing regulatory conditions on the entry and exit of broadcast television stations, license transfers, and media ownership limits).

202. COASE, MECKLING & MINASIAN, *supra* note 37, at 23.

has argued that the mobile market has developed with a great number of efficiencies due to the nature of the spectrum rights issued.²⁰³ Second, virtually all governments have moved decisively away from the traditional licensing model in mobile services. While some countries have retained technology restrictions (e.g., mandating GSM in 2G licenses in the European Union, while prohibiting broadband), the discretion afforded licensees is far beyond what is delegated in the operating permits issued prior to the advent of mobile services.²⁰⁴

Third, for decades the mobile phone market was governed by traditional licenses, and it sputtered. This was because the “land mobile” licenses, narrowly crafted, did not permit cellularization. The 64-kb/s coding rate for twenty-four voice circuits on a standard T-1 wire used in a major city, such as St. Louis, would accommodate just twenty-four phone calls at one time.²⁰⁵ While cellular technology was developed at Bell Labs in 1947, it took until 1989 for U.S. regulators to authorize the technology.²⁰⁶ However one calibrates the delay, it was profound, and forms an oft-used reference point in gauging regulatory lag in spectrum allocation.²⁰⁷ Even during the latter years of this process, when cellular systems were actively considered by policy makers, regulators proceeded cautiously, at least in part, because they believed that the service in question would never be of interest to mass market consumers. They saw it, wrongly, as a niche status symbol, of interest only to luxury car buyers and salesmen.²⁰⁸ Market planning under standard allocation rules, in short, exhibited non-market failure.

Fourth, those countries that have gone most decidedly towards deregulation in mobile spectrum allocations by attempting to issue de facto spectrum licenses have emerged with relatively competitive, well-developed

203. *Wireless Bureau Chief Daniel Phythyon Hails Success of Market-Based Spectrum Policies*, *supra* note 20.

204. Fifteenth Annual Competition Report, *supra* note 20, at 9734 ¶106 n.303 (describing the European mandate on a single harmonized standard in GSM in 2G services in the early 2000s).

205. See GEORGE CALHOUN, DIGITAL CELLULAR RADIO 178, 194 (Artech House 1988). The digital coding rate is near 4.8 kb/s today. *Id.* at 194. Some voice codecs use 5.3 kb/s. See Bur Goode, Voice Over Internet Protocol (VOIP), 90 Proceedings of the IEEE 1495, 1497 (2002), available at <http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=1041060>.

206. For a history of the cellular allocation, see CALHOUN, *supra* note 205. For a discussion of the license assignment process in cellular, see Thomas W. Hazlett & Robert J. Michaels, *The Cost of Rent Seeking: Evidence from the Cellular Telephone License Lotteries*, 59 SO. ECON. J. 425 (1993).

207. NBP 2010, *supra* note 3, at 79, Exhibit 5-C.

208. See MURRAY, *supra* note 124, at 50, 277.

mobile markets.²⁰⁹ This suggests that there are no significant advantages associated with whatever remains of top-down regulatory rules.

Finally, both the aggregate levels of liberally licensed spectrum and marginal increments have been found to generate extremely high social value. The spectrum used in the United States for mobile services in 2008, about 200 MHz, produced annual consumer surplus in excess of \$200 billion.²¹⁰ And additional bandwidth, allocated via mobile licenses, would produce substantial new value. A 30 MHz increment in the United States from 1997 to 2003 (during which the “PCS C-block fiasco” blocked the use of a similar amount) would have generated an extra annual consumer surplus of approximately \$10 billion (in constant 2000 dollars).²¹¹ That these bandwidth additions support new and highly valuable services implies that it is the liberal license regime that best supports investment.

Liberal usage rights are effective social tools for discovering the optimal mixes for spectrum. Not just of basic structures or business models, but in foraging through the countless tradeoffs: when to route traffic through fixed, wired networks versus mobile wireless; when to upgrade technologies versus adding queuing times or splitting more cells; when to allocate more bandwidth to data services; reducing the quality of some voice calls during high-use peak times but increasing the speed of online access; and how much to subsidize new handsets, which is expensive for the network but helpful both in improving users’ experience and in reducing network congestion, since the newer technologies generally use lower power levels and have better processing techniques.

It is often argued that the spectrum allocation system should provide for a mix of services and business models and market structures.²¹² Of course. But the implication drawn is that the government needs to impose this mix, and thereby determine the relative size (and rules) of the alternative

209. Four such countries (New Zealand, Australia, Guatemala and El Salvador) are discussed in Hazlett, *Wireless License Values*, *supra* note 72. See also Thomas W. Hazlett, Giancarlo Ibarguen & Wayne A. Leighton, *Property Rights to Radio Spectrum in Guatemala and El Salvador: An Experiment in Liberalization*, 3 REV. L. & ECON. 437 (2007).

210. Hazlett, Muñoz & Avanzini, *supra* note 6, at 100 tbl.1 (noting \$211.8 billion in 2009 in 2008 dollars).

211. *Welfare Analysis*, *supra* note 68, at 436 tbl.5.

212. Ellen P. Goodman, *Spectrum Rights in the Telecosm to Come*, 41 SAN DIEGO L. REV. 269, 384 (2004) (noting a mixed regime of uses in the spectrum telecosm, with diversity in possible usage patterns, ranging from conservation, individual ownership, community ownership, widely-shared use, narrowly-shared use, limited access, and privately licensed bands).

allocations. That is clearly the systemic source of the problems, which, by consensus, plague the spectrum allocation process.²¹³

The relevant public policy question is precisely how the optimal mix of models should be determined. Administrative allocation rests choices with dispassionate agents who prosper according to how well they play a political, bureaucratic game, a contest that rewards incumbent protection or stasis more frequently than it welcomes entrepreneurial innovation. Coase argued that responsible economic agents with strong incentives to internalize costs and benefits are theoretically attractive actors for this task. By revealing how actual markets operate (with real-world rigidities, transaction costs, and regulatory constraints), market evidence now tends to confirm Coase's hypothesis. Current policy analysis also fails to appreciate that "spectrum commons" allocations (a euphemism given to unlicensed bands that are effectively controlled not by a group of owners, as in a commons, but by state regulators) do not uniquely rely on the case-by-case allocation system in place in the United States since 1927. A market-oriented policy that generally liberalized exclusive spectrum usage rights and distributed them to the market would have ample scope for bandwidth regulated in the way that unlicensed bands are overseen today. The government (or other for-profit or non-profit agency) would acquire exclusive rights in market transactions; the state has the added advantage of potentially using condemnations or gifts to reorganize spectrum. The government acquires land for public parks in this manner. Indeed, public parks are given as explicit analogies to unlicensed bands, but such "commons" are constrained in form and function by land markets relying on private ownership rights.²¹⁴ These markets afford price information that greatly rationalizes purchases or sales of land by public agencies and has the overwhelming advantage that productive deployment of the underlying resource does not get held up while the Federal Land Commission deliberates, years on end, about each proposed use of real estate.²¹⁵ Applying this market-oriented approach to spectrum usage rights would promote similar efficient and effective use.

213. See NBP 2010, *supra* note 3, at 79, Exhibit 5-C (representing the state of spectrum allocation and historical time lags in administrative proceedings).

214. The "government may also wish to promote the important efficiency and innovation benefits of a spectrum commons by allocating spectrum bands for shared use, much as it allocates land to public parks." SPTFR 2002, *supra* note 1, at 38. The error in this analogy is that the government, with a "public interest" allocation regime for all spectrum uses, pointedly does not allocate unlicensed bands "much as it allocates land to public parks."

215. Coase used the FLC example:

B. LICENSED AND UNLICENSED DEPLOYMENTS IN AN EFFICIENT MIXED USE REGIME

It is sometimes suggested that unlicensed allocations have produced relatively high social value, as well. The use of unlicensed bandwidth indeed provides valuable services, for example, the 2.4 GHz frequencies that support wi-fi connections in wireless local area networks (“WLAN”). Studies have been conducted, including one sponsored by Microsoft in 2009, which purport to estimate the value of the unlicensed allocation by assessing the combined value of the services observed using allocated bandwidth.²¹⁶ This methodology is inappropriate because it fails to consider alternatives—either the opportunity cost of the spectrum allocated or the supply of competitive services using a different allocation system.

Perhaps the easiest way to see the problem is to consider the case of broadcast television. The TV Band in the United States today consists of forth-nine channels, or 294 MHz as each channel is allocated 6 MHz by the FCC. Broadcast television stations report annual revenues of about \$22 billion, and annual expenditures on TV sets are about \$20 billion.²¹⁷ As a thought experiment, let us assume that consumer surplus associated with the use of TV sets and broadcast TV content are equal to those amounts: \$60 billion annually. But the values of terrestrial TV broadcasting as a distribution

That such would be the consequences of the establishment of a Federal Land Commission is not, I think, open to serious doubt. It is my contention that similar consequences have resulted from the establishment of the FCC. The most detailed enquiries are conducted before a grant is made of a license for the operation of a broadcasting station. The procedures are costly and time-consuming. This is particularly true in comparative hearings in which the FCC often has to choose between claimants, each of whom seems to be about equally well qualified, and between whom therefore the choice has to be based on some quite trivial or even dubious consideration.

Coase, *supra* note 64, at 163.

216. *See generally* RICHARD THANKI, PERSPECTIVE ASSOCIATES, THE ECONOMIC VALUE GENERATED BY CURRENT AND FUTURE ALLOCATIONS OF UNLICENSED SPECTRUM (2009) (study was supported by funding from Microsoft), <http://apps.fcc.gov/ecfs/document/view?id=7020039036>; Letter from Crag Mundie, Chief Research & Strategy Officer, and Anoop Gupta, Corp. Vice President—Technology Policy & Strategy, Microsoft Corp., to Julius Genachowski, Chairman, FCC (Sept. 21, 2009), *available at* <http://apps.fcc.gov/ecfs/document/view?id=7020039035>.

217. Fourteenth Video Competition Report, *supra* note 201, at 8708 ¶ 215 tbl.17; Press Release, Consumer Electronics Association, CE Industry to Reach Record-High Revenues in 2012, According to CEA (Jan. 10, 2012), <http://www.ce.org/News/News-Releases/Press-Releases/2012-Press-Releases/CE-Industry-to-R Reach-Record-High-Revenues-in-2012.aspx> (estimating 2012 U.S. sales of \$17.2 billion for HDTV display units, \$7.5 billion for 3DTV, \$7.7 billion for internet-connected displays).

system and the allocation of spectrum for this service are nowhere near \$60 billion. Neither the revenues nor the consumer surplus associated with them depend uniquely on the TV Band's 294 MHz.²¹⁸ Over 90% of U.S. TV viewers live in households subscribing to cable or satellite TV, thereby paying for a substitute video delivery platform (which also retransmits off-air broadcast signals), and more importantly, close to 100% of viewers *could subscribe* to cable or satellite TV.²¹⁹ Because these competing video distribution systems and emerging broadband networks can replace over-the-air ("OTA") terrestrial television station broadcasts, the social value of the TV Band reflects this prospect of more cost effective alternatives.

The opportunity cost of devoting 294 MHz of prime VHF and UHF spectrum to something other than mobile services easily exceeds \$1 trillion.²²⁰ Meanwhile, the ten million U.S. households without cable or satellite subscriptions could be supplied with these technologies²²¹ at a one-time cost equal to about \$3 billion.²²² Hence, even if the consumer surplus generated by TV sets and TV broadcast content is the assumed value of \$60 billion annually, the value supplied by the FCC's TV Band spectrum allocation is quite negative.

Similarly, the 2.4 GHz band is useful given its lock-in by regulators, but there are costs to the set-aside. These costs include the social value that could be gained using the 83.5 MHz allocated to the Industrial, Scientific, Medical ("ISM") Band, or some fraction of it, in different ways than those specified

218. In fact, the TV Band was originally 81 channels and 486 MHz between 1939 and 1953, with the digital transition reallocating 108 MHz through auction. See Thomas W. Hazlett, *Tragedy T.V.: Rights Fragmentation and the Junk Band Problem*, 53 ARIZ. L. REV. 83, 105, 109 (2011). The white spaces, however, between the remaining channels has not yet changed, pending the incentive auction proceedings authorized in 47 U.S.C. §§ 1451–1452 (2011) (authorizing the FCC to conduct incentive auctions with forward and reserve auction components for broadcast TV spectrum, with a statutory deadline for completion in 2022).

219. Fourteenth Video Competition Report, *supra* note 201, at 8624 ¶ 37 ("By 2010, cable MVPD service was available to 128.8 million homes (98.5% out of 130.8 million U.S. homes). We assume that DBS MVPDs are available to all homes, but recognize that this slightly overstates the actual availability of DBS."). This proportion is far higher than the percentage of viewers that could obtain their signals directly over the air from a terrestrial station broadcast.

220. See Hazlett, *A Proposal for an Overlay Auction*, *supra* note 50, at 5.

221. There is an argument for public safety emergency alerts; the existence of over-the-air ("OTA") broadcasting as a distribution system is not in question as a matter of category, but as a matter of degree and scale. This Article explicitly does not argue for the end of OTA, but the data by which administrative decision makers reallocate the airwaves to alternative uses in a cost-benefit manner.

222. See Hazlett, *A Proposal for an Overlay Auction*, *supra* note 50, at 7, with a high estimate that 9.6% of households in 2011 relied solely on OTA television without cable or DBS service. Fourteenth Video Competition Report, *supra* note 201, at 8705–06 ¶ 211.

by the FCC.²²³ Moreover, instead of providing additional bandwidth for unlicensed devices, WLANs could be efficiently allocated in spectrum markets. Of course, this alternative relies on the existence of flexible licenses that would permit WLAN technologies to use those airwaves. In this environment, which exists currently for CMRS licenses, spectrum rights could be “bid into” WLAN employments just as such rights allocated bandwidth for cellular networks—Wireless Wide Area Networks (“WWAN”). Indeed, such opportunities are currently available in the market. Instead of lobbying the FCC to put additional bands aside, as did Apple Computer in pushing for a large unlicensed PCS allocation in the 5 GHz band in the early 1990s,²²⁴ device manufacturers could buy (directly, through secondary market deals, or by forming consortia) the spectrum usage rights they demand. This would incorporate price data into the allocation process, with such firms internalizing opportunity costs rather than socializing them via FCC regulation.²²⁵

In fact, there is little interest in expanding unlicensed bands with liberal frequency rights purchased in the market; the alternatives are too valuable.

223. Given the large number of devices and WLANs using ISM frequencies, this is not a proposal but a thought experiment. Once investments have been built around particular rules, transaction costs to transition to alternative allocations also become part of the analysis.

224. See generally *In the Matter of Amendment of the Commission’s Rules to Provide for Unlicensed NII/SUPERNet Operations in the 5 GHz Frequency Range*, Notice of Proposed Rule Making, 11 FCC Rcd. 7205 (1996), available at http://transition.fcc.gov/Bureaus/Engineering_Technology/Notices/1996/fcc96193.txt.

225. See Kwerel & Williams, *supra* note 42, at 31. Because the parties would internalize transaction costs—licensees, equipment makers, and consumers—incentives would be strong to avoid cumbersome payment systems that are more trouble than they are worth:

Future expansion of dedicated spectrum for unlicensed use could be obtained through negotiation between the manufactures of such devices and spectrum licensees. One possible arrangement would be for a licensee or group of licensees covering a particular band throughout the United States to charge manufacturers a fee for the right to produce and market devices to operate in that band. Such contracts could provide different grades of access for different fees, thus providing for a wider range of uses than are possible under the current rules. Competition between licensees would ensure that fees reflect the opportunity cost of the spectrum. Alternatively, manufacturers of low power devices might form a bidding consortium to acquire additional spectrum in our auction. If there is a continued desire as a matter of public policy to provide spectrum for such devices on a “free” basis, the FCC itself might purchase the spectrum in the auction, essentially reducing overall proceeds to the Treasury. This would have the advantage of making the opportunity cost of such allocations.

Id.

This is important information, more reflective of consumer welfare trade-offs than the evidence proffered in support of additional regulatory set-asides of unlicensed spectrum by firms that hope to use such resources for a price of zero. Indeed, the FCC has made a pronounced policy swing over the past fifteen years, aggressively allocating hundreds of MHz for new unlicensed allocations.²²⁶ But relatively little economic value appears to be generated in these increments,²²⁷ as Coase, Meckling, and Minasian may have anticipated:

The absence of a market price (which measures the value of a frequency to another user or in another use) means that a user has little idea of when he is using a frequency “wastefully” and no financial incentive to find out. Obviously, a frequency should not be used for a particular purpose if it prevents the accomplishment of greater value It is clear that such wasteful use must be very common with the existing system.²²⁸

Hence, it is inappropriate to attribute all of the gross economic activity associated with the use of unlicensed devices solely to the existence of unlicensed allocations. The marginal benefits created by wi-fi connections or cordless phones are properly attributed, in large measure, to the fixed networks to which they connect. The net value that is generated, correctly calculated, reflects both spectrum opportunity costs as well as systemic costs, which arise from the fact that by relying on administrative set-asides of dedicated spectrum blocks for unlicensed uses, the regulatory process is intrinsically tied to a case-by-case allocation system wherein political appointees consider the best way to issue spectrum usage rights. This system, by consensus, produces severe spectrum misallocation, regulatory lags, and disappointing welfare outcomes for wireless consumers.

The complexity of coordinating economic activity over mobile links, where thousands or millions of real property owners may be involved and economies of scale may be national or international in scope, produces a

226. See Hazlett & Leo, *supra* note 97, at 1047. Unlicensed allocations currently out-allocate licensed ones: “[b]y 2004, the FCC had allocated approximately 665 MHz of spectrum in the same frequency range to unlicensed use. In comparison, as of that same date, about 385 MHz in this range had been allocated to liberal licenses—an unlicensed-to-licensed ratio of 1.7.” *Id.* at 1049. “[In December 2008, 240 MHz additional unlicensed spectrum] [b]rought the total unlicensed allocation to 955 MHz. By comparison, as of year-end 2008, approximately 422 MHz had been allocated to liberal licenses, bringing the ratio of unlicensed to liberal-license spectrum to about 2.3 [to] 1.” *Id.*; see also *id.* at 1048 n.44; Thomas W. Hazlett, *Spectrum Tragedies*, 22 YALE J. ON REG. 242, 258 fig.2 (2005) (“comparing 648.5 MHz of unlicensed spectrum to 189 MHz of ‘flexible use’ licensed spectrum”).

227. See Hazlett & Leo, *supra* note 97, at 1055 (discussing the Starbucks Fallacy).

228. COASE, MECKLING & MINASIAN, *supra* note 37, at 88.

different conclusion for liberal licenses. It is clear that mobile markets could not exist in any form nearly so valuable without “exclusive use” licenses. Indeed, the expression “exclusive use” is a regulatory term²²⁹ revealing deep confusion that illuminates this discussion. The licenses affording exclusive control over frequency spaces give rise to the most valuable wireless services now offered. The spectrum is intensely used and shared by millions of subscribers. Given exclusivity, operators are able and willing to invest aggressively in complementary capital to build platforms for consumer interactivity. This is the standard argument for private property rights: to be optimally cultivated, land requires an owner. The costs of productive investments can now accrue returns for the investors who make them. Network infrastructure, including handsets,²³⁰ is developed with billions of dollars in financial backing to exploit and share the value created via property rights afforded licensees. Exclusive use facilitates spectrum sharing at industry scale.

