

31:1 BERKELEY TECHNOLOGY LAW JOURNAL

2016

Pages

1

to

362

Berkeley Technology Law Journal

Volume 31, Number 1

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 © 2016 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 31

NUMBER 1

2016

TABLE OF CONTENTS

ARTICLES

A FEW KIND WORDS FOR ABSOLUTE INFRINGEMENT LIABILITY IN PATENT LAW.....	1
<i>Robert P. Merges</i>	
STRATEGIC DECISION MAKING IN DUAL PTAB AND DISTRICT COURT PROCEEDINGS	45
<i>Saurabh Vishnubhakat, Arti K. Rai & Jay P. Kesan</i>	
CHILLING EFFECTS: ONLINE SURVEILLANCE AND WIKIPEDIA USE.....	117
<i>Jonathon W. Penney</i>	
FACILITATING COMPETITION BY REMEDIAL REGULATION	183
<i>Kristelia A. García</i>	
THE INFORMATIONAL VALUE OF PATENTS	259
<i>Clark D. Asay</i>	
GAGS AS GUIDANCE: EXPANDING NOTICE OF NATIONAL SECURITY LETTER INVESTIGATIONS TO TARGETS AND THE PUBLIC	325
<i>Rebecca Wexler</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., 20th ed. 2015). Please cite this issue of the *Berkeley Technology Law Journal* as 31 BERKELEY TECH. L.J. ____ (2016).

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*; the *Bolt*, a collection of short comments and updates about new developments in law and technology written by BTLJ members; and *BTLJ Commentaries*, an exclusively online publication for pieces that are especially time-sensitive and shorter than typical law review articles.

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., 20th ed. 2015).

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

ORRICK, HERRINGTON &
SUTCLIFFE LLP

Benefactors

COVINGTON & BURLING LLP

SIDLEY AUSTIN LLP

FISH & RICHARDSON P.C.

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

KASOWITZ BENSON
TORRES & FRIEDMAN LLP

VAN PELT, YI & JAMES LLP

KIRKLAND & ELLIS LLP

WEIL, GOTSHAL & MANGES LLP

LATHAM & WATKINS LLP

WHITE & CASE LLP

MCDERMOTT WILL & EMERY

WILMER CUTLER PICKERING
HALE AND DORR LLP

MORRISON & FOERSTER LLP

WILSON SONSINI
GOODRICH & ROSATI

WINSTON & STRAWN LLP

Members

BAKER BOTTS LLP

KILPATRICK TOWNSEND &
STOCKTON LLP

BAKER & MCKENZIE LLP

KNOBBE MARTENS
OLSON & BEAR LLP

DURIE TANGRI LLP

LEE TRAN & LIANG LLP

GJEL ACCIDENT ATTORNEYS

MUNGER, TOLLES & OLSON LLP

GTC LAW GROUP LLP &
AFFILIATES

O'MELVENY & MYERS LLP

GUNDERSON DETTMER STOUGH
VILLENEUVE FRANKLIN &
HACHIGIAN, LLP

PAUL HASTINGS LLP

HAYNES AND BOONE, LLP

ROPES & GRAY LLP

HOGAN LOVELLS LLP

SIMPSON THACHER & BARTLETT
LLP

IRELL & MANELLA LLP

TURNER BOYD LLP

KEKER & VAN NEST LLP

WEAVER AUSTIN VILLENEUVE &
SAMPSON, LLP

BOARD OF EDITORS 2015–2016

Executive Committee

Editor-in-Chief
JOSHUA D. FURMAN

Managing Editor
MISHA TSUKERMAN

Senior Articles Editors
RAVI ANTANI
ZACHARY FLOOD
KELLY VARGAS

Senior Executive Editor
DIANA OBRADOVICH

Senior Scholarship Editor
CAROLINA GARCIA

Senior Annual Review Editors
GINNY SCHOLTES
SORIN ZAHARIA

Senior Online Content Editor
NOAH DRAKE

Editorial Board

Commentaries Editors
SWAROOP POUDEL
YU TANEBE
BONNIE WATSON

Production Editors
ROXANA GUIDERO
DUSTIN VANDENBERG

Technical Editors
KRISTINA PHAM
CHRISTOPHER YANDEL

Annual Review Editors
CYNTHIA LEE
BENJAMIN LI

Notes & Comments Editors
WAQAS AKMAL
ERICA FISHER

Symposium Editors
TOMMY BARCZYK
JOHN RUSSELL

Submissions Editors
PHILIP MERKSAMER
MAYA ZIV

Web & Technology Editors
JOE CRAIG
LEIGHANNA MIXTER

External Relations Editor
SIMONE FRIEDLANDER

Member Relations Editor
STEPHANIE CHENG

Alumni Relations Editor
GOLDA CALONGE

Web Content Editor
FAYE WHISTON

JESSICA ANNIS
JOEL BROUSSARD
CHRISTIAN CHESSMAN
DANEILLE DEVLIN
KELSEY GUANCIALE

Articles Editors
YESOL HAN
CASSY HAVENS
MARK JOSEPH
BILAL MALIK

CHRIS NORTON
LIDA RAMSEY
ERIC RIEDEL
ARIEL ROGERS
MAX SLADEK DE LA CAL

MEMBERSHIP

Vol. 31 No. 1

Associate Editors

WILL BINKLEY	PAMUDH KARIYAWASAM	ALEJANDRO ROTHAMEL
BRITTANY BRUNS	HILARY KRASE	LUKAS SIMAS
DAPHNE CHEN	JOYCE LI	RICHARD SIMS
YUHAN (ALICE) CHI	MATTHEW MCCLELLAN	FERNANDA SOLIS CAMARA
KEVIN CHIU	GAVIN MOLER	SARAH SUWANDA
SARA CHUGH	CATALINA MONCADA	JON TANAKA
NOA DREYMAN	SHELBY NACINO	VALERIE TRUONG
RAN DUAN	JUAN NAZAR	RAOUL GRIFONI-WATERMAN
JORDAN FRABONI	ROBERT OLSEN	REID WHITAKER
JEREMY ISARD	EUNICE PARK	
NATASA JANEZ	JAIDEEP REDDY	