The nature of spectrum sharing is commonly misconstrued. The unlicensed bands are said to be “[bands that] no one controls,”²³¹ and facilitate a “spectrum commons.”²³² In fact, open access is not the regime governing these bands. Regulators assert authority, licensing the radios permitted to use these frequencies. These restrictions severely limit spectrum sharing, which would, if unrestricted, lead to dissipation of the value of the resource: the tragedy of the commons.²³³ The primary regulatory devices are

229. *SPTFR 2002*, *supra* note 1, at 5 (“Exclusive use model. A licensing model in which a licensee has exclusive and transferable flexible use rights for specified spectrum within a defined geographic area, with flexible use rights that are governed primarily by technical rules to protect spectrum users against interference.”).

230. Charles L. Jackson, *Wireless Handsets Are Part of the Network*, CTIA (Apr. 27, 2007); Appendix C to Comments of CTIA—The Wireless Association, Skype Communications S.A.R.L., RM-11361 (Apr. 30, 2007), http://files.ctia.org/pdf/Comments_CTIA_Skype_Opposition_Complete_43007.pdf.

231. Yochai Benkler, *Overcoming Agoraphobia: Building the Commons of the Digitally Networked Environment*, 11 HARV. J.L. & TECH. 287, 360 (1998).

232. *SPTFR 2002*, *supra* note 1, at 36, 40.

233. The name is a misnomer, as the paradigmatic 1968 article that gave us this term described it as a “tragedy of open access.” Hardin, *supra* note 12. Eggerston explained the confusion:

In retrospect, the confusion over the nature of common property probably was caused substantially by a mix-up of proper names and theoretical categories. In the field, resources that are governed by open access arrangements often are locally known as “the commons” (or have the word “commons” in their name) because previously they were exclusive common property. The confusion is facilitated by the proximity of open access and common property on the privatization scale. Relatively

(a) power limits, and (b) technology constraints. There are also, in some instances, (c) business model restrictions.²³⁴ The policy strategy is to mitigate potential conflicts via in personam rules, regulations that define what various parties may do with respect to the resource. This is a policy approach distinct from in rem spectrum rights, formulated in liberal licenses, which define a given frequency space and then delegate usage choices to designated rights holders.²³⁵ In particular, the effect of the power limits is to geographically separate users, reducing airwave conflicts by localizing transmissions. By spacing out radios in physical space, these rules create white spaces. In the same way, custom crafted in personam rules necessarily limit the universe of possible sharing combinations.

Of course, wireless networks built by mobile operators—licensees possessing in rem rights—are faced with similar challenges and deploy parallel limiting strategies. They dwell in a maelstrom of conflicting activities that both exacerbate and relieve the basic problem of spectrum scarcity. On the one side, carriers undertake to increase demand for access to the spectrum they control, building networks, distributing or assisting the distribution of mobile handsets, content, and applications—products that add value to the wireless user experience. On the other hand, carriers must ration network access, mitigating congestion while capturing revenues for suppliers, including their own network. This rationing is achieved by a complex set of pricing, service menu, technology and network architecture decisions. For example, carriers may invest in a technology upgrade or network rebuild that triples effective bandwidth for subscribers. But the upgrade is costly and will force the operator to project (a) the net cost of the upgrade relative to the next best capital infrastructure path; and (b) how much customers would pay for the incremental gains, embedding forecasts of price changes, subscriber growth, as well as ancillary revenue streams perhaps made newly available. Competitive market forces drive the operator to seek these particular choice variables, to implement all plans with respect

small changes in the economic environment can push a common property regime into open access or vice versa.

Eggertsson, *supra* note 12, at 75–76; see also David D. Haddock, *Force, Threat, Negotiation: The Private Enforcement of Rights*, in PROPERTY RIGHTS: COOPERATION, CONFLICT, AND LAW 168, 185–86 (2003) (discussing the role of government force to enforce rights when private rights are unavailable).

234. Many countries (e.g., Mexico, the United Kingdom) have limited wi-fi use in the 2.4 GHz band, e.g., to noncommercial operations. Widespread noncompliance resulted from this restriction.

235. See generally Thomas W. Merrill & Henry E. Smith, *What Happened to Property in Law and Economics?*, 111 YALE L.J. 357 (2001).

to the most efficient pathways, and to execute only those choices that promise to produce net gains. In a dynamic world, the optimization parameters continually change, making the search for “optimal combinations” an endless pursuit.

C. RULES TO ENABLE MARKETS

Coase argued, *ex ante*, that spectrum ownership, as opposed to administrative allocation, would steer this “optimal combinations” process in a more efficient direction.²³⁶ Current data from the performance of mobile markets, using liberal licenses and *de facto* spectrum ownership rights issued long after Coase wrote, coupled with ongoing stasis in traditionally regulated markets such as radio and TV broadcasting, strongly endorses this positive prediction and normative proposal. But, the more subtle implication relates to the definition of harmful interference in radio spectrum rights. Indeed, Ronald Coase’s core insight in his 1959 and 1960 papers was that interference between resource users, including wireless operators, was the primary concern.²³⁷ Conflicts were ubiquitous in the economy—Coase did not use the phrase “externality,” popular with economists when discussing the topic—and the generic question did not change with so-called “spillovers.” That question was: how to select which resources have the most value given conflicting and mutually exclusive demands? Government administrators enjoyed comparative advantage in some social activities, but not in discovering or acting upon the information needed to make choices between alternative uses for spectrum. The task, therefore, was to secure rules governing radio spectrum use that would allow private, profit-seeking enterprises to improve upon the choices state allocation authorities would make. This task included questions governing “interference.”

The rules shifting resource allocation choices from public to private actors need not be perfect. Since administrative allocation is a costly system of creating use rights, the process delegating such choices instead to markets may also prove costly, and yet remain superior to government allocation. As the late, great, regulatory economist Alfred Kahn once said: “[W]herever it seems likely to be effective, even very imperfect competition is preferable to regulation.”²³⁸ Law professor Richard Epstein puts the point more broadly:

236. COASE, MECKLING & MINASIAN, *supra* note 37, at 23.

237. *See* Coase, *supra* notes 44, 156.

238. 1 ALFRED E. KAHN, *THE ECONOMICS OF REGULATION: PRINCIPLES AND INSTITUTIONS* xxiii (1988) (citing 2 ALFRED E. KAHN, *THE ECONOMICS OF REGULATION: PRINCIPLES AND INSTITUTIONS* 112, 328–29).

“The study of human institutions is always a search for the most tolerable imperfections.”²³⁹

The relevant task is, then, to enable wireless markets to maximize output rather than to minimize interference.²⁴⁰ In this effort, we should expect there to be many conflicts between users, and even potential users. But, in defining resource rights such that utilization choices are made by responsible economic agents, greater efficiencies result than in alternative regimes. Today, the rules for such a system need not be conjured from whole cloth. They are best designed by building upon existing frameworks. Indeed, this is an advantage that scholars such as Coase (1959), De Vany et. al. (1969), and Minasian (1975)²⁴¹ did not enjoy.

As a technical legal matter, rights to radio spectrum cannot be *owned*—either by the government or by private parties. The airwaves are said, as per the 1927 Radio Act, to belong as public convenience, interest, or necessity requires. As such, the Federal Radio Commission and, since the 1934 Communications Act, the Federal Communications Commission (“FCC”), have been given authority to regulate all uses of spectrum so as to protect the “public airwaves.”²⁴² Yet, by moving from operating permits to defined spectrum spaces, authorizations that permit flexible use of the airwaves allocated the license, U.S. regulators create de facto private property rights in spectrum.²⁴³ Since the 1993 federal budget authorized the use of competitive bidding for most FCC licenses, the liberalized rights to exclusively control frequency spaces have been auctioned; since 1994, over 30,000 licenses have

239. SOWELL, *supra* note 14, at 96 (quoting RICHARD A. EPSTEIN, OVERDOSE: HOW EXCESSIVE GOVERNMENT REGULATION STIFLES PHARMACEUTICAL INNOVATION 15 (2006)). Of course, this view could be appropriately cited to Harold Demsetz, *Information and Efficiency: Another Viewpoint*, 12 J.L. & ECON. 1, 1–4 (1969), which employed the descriptive phrases “nirvana view” or “nirvana approach” to clarify the situation. Today, the phrase “Nirvana Fallacy” merits its own Wikipedia entry. See *Nirvana Fallacy*, WIKIPEDIA, http://en.wikipedia.org/wiki/Nirvana_fallacy (last visited July 15, 2012).

240. “It is sometimes implied that the aim of regulation in the radio industry should be to minimize interference. But this would be wrong. The aim should be to maximize output.” Coase, *supra* note 156, at 27. Taken out of context, this passage has often been misinterpreted. Coase makes clear, both in this paper and in his “Social Cost” article the following year, that economic policy must aim not at determining the right particulars in a given case, but to devise institutions that most reliably discover and implement such solutions over time.

241. Coase, *Federal Communications Commission*, *supra* note 156; De Vany, *supra* note 38; Minasian, *supra* note 72.

242. *Wireless Craze*, *supra* note 30, at 359–61; Radio Act of 1927, Pub. L. No. 69-632, 44 Stat. 1162 (1927).

243. These rights are labeled “exclusively assigned, flexible-use spectrum” (EAFUS) rights in Hazlett & Spitzer, *Advanced Wireless Technologies*, *supra* note 196, at 596.

been assigned and some \$52 billion collected by the U.S. Treasury in winning bids.²⁴⁴

In thinking about defining rights to best handle the problem of harmful interference, these newer regulatory models are of prime importance. In general, property rules evolve over time, shifting in response to changes in demand, technology, population, and resource scarcity.²⁴⁵ Indeed, when resources have limited value, it is often not worth defining rights—whether for state ownership, private property, or a commons—if the process of creating and enforcing such rules entails significant cost. As circumstances change, these trade-offs will likewise change, such as when the introduction of barbed wire in the 1870s allowed cheaper enforcement of private land rights on the Great Plains.²⁴⁶ This precipitated a trend by homesteaders to use fencing material to separate crops from “invading cowboys and their herds,” and by cattlemen, to enclose their accustomed range.²⁴⁷ The definition of land for livestock grazing, crop cultivation, and later, the transcontinental railroads, would transfer much of the government “unclaimed communal property” to “private hands.”²⁴⁸

244. Thomas W. Hazlett, David Porter & Vernon Smith, *Radio Spectrum and the Disruptive Clarity of Ronald Coase*, 54 J.L. & ECON. S125 app. tbl. A1 (2011) [hereinafter Hazlett et al., *Disruptive Clarity*].

245. See generally Harold Demsetz, *The Exchange and Enforcement of Property Rights*, 7 J.L. & ECON. 11 (1964); Demsetz, *supra* note 60.

246. Terry Anderson & Peter J. Hill, *The Evolution of Property Rights: A Study of the American West*, 18 J.L. & ECON. 163, 172 (1975).

247. *Id.* at 172 (describing how in 1882 the Frying Pan Ranch in Panhandle “spent \$39,000 erecting a four-wire fence around a pasture of 250,000 acres.”); see also *Frying Pan Ranch*, TEXAS STATE HISTORICAL ASSOCIATION, <http://www.tshaonline.org/handbook/online/articles/apf03> (last visited July 25, 2012) (describing how in 1887 the Fort Worth and Denver City Railway ran “diagonally through the ranch from southeast to northwest”).

248. Anderson & Hill, *supra* note 246, at 172. The increase in the value of land and property rights definitions increased along with productive and population activity:

But the growing demand for land by cattlemen, sheepherders, and grangers eventually caused the value of land to increase and hence increased the benefits from definition and enforcement activity [in range rights] These rights provided some exclusivity over use in land, but as population increased (see Table 1) [Population of the Great Plains 1850–1900, from 274,139 to 7,377,091], settlement became more dense and land values rose even more. Individuals and groups began devoting more resources toward the definition and enforcement of private property rights.

Id. at 170 (footnote and table omitted).

The development of property rights for private cultivation is incremental, just as it was in the American West.²⁴⁹ Some see the rights definition problem in spectrum as a task that will create a whole new policy regime—or accuse unnamed others of seeing things this way. John Berresford and Wayne Leighton, two FCC analysts writing in 2004, stated:

We agree that if property law is in fact more efficient than spectrum law, then making spectrum more property-like will improve efficiency. To talk of scrapping present spectrum law and replacing it with “property rights,” however, is to throw the baby out with the bathwater.²⁵⁰

But no babies are being tossed. Whether existing spectrum law could be wholly replaced by an entirely new system of rights is an interesting question to ponder. It is not, however, a question worth expending resources upon. Given the broad push towards liberal licenses, existing templates for de facto spectrum ownership are already available in the United States and elsewhere. There are varying programs for further liberalization that might be considered, none of which entail destroying existing rights in order to start from scratch—a process that entails vast disruptions and, in any case, is very unlikely to happen. The contemplated programs of liberalization include regulator-driven initiatives, as undertaken by Ofcom in the United Kingdom in 2002–2004;²⁵¹ legislative reforms as instituted in Australia, New Zealand, Guatemala, and El Salvador in the 1990s;²⁵² and an extension of liberal

249. *Id.* at 172 (“During the 1870’s and 1880’s many acres were privately claimed under the homestead, preemption, and desert land laws. And finally, land was granted outright to the transcontinental railroads who in turn transferred much of it into private hands.”).

250. John Berresford & Wayne Leighton, *The Law of Property and the Law of Spectrum: A Critical Comparison*, 13 COMMLAW CONSPECTUS 35, 37 (2004). We are not aware of any scholars who argue that the entire legal/regulatory regime for radio spectrum should be scrapped, and none are cited in the article cited. This includes Peter W. Huber, author of *LAW AND DISORDER IN CYBERSPACE: ABOLISH THE FCC AND LET COMMON LAW RULE THE TELECOSM* (1997). It also includes, obviously, Thomas W. Hazlett, the co-author of this Article, author of *Optimal Abolition of FCC Spectrum Allocation*, 23 J. ECON. PERSP. 103 (2008) [hereinafter Hazlett, *Optimal Abolition*]. Both of these works explicitly rely on existing de facto property rights exercised by spectrum users, and the legal templates defining them, as the basis for moving forward with further reforms.

251. Martin Cave, *Independent Audit of Spectrum Holdings: Final Report, An Independent Audit for Her Majesty’s Treasury*, SPECTRUMAUDIT (2005), available at <http://www.spectrumaudit.org.uk/pdf/caveaudit.pdf>; Ofcom, *The Communications Market 2006 Overview*, at § 1.6.4, (2006), <http://stakeholders.ofcom.org.uk/market-data-research/market-data/communications-market-reports/cm06/overview06/spectrum/> (authorizing spectrum trading in 2004, with the first spectrum trades in 2005–2006); Ofcom, *Simplifying Spectrum Trading*, <http://stakeholders.ofcom.org.uk/consultations/simplify/statement/> (refining proposals from September 2009 consultation on how to streamline market liberalization).

252. Hazlett, *Wireless License Values*, *supra* note 72, at 582–86.

models, such as overlays rights, developed by regulators but deployed parsimoniously, which form a central feature of the policy solution proposed in this Article.²⁵³

D. OBSERVED EFFICIENCIES IN EMERGING SPECTRUM MARKETS

In defining property rights, various methods of social coordination immediately come into play. Indeed, a basic advantage resulting from the creation of *in rem* rights is the production of low-cost enforcement information for all members of society—both owners and non-owners. The rights and responsibilities associated with ownership are fairly well understood, such that individuals can ascertain how to interact, either from cooperating in a trade or in respecting each other’s rights by not trading. Thomas Merrill and Henry Smith find that shared understandings of ownership help explain why property forms are limited in number, while contractual terms are essentially infinite—agreements specifying anything signatories to the bargain decide to include.²⁵⁴ The information contained in the former must spread to large numbers of parties, while the latter are relevant to only a small number of well-informed parties.²⁵⁵

The essential reason for creating a regime of property-like ownership rights in spectrum, perhaps better seen as a transition from *in personam* to *in rem* rights, or from “governance” to “exclusion,”²⁵⁶ is to better control harmful interference. Here we diverge from the usual case for this reform, which is to expand the economic benefits generated from wireless services. Yet the original proposal for this reform, from Ronald Coase, was developed to resolve the problem that we deal with here: how to best define usage rights in spectrum so as to mitigate harmful interference.²⁵⁷ The original justification for administrative allocation, as retold by Coase, was that, without administrative control of all airwave use, endemic conflicts would destroy the value of radio spectrum, unleashing “chaos,”²⁵⁸ “ethereal bedlam,”²⁵⁹ or a “cacophony of competing voices.”²⁶⁰ Coase saw that this was

253. See *infra* Section IV.K.

254. See Merrill & Smith, *The Property/Contract Interface*, *supra* note 18, at 778.

255. *Id.*

256. Henry E. Smith, *Exclusion Versus Governance: Two Strategies for Delineating Property Rights*, 31 J. LEGAL STUD. 453, 456 (2002) (explaining that “[f]or low levels of precision, the marginal cost of exclusion is low but increases rapidly with higher levels of precision; the marginal cost of governance rules may be higher than for exclusion rules but may increase less rapidly.”).

257. Coase, *supra* note 156, at 34–35.

258. *Nat’l Broad. Co. v. United States*, 319 U.S. 190, 213 (1943).

259. Coase, *supra* note 156, at 2 (citing 2 S. Rep. No. 659, at 4 (1910)).

260. *Red Lion Broad. Co. v. FCC*, 395 U.S. 367, 376 (1969).

not the case. The cited interference issues were solved by a system of use restrictions, which could either be applied top-down by regulators or, alternatively, by delimited rights issued to private parties. Coase's preference for market competition and its efficiency properties led him to favor the latter approach—subject to empirical verification that its net benefits did, in fact, exceed those of administrative allocation.²⁶¹

So, the “choice of the optimal regime” is, in essence, just the policy margin to which investigating how to resolve harmful interference leads. It should be noted that Coase partially changed his mind. While Coase argued initially for private ownership of spectrum in his 1959 article, analogizing spectrum to land and markets in real estate,²⁶² he later detoured:

What does not seem to have been understood is that what is being allocated by the Federal Communications Commission, or, if there were a market, what would be sold, is the right to use a piece of equipment to transmit signals in a particular way. Once the question is looked at in this way, it is unnecessary to think in terms of ownership of frequencies or the ether.²⁶³

As Merrill and Smith noted in 2001,²⁶⁴ this approach is highly inconsistent. The very process whereby the regulator issues parsimonious use rights, determining how “a piece of equipment [may] transmit signals” is what nests choices regarding economic trade-offs with dispassionate regulators instead of profit-seeking entrepreneurs.²⁶⁵ The costs and benefits that Coase aims to internalize via property institutions are truncated; more efficient technologies, services, or business models are deployed only at the discretion of regulators. To argue for market allocation of spectrum is to bypass this system of restrictions and obviate the need for bureaucratic permissions.

More poignantly, the problem with in personam rights in spectrum is that the process by which the regulator defines each new right under a “bundle of

261. At the time, Coase was convinced not by empirical evidence, but by his understanding of Adam Smith's “invisible hand” (tending to recommend market competition), and by the “incredibly feeble” arguments put forward against it by defenders of the existing regime. See Ronald H. Coase, *Law and Economics at Chicago*, 36 J.L. & ECON. 239, 249 (1993).

262. “We know from our ordinary experience that land can be allocated to land users without the need for government regulation by using the price mechanism. But if no property rights were created in land, so that everyone could use a tract of land, it is clear that there would be considerable confusion” Coase, *supra* note 156, at 14.

263. Coase, *supra* note 156, at 33.

264. Merrill & Smith, *supra* note 235.

265. Coase, *supra* note 156, at 33.

sticks” approach, such as when a new wireless technology becomes available, is expensive. Administrative procedures at the FCC are famous for putting the burden of proof on competitive entrants, for sponsoring rent seeking rivalries that create and distribute supra-competitive profits on behalf of politically influential interests or constituencies, and for reliably sacrificing the consuming public’s interest in technological innovation and business competition. By shifting from a defined bundle of explicit spectrum use rights to in rem ownership, economic incentives are redirected to discover and deploy the very options that cannot be known *ex ante*. Such spaces will no doubt be imperfectly defined, but the prospect that improvements will flow from expanding the bundles of rights to include *all* the productive activities that licensees can deploy within such definitions is the proper goal and the observed outcome of liberalization that takes spectrum rights from “governance” to “exclusion.”²⁶⁶

There are no pure property rights.²⁶⁷ Property borders are always subject to some degree of fuzziness, and enforcement of defined interests are inevitably subject to some level of stochastic deviation. Contracts, in which property is reconfigured and rights are traded, are likewise incomplete, reflecting efficiencies internalized by the parties to the contract.²⁶⁸ The law allows that landowners must accept trespass in some situations, for example, when an emergency drives an incursion that is incidental and does not allow for a properly negotiated agreement.²⁶⁹ When air travel developed in the early twentieth century, the rights of landowners to enforce their equity interests

266. See Smith, *supra* note 256, at S462. Smith discussed the Demsetz thesis in the open-field period of England: “Under the Demsetz thesis, the trend toward enclosure reflects increased land value leading to more exclusive rights.” Smith continues his discussion of the Demsetz model: “One initial take on the evolution of property rights is to focus on the costs and benefits of defining and enforcing them. Demsetz proposed that property rights are devices to internalize externalities and will develop when the gains of internalization outweigh its costs.” *Id.*

267. Smith, *Property as the Law of Things*, *supra* note 18, at 1706 (“More generally, situations between the fully in rem and the fully in personam present themselves, and a preliminary inquiry reveals that intermediate situations are handled with less formalism and less rigid standardization than in rem situations but do not allow the degree of customization possible in contract law. As Thomas Merrill and I have shown, in rem rights avail against many parties, and those duty holders tend to be anonymous or indefinite.”); *id.* at 1707 (“In situations falling between in personam and in rem, we tend to find intermediate levels of mandatoriness and standardization.”).

268. RICHARD A. POSNER, *ECONOMIC ANALYSIS OF LAW* 118–19 (8th Ed. 2011) (explaining that “[a]djudicative gap-filling is a particularly economical method of dealing with contingencies that, even if foreseeable . . . are so unlikely . . . that the costs of careful drafting to deal with them exceed the benefits”).

269. *Id.*

high in the sky was truncated by legislation. The solution was efficient. In determining what rights are properly included in the ownership package, priority is given to highly complementary rights that are most productively used in an integrated fashion.²⁷⁰ The fly-over routes of airplanes are inefficient, and cannot be efficiently integrated with the land parcels underlying them, given the large number of property owners involved.²⁷¹ The rights create inefficiencies, and therefore, they are split.²⁷²

Consider the case of property owners sitting atop petroleum reserves. While these landowners benefit society in owning such resources, and thereby seek to economically supply them to consumers, it is prohibitively costly to define the precise, or even approximate, parameters of the underground pools and attribute the subsurface resource rights to landowners (or other owners). Instead, such resources are discovered over time. When oil is found, reserves must be jointly pumped—in any situation where multiple surface owners are located over the pool—with explicit coordination between owners. Otherwise, an inefficient race to dissipate the resource is ignited. Property rules, called “unitization,”²⁷³ developed to define the private property rights to petroleum.²⁷⁴

With radio spectrum, one must first ask whether spectrum rights should be separate from real property. A landowner might claim ownership of the airwaves in the aerial cone defined by land surface rights. Some spectrum analysts have actually argued for such a system.²⁷⁵ The approach would clearly deprive society of highly valuable opportunities for communications services. Were wide-area networks (“WAN”), such as mobile phone systems,

270. Smith, *Property as the Law of Things*, *supra* note 18, at 1703 (“Property clusters complementary attributes—land’s soil nutrients, moisture, building support, or parts of everyday objects like chairs—into the parcels of real estate or tangible and intangible objects of personal property. It then employs information-hiding and limited interfaces to manage complexity.”).

271. Smith, *supra* note 65, at 1746 n.6. (“*cujus est solum, ejus est usque ad coelum et ad inferos* (he who owns the soil owns also to the sky and to the depths)”).

272. See *infra* note 287; Smith, *Property as the Law of Things*, *supra* note 18, at 1721–22 (citing Henry Hansmann & Reinier Kraakman, *The Essential Role of Organizational Law*, 110 YALE L.J. 387, 390 (2000)) (discussing asset planning).

273. Gary D. Libecap & James L. Smith, *The Economic Evolution of Petroleum Property Rights in the United States*, 31 J. LEGAL STUD. S589 (2002) (examining how property rights develop when the benefits exceed the costs of doing so, in the case of unitization, a common-pool property arrangement in subterranean oil and gas extraction).

274. Gary D. Libecap, *Contracting for Property Rights*, in PROPERTY RIGHTS: COOPERATION, CONFLICT AND LAW 142–67 (2003).

275. J.H. Snider, *FCC Lets the Telecom Giants Steal from You, Via Eminent Domain, Fat Cat Donors Get Airwaves—Worth Billions—In Our Homes*, SACRAMENTO BEE, Apr. 7, 2002, <http://www.newamerica.net/node/6856>.

forced to negotiate with thousands of landowners in order to provide regional or national wireless services, it would undermine productive activity, similar to the airline fly-over example. It would be far better if the rights for spectrum are, in general, split from the control of landowners.

However, there are spectrum cases where the rights of real property owners should matter. For extremely low-powered emissions, usage is necessarily quite localized as weak signals do not tend to travel far. Coordinating with conflicting users is not so complex as with higher-powered, longer-distance emissions.²⁷⁶ On the other hand, defining and enforcing wide-area rights down to such levels is expensive. At some very low power level, the effort of defining these rights is not likely to be worth the cost. Radio-based networks are engineered to avoid, by and large, use of the “noise floor” which hosts such very low-powered emissions and a host of atmospheric and topographic obstacles.²⁷⁷ So, the efficient policy is to define exclusive ownership rights over higher power levels across wide geographic areas (local, regional or national), accommodating the emissions generally deployed in such networks while ceding property owners control over very low-powered emissions that peacefully co-exist beneath the wide-area communications traffic. Such use of localized, very low-powered wireless devices need not be formally recognized by easements, but simply embedded via limitations placed on the rights authorized for wide-area licenses.

There is considerable interest about the need to construct or assign spectrum rights such that transmission rights can be assigned with real-time coordination.²⁷⁸ This interest is often driven by the emergence of cognitive radios capable of scanning the radio environment and opportunistically using frequency spaces that are temporarily vacant. This desire to utilize under-utilized spectrum space is quite understandable given the nature of the spectrum allocation system, a regime that routinely blocks productive wireless activity. But the approach taken, nesting rules for opportunistic spectrum sharing in the same regulatory system that has under-allocated rights to begin with, is dubious given that there is a demonstrably superior mechanism for rights allocation.

That mechanism is the market allocation of spectrum capacity found in the mobile market. With liberal licenses, carriers routinely allocate spectrum minutes to a large variety of networks, customers, and application providers.

276. SPTFR 2002, *supra* note 1, at 19.

277. *Id.* at 25 n.47 (describing the engineering of radio devices with operational specifications of signal-to-noise ratios).

278. *See, e.g.*, Forde & Doyle, *supra* note 41.

Retail subscribers purchase access to the carrier's spectrum along with network access for a monthly fee (post-paid contracts) or by purchasing minutes of use (pre-paid contracts).²⁷⁹ Wholesale subscribers, including MVNOs or system aggregators facilitating M2M applications, purchase spectrum and network access for millions of minutes of monthly usage.²⁸⁰ Application providers, including handset vendors, strike contractual bargains with networks allowing their customers to obtain wireless connectivity from a particular carrier.²⁸¹

In this manner, spectrum is intensely shared, but the nature of the contractual forms shows how efficient these forms are. Deals are, with an important exception, universally *bundled*: spectrum access is sold in a package with network access. Naked spectrum access—possibly with the set-aside of a certain frequency band for use by devices that do not access the mobile operator's network—is not sold directly to the consumer. The exception is that naked, unbundled spectrum rights do trade, actively, in license markets. These are often called “secondary markets.” However, the term can be applied far more broadly in that operators are in the business of (a) acquiring bandwidth via wireless license rights, aggregating those rights into efficient combinations through initial bids in auctions and then purchases in secondary markets; (b) cultivating the value of the acquired bandwidth via the construction of complementary networks and mobile services ecosystems; and then (c) reselling the spectrum rights, bundled with network and ecosystem access.²⁸²

Are spectrum rights sold dynamically? Yes and no. Retail and wholesale customers typically access networks on demand, using long-term contracts. Quality of Service (“QoS”) is a key competitive metric; when network services are unavailable (e.g., blocked or dropped calls, slow data throughput) carriers tend to lose market share to rivals, enforcing competitive incentives to improve QoS.²⁸³ Many of the services offered are “free” on a per-unit basis. After paying a fixed entry fee (monthly subscription price), on-net or

279. Fifteenth Annual Competition Report, *supra* note 20, at 9767 ¶ 167.

280. *Id.* at 9756 ¶ 150, 9785 ¶ 195.

281. *See* Hazlett & Leo, *supra* note 97, at 1089.

282. *See, e.g.*, Press Release, Inmarsat, N. America Spectrum Deal Signed with SkyTerra and MSV (Dec. 21, 2007), *available at* http://www.sec.gov/Archives/edgar/data/756502/000114420407068694/v097951_ex99-1.htm (describing the “re-banding and efficient reuse” of licensed L-Band spectrum, with “Coordination parameters for the parties’ next generation satellites covering North America, both the new Inmarsat-4s and the new MSV1 and MSV2 satellites, in a manner designed to increase spectrum efficiency.”).

283. *See* Fifteenth Annual Competition Report, *supra* note 20, at 9758 ¶ 156.

off-peak phone calls or text messages are uncharged.²⁸⁴ Mobile network data consumption and peak voice calls are usually “free” to a cap, then limits take effect or overage charges apply.²⁸⁵

Were there sufficient demand for commoditized spectrum access for opportunistic devices to justify its opportunity cost, networks would sell it and customers would buy it. Indeed, many of the devices that currently access networks—from Kindles (for book downloads) to OnStar (for emergency phone calls) to dongles (yielding broadband connections for notebooks)—are analogous from the consumer’s perspective to plug-and-play radios that operate via direct access to broadcast spectrum.²⁸⁶ It is perfectly legal to extend the carriers’ models to accommodate cognitive radio access directly by consumers. A compelling hypothesis explaining why such deals have, in the main, not emerged is this: There are great efficiencies in vertically integrating wireless networks with spectrum ownership rights. While thousands of investors and millions of customers may pay to create and then operate such systems, leaving control in the hands of one for-profit enterprise reduces complexity. Borders are reduced, and border disputes eliminated. Instead of a firm aggregating a national swath of radio spectrum rights and then selling access to multiple parties in the form of naked spectrum rights, the market iterates on a model that cross-subsidizes spectrum and networks under common ownership, and then divvies up total value into discrete bundles of spectrum and network service.

In this sense, commoditization of spectrum rights is not an efficient outcome, as opposed to bundled rights. Commoditized spectrum, except perhaps in very localized applications best served by unlicensed spectrum set-asides, has little value to users relative to spectrum inputs with network access. The difference is large enough to provide an incentive for carriers to create and build those networks; it would not be their first choice, in terms of profit maximization, to construct such costly accouterments.²⁸⁷ Because competitive forces constrain the networks to efficient operations, they are not able to command profit-maximizing returns without integrating into both ends of the business and selling both wholesale and retail customers bundled packages.

284. *Id.* at 9724–25 ¶ 82.

285. *Id.* at 9725 ¶ 83.

286. *Id.* at 9754–55 ¶ 146.

287. Smith, *Property as the Law of Things*, *supra* note 18, at 1721–22 (discussing bundles of abstract rights in entity property of sophisticated contracts, equitable property as trusts, and financial instruments, with the efficiencies of asset partitioning as a higher-level modularization of resources).

The important implication is that exclusive spectrum rights should not be over-defined. To help create secondary markets, some advocate that spectrum rights be divided into tiny increments²⁸⁸ in frequency space, area, time, or in additional dimensions such as altitude (height above adjacent terrain) or directionality (signals beamed at different angles producing distinct, non-interfering communications links), as discussed in the next section. Indeed, there are an infinite number of possible spectrum packages that could be devised. But the goal of rights definition is not to maximize the customization of possible parcels, but to simplify the process wherein rights are transferred to parties who can best maximize social value—a rule that also applies when seeking the parties in the best position to design the packages.²⁸⁹ Should the state have a comparative advantage in this respect, it would make economic sense to have regulators supply this design. But the evidence is overwhelming that this is not the case. Government usage definitions have historically been poorly designed and, even worse, rigid. Markets, given a chance to reallocate the spectrum allocated to mobile licenses, have proven adept at devising efficient models to create new wireless services, and at adjusting business models as innovative applications and more advanced technologies become available.

However, one very important point is made in the dynamic access model presented by engineers Tim Forde and Linda Doyle. Arguing for the release of small, flexible-use spectrum blocks, they limit the definitions of these blocks to three dimensions, adopting the TAS rights of De Vany et al. (1969).²⁹⁰ They also include, as an integral part of their proposal, an auction “for the IPO-like release of spectrum licenses,”²⁹¹ giving market competitors an opportunity to aggregate such rights into far larger blocks, presumably (from observing how mobile markets work) dictated by efficiency. Indeed, they propose a combinatorial clock auction,²⁹² a mechanism designed to

288. MATHESON & MORRIS, *supra* note 56, at 8; *see supra* note 74 and accompanying text; Coase, *supra* note 156, at 33.

289. Smith, *Property as the Law of Things*, *supra* note 18, at 1708 (citing Smith, *Standardization in Property Law*, in RESEARCH HANDBOOK ON THE ECONOMICS OF PROPERTY LAW) (“Modularity and its standardization of the ‘outsides’ of property packages allow achievement of a wide range of objectives (lowering frustration costs), while keeping information costs under control, relative to a system of more tailored packages.”).

290. De Vany et al., *supra* note 38, at 1501.

291. Forde & Doyle, *supra* note 41, at 329.

292. *Id.* at 329 fig.3 (describing the general clock mechanism with the following auction components: reserve, post prices, bid, excess, tick, and assignment). They propose a combinatorial clock auction for network-based access. *Id.* at 332. Take note that the authors suggest the clock auction “is not itself ‘cognitive’ [but] the mechanism can be optimally exploited by cognitive participants.” *Id.* at 332–33. This result is implicated by the exclusive

facilitate the efficient valuation and aggregation of highly complementary rights, applied perfectly to the assets being distributed here.²⁹³

The elements of efficient spectrum markets are thus found in the Forde-Doyle proposal: liberal spectrum licenses, allowing responsible economic parties to internalize the costs and benefits of alternative spectrum use, are a ready mechanism for overcoming inefficient rights fragmentation in the initial creation or distribution of such rights.²⁹⁴ The key policy issue revolves around the cost of government parcelization of spectrum blocks in combinatorial auctions that allow bidders to selectively aggregate rights into efficiently-shaped packages that winning bidders are then free to disaggregate, versus a more cumbersome approach where packages are defined by regulators.²⁹⁵ That U.S. regulators have proven unable to conduct

rights foundation for bidders. *Id.* For a general overview of combinatorial clock auctions, see David Porter & Vernon Smith, *FCC License Auction Design: A 12-Year Experiment*, 3 J.L. ECON. & POL'Y 63 (2006). Combinatorial auctions allow a licensee to bid for “packages of licenses,” *id.* at 74, rather than pre-arranged combinations:

[B]idders are not able to make bids such as ‘I want License B *and* C together or neither.’ Not allowing “*and*” bids handicaps bidders who have regional or national business plans. Combinatorial auctions allow for such bidding possibilities. In addition to *and* bids, combinatorial auctions allow for “*or*,” “*only if*” and other logical bid constraints.

Id. (emphasis in original). Porter & Smith also provide revenue per auction per year data from FCC spectrum auctions that did not apply combinatorial bidding and had varying results. *Id.* at 78 figs.1–2. Strategic bidding behaviors included new terms as jump bidding, up yourself, retaliatory bids, parking, eligibility management, lateral hand-off, bid and waive, trailing digits, and budget bluffing. *Id.* at 71–72. Recommended improvements to auction design included click-box bidding, limit withdrawals, increment smoothing, among structural changes by way of combinatorial auctions which would allow bidders to select the complementary bundles of licenses. *Id.* at 73–74.

293. As Ford & Doyle explained:

Combinatorial auctions are very useful when there are a range of items on sale which may be logically grouped together into many different packages to suit either the buyer, the seller or both. One of the main benefits of combinatorial auctions is that they reduce the financial exposure of the traders, i.e. a buyer will either get to buy all of its target package or none of it.

Ford & Doyle, *supra* note 41, at 331. The FCC endorsed the combinatorial bidding format in a May 2000 conference of prominent economists. FCC, *May 5–7, 2000 Combinatorial Bidding Conference*, http://wireless.fcc.gov/auctions/default.htm?job=conference_agenda&y=2000 [hereinafter *2000 Combinatorial Bidding Conference*].

294. Forde & Doyle, *supra* note 41, at 332 (comparing the access mechanism as more like a “commodities market” with session to session trades, rather than an “initial-public-offering style trading”).

295. *2000 Combinatorial Bidding Conference*, *supra* note 293; Porter et al., *supra* note 144; COMBINATORIAL AUCTIONS (Peter Cramton, Yoav Shoham & Richard Steinberg, eds., 2006) [hereinafter COMBINATORIAL AUCTIONS].

auctions with serious combinatorial bidding options, despite having professed interest in adopting the approach for well over a decade,²⁹⁶ suggests that administrative costs today may be relatively high.

E. MOBILE OPERATORS AS SPECTRUM OWNERS AND NEIGHBORS

The solution to harmful interference is worked out with little fanfare in markets where liberal licenses are issued and firms are free to aggregate rights by secondary market trading. The natural tendency is to then use mergers to eliminate borders. Without the 40,000-plus CMRS licenses owned by Sprint Nextel, wireless services provided by the carrier or carriers deploying the same bandwidth would have to account for the spillovers that might occur across boundaries, both in the geographic and frequency dimensions.²⁹⁷ Without aggregated licenses, coordination would be extensive and complex. Moreover, economies of scale in provisioning networks—from volume discounts in the purchase of infrastructure and handsets, to construction of a national network complete with roaming capability for subscribers, to marketing in national media outlets—would be sacrificed. All of these forms of harmful interference are overcome via rights aggregation.

Merger is not free, however. It costs a firm to manage a larger asset base, and not all such costs will be subject to scale economies. Firms tend to engage in such mergers only when the advantages of a merger exceed the disadvantages. Profit incentives, including capital market liquidation of poorly managed firms that botch the cost-benefit calculus, govern the process. No such feedback loop is available in the public sector. Regulators routinely disaggregate licenses, producing band plans with expensive borders, managing spillovers by “guard bands,” technology, or service restrictions.²⁹⁸ Guard bands leave spectrum fallow to protect traffic in neighboring frequencies. The overuse of such techniques is well documented, most dramatically in the use of “taboo” channels that have dominated the TV

296. *2000 Combinatorial Bidding Conference*, *supra* note 293. Through 2011, however, the only combination bidding allowed in an FCC auction occurred in Auction 73 (for 700 MHz licenses) and only allowed package bids that aggregated the largest (Regional Economic Area Grouping REAG licenses) in Block C. *See* Hazlett et al., *Disruptive Clarity*, *supra* note 244, at S145 n.39.

297. *See* discussion *supra* Section III.E (on Nextel Spectrum Swap); FCC Adopts Solution to Interference Problem Faced by 800 MHz Public Safety Radio Systems, News Release, FCC, July 8, 2004; *Fifteenth Wireless Competition Report*, *supra* note 20, at Appendix A ¶ 6, available at http://hraunfoss.fcc.gov/edocs_public/attachmatch/DOC-249414A1.pdf.

298. *Fifteenth Wireless Competition Report*, *supra* note 20, at Appendix A ¶ 10 n.26 (citing 700 MHz Second Report & Order, 22 FCC Rcd. at 15,292–93, ¶ 3) (describing the modification of 700 MHz guard bands, to reallocate uses to commercial spectrum) (Appendix text only available at FCC website).

Band for its entire existence.²⁹⁹ Given that regulators bear no cost for under-allocations yet receive strong political support for overprotecting incumbents (both from the licensees and from politicians who see the regulation-created rents as a platform for launching various demands for quid pro quo commitments by licensees), economists are not surprised by the scope of the anti-consumer outcomes.³⁰⁰

While most of the “interference” problem in liberally allocated spectrum is efficiently remedied via rational choices enabled by flexible-use and secondary market options, the borders that remain, even following license aggregation, are typically well behaved. That is, when two wireless operators, each possessing liberal spectrum rights, share a frequency border, the outcome does not generate much legal or administrative conflict. With flexible use rights, parties internalize the costs of both signal degradation and haggling. In any given interference issue, the parties involved are highly motivated to fix the problem, cooperating to create gains, which are then split between them by agreement.³⁰¹ Of course, there can always be acrimony over such skirmishes: who incurs costs, and who then pays what to compensate.

Again, mobile markets supply the relevant empirical data. In an interesting analysis, Verizon Wireless executive Charla Rath has noted that her firm owns “more than 1500 mobile licenses, not to mention thousands more microwave licenses.”³⁰² While this eliminates countless potential border disputes, it does not eliminate the possibility of such disagreements entirely. Indeed, given the large stakes involved in serving “nearly 100 million customers,” in theory, the borders Verizon Wireless continues to share with

299. NBP 2010, *supra* note 3, at 79 (“Legacy ‘command and control’ rules, high transaction costs and highly fragmented license regimes sometimes preserve outmoded band plans and prevent the aggregation (or disaggregation) of spectrum into more valuable license configurations.”).

300. The observation of overprotection and wasteful spectrum management dates back many decades. *See, e.g.*, ROGER G. NOLL, MERTON J. PECK, & JOHN J. MCGOWAN, ECONOMIC ASPECTS OF TELEVISION REGULATION 4 tbl.1-1 (1973) (explaining that in 1971, 37% of commercial broadcast spectrum assignments were unclaimed, and 66% of noncommercial assignments were unclaimed); ITHIEL DE SOLA POOL, TECHNOLOGIES OF FREEDOM 139 (1983) (“The present system makes for inefficient use of spectrum and thus causes its scarcity, whereas a market system achieves equilibrium by both reducing demand for and increasing supply of usable bandwidth.”).

301. *See, e.g.*, Adam Santariano & Peter Burrows, *Apple, Verizon Took Years to Overcome Their iPhone Differences*, BLOOMBERG (Jan. 12, 2011), <http://www.bloomberg.com/news/2011-01-12/apple-verizon-took-years-to-overcome-their-iphone-differences.html>.

302. Rath, *supra* note 28, at 528.

other wireless licensees are subject to compromise by actions taken by rivals innocently, strategically, or maliciously.

That has not happened. “It is critical to our business that we are able to negotiate and resolve quickly most, if not all, rights and interference issues without seeking intervention or assistance of the Federal Communications Commission,” writes Rath.³⁰³ The regulatory quagmires seen in the Nextel-Public Safety dispute, or the decades-long misallocation of spectrum seen in the TV Band, do not materialize in the mobile markets. That is because the parties involved would not gain from such an outcome, and rationally take measures to avoid it. Rath continues:

Wireless carriers’ thousands of licenses and thousands of miles of adjacent and co-channel boundaries create a laboratory for evaluating whether this successful approach to interference “rights” negotiations is pertinent to a larger radio operating rights framework.

Under current rules, licensees negotiate to extend rights into each others’ licensed spectrum on a daily basis. These . . . involve hundreds of individual negotiations between companies’ engineers who are tasked with the day-to-day operations of the network [C]arriers (including Verizon Wireless) do not always achieve their goals. That said, because the rights of both licensees are clear, there is no benefit to seeking regulatory redress. Instead, we manage the process in the market and look to other ways to gain the rights to spectrum we need to operate—typically through spectrum purchase or lease.³⁰⁴

This is the way “interference” disputes exist in liberal markets. They are simply a part of doing business, with remedies—including situations where one company will pay another to redirect a transmitter or move a base station—occurring not to eliminate conflicts but to economically manage them. CMRS licensees believe that “the rights of both licensees are clear,” meaning that they are clear enough.³⁰⁵ These approximate rules sufficiently resolve the interference problem when licensees have flexibility to switch network deployments or to trade rights, which typically happens after stalled negotiations are remediated “typically through spectrum purchase or lease.”³⁰⁶

303. *Id.*

304. *Id.* at 529.

305. *Id.*

306. *Id.* at 530.

This experience is of enormous importance in informing policy makers as to how to most effectively deal with the interference problems in wireless. Phil Weiser and Dale Hatfield, however, take issue with this conclusion. They follow extant FCC commentary by arguing that the relatively harmonious nature of CMRS bands is due to the “stable and ‘repeat players’”³⁰⁷ that dwell there. There are several problems with this argument. First, this assumes the players to be exogenous; in fact, the liberal rules have allowed national networks to aggregate licenses and emerge as enterprises, holding spectrum rights portfolios quite distinct from what regulators initially licensed. Second, were the attribution correct that stable, repeat players are relatively successful in solving border disputes, the gains from this form of organization would again flow to the market forces that create such relationships. But third, the empirical claim has substance only in those instances where liberal spectrum rights have allowed the right kind of negotiations and adjustments by stable, repeat players. When the FCC has put different types of licenses into play—as in the Nextel-Public Safety 800 MHz band dispute, when fragmentation through interleaving of non-commercial, non-profit licensees froze the “interference” problem in place—the fact that there were stable, repeat players did not prevent gridlock.³⁰⁸ Nor was gridlock prevented by the professionalism or intelligence of the engineers, many of whom work for cellular carriers that solve problems daily in negotiations with rival firms or are highly-trained experts dedicated to public service.³⁰⁹

Even more important is that the Nextel-Public Safety interference dispute was not remedied by the engineers, but by license trades. Greater specificity by incorporating rules into Nextel’s SMR licenses or the police and fire departments emergency radio authorizations that would narrow

307. *Id.* at 529 (discussing Weiser & Hatfield, *supra* note 17, at 589).

308. Similarly, the interference problem in TV broadcasting has been handled in a hugely wasteful manner—setting aside the great majority of TV channels as taboo for at least six decades. Weiser & Hatfield, *supra* note 17, at 559. These taboos serve as guard bands, idle spectrum set aside to protect reception of adjacent signals. That station engineers might handle problems between them, fixing some residual conflicts, does not suggest that the stable, repeat game engaged for over sixty years by the TV stations has had any significant impact in reducing the social cost of interference control. That is because the narrowly specified transmission rights in TV broadcasting prohibit more productive activity, a problem that cooperating engineers cannot fix.

309. Indeed, the widely acknowledged waste of spectrum allocated to wireless deployments in the military occurs not due to the lack of knowledge or dedication of American soldiers or U.S. Defense Department employees, but due to the institutional basis on which the government agency makes its resource choices.

opportunities for spillovers³¹⁰ was not the answer. What would, in other instances, have been secondary market transactions were here executed in the FCC's "spectrum swap," a proposal first advanced by Nextel. That a regulatory proceeding lasting about a decade was required to fix the interference problem was due to the sharp preemption of markets by FCC spectrum allocation policy. Mobile markets, with liberal licenses held by responsible, profit-seeking agents, do not have to wait for regulatory interventions to defragment regulated uses. Not having to ask regulatory permission for a spectrum swap generates efficiencies, including stable, repeat players who behave rationally with respect to cost-benefit calculations.

F. THE COSTLY AND UNPRODUCTIVE SEARCH FOR ADDITIONAL SPECIFICITY

[R]eform may most need to focus on the precise definition of rights to the spectrum. These rights were ill-defined almost 80 years ago and, despite the emergence of relatively efficient institutions for addressing spectrum use, remain in need of reform today. Finally, such reform must also recognize that, much as defining rights to land has not been simple, clarifying the rights to spectrum will be a complex task.³¹¹

This Article suggests that this conventional wisdom to "focus on precise definition" is misguided. Well-established reforms have produced de facto spectrum ownership rights that efficiently promote economizing in the provision of wireless markets, including large-scale investments in complementary wireless networks, devices, and applications. Such reforms are proven to scale extremely well, meaning that they can generate large new gains in economic welfare if extended and accommodate a diverse variety of technologies, business models, and economic structures, producing the price data key to rationally evaluating relevant tradeoffs in the allocation of radio spectrum.

Greater precision in usage rights can be helpful, but such improvements are not costless. To the extent that further clarifications are a complex task, they invoke the prospect of significant delays and rent seeking expense, not to mention anti-competitive outcomes when such technical specifications restrict innovation (a common regulatory approach to mitigating

310. See 800 MHz R&O, *supra* note 128; see also MURRAY, *supra* note 124 (discussing Nextel SMR as an example of unwinding and re-banding fragmented rights allocated by sequential regulation across private and public licensees).

311. Berresford & Leighton, *supra* note 250, at 37.

interference) or are steered to protect favored interests, restricting competition (a standard regulatory outcome).

Extremely large welfare increases are associated with the provision of additional bandwidth injected into the market via existing CMRS templates. This can be seen not only in the analysis offered in this Article, but in the conclusion of the FCC's 2010 National Broadband Plan: in emphasizing the prime importance of wireless services to the development of both broadband markets and the U.S. economy, it ignored the issue of rights clarification.³¹²

Many scholars see things differently, reversing the order of operations, describing technical specificity by which efficiency occurs, over the outer boundaries of complementary rights bundles sought by capital-intensive network enterprises. Matheson & Morris outline a seven-dimension paradigm to define "electrospace."³¹³ The approach builds on the "TAS" template put forward by De Vany et al. (1969), which defined Time, Area, and Spectrum in four dimensions (longitude and latitude being required for Area)³¹⁴ and which has been effectively implemented in the form of CMRS licenses. Matheson & Morris expand the format, as illustrated in Figure 5, *infra*, by specifying altitude, azimuth and elevation angle.³¹⁵

The reason for this is that multiple wireless communications may take place at different altitudes, particularly when using fixed point-to-point links connecting with beam transmissions. The same basic idea allows

312. None of the seventeen high-level spectrum policy recommendations in the NBP propose greater delineation of border contours for licenses. They focus, rather, on moving 500 MHz of CMRS spectrum into the market by 2020. *See 2010 NBP, supra* note 3, at 10. This reveals an appropriate emphasis on bandwidth using definitions already available. It does not settle all remaining definitional questions, of course, and there are likely to be incremental changes that could improve the utility of spectrum use rights. For instance, streamlined dispute resolution processes have been previously suggested. *See Wireless Craze, supra* note 30, at 461–62. Some countries have experimented with new institutions to expeditiously resolve interference disputes. Charles Jackson describes New Zealand's system, which calls for binding arbitration before a tribunal composed of government officials to make judgments based on listed criteria. *See* Charles L. Jackson, *Spectrum Markets: Challenges Ahead*, Presentation at Northwestern University, Kellogg School of Management (June 2–3, 2011) (presentation slides available at <http://www.kellogg.northwestern.edu/meds/spectrummarkets/slides/Jackson%20Thoughts%20on%20Spectrum%20Property%20Rights%20and%20Spot%20Markets.ppt>). Guatemala's 1996 reform legislation mandates private mediation to resolve conflicts; if such a process fails, the telecoms regulator is then mandated to decide the issue. *See* Giancarlo Ibarquen, *Liberating the Radio Spectrum in Guatemala*, 27 TELECOM. POL'Y 543, 546 n.17 (2003).

313. MATHESON & MORRIS, *supra* note 56, at 7 (defining "electrospace" as an alternative term to include features of "frequency, time, space, direction, and other directions" in the electrospace dimensions of spectrum "location").

314. De Vany et al., *supra* note 38, at 1501.

315. MATHESON & MORRIS, *supra* note 56, at 9.

communications traffic to overlap in TAS cubes, for example, by directionality, altering the angle of signals across the horizon (azimuth), or by elevation.

Figure 5: Defining Electrospace Parcels as per Matheson-Morris (2011)³¹⁶

QUANTITY	UNITS	# OF DIMENSIONS
Frequency	kHz, MHz, or GHz	1
Time	seconds, hours, years	1
Spatial Location	latitude, longitude, altitude	3
Direction of Travel	azimuth, elevation angle	2

Within coordinated radio networks, the technical possibilities are boundless.³¹⁷ Consider just the seemingly uncontroversial Time dimension: It is evident that there is no limit to the number of periods that can be defined as the relevant term for a given spectrum right. One can simply divide the Time unit by progressively higher integers. Yet the exercise instantly becomes irrelevant for public policy. The task in fashioning property rights is not to determine the smallest units that are possible to define, but to determine the size and shape of rights that can most efficiently be (a) defined by regulators, and (b) transferred into productive use. The latter process seeks to discover new and more useful ways of sharing the resource. The primary task in the “IPO-like release” of spectrum rights to the market is to package authorizations that assist, or at least do not hamper (as with crippling fragmentation), those productive activities by forcing costly transactions on the market to reassemble complementary rights.

As David Friedman posits:

316. Reproduced from MATHESON & MORRIS, *supra* note 56, at 8 tbl.1.

317. *Report of the Interference Protection Working Group, FCC Spectrum Policy Task Force*, Nov. 15, 2002, at 5, <http://www.fcc.gov/sptf/files/IPWGFinalReport.doc> [hereinafter *IPWG Final Report*], cited in Goodman, *supra* note 212, at 300 n.92 (quoting the Interference Protection Workshop remarks of Paul Steffes):

Obviously, the number of users and the management of the problem becomes dramatically enhanced [as spectrum use becomes more complex. Consideration of interference is] at least a six dimensional problem, meaning spatial, x-y-z, frequency, time and waveform, and of course since the wave form can be infinitely complicated, you can make it an n-fold problem, which basically has more variables than you have numbers.

Id.

When constructing bundles of rights the first question becomes “Which rights belong together?” If I own the right to farm the land, the right to walk on the land is worth more to me than to anyone else, so the two belong in the same bundle. . . . The right to forbid radio waves from passing over my property, on the other hand, is very little use to me. If every property owner had that right, setting up a radio station would require unanimous consent from every owner within range of the broadcast, making a transfer of the right from the owners to the person to whom it is of most value a prohibitively difficult transaction.³¹⁸

The task of drawing boundaries around property rights is not to define all possible configurations but to provide a cost-effective means for resource optimization. In assigning spectrum rights to private actors who will then organize wireless services using them, the government need only define those parameters that it has a comparative advantage in designating. Clearly, a too parsimonious definition that does not establish sufficient authority for licensees to invest in creating networks, such as a TAS license that omitted the A, would fail to provide the legal inputs necessary for economic maximization. By the same token, a rights definition that over-specifies rights can fragment ownership so as to undermine efficiency. These policy errors give rise, respectively, to the twin tragedies of the commons and the anti-commons.

TAS holders are capable of designing sharing arrangements for directional signals and transmissions at different heights above surrounding terrain. Indeed, mobile networks already pack their CMRS license-defined spaces with intense traffic exploiting these dimensionalities, which is why Matheson and Morris, having observed market development of the capacity-increasing technologies that exploit such dimensionality, are aware of and keen to define them.³¹⁹ These strategies have already been created and productively deployed via the licenses in use.³²⁰ Lacking economic gains from further specification by the state, there is no reason to shift organizational authority for partitioning such spaces back to regulators.

Indeed, the TAS rights proposed by De Vany et al. can be criticized for including Time. Matheson and Morris approve of and expand upon the inclusion of this variable in the contours that they recommend, commenting: “Time can be subdivided over a wide range of increments.”³²¹ Yes, it can. It

318. DAVID D. FRIEDMAN, *LAW’S ORDER: WHAT ECONOMICS HAS TO DO WITH LAW AND WHY IT MATTERS* 113 (2000).

319. MATHESON & MORRIS, *supra* note 56, at 14.

320. *Id.* at 15.

321. *Id.* at 9.

can be so subdivided by legal authorities in issuing initial rights and by private rights holders who obtain such rights and then seek to optimize their use. The question at hand is not “What can be defined?” but “What need be defined by the state such that the optimization process can most effectively proceed from there?”

Because there is high complementarity between the productive use of a given frequency space from one minute to the next and from one decade to the next over networks of high fixed cost, infrastructure, and content agreements,³²² bundling rights over Time is the best method. As networks are planned, deployed, operated, and upgraded over long time horizons (measured in decades), continuity of spectrum rights proves efficient. Breaking frequency access rights into tiny nanoseconds³²³ would impose needless transaction costs as firms would engage in market trading to reassemble what public policy has arbitrarily disaggregated. Indeed, re-auctioning license rights over a large time period of ten or twenty years would dissipate value by discouraging ongoing investments in the networks. Should licensees have to re-bid to retain license rights, the regulator would have to provide those licensees with incentives to curtail maintenance or upgrades several years prior to the re-auction date.

U.S. regulators have avoided this cost of variation on time to return investment by defining rights indefinitely. While wireless licenses have explicit terms (usually 10 years), they come with an explicit “expectation of renewal” in perpetuity.³²⁴ Internationally, there has been some confusion on this issue. For instance, New Zealand began auctioning wireless licenses in 1989 and was to re-auction licenses after the 20-year license terms began expiring in 2010.³²⁵ This approach, which would have induced carriers to

322. Online video over mobile broadband introduces new economies of scale to the mobile equation, carried over from traditional media business models. *See generally* JEFF ULIN, *THE BUSINESS OF MEDIA DISTRIBUTION: MONETIZING FILM, TV AND VIDEO CONTENT IN AN ONLINE WORLD* (2010).

323. As has been suggested in Eli Noam, *Spectrum Auctions: Yesterday's Heresy, Today's Orthodoxy, Tomorrow's Anachronism. Taking the Next Step to Open Spectrum Access*, 41 J.L. & ECON. 765, 779 (1998) (on access codes at real-time demand).

324. The FCC has addressed consistent “renewal expectancy” across commercial mobile broadband licenses. In the Matter of Amendment of Parts 1, 22, 24, 27, 74, 80, 90, 95, and 101 To Establish Uniform License Renewal, Discontinuance of Operation, and Geographic Partitioning and Spectrum Disaggregation Rules and Policies for Certain Wireless Radio Services, Notice of Proposed Rulemaking and Order, 25 FCC Rcd. 6996 (2010), *available at* <http://apps.fcc.gov/ecfs/document/view?id=7020520540>.

325. New Zealand Ministry of Economic Development, Radio Spectrum Management, *The Reallocation of Commercial Spectrum Rights at Expiry, Background Information*, <http://www.rsm.govt.nz/cms/policy-and-planning/projects/recently-completed-work/rights-at-expiry/background-information> [hereinafter *Reallocation of Commercial Spectrum Rights*];

inefficiently depreciate their capital, was circumvented with new rules promulgated between 2003 and 2005.³²⁶ The regulator assessed renewal fees instead.³²⁷ All three existing wireless networks retained their spectrum licenses and stayed in business without an auction.³²⁸

The outcome is presumptively efficient. Had these incumbent firms not been value-maximizing rights holders, they would have previously sold to higher bidders. Moreover, the spectrum rights developed for the market in one period complement the supply of similar services offered in the next. In shifting ownership back to the state (as auctioneer) at specified intervals, re-auctions disrupt this complementarity. While other property rights, such as those for intellectual property, are properly time-limited insofar as the creative efforts being protected have already occurred, the productive use of spectrum relies on *continuing* licensee investments in networks and applications that exclusive rights encourage.³²⁹

Of course, where transaction costs are assumed to be zero, such considerations are irrelevant. Rights are reassembled easily, flawlessly, and instantly. But, the exercise of rights definition can only have meaning under the assumption of costly transactions. The basic strategy is to bundle packages of rights such that transactions, following rights assignments, are economized. That does not imply that rights should, going forward, be restricted. At the point where contracting parties internalize transaction costs, economic incentives police the partitioning. Indeed, with flexible-use rules in place, rights holders reveal efficiencies by adjusting to changing demands and opportunities.

Observation of wireless markets yields insights that inform efficient rights definition. Licensees deeply and finely divide rights to use radio spectrum. Very large quantities of voice minutes (to MVNOs) or data traffic (to M2M aggregators) are traded.³³⁰ Tiny increments of time for single voice

Spectrum Auctions, <http://www.rsm.govt.nz/cms/licensees/spectrum-auctions> (citing New Zealand Radiocommunications Act of 1989, Pub. Act 1989 No. 148, <http://www.legislation.govt.nz/act/public/1989/0148/latest/DLM195576.html>).

326. See *Reallocation of Commercial Spectrum Rights*, *supra* note 325. These include policy changes in May 10, 2003, for effect in 2005, to determine price-setting formulae. *Id.*

327. See *Reallocation of Commercial Spectrum Rights*, *supra* note 325.

328. Simon Davies, *New Zealand Networks Renew Radio Spectrum Licenses*, CELLULAR NEWS (Mar. 29, 2012), <http://www.cellular-news.com/story/53714.php>.

329. This confusion about time limiting spectrum rights, analogous to patents, is found in WILLIAM J. BAUMOL & DOROTHY ROBYN, *TOWARD AN EVOLUTIONARY REGIME FOR SPECTRUM GOVERNANCE: LICENSING OR UNRESTRICTED ENTRY?* (2006) (suggesting spectrum license term limits akin to patent term limits).

330. Fifteenth Annual Competition Report, *supra* note 20, at 9672 (with an increase in wholesale subscribers in 2009).

calls or text messages are also offered, although minimum bundles (for example, thirty minutes for a low-cost, pre-paid phone card) are routine, and price-per-unit falls dramatically as larger volumes are purchased.³³¹ Naked spectrum deals are also routine, but only in transactions where spectrum inputs are acquired by networks.³³²

Having aggregated bandwidth through license purchases, mobile carriers then bundle spectrum access with network access. The aggregation of licenses, as well as the bundling of service with spectrum, serves to eliminate “interference” by eliminating borders. There is no evidence that markets would prefer to organize spectrum sharing by dividing parcels with respect to time or directionality of signals. The TAS format developed even prior to the advent of CMRS licenses appears to work well—so well, in fact, that the “T” can be deleted;³³³ just Area and Spectrum need be specified. Basic border spillover rules, including the maximum radiation received at the edges of the contour, are in place for CMRS³³⁴ and can reasonably be used for allocations in the future.

The Matheson and Morris approach is, in our view, properly motivated:

We present a way to express the rights to use spectrum that is not tied to any specific service or technology. It would allow market forces to allocate spectrum such that new radio technologies and applications can be rapidly accommodated with minimal regulatory oversight.³³⁵

The construction focuses single-mindedly on the wrong margins, however:

[T]he physics of radio signal propagation that underlie any spectrum applications and introduces the seven-dimension “electrospace” approach to describing radio signals and the rights

331. *Id.* at 9677 (where price-per-text has dropped 25% for the fifth consecutive year, according to an analyst estimate).

332. *Id.* at 9716 ¶ 62.

333. Regarding the TAS bundle, see De Vany et al., *supra* note 38, at 1517–18. See Section III.F, *supra*, for a discussion on time as a resource appropriately included in an exclusive use bundle managed by a private agent. The time dimension relates to the capacity utilization choice of the agent. Peak and non-peak usage patterns of spectrum is a dynamic feature of wireless service, along with various fixed and variable costs calculated by spectrum licensee.

334. CMRS operates under Part 22 and Part 90, with borders that share with public safety, with “analog AMPS and four different digital technologies, TDMA, CDMA, GSM, and iDEN.” See *IPWG Final Report*, *supra* note 317, at 4.

335. MATHESON & MORRIS, *supra* note 56, at 5.

to emit them. We argue that increased exploitation of these dimensions will be central in improving spectrum capacity.³³⁶

It is absolutely beneficial to more intensively and productively use the dimensions the authors define, and even dimensions too complicated for the authors to define³³⁷—but it is not the case that regulators should be charged with defining these dimensions.³³⁸ Indeed, charging the administrative process with too much precision has reliably led to non-market failure.³³⁹ The costs generated by complex tasks will tend not to deter public sector agents who arguably prefer problems that generate administrative process and delay

336. *Id.*

337. This is their approach to polarization and modulation, which could open new definitional boundaries: “The high degree of necessary cooperation between users of different polarities and modulations suggest that establishing regulatory boundaries between access rights across those dimensions would not be particularly useful, at least with current technology.” MATHESON & MORRIS, *supra* note 56, at 9. Of course, with more advanced technology, liberal license holders would be free to carve out their own new subdivisions, and—were such partitioning to yield new efficiencies—would have incentives to discover and deploy such advances. The analysis is precisely grounded, and its implications are relevant not only to what MATHESON & MORRIS exclude from their definitions, but much of what they include.

338. The Interference Protection Working Group for the FCC Spectrum Policy Task Force Report critiqued interference definitions as technical and idiosyncratic:

Finally, inconsistencies may appear to exist where the FCC’s rules contain different expressions of engineering units that have been derived from different considerations of power, relative power, or electric field. These differences reflect the different approaches to describing and characterizing interference and its impact that are peculiar to the specific situation such as for broadcast or mobile services, for example, D/U and C/I, or dBm and dBmV/m.

IPWG Final Report, *supra* note 317, at 31.

339. *Id.* “The same variability in language on interference that pervades the FCC’s rulemaking proceedings and the resulting rules also gives rise to inconsistent discussion of the impact of interference from a non-technical perspective.” *Id.* The critique focuses on the consistency and standardization of language. Although “clarity” is a requested trait for form, the object is communicative and negotiable simplicity rather than technical finesse. *Id.* at 33. It separates “distinctions between of levels of interference” from information-cost aspects of simple rights definition, for “consistent and appropriate” applicability to “remove some ambiguity” for enforcement. *Id.* “Clarity” is applied to the communicative aspect of rulemaking.

[A]ctual interference rules themselves are not easily identified and isolated in the context of all the rules governing a particular service. For example, for certain services interference may be indirectly governed by specifying minimum separation distances or limitations on transmitter power and antenna height—without ever mentioning the word “interference” or referring to levels of interference.

Id.

controversial, albeit more fiercely competitive, outcomes. Profit-seeking entrepreneurs, however, aggressively search for promising opportunities to open up new lines of commerce, as are facilitated by exploiting additional dimensionality in the supply of services, either directly through vertical integration or via sales to customers.

One clear driver of this attempt to draw such precise rights in law is provided by the idea that cognitive radio, or other devices adept at opportunistically deploying radio spectrum, needs tiny slices of bandwidth to be defined for its use. Dynamic access is the explicit concern of many such definitional exercises.³⁴⁰ But the use of such slivers does not require that regulatory definition. Contracts that provide for dynamic access to licensed spectrum are quite capable of providing such inputs, as observed in the market. Intense use is made of “white spaces” with devices that seek under-utilized frequency spaces, both in the subscriber devices purchased through carriers (or their partner firms), and via virtual networks accessing spectrum via wholesale market transactions.³⁴¹ The organizational patterns of such services, however, often conflict with the use models assumed by analysts, which typically behave more like spot markets than long-term relationships. As noted, the observed business models rely on bundled access (spectrum plus network services) and are marketed in ways that appear distinct from those used in, say, unlicensed spectrum.³⁴² This is asserted to be a bug by advocates of unbundled access models. It is, in fact, a feature. The parties to the observed contracts internalize costs and benefits, including organizational costs and marketing or billing expenses, and so tailor deals that incorporate such factors. Alternative options are available, but rejected in favor of models considered more valuable by the parties directly involved.

In fact, the fragmented “commoditization” of spectrum rights is not a categorically superior form of market organization. How a resource is packaged and traded depends on various cost and demand factors specific to that market.³⁴³ In spectrum, it is quite clear that substantial costs accompany

340. See, e.g., Timothy Forde & Linda Doyle, Towards a Fluid Spectrum Market for Exclusive Usage Rights, 2nd IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks 620 (2007), available at <http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=4221548>; DE VRIES & SIEH, *supra* note 57.

341. See discussion *supra* Section IV.E (discussing mobile operators as spectrum owners and neighbors, with reference to Rath, *supra* note 28, at 529).

342. Rath, *supra* note 28, at 529.

343. Forde & Doyle explain that:

The key characteristic of the clock auction is the relative simplicity of the price discovery phase, i.e. the clock phase. Firstly, its simplicity means that the burden on both the bidders and the auctioneer is low. The bidder

disaggregation, particularly with respect to defining and policing boundary lines. This is the central problem tackled in regulatory proceedings focused on determining interference rules, and yet the fact that the most commonly adopted and effective remedy is to eliminate such fragmentation seems to elude many policy makers and analysts.

However, there is one very promising avenue of inquiry that properly targets mechanisms for assigning rights in a dynamic access environment. Tim Forde and Linda Doyle deal explicitly with the cost of borders, and note that it will tend to become more severe as digital wireless technologies advance; this reasoning represents a needed correction to the conventional wisdom that such systems are categorically making spectrum sharing easier:

In an increasingly fragmented and less-regulated spectrum landscape, there will be an inevitable chafing of rights at the boundaries between the users of neighbouring blocks of spectrum; the spectral activities of one network may impinge on the ability of a neighbouring network to extract maximum value from its exclusively assigned spectrum.

Nonetheless, it is argued that it is possible to reduce, or eliminate, the existence of such externalities through the adjustment and exchange of rights through value extraction-focused bargaining. In essence, this represents a shift from a mindset of unilaterally enforcing rights to one of mutual remediation of any encroachment. Such an approach may . . . [allow] neighbours to find a tolerable level (a balance) of interference.³⁴⁴

Forde and Doyle then present a proposal for how exclusive spectrum rights could be auctioned such that complementary rights—which, in particular, exhibit spillover effects between them—can be discovered and assembled into bundles. This will “address a variety of coexistence issues.”³⁴⁵

G. THE AM RADIO CRITIQUE

Radio engineering expert Charles Jackson, while agreeing that exclusive spectrum rights work well for services such as mobile telephony (cellular and

simply has to sum the cost of any items it wants to package, ascertain that they are within its budget for that resource and make its bid.

Forde & Doyle, *supra* note 41, at 331.

344. Timothy K. Forde & Linda E. Doyle, *Exclusivity, Externalities & Easements: Dynamic Spectrum Access and Coasian Bargaining*, Proceedings of the IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks, Dublin, Ireland, at 303–15 (Apr. 2007), <http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=04221510>.

345. *Id.*

PCS), suggests that they work poorly for certain other technologies.³⁴⁶ In particular, he argues that AM radio, the initial mass-market service that drove policy makers to assemble a spectrum regulatory agency, is problematic. AM radio, a one-way broadcasting service, operates at low frequencies that are not well behaved over geographic distances.³⁴⁷ In particular, they tend to unexpectedly blanket larger areas than they are designed to serve, destroying rival communications beyond their markets. These problems, in Jackson's view, are responsible for the creation of administrative allocation of radio spectrum.³⁴⁸

Jackson is correct that the federal government was involved, from the earliest days of AM radio, in coordinating broadcasting rights. The U.S. Department of Commerce—under the Radio Act of 1912, but with limited authority to assign wireless rights for the purpose of “minimizing interference”³⁴⁹—employed what Senator C.C. Dill called the common law “right of user” rule to define and enforce rights.³⁵⁰ This regime did not give scope to advance “public interest, convenience or necessity,” which is why policymakers and incumbent commercial broadcasters favored new legislation.³⁵¹ Common law rules limited regulatory discretion and, pointedly, would predictably allow for competitive entry. “Public interest” licensing erected barriers that protected incumbents. Absent such common law underpinnings, the legislation itself remains a mystery.³⁵²

346. Charles Jackson, *Limits to Decentralization: the Example of AM Radio Broadcasting Or Was A Common Law Solution to Chaos in the Radio Waves Reasonable in 1927?*, at 30, Telecommunications Policy Research Conference (2005) (copy on file with authors).

347. *Id.* at 24.

348. *Id.* at 31.

349. 62 P.L. 264; Radio Act of 1912, Pub. L. No. 264, 37 Stat. 302; *see also* MARVIN R. BENSMAN, *THE BEGINNING OF BROADCAST REGULATION IN THE TWENTIETH CENTURY* 8–9, 34 (Jefferson, North Carolina: McFarland 2000) (describing the general powers in the Department of Commerce to “follow closely scientific discovery and invention in the principles, methods and instruments of radio communication, all of which are subject to rapid change.”).

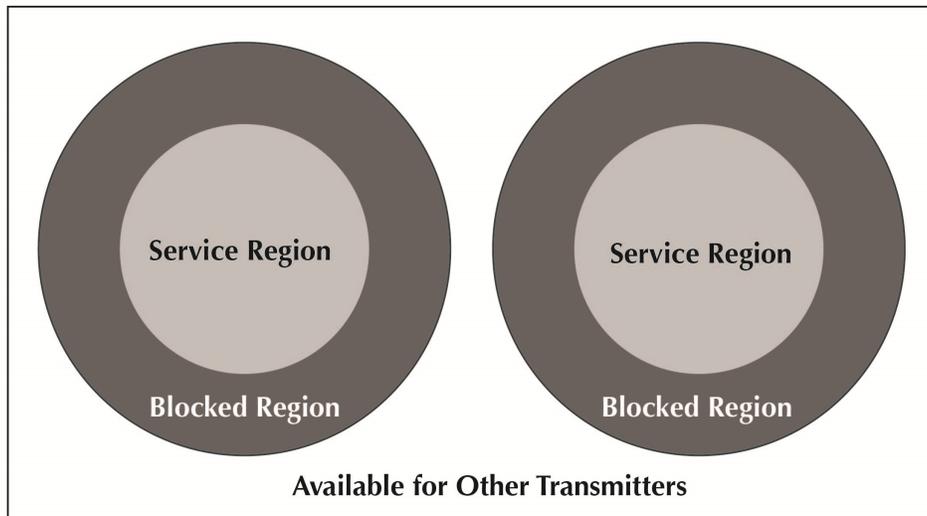
350. CLARENCE C. DILL, *RADIO LAW, PRACTICE AND PROCEDURE* 77–78 (Washington, D.C.: National Law Book Company 1938) (citation omitted).

351. *Id.* (“The most important of these which the radio statute [Radio Act of 1927] sets aside is the principle of acquiring a certain property right by user Congress wrote into the radio law the provision that user should have no effect upon the right of the Commission to provide for the use of any wave length by a new and different person if the public interest would be served thereby.”); *see Wireless Craze, supra* note 30, at 360–73.

352. The traditional explanation of the 1927 Act, repeated in Jackson's paper, is that the Secretary of Commerce had no authority to further enforce broadcasting rules as per the 1926 *Zenith* opinion issued by a federal district court. This is wrong. A 1923 federal district court ruling, in *Intervity*, held that the Commerce Department *did* have the authority to issue

Jackson argues that when wireless contours are clear and easy, private ownership makes sense. When rights are more complex, they are difficult to define, and poor candidates for private ownership. He offers, as an example of the first case, the situation in Figure 6. There, on a given frequency, geographic separation between two broadcast stations yields opportunities for other transmissions. Since the emissions of the two stations do not overlap, a party desiring to use the “white space” (free of interference) would easily be able to obtain the relevant rights, presumably via initial (regulatory) assignment or through negotiation with a spectrum owner.

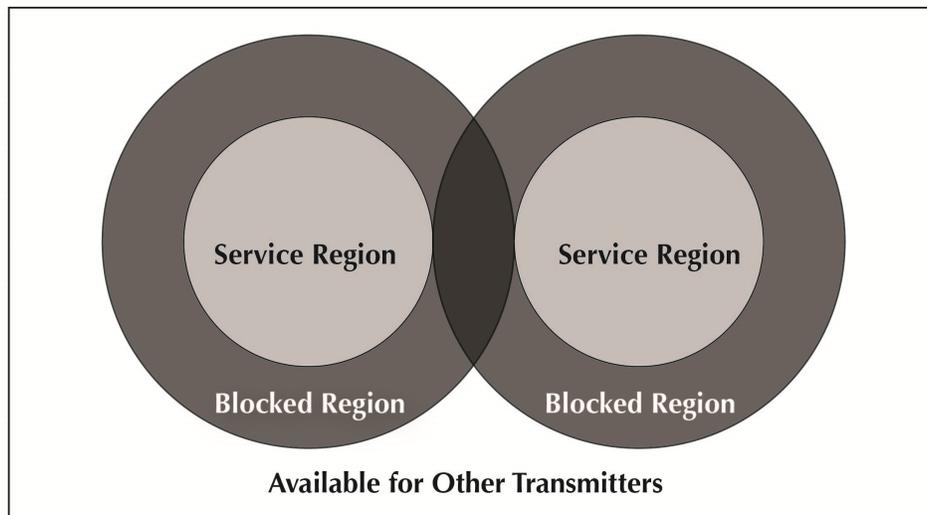
Figure 6: Easy Spectrum Contours for Broadcasters³⁵³



However, this happy equilibrium is said to collapse when the picture changes to reflect the complexities of AM broadcasting. Then service contours overlap, as in Figure 7:

licenses so as to restrict conflicts. It was the political choice of Commerce Secretary Herbert Hoover—curiously rejecting the verdict yielding his agency more power—that led to the “period of the breakdown of the law.” To remedy this, all a statute need do was clarify that the earlier opinion (*Intervity*) had been correct and that the Government’s power to enforce rules to “minimize interference” were intact. The legislation that resulted rejected this limited enforcement patch, creating a wholly new regulatory structure. As explained by the participants, including Hoover and Senator Dill, the goal was not to restore interference rules but to extend political control over spectrum. Major commercial broadcasters were themselves in favor of this shift in property rights, as they realized rents from enhanced barriers to entry. This is explained in much greater detail in my articles, Thomas, W. Hazlett, *Rationality of Regulation, Physical Scarcity, Rent Seeking and the First Amendment*, 97 COLUM. L. REV. 905 (1997), and *Wireless Craze*, *supra* note 30.

353. Figure 6 reproduced from Jackson, *supra* note 346, at 14 fig.2.

Figure 7: Not So Easy Spectrum Contours for Broadcasters³⁵⁴

Jackson argues that in the region where the interference of the two broadcasters overlaps, a difficult contracting problem results. Not one, but two parties must agree in order for an entrant to gain the right to reuse the spectrum space. In his words, “this splitting of property rights is . . . maybe not fatal, but certainly debilitating.”³⁵⁵ Jackson suggests that administrative allocation is therefore superior to private property in this situation.³⁵⁶

Jackson’s analysis is lacking in this regard. Coordinating the two activities is a challenge that exists for state agents as well as for private resource owners.³⁵⁷ Stochastic changes add to complexity, but do not change the calculus of the institutional choices for rights definition. The history of the FCC shows that administrative allocation is generally an expensive form of dispute resolution.³⁵⁸ Typically, administrative allocation unduly constrains wireless applications, an overly conservative approach that lowers social welfare. Indeed, radio and television stations have been routinely under-allocated, lowering consumer welfare.³⁵⁹ The results are general.³⁶⁰ In UHF

354. Figure 7 reproduced from *id.* at 14 fig.3.

355. Jackson, *supra* note 346, at 14.

356. *Id.*

357. We here ignore the availability of standard contracting strategies that would remedy hold-out problems, including contingent contracts (payments to one station are made if and only if the other party agrees to the same terms) and/or the use of rival bids across multiple spectrum spaces, inducing a reverse auction.

358. See discussion *supra* Section I.A; NBP 2010, *supra* note 3 (citing the six- to ten-year reallocation process).

359. Thomas W. Hazlett & Bruno Viani, *Legislators v. Regulators: The Case of Low Power FM Radio*, 7 *BUS. & POLITICS* 1, 9 (2005).

TV, there are five taboo (unused) channels for each allocated channel per market;³⁶¹ in FM radio, three taboo channels are set aside.³⁶² This has proven highly wasteful.³⁶³ It is precisely the sort of interference control that one would associate with administrative allocation—it reduces total output but the result looks harmonious, as the most visible conflicts are reduced. The wasteful outcome differs with private spectrum ownership rights, however, as the licensees—when permitted to use spectrum for whatever they determine is its highest and best use—would increase profits by controlling interference in a lower cost manner than regulation. And the issue is not whether private rights holders will do this imperfectly, but whether they can outrun the bear—the state property alternative.

In asking whether administrative allocation or private rights holders provide the best means to use spectrum, Jackson's example is better than he imagines. In positing two independently owned stations with overlapping (deterministically or stochastically) emission contours, he addresses rights definition as the FCC often does. The agency creates fragmentary spectrum use rights, and disperses them, even while they are highly complementary.³⁶⁴ That makes the creation of efficient wireless operations that require a high degree of coordination needlessly difficult. The FCC then attempts to supply its own form of coordination by further “command and control” allocation measures, on the grounds that market failure occurs due to high transaction costs.

In fact, the “debilitating” problem would be avoided by an efficient assignment of legal rights: joint ownership of the frequency space within the box. Placing both station contours within one ownership bundle does not magically eliminate radio interference, but does reliably resolve the proffered

360. Hazlett et al., *Optimal Abolition*, *supra* note 250, at 103–04.

361. REPORT OF THE SPECTRUM EFFICIENCY WORKING GROUP 24 (2002) (citing S. Merrill Weiss, Address at the Public Workshop on Spectrum Efficiency (Aug. 5, 2002)), available at http://transition.fcc.gov/sptf/files/SEWGFfinalReport_1.pdf; Michael Marcus, *Comments re: Ex Parte Comments of Sports Technology Alliance*, ET Docket No. 04-186, at 2 (May 5, 2008), available at <http://fjallfoss.fcc.gov/ecfs2/document/view.action?id=6520008054>.

362. Hazlett & Viani, *supra* note 359, at 4–5.

363. Dale N. Hatfield, *The Challenge of Increasing Broadband Capacity*, 63 FED. COMM. L.J. 43, 57 (2010) (“The lack of perfect transmitters and receivers inevitably results in the loss of some capacity between bands due to the need to, for example, provide a buffer or guard band between the two bands to supply the necessary isolation and thereby reduce the associated interference to an acceptable level. Excessively fragmenting the spectrum among different bands reduces the overall technical efficiency of spectrum utilization for this reason.”).

364. See Porter et al., *supra* note 144 (discussing the efficiencies of combinatorial auctions).

disputes. Spectrum sharing is optimized because marginal costs and marginal benefits flow to the (integrated) rights holder. This ownership pattern would naturally emerge in the absence of the transaction costs imposed by inefficient rights distributions, as both licenses are together worth more due to the coordination costs that Jackson wrongly associates with market failure.³⁶⁵

Forde and Doyle explicitly feature such rights aggregation as a central aim of an efficient regulatory system. The model they develop would allow markets to avoid the “externality” problems that Jackson considers. They note that there are many ways that rival radio stations could peacefully co-exist: investing in better transmitters or receivers; moving or repositioning certain equipment; splitting time; altering power levels.³⁶⁶ But a general fallback will be that “certain frequencies will end up being sold in combinations . . . and never in individual blocks.”³⁶⁷ With integrated control, managers will seek to maximize the value of both broadcasting opportunities jointly, a path that will logically lead to the most efficient resolution of the interference spillover.³⁶⁸ The importance of integrated ownership is that the profit-seeking licensee will search for the compromise yielding the highest output, and implement that solution without being deterred by the need to transact with other rights-holders.

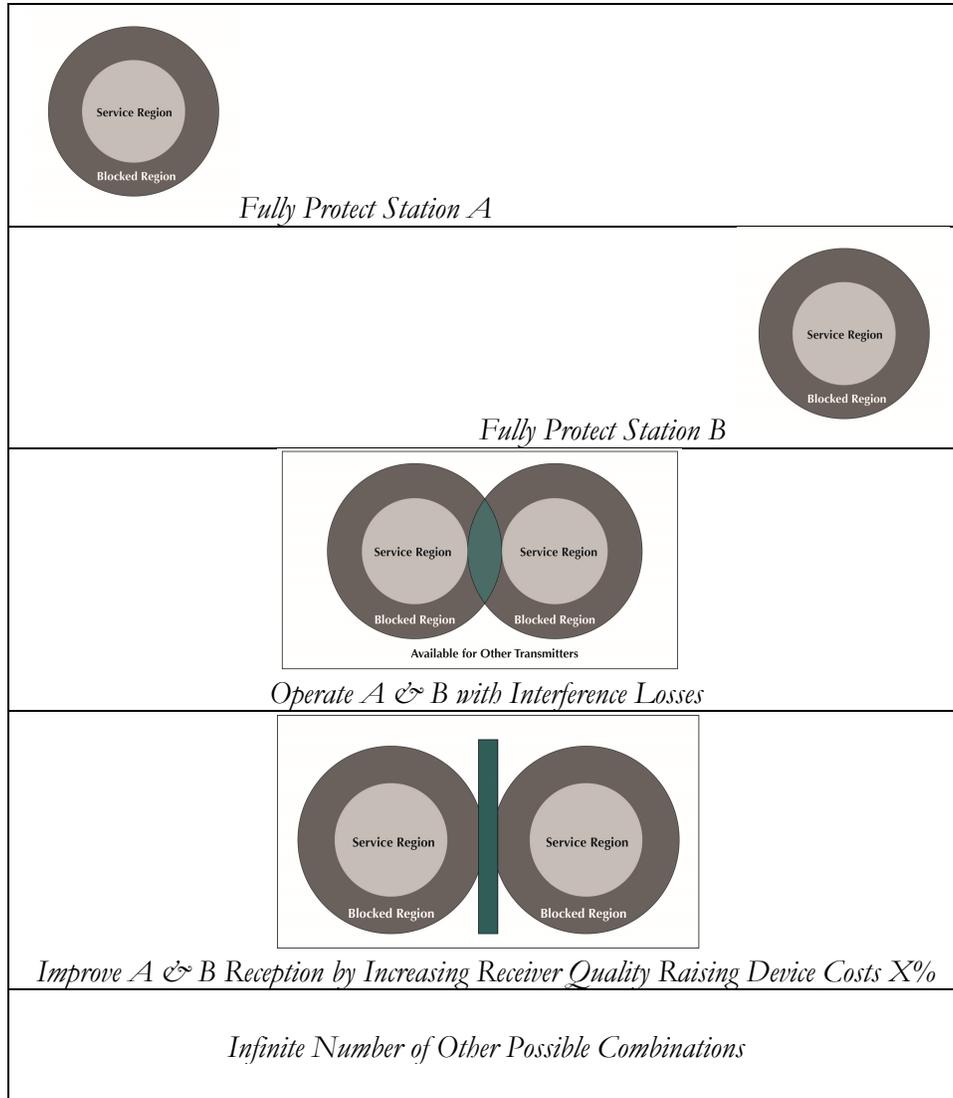
365. Jackson, *supra* note 346, at 14.

366. Forde & Doyle, *supra* note 41, at 331.

367. Forde & Doyle, *supra* note 340, at 628.

368. *See infra* Figure 8.

Figure 8: Options When Interfering Stations Are Under Integrated Ownership



H. LIABILITY RULES VERSUS PROPERTY RULES

Pierre de Vries and Kaleb Sieh also argue for “unambiguous” determination of spectrum use rights, largely ignoring the costs of this regulatory enterprise and underestimating (or ignoring) the degree of utility of existing liberal license templates assigned to responsible economic agents.³⁶⁹ This is a key omission, particularly as they take up the question as

369. DE VRIES & SIEH, *supra* note 57.

to whether it is best to use property rules or liability rules in defining spectrum uses. They conclude that, because the literature is undecided on which set of rules is optimal, it is best to be agnostic on this question.³⁷⁰ This is a seriously flawed conclusion.

There are many dimensions in which to see property ownership as a useful social institution.³⁷¹ One is that owners take care to conserve, nurture, and grow their resources, taking appropriate care—whenever it is worth the cost—to protect or improve it.³⁷² Exclusive control, as afforded by property ownership, does not mean that resources are reserved for “exclusive use,” but that the owners will attempt to maximize their value, using them to do whatever yields the highest return. A frequency space worth \$1 million as an input into a particular wireless service is not, alternatively, used to provide \$10,000 worth of value as a guard band to a technically lagging neighbor.

Liability rules are substitutes for property rules when the latter are prohibitively expensive to use. The standard case regards auto collisions. Were transaction costs zero, people who might be involved in traffic accidents could contract, agreeing to terms regarding payments (perhaps specified by an arbitration service) following any damage-inflicting incident. Yet, transaction costs are not zero in the real world. In fact, the costs of contracting with all potential parties to a traffic accident are prohibitive.

Hence, a second best solution is called for. That is where liability rules come in. If Party A drives recklessly, causing an accident and destroying the automobile of Party B, we are already beyond property rules. Under the ownership system, A could have taken (or destroyed) B’s auto, but only by paying a price acceptable to B to buy the car. Because the accident has occurred, and because A and B have not contracted over the price of the car, the damages must now be set by law. This is not a perfect system, as the value of the car to B must be calculated by the court (instead of established in a mutually beneficial trade), but it does lead to the best possible outcome under the circumstances.

For voluntary transactions where bargaining can occur, however, liability rules are socially useful. The process ensures that resources are used in their

370. *Id.*

371. For a general overview of the law & economics literature on property rights, see RICHARD A. POSNER, *ECONOMIC ANALYSIS OF LAW* 218–19 (New York: Aspen 8th ed. 2011); *see also* PROPERTY RIGHTS: COOPERATION, CONFLICT, AND LAW 282, 287 (Terry L. Anderson & Fred S. McChesney eds., Princeton University Press 2003). For the role of technology in resource use, *see generally* THE TECHNOLOGY OF PROPERTY RIGHTS (Terry L. Anderson & Peter J. Hill, eds., 2001).

372. *See* sources cited *supra* note 371.

highest valued employments. That is precisely what we should aim to achieve in radio spectrum usage, where the information known to regulators is sparse, and where incentives to discover or implement optimal employments is relatively weak. Liability rules vest regulators or courts to make valuation choices, and so continue the administrative allocation process, albeit perhaps in a different format. Indeed, some FCC allocations do apply liability rules—in particular, unlicensed “Part 15” regulations that authorize certain low power radio devices to operate, but under the condition that they accept all interference from other devices (whose users are relieved of liability) and to refrain from interfering with any licensed service (imposing full liability on the unlicensed device user).³⁷³ This approach is used because of the non-exclusive spectrum rights associated with unlicensed operation; property rules are not an option here. But, having assigned exclusive rights to a liberal licensee, the process by which the rights holder bargains with others is extremely useful, revealing what other opportunities are available and how they stack up against the costs they impose. To allow others to use these spectrum spaces without authorization, forcing the rights holder to pursue liability *ex post*, not only imposes costly litigation but puts the determination of valuation back in the public sector. This surrenders the optimizing properties gained in the assignment of liberal rights, and turns choices between alternatives back into a legal-administrative question.

Hence, the following agnostic conclusion rendered by De Vries and Sieh is incorrect:

There is an extensive literature on whether liability or property remedies are preferable. The various conclusions depend in no small part on the assumptions made in each case and there seems to be no consensus. Some regard liability remedies as the most suitable in externality cases, though different assumptions can lead to the conclusion that injunctive relief [won under property rules] is no worse, and may be better, than damages.

More work is therefore needed before making recommendations on whether damages or injunctions are the preferable remedy in radio regulation³⁷⁴

Indeed, conclusions change with assumptions on when liability or property rules are most appropriate. Summarizing the classic contribution in the field, an essay by Guido Calabresi and Douglas Melamed published in 1972, Keith Hylton writes: “Property rules are generally preferable Liability rules are

373. See 47 C.F.R. pt. 15 (2011).

374. DE VRIES & SIEH, *supra* note 57, at 8.

used instead of property rules largely because high transaction costs make it infeasible to have property rules”³⁷⁵ So the differentiation, rather than rendering a choice impossible, facilitates straightforward application to the issue at hand. The conditions prevailing in wireless markets demonstrate that bandwidth stewards are extremely important in coordinating efficient economic activity. Moreover, transactions between networks, suppliers, and customers provide robust market outcomes. Thousands of licenses have traded; scores of virtual networks have formed; hundreds of carrier-to-carrier roaming agreements have been executed; millions of customers have subscribed. All of these contracts have transferred spectrum rights on the property model *without* using liability rules.

I. OPTIMAL SPECTRUM RIGHTS DEFINED

The goals in defining exclusive spectrum rights were presented by De Vany et al in 1969 and have yet to change, despite gains in information transmitting capacity of radio transmitters and receivers.

Ideally, the boundaries and field-strength limits of the initial TAS packages would be economically optimal—that is, they would be perfectly suited (in each dimension) to the use that would result from existing market and technological conditions. Initial optimality would save subsequent exchange costs that otherwise would be incurred in restructuring rights into optimal configurations. However, the information required for achieving initial optimality is not now available; obtaining the information could cost more than the resulting increment in value, and in any case this cost is likely to exceed the cost of achieving optimality through an initial round of exchanges. Some skillful guesses on original packaging should therefore be made, leaving the correction of mistakes to subsequent rights exchanges in the marketplace. In addition, even if initially optimal TAS packages could be defined, they would not remain optimal over time.³⁷⁶

This remains an excellent guide to policy. Hence defining contours can be made simple by reference to the CMRS rules already in place in the United States, currently a leading template for the use of spectrum integrated with mobile handsets and mobile services of broadband connectivity. Other regulators have their own templates that may also prove useful. Australia, for

375. Keith N. Hylton, *Calabresi's Influence on Law and Economics*, in PIONEERS IN LAW AND ECONOMICS 224, 230 (Lloyd R. Cohen & Joshua D. Wright, eds. 2009) (summarizing Guido Calabresi & A. Douglas Melamed, *Property Rules, Liability Rules, and Inalienability: One View of the Cathedral*, 85 HARV. L. REV. 1089 (1972)).

376. De Vany et al., *supra* note 38, at 1517.

example, defines “spectrum cubes” that have flexible use and can be easily aggregated in initial assignments.³⁷⁷

As noted in Section IV.F, *supra*, the TAS packages offered by De Vany et al. are probably over-defined; time is superfluous. The rights to use radio spectrum are efficiently defined as *perpetual* in duration.³⁷⁸ The remaining dimensions, area and frequency, are obviously needed, however. The question that then arises relates to the definition of de minimus interference. Given that radio signals attenuate over geographic distance and frequency space, some detectable radio emissions will naturally leak into various other parcels. CMRS rules define these levels. There are small variations from country to country. Charles Jackson notes that the United States maintains a “flexibly enforced” boundary limit of -47 dBW.³⁷⁹ New Zealand maintains a slightly quieter standard: -52 dBW.³⁸⁰ In practice, they are probably indistinguishable aside from major step functions in available technology limits. Additional specification is likely superfluous, although disputes will inevitably arise. What can be added to the process is, as Jackson notes, a binding arbitration system—deployed by New Zealand—that can expeditiously resolve such conflicts.³⁸¹

J. AUCTIONS AS INTERFERENCE CONTROL MECHANISMS

Rights assignments are central to the control of harmful interference. This seems counterintuitive to those familiar with the standard bifurcation of wireless policy: spectrum allocation establishes use rights; license awards then distribute those use rights.³⁸² In this scenario, rules limiting conflicts are the sole purview of the allocation process. License awards, whether by beauty contest, lottery, or auction,³⁸³ simply transfer those rights among potential claimants.

In fact, the manner in which such transfers occur is key in determining harmful interference. It does overstate the matter to make the bold claim that most—indeed, an *overwhelming majority*—of radio interference disputes are

377. Hazlett, *Wireless License Values*, *supra* note 72, at 577 tbl.4 (describing “Spectrum Trading Units” of Australia where the regulator identifies blocks of such units).

378. All rights can be taken in a condemnation by the state, which involves the payment of just compensation. That would presumably apply with radio spectrum rights.

379. Jackson, *supra* note 312, at 10.

380. *Id.*

381. *Id.* at 14.

382. See, e.g., John O. Robinson, *Spectrum Management Policy in the United States: An Historical Account* (FCC, Office of Plans & Policies, Working Paper No. 15, Apr. 1985).

383. Or, for that matter, via unlicensed use rights, where devices are licensed by the regulator.

resolved not by formal regulations or market-based spectrum contracts, but by merger. By integrating usage rights, firms internalize conflicts and deploy value-maximizing spectrum sharing rules in the natural matter of their network management. Nationwide mobile networks aggregate use rights across time, geography, and spectrum, and regulators “re-band” allocations to replace “interleaving” with contiguous blocks (as in the Nextel-public safety “spectrum swap”).³⁸⁴ Technical specificity often causes such border conflicts. Eliminating such borders via integration makes them disappear.

The merger remedy is not costless. Management expense typically mounts with scale and complexity; license aggregation tends to produce costs that at least partly offset the gains from creating a more efficient organization of integration. Institutions to accommodate this balancing are thereby essential to the control of harmful interference. As there is not a fixed, obvious level of aggregation, or categorically dominant structure of rights partitioning (sometimes called spectrum sharing), the challenge is to discover, implement, and then manage wireless networks over time, as circumstances change. This drives the argument for spectrum markets.

“Given a combination of De Vany-like flexible rights,” wrote Timothy Forde and Linda Doyle in a 2007 IEEE conference paper, “DSA [dynamic spectrum access] technology and rational profit-seeking firms competing in a market, there are a number of ways in which these three elements could combine to deal with presence of interference-related externalities.”³⁸⁵ This is a modern restatement of the 1962 Coase, Meckling, and Minasian explanation of how economic incentives and property rights enable the discovery of “optimal combinations.”³⁸⁶ The efficiency logic of the statement works whatever the state of technology options. Indeed, the availability of robust market forces—flexible spectrum use rights held, in efficient bundles, by for-profit rights holders—will itself facilitate the introduction of innovative services and advanced technologies that the process of “creative destruction” uniquely produces. This process relies on these key ingredients:

- (a) broad, exclusive rights to control spectrum access;
- (b) for-profit licensees with strong incentives to discover the “optimal combinations,” internalizing costs and benefits; and
- (c) efficient bundles of rights, assembled initially by regulators, in initial auctions, or via secondary markets.

384. See *supra* note 155 and accompanying text (discussing the Nextel spectrum swap).

385. Forde & Doyle, *supra* note 344, at 311.

386. COASE, MECKLING & MINASIAN, *supra* note 37, at 76–77.

The prescription includes the freedom to aggregate or disaggregate exclusive rights, and is strongly supported by competitive bidding for licenses in auctions that include combinatorial bids. Forde and Doyle appropriately focus on the legal and regulatory rules to reduce barriers to efficient market organization when initially issuing rights.³⁸⁷ They model a market in which dynamic spectrum access—short, numerous bursts—can be accommodated on licensed bands with exclusive rights, with rights aggregation serving to coordinate potentially conflicting activities.³⁸⁸ They argue that to counter the “inevitable chafing of rights at the boundaries . . . it is possible to reduce, or eliminate, the existence of such externalities through the adjustment and exchange of rights.”³⁸⁹ This leads them to construct auction mechanisms that more easily permit rights to be bundled via the bids of efficiency-motivated agents.³⁹⁰ Numerous such combinatorial auction mechanisms have been developed in wireless license auctions and in other bidding problems.³⁹¹ They form an integral part of an efficient policy response to controlling harmful interference.

This is not generally appreciated. For instance, De Vries and Sieh, concerned with creating “unambiguous” spectrum use rights to counter disputes over airwave interference, propose that FCC rules define probabilistic spillovers.³⁹² These occur when borders of TAS rights are encroached due to stochastic events, such as atmospheric conditions. But the cost of increasing the complexity of the rights definition process does not influence their policy choice. The alternative to creating more regulatory process to insert, *ex ante*, probability distributions into defined rights, is to delegate such issues to market packaging. The most damaging and contentious borders, especially those complicated by stochastic natural effects, will tend to be integrated. Less problematic borders—and many fewer of them, to the degree that probabilistic spillovers are an issue—will exist. This general phenomenon has been seen in wireless markets, both with the aggregation of liberal licenses and the gridlock exhibited throughout the economy with interleaving involving illiberal (or government-owned) licenses.³⁹³ A major theme of business structure research is the extent to which firms integrate to overcome the costs of using “the price

387. Forde & Doyle, *supra* note 344, at 311.

388. *Id.*

389. *Id.* at 303.

390. *Id.*

391. Porter et al., *supra* note 144; *see also* COMBINATORIAL AUCTIONS, *supra* note 296.

392. DE VRIES & SIEH, *supra* note 57.

393. Hazlett & Leo, *supra* note 97.

mechanism.”³⁹⁴ The central logic is that companies systemically attempt to build efficient ecosystems around their core businesses. This drives them to coordinate production with complex complementarities *internally*, leaving other firms—operating on fairly simply defined modular interfaces—to provide other inputs or complements *externally*. The competitive process is continually searching for new structural efficiencies, which redefine these product borders.³⁹⁵

K. OVERLAYS

The prime task, however, is not to define rights more carefully, but to devise allocation strategies that can allow rational reconfiguration of spectrum usage. Forde and Doyle have made repeated efforts to do that, focusing on devising auction mechanisms that steer initial rights assignments to the parties that can most productively use them.³⁹⁶ Their approach focuses on the issuance of liberal, flexible-use rights, and efficient distribution using formats like combinatorial auctions.³⁹⁷ Their approach is to be applauded, as it devotes attention to the relevant margins, delegating the most complex decision-making to competitive, for-profit enterprises.³⁹⁸

Regulators must be able to define spectrum rights in the parsimonious way that auctions require, however. In situations where incumbent licensees are already in place, even the availability of abundant unused spectrum resources presents difficulties that have proven insurmountable for administrative allocation. For instance, the TV Band, while little used in the United States since its inception, continues to block very productive opportunities because regulators cannot effectively delineate between the rights of current broadcasters and the (potentially more valuable) rights of new TV Band users.³⁹⁹

One set of options is being pursued by the Federal Communications Commission in the form of “incentive auctions.” Initially proposed in 2002 by FCC experts Evan Kwerel and John Williams, the basic idea is to

394. The term was used in Ronald Coase, *The Nature of the Firm*, 4 *ECONOMICA* 386–405 (1937). In this instance, a firm’s cost of using the market—what happens when adjacent bandwidth is controlled by regulators and/or rivals, rather than by the firm itself—can be avoided by buying (liberal) licenses.

395. See, e.g., Carliss Y. Baldwin & Kim B. Clark, *Managing in an Age of Modularity*, 75 *HARV. BUS. REV.* 84 (1997).

396. Forde & Doyle, *supra* note 41.

397. *Id.* at 329.

398. *Id.* (recognizing the computational complexity of an auction mechanism responsive to dynamic adjustments of cognitive participants).

399. Hazlett, *Tragedy TV*, *supra* note 76, at 84.

reallocate some of the television band to mobile services via a pair of auctions.⁴⁰⁰ The first would be a reverse auction in which existing TV station owners would state offer prices—what they would accept in payment to exit. With this information, the FCC then proceeds to remove the lowest-priced stations (as per the offers stated in the reverse auction), and to relocate the remaining stations, packing them onto new channels. This process is designed to leave contiguous television spectrum vacant (clear of TV broadcasting), making up to 120 MHz of the 294 MHz available for reallocation to CMRS licenses. These are then sold in a standard forward auction.⁴⁰¹

The National Broadband Plan, which found that making additional bandwidth available for wireless networks was essential for the U.S. economy, made this policy strategy a key initiative for the FCC with its Report in March 2010.⁴⁰² The first hurdle was to obtain a new federal statute empowering the FCC to hold the reverse auction, and to designate the bids from the forward auction to compensate TV stations whose bids were accepted. Congress passed legislation enabling this process in February 2012.⁴⁰³ The agency is expected to issue its first public notices, with initial discussion of how it intends to structure the upcoming auctions, in late 2012.⁴⁰⁴ Actual reallocation of TV Band spectrum is not expected to occur before 2017.⁴⁰⁵ The legislation gives the FCC until 2022 to conduct an auction.⁴⁰⁶

An alternative proposal to reverse auctions was the use of an “overlay” auction.⁴⁰⁷ While it was rejected by the FCC,⁴⁰⁸ the Commission, particularly

400. See Kwerel & Williams, *supra* note 42, at 33.

401. For more on issues related to a reverse auction and forward auction stages of the incentive auction, see HAZLETT ET AL., INCENTIVE AUCTIONS PAPER, *supra* note 49.

402. NBP 2010, *supra* note 3, at 75.

403. MIDDLE CLASS TAX RELIEF ACT AND JOB CREATION ACT of 2012, P.L. 112-96, 126 STAT. 156, 47 U.S.C. § 1451 (authorizing the FCC to conduct incentive auctions with forward and reserve auction components for broadcast TV spectrum).

404. *Id.*; see generally FCC, *Incentive Auctions* <http://www.fcc.gov/topic/incentive-auctions> (last visited Oct. 18, 2012).

405. FCC, *Incentive Auctions*, *supra* note 404; 2012 Incentive Auction NPRM, *supra* note 49. Reply comments to the Notice for Proposed Rulemaking are due February 2013, when the FCC will release a Report & Order estimated in 2014, with implementation by 2017. 2012 Incentive Auction NPRM, *supra* note 49.

406. 2012 Incentive Auction NPRM, *supra* note 49, at 12,371.

407. Hazlett, *A Proposal for an Overlay Auction*, *supra* note 50. A recent proposal in the European Union advocates a similar policy approach. PIER LUIGI PARCU ET AL., AUTHORIZED SHARED ACCESS: AN INNOVATIVE MODEL OF PRO-COMPETITIVE SPECTRUM MANAGEMENT (Studio Economico Parcu & Associates 2012), available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2174518.

in the use of PCS and Advanced Wireless Services (“AWS”) licenses, has employed this method previously.⁴⁰⁹ In the use of PCS, regulatory gridlock deterred 2G digital phone services in the United States because, while the FCC had selected a band to be used for the new service, the chosen band was littered with about 4,500 microwave licensees.⁴¹⁰ These operations, which provided point-to-point communications for public utilities, off-shore oil drilling rigs, railroads and other users, consumed very little spectrum.⁴¹¹ But they provided important services related to public safety, and thus any regulatory action to relocate such communications (either by substituting other spectrum for their wireless links, or by switching to fiber optic lines) was greeted with intense opposition.⁴¹² The allegation was that relocation would directly impact operations involving “safety of life and property” if the incumbent uses were compromised in any way.⁴¹³

The solution, after years of regulatory stasis, was to assign PCS overlay rights by competitive bidding,⁴¹⁴ providing new licensees with flexible use of the allocated spectrum (30 MHz for A, B, and C block licenses; 10 MHz for E, F, and D; 120 MHz in total).⁴¹⁵ But existing microwave users were vested,⁴¹⁶ and could continue operating as they had been. Meanwhile, the new

408. While the NBP explicitly endorsed the “incentive auction” policy as its first choice, it recommended overlays as a possible back-up plan: “Explore alternatives—including changes in broadcast technical architecture, an overlay license auction, or more extensive channel sharing—in the event the preceding recommendations do not yield a significant amount of spectrum.” NBP 2010, *supra* note 3, at 88.

409. Overlay auctions have also been considered in the United Kingdom. *See Cave, supra* note 251, at 81–82.

410. *See* 1993 Broadband PCS Second R&O, *supra* note 104; *see also* Cramton et al., *supra* note 122.

411. *See* 1993 Broadband PCS Second R&O, *supra* note 104.

412. Cramton et al., *supra* note 122, at 17.

413. *See, e.g.,* Reply Comments of Association of Public-Safety Communications Officials-International, Inc., Amendment to the Commission’s Rules Regarding a Plan for Sharing the Costs of Microwave Relocation, WT Docket No. 95-157, at 7 (Jan. 1, 1996), <http://apps.fcc.gov/ecfs//document/view.action?id=1537490001> (“Nor should public safety incumbents ever be exposed to the potential of being forced to relinquish critical communications frequencies without assurances of receiving comparable facilities at no cost to taxpayers.”).

414. *See* Cramton et al., *supra* note 122, at 660–61.

415. *See generally* 47 C.F.R. 24.701 (2000) (broadband PCS subject to competitive bidding); In the Matter of Amendment of Part 1 of the Commission’s Rules—Competitive Bidding Procedures, Order on Reconsideration of the Third Report & Order, Fifth Report & Order, and Fourth Further Notice of Proposed Rulemaking, 15 FCC Rcd. 15,293, 15,331 (2000), *available at* <http://www.fcc.gov/Bureaus/Wireless/Orders/2000/fcc00274.doc>.

416. *See* 1993 Broadband PCS Second R&O, *supra* note 104, at 7706.

PCS licensees were obligated to avoid creating harmful interference that would damage their operations.⁴¹⁷

This created a new dynamic. Instead of maintaining a political position that no change was possible, incumbents now faced real opportunity costs. If they relocated, they would provide valuable new bandwidth to the PCS licensee authorized to use it; maintaining their intransigence would no longer be free. On the other hand, the overlay licensees—who had secondary rights to use the spectrum allocated through the incumbents' licenses—also had straightforward economic choices. Either the overlay licensees could work around the incumbents' transmissions, or else buy them out. Either path was costly, and overlay licensees naturally preferred the least costly path. Hence, both sides had proper incentives to find efficient solutions.

The fact that many of the microwave incumbents were non-profit or rate-of-return regulated enterprises complicated the negotiations between incumbents and overlay licensees, because the incumbents were less motivated by cash payments than businesses seeking to maximize profits.⁴¹⁸ To offset this lack of economic motivation, regulators alertly supplied additional incentives, mandating good faith negotiations for relocation, placing responsibility on overlay owners to pay for incumbents' moving costs, and then imposing time limits.⁴¹⁹ A similar set of overlays and circumstances was obtained in 2006 with the auctioning of AWS licenses. In this case, incumbent wireless users were largely government agencies,⁴²⁰ and hence, market deals were simulated. Overlay (AWS) licensees were mandated to pay the full costs of relocating incumbents, and the FCC adjudicated disputes over terms.⁴²¹

In both instances, the process worked—not perfectly, but ostensibly better than administrative allocation. Without overlay owners initiating deals and paying relocation costs, it has proven virtually impossible for the FCC to reallocate large numbers of incumbent wireless users to clear broad swaths of spectrum.⁴²² Reallocation has costs of its own, including those associated with regulatory gridlock.

417. 1993 Broadband PCS Second R&O, *supra* note 104, at 7755, 7767.

418. *See generally* 1993 Broadband PCS Second R&O, *supra* note 104.

419. *See* 2004 U-PCS NPRM, *supra* note 102, at 5126–27 (describing UTAM, Inc. case, where a separate entity facilitated negotiation among parties).

420. *See* NTIA, Relocation of Federal Systems in the 1710–1755 MHz Frequency Band, *infra* note 435, at 2.

421. *See id.*

422. NBP 2010, *supra* note 3, at 90 (recommending Congressional authority for incentive auctions). The FCC could not otherwise reallocate the TV band without a legislative grant of authority. *See* 2012 Incentive Auction NPRM, *supra* note 49 (describing

TV Band overlays could improve upon both past and current spectrum allocation efforts. By enabling private, for-profit enterprises with the rights to claim the residual value of whatever new opportunities they create, they will seek to develop efficient, cooperative bargains with TV broadcasters. In securing alternative video distribution platforms for their content, finding innovative ways to compress broadcast TV signals, and accommodating more intense sharing of channels, overlay owners would stand to reap very substantial rewards—as much as \$50 billion or more—by reducing low-valued uses of the TV Band in favor of higher-valued services.⁴²³ Of course, the government could capture much of this gain in overlay license auctions, but the social welfare gains are likely to be many times the magnitude of the producers' surplus.⁴²⁴

The FCC chose a different policy approach based on their view of transaction costs:

Incentive auctions present a more efficient alternative to the FCC's overlay auction authority, in which the FCC auctions encumbered overlay licenses and lets the new overlay licensees negotiate with incumbents to clear spectrum. These piecemeal voluntary negotiations between new licensees and incumbents introduce delays as well as high transaction costs as new licensees contend with holdouts and other bargaining problems. Anticipating these delays and negotiating costs, bidders typically pay significantly less for encumbered spectrum. The value of spectrum that must be cleared through such a voluntary process is reduced even more by uncertainty about the final cost of clearing.⁴²⁵

However, the FCC's argument is both incomplete and not compelling. Private market negotiations entail costs just like government reallocation efforts. The latter are not considered in any serious way, an approach that smacks of the Nirvana Fallacy.⁴²⁶

the incentive auction NPRM); MIDDLE CLASS TAX RELIEF ACT AND JOB CREATION ACT of 2012, P.L. 112-96, 126 Stat. 156, 47 U.S.C. § 1451.

423. *Tragedy TV*, *supra* note 76, at 112 n.154 (“The Federal Communications Commission estimates the spectrum’s value at between \$20 billion and \$132 billion.”).

424. See discussion *supra* Section III.E (discussing consumer surplus).

425. *NBP 2010*, *supra* note 3, at 82.

426. Demsetz writes:

The view that now pervades much public policy economics implicitly presents the relevant choice as between an ideal norm and an existing “imperfect” institutional arrangement. This *nirvana* approach differs considerably from a *comparative institution* approach in which the relevant choice is between alternative real institutional arrangements. In practice, those who adopt the nirvana viewpoint seek to discover discrepancies

Two obvious heavy burdens are already visible, even if they have not yet become apparent to the FCC. First, the incentive auction program is regulation-intensive. A two-year delay⁴²⁷ was incurred as new authority was sought from Congress following the issuance of the National Broadband Plan. The eventual authorization includes provisions that the FCC did not ask for and did not want,⁴²⁸ as it will reduce the amount of spectrum that can be reallocated while simultaneously increasing the probability that litigation from disgruntled broadcast station owners will mire the entire transition in legal process.⁴²⁹ Overlays, in contrast, are standard forms that could have been defined, allocated, and auctioned years ago, without new Congressional authorization.

Second, despite whatever impediments negotiations between incumbents and overlay licensees may entail, the FCC implicitly concedes that its alternative path of incentive auctions is strewn with obstacles. This is the clue embedded in the “incentive auction” plan to reallocate approximately 40% of the TV Band, or 120 MHz of the 294 MHz.⁴³⁰ A more sweeping transition involving all 294 MHz could be priced by overlay licensees based on relevant market data. But the FCC policy will continue to use spectrum having very high social opportunity costs to park off-air broadcast stations, a service that provides little if any incremental value beyond what cable, satellite and broadband platforms can deliver on their own. These costs, unseen as

between the ideal and the real and if discrepancies are found, they deduce that the real is inefficient. Users of the comparative institution approach attempt to assess which alternative real institutional arrangement seems best able to cope with the economic problem

Demsetz, *supra* note 239, at 1 (emphasis in original).

427. The delay is actually much longer, given that a reallocation via overlays could have begun in 2009 (when the NBP was being written) or sooner.

428. Indeed, the head of the NBP, Blair Levin, has argued that the provisioning legislation was a disaster that threatens the entire program.

“The legislation ties the FCC’s hands in a variety of ways,” said Levin, who left the FCC following release of the broadband plan and is now attached to the Aspen Institute. “It opens it up to litigation risk, which then, in conjunction with the other handcuffs, makes it difficult to pull off a successful auction. The nature of the bill dramatically increases the probability that there will be less spectrum recovered and less money for the [U.S.] Treasury.”

Kim McAvoy, *Levin: TV Spectrum Auctions Likely Doomed*, TV NEWSCHECK (Jan. 5, 2012), <http://www.tvnewscheck.com/article/56476/levin-tv-spectrum-auctions-likely-doomed>.

429. See Hazlett et al., *Incentive Auctions Paper*, *supra* note 49, at 4, 6 (on the repacking compensation determined in administrative process, with anticipation of results to be “vigorously challenged” by stakeholders).

430. Hazlett, *A Proposal for an Overlay Auction*, *supra* note 50; *Incentive Auctions*, *supra* note 49; 2012 Incentive Auction NPRM, *supra* note 49.

unrealized opportunities, exhibit the classic regulatory bias in favor of “Type II error.”⁴³¹

The evidence that the Commission draws on for its conclusion is lacking. That Commission draws attention to “holdouts” and “bargaining problems,” but does not recognize the similar issues in its own regulatory actions. Moreover, it fails to understand that market actors have been quite successful in overcoming similar obstacles in the past. The FCC has, after all, fragmented mobile license rights beyond imagination, issuing *well over 50,000 licenses*.⁴³² These rights have been aggregated into somewhere between four and seven national wireless networks over the past several years. Much of the reassembly was accommodated by the adoption of license auctions, for which regulators—at least those sufficiently farsighted to push Congress to enable this policy as it did in 1993—deserve credit. Yet, the great majority of the consolidation took place in secondary market transactions where every merger overcame a potential “holdout” problem. Hence, the evidence presented does not support the FCC’s position.

The general utility of the overlay approach recommends it highly. Overlays may be issued for bands with existing users, including for-profit and non-profit (including government) licensees. These new rights can and generally should be auctioned, thereby introducing responsible economic agents into the markets in the form of rights holders positioned to efficiently reallocate spectrum in the band. Users and their business models can be reorganized, frequency assignments may be switched, new technologies may be deployed, and ultimately the optimal combinations for wireless activities that reflect current conditions and opportunities can be discovered. The “relocations” that occur will be strictly voluntary, mutually beneficial contractual bargains that generate net benefits for all parties involved. Cooperation replaces conflict; spectrum repurposing moves forward; regulatory quagmires are avoided.

This process has been shown to work using standardized rights templates across diverse situations. Scores of TV stations went dark early (prior to the June 2009 analog station switch off) to accommodate Qualcomm’s use of TV spectrum for a new service, MediaFlo, in 2006 to 2008.⁴³³ The approach

431. *Wireless Craze*, *supra* note 30, at 380–82 & fig.5 (discussing social costs and benefits of spectrum use, type II error of underuse, type I error of overuse, and the tragedy of the commons).

432. *Federal Preemption*, *supra* note 29, at 194, 201.

433. Thomas W. Hazlett, *A Law and Economics Approach to Spectrum Property Rights: A Response to Professors Weiser & Hatfield*, 15 GEO. MASON L. REV. 975, 999–1002 (2008). In fact, even though MediaFlo resided on Channel 55, estimates of as many as one in six terrestrial broadcasting stations were reconfigured, including in neighboring Channels 54 and

worked very well in PCS licensing, where the overlay concept rescued government spectrum allocation policy that had been mired in a long-running dispute over PCS-microwave interference. And, this approach also worked for AWS licenses, where public agencies had delayed, since at least 1992,⁴³⁴ in making room for new wireless communications. After licenses were auctioned in September 2006, high bidders—including T-Mobile, which purchased the largest package of rights—emerged to push reallocation forward.⁴³⁵ These new overlay licensees were obligated to pay for incumbent relocation, and had adjusted their auction bids to reflect anticipated costs.⁴³⁶ But, as residual rights holders in the band, they actively pushed the process forward.⁴³⁷ Even with government agencies providing little enthusiasm for the needed changes (which was not surprising given their lack of an

56. *Id.* at 1002; *Tragedy TV*, *supra* note 76, at 110. Such private reconfiguration occurred prior to mandated deadlines.

434. This is when federal legislation designating that the frequencies used for Advanced Wireless Services were authorized for reallocation to commercial mobile services. In 2001, a notice for proposed rulemaking for AWS below 3GHz was released by the FCC, following proposals and requests in the 1992 *Emerging Technologies Proceeding* (ET Docket No. 92-9), which identified possible bands for advanced wireless devices. See Press Release, FCC, FCC Looks to Allocate Additional Spectrum for New Advanced Wireless Systems (Jan. 4, 2001), http://hraunfoss.fcc.gov/edocs_public/attachmatch/DOC-208768A1.pdf. In November 2003, new service rules were promulgated for this AWS spectrum, auctioned in Auction #66 in 2006. However, AWS spectrum slated for auction in 2011 has missed deadlines, and statutorily must be auctioned by 2015. See MIDDLE CLASS TAX RELIEF ACT AND JOB CREATION ACT of 2012, P.L. 112-96, 126 STAT. 156, 47 U.S.C. § 1451 (2011) (setting a 2015 deadline for auction for AWS-2 (1915–1920 MHz, 1995–2000 MHz), AWS-3 (2155–2180 MHz), provided that the bands can be “used without causing harmful interference to commercial mobile licenses” in 1930–1995 MHz).

435. Comments of T-Mobile USA, Inc., Relocation of Federal Systems in the 1710–1755 MHz Frequency Band: Review of the Initial Implementation of the Commercial Spectrum Enhancement Act, Docket No. 0906231085-91085-01, at 2 (Nat’l Telecomms. & Info. Admin., Aug. 21, 2009), http://www.ntia.doc.gov/files/ntia/t-mobile_csea_noi_comments_8-21-09_0.pdf.

436. *Id.* at 8 (describing the quality of information provided by federal agencies on incumbent uses: “Because several agencies underestimated the cost and time involved in relocation, prospective bidders, including T-Mobile, received inaccurate projections about when the spectrum would be commercially available.”); *id.* at 9 (“Prior to the auction, AWS bidders were given only a limited amount of information about the incumbent systems: agency name, center frequency, system type and name, and coordinates for transmitters and receivers.”); *id.* at 10 (“As noted above, much of the data compiled by federal agencies and provided to prospective licensees was incomplete or inaccurate and impeded licensee requests for early access to spectrum. In some cases, initial cost-estimates were so off the mark that the additional requests for funds from OMB triggered Congressional review, stalling efforts for early access to the spectrum and further delaying service to consumers.”).

437. See *id.* at 10 (“It was only after the auction that carriers learned these 12 assignments barred use of the entire AWS and, causing significant delays in launching commercial service.”).

economic stake in the outcome), the presence of interested parties with important assets at stake produced progress. By the end of 2008, T-Mobile offered 3G services to consumers using the AWS spectrum.⁴³⁸

The overlay approach delegates the process of reallocation to specialized agents. In lieu of disinterested parties attempting to liquidate existing band users to make way for new uses, private parties received a claim to their newly created values. The normal economic incentives drive these parties to obtain information, innovate in the technologies and market organization employed, and discover new ways of lowering the exit costs for incumbents. In these pursuits, parties are not hamstrung by bureaucratic procedures, but can avail themselves of the full line of efficiency-enhancing institutions created for such purposes—including capital markets. With ready funding for value-creating projects, financial investors can pay to liquidate the obstacles that need to be removed to make way for greater economic gains. With “incentive auctions,” policy makers now attempt to mimic the market. The success or failure of incentive auctions will be a fitting experiment.

V. CONCLUSION

It is clear that existing spectrum allocation rules deter the productive use of a key natural resource in the information economy. Yet a tantalizing series of natural experiments shows a more productive path for allocating spectrum. Amazing new wireless applications have emerged using the modest bandwidth allotted to liberal licenses. With these airwaves, used according to market demands, networks have formed that are revolutionizing communications, forging new paths to economic development, and discovering geysers of value in emerging social media. But the success of such wealth-creating innovations frustrates policy makers, tasked with the challenge of supplying spectrum inputs to this rapidly scaling marketplace.⁴³⁹

438. Press Release, T-Mobile, T-Mobile USA Announces Commercial 3G Network Availability in 21 Markets by Mid-October (Sept. 18, 2008), *available at* <http://newsroom.t-mobile.com/articles/t-mobile-3G-network-availability> (describing deployment of 3G service to twenty-seven major markets, on UMTS/HSDPA technology). Regarding AWS, “T-Mobile and the U.S. government, namely the Department of Commerce, Department of Defense and the Department of Justice, continue to work closely and effectively together to make available AWS spectrum that will give our customers access to T-Mobile’s 3G network.” *Id.*

439. “[B]y estimating various factors . . . mobile broadband is likely to entail economic value of at least \$100 billion in the next five years.” FCC, MOBILE BROADBAND: THE BENEFITS OF ADDITIONAL SPECTRUM 5 (Oct. 2010), *available at* <http://download.broadband.gov/plan/fcc-staff-technical-paper-mobile-broadband-benefits-of-additional-spectrum.pdf> (listing

While virtually swimming in little-used, under-deployed frequency bands, each and every attempt spectrum regulators make to move additional bandwidth into more productive use is met with fierce resistance. The prospect of harmful interference inevitably looms whenever the effort is to increase wireless deployments; the additional traffic yields benefits (new uses) but also costs (crowding old uses). Under the administrative allocation system defined by the 1927 Radio Act, regulators are charged with determining the proper balance. The general scholarly dissatisfaction with the operation of this system is that regulators, on net, are decidedly prone to protect the existing order. Change is too difficult. The beneficial wireless uses that are blocked in order to prevent possible airwave conflicts with existing services errs far too much on the side of silence.

The dysfunctional nature of the Federal Communications Commission—documented by the Commission itself, in charting the socially expensive delays incurred in allocating or reallocating spectrum—has given rise to academic discussion as to how best to break the logjam. Many analysts (and the Commission) point out that spectrum markets, to be efficient, depend on spectrum use rights that are defined “clearly and exhaustively.” They bemoan the fact that this has never been done by regulators. From there, they conclude that this task must be completed and that spectrum markets demand greater specificity in the delineation of property rights. They proceed to blame long-standing interference disputes, arguments that have seemingly delayed valuable reallocations of bandwidth for years, on the lack of full, precise, unambiguous delineations of the boundaries of spectrum use rights.

This policy response, as we have attempted to show, is wrong. Exclusive use rights for mobile service licensees are not defined precisely and do not preempt all possible disputes—not nearly. But they have proven “good enough.” They are sufficient to organize very active markets, filled with complex economic relationships and innovative forms of social coordination, and do so in a way that could be accomplished via alternative rules—most pointedly, administrative allocation of radio spectrum. The liberal licenses used in these markets encourage and enable the creation of complementary assets making the spectrum used far more valuable than it would otherwise be. Yet the ownership rights are not exact. Indeed, they cannot be exact; fortunately, they do not need to be. They must simply beat the alternatives.

One alternative is for spectrum allocation to proceed, blocking current deployments, as regulators search for more exact rights specifications. This

mobile statistics such as economic impacts, tablet sales, mobile application sales, mobile online commerce, and social value in education and health care).

has happened in many markets, with generally disastrous consequences for consumer welfare. Take the L Band, used for satellite services.⁴⁴⁰ An operator there, a decade ago, sought to increase the value of services provided by rearranging the band and supplying terrestrial mobile voice and data connections in addition to satellite links. The extra traffic would make noisy what was very quiet. Objections were heard. A large number of technical reports, some reporting results of field tests, evaluated how new L Band traffic would impact the reception of GPS signals using adjacent frequencies. The decisive moment arrived when the U.S. Department of Commerce sent the FCC a letter stating that the performance of GPS receivers—some used for emergency navigation warnings on airplanes, other used by the U.S. military—would be, on some occasions, degraded. Commerce strongly argued that no terrestrial L Band operations be permitted, a position backed by the Federal Aviation Administration and the Department of Defense.⁴⁴¹ The FCC immediately determined that harmful interference would best be avoided by leaving the L Band quiet, killing the authorization for terrestrial service and ending what was to be a \$14 billion mobile communications network.⁴⁴² Disinterested observers appalled by the decision and the process dismissed the regulatory choice as political.⁴⁴³

That tempest is a microcosm of FCC dysfunctionality. An intense debate looms among the technical experts. But their voluminous engineering data decide nothing. They do not—and cannot—determine the “proper” spectrum use, which depends on competing values. (*How much would entice GPS users to tolerate 8 dB of signal interference? Perhaps free new GPS receivers plus LTE 4G unlimited data subscriptions?*) The decision as to resource use is then made by administrators who, after reading technical reports, make a political choice—but in actuality render an *economic decision*.

440. The fascinating and quite complicated story of the “LightSquared debacle” is discussed in detail in Thomas W. Hazlett & Brent Skorup, *Tragedy of the Regulatory Commons: LightSquared and the Missing Spectrum Rights*, 11 DUKE L. & TECH. REV. (forthcoming 2013).

441. Letter from Lawrence E. Strickling, Assistant Secretary for Communications and Information, to Julius Genachowski, Chairman, Federal Communications Commission, U.S. Dep’t of Commerce (Feb. 14, 2012).

442. Statement From FCC Spokesperson Tammy Sun on Letter from NTIA Addressing Harmful Interference Testing Conclusions Pertinent to LightSquared and Global Positioning Systems (Feb. 14, 2012). The \$14 billion expenditure included construction costs for physical infrastructure and spectrum acquisition costs, including payments to other satellite licensees in the L Band. See Statement from Sanjiv Ahuja, CEO of LightSquared (Dec. 14, 2011), available at <http://www.prnewswire.com/news-releases/statement-from-sanjiv-ahuja-ceo-of-lightsquared-35621603.html>.

443. Telecoms in America, *A Dark Day for Lightsquared, Plans for a New National Wireless Network Hit a Regulatory Wall*, ECONOMIST, Feb. 18, 2012, <http://www.economist.com/node/21547813>.

To improve social welfare, public policy must be restructured. Technical specificity is worth including when it is worth its cost, but extremely damaging to pursue when it is not. Moreover, it both binds and blinds the institutional process. Great activity occurs of an analytical nature—indeed, it looks like administrative fact-finding to courts sitting in review. But greater exactitude in rights delineation is rarely the problem in spectrum allocation. The key issues do not involve what signals interfere but what economic plans interfere, and there exist myriad ways in which to make the conflicting plans coincide. But these win-win pathways are truncated by the lack of ownership rights that regulators—instead of pursuing fruitless harmful interference dockets—should be defining and setting adrift, delegating spectrum use decisions to responsible economic agents who experience opportunity costs in real-time, who prosper when they are right and suffer when they are wrong.

When the margins of use rights are in doubt, conflicts are generally solved—in the market or by regulators—not with greater exactitude but with reconfiguring ownership rights. This eliminates fragmentation, such as the notorious “interleaving” allocations, so as to “reallocate” spectrum. Withholding rights from the market, by issuing “operating permits” in lieu of “spectrum licenses,” ensures that a great deal of productive spectrum activity is little more than alternative applications locked in the vaults of regulators. In this situation, licensees do not strive to maximize the use of the spectrum allocated, in that they incur no opportunity costs from doing what is socially inefficient. They cannot be paid to perform any better. By eliminating these anti-social rights truncations, and putting all flexible-use rights into the bundles managed by licensees, liberalization enables activity by way of privatized coordination in wireless. When the right of transfer is properly included, secondary markets are enabled and market forces naturally resolve the problems of fragmentation.

Of course, to mitigate transaction costs, regulators should properly assign rights in the first place. In that regard, the introduction of competitive bidding to assign licenses—adopted in dozens of countries over the past quarter-century—are a welcome policy innovation. Including combination bids would improve that mechanism. In general, aggregations of license rights provide for rational choices to be made about “interference,” which is internalized, as well as for partitioning, which should be a permissible activity.

There is no need to reinvent the wheel. Reasonable packages of spectrum use rights have been created by regulators and deployed in the market, enabling efficient service delivery. They have not only accommodated investment and commerce, but waves of innovation. The “smartphone

revolution” has been launched using liberal mobile phone licenses, and has in turn introduced mobile phone application platforms that disrupt old business models by offering greatly expanded sources of services and content. No new government authorizations or requests for permission for new network applications were needed. New applications could have created harmful interference between mobile phones—indeed, they certainly did when crowded networks slowed to a crawl at peak times. But carriers compete to manage these flows, to protect their quality of service, and to welcome fancy new consumer-pleasing gadgets with sophisticated computing bundled within.

Economics governs spectrum. That simple maxim can, and should, better inform our governing choices.