Members

YASMINE AGELIDIS	NOAH GUINEY	MICHELLE PARK
NIKI BAWA	BRIAN HALL	DYLAN PETERSON
KATE BRIDGE	ANDREA HALL	KEVIN PORMIR
JESSICA BRODSKY	JESSICA HOLLIS	CHRISTELLE PRIDE
KELSEA CARLSON	JENNIFER HSU	JING XUN QUEK
STEPHEN CHAO	KENSUKE INOUE	LEE REDFEARN
JOSEPH CHRISTIE	HYE JIN KIM	MATT RICE
RACHEL CORRIGAN	RITHIKA KULATHILA	FAITH SHAPIRO
MICHAEL DEAMER	SARAH KWON	DARINA SHTRAKHMAN
THOMAS DEC	TIFFANY LEUNG	JOSHUA STEELE
DARIUS DEGHAN	MEI LIU	EVE TAI
ELENA FALLOON	STASHA LOEZA	MY THAN
MEGHAN FENZEL	SARAH MULLINS	CASEY TONG
EVAN FERGUSON	ELI NESS	ELISSA WALTER
WHITNEY FLORIAN	NATE NGEREBARA	MELISSA WEE
ETHAN FRIEDMAN	PEGGY NI	TAMARA WIESEBRON
ANDREY GAVRILENKO	JESSICA OGLESBEE	SHONG YIN
KAN GU	BARCLAY OUDERSLUYS	YU ZHAO
	ROBERT PARIS	

BTLJ ADVISORY BOARD

JIM DEMPSEY

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

ROBERT C. BERRING, JR.

*Walter Perry Johnson Professor of Law
U.C. Berkeley School of Law*

MATTHEW D. POWERS

Tensegrity Law Group, LLP

JESSE H. CHOPER

*Earl Warren Professor of Public Law
U.C. Berkeley School of Law*

PAMELA SAMUELSON

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

PETER S. MENELL

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

LIONEL S. SOBEL

*Visiting Professor of Law
U.C.L.A. School of Law*

ROBERT P. MERGES

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

LARRY W. SONSINI

Wilson Sonsini Goodrich & Rosati

REGIS MCKENNA

*Chairman and CEO
Regis McKenna, Inc.*

MICHAEL STERN

Cooley LLP

DEIRDRE K. MULLIGAN

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

MICHAEL TRAYNOR

Cobalt LLP

JAMES POOLEY

*Deputy Director General of the
World Intellectual Property Organization*

THOMAS F. VILLENEUVE

*Gunderson Dettmer Stough Villeneuve
Franklin & Hachigian LLP*

BERKELEY CENTER FOR LAW & TECHNOLOGY 2015–2016

Executive Director

JIM DEMPSEY

Faculty Directors

KENNETH A. BAMBERGER	PETER S. MENELL	PAMELA SAMUELSON
CATHERINE CRUMP	ROBERT P. MERGES	PAUL SCHWARTZ
CHRIS HOOFNAGLE	DEIRDRE MULLIGAN	JENNIFER URBAN
SONIA KATYAL	MOLLY S. VAN HOUWELING	

Staff Directors

LOUISE LEE

RICHARD FISK

CLAIRE TRIAS

A FEW KIND WORDS FOR ABSOLUTE INFRINGEMENT LIABILITY IN PATENT LAW

Robert P. Merges[†]

ABSTRACT

The goal of this Article is to partially defend patent law's absolute liability rule. Most scholars question the utility of absolute liability, which makes it irrelevant whether an infringement defendant copied from the patentee or independently invented the patented invention. This Article draws from two literatures in making a defense. First, it looks to studies of how technological information is communicated or "diffused." These studies, together with research by psychologists on "inadvertent copying," demonstrate that ideas are sometimes copied in obscure and subtle ways, leaving little or no evidence that copying has indeed occurred. Next, the Article turns to the literature on optimal standards of care in tort law, which is used to describe what would happen if U.S. law changed to require patentees to show copying. Potential patent infringement defendants—a class that includes virtually all people and companies performing research and development—might well impose strict limitations on receipt of technological information. This would help rebut allegations of copying, reducing the risk of legal liability. Such strict limitations would stifle innovation because technological communities thrive on ubiquitous and unregulated communication. Patent law, as it stands, encourages communication by making proof of copying irrelevant in patent cases. As a consequence, researchers (as potential patent infringement defendants) currently have no reason to restrict their access to technical communications.

To further support the case for absolute liability, this Article looks to both copyright law and common law rules on the theft of ideas. Both require proof of copying, and both have led potential defendants to invest in restrictive measures designed to prove that they had no access to third party ideas.

By requiring only proof that the infringer is using an invention covered by the patentee's claims, patent law dispenses with the need to prove that the infringer copied from the patentee. This means that researchers have very little incentive to protect against receipt of outside information. Potential infringers get no payoff from disproving contact with a patentee because the reduced incidence of copying that would follow from lack of

DOI: <http://dx.doi.org/10.15779/Z38285T>

© 2016 Robert P. Merges.

[†] Wilson, Sonsini, Goodrich & Rosati Professor of Law, University of California, Berkeley, School of Law. Thanks to participants in Professor Rebecca Eisenberg's IP Scholarship Seminar at the University of Michigan in 2012; to participants in the IP Scholars Conference at Berkeley in August, 2014; and, as always, to the excellent editorial assistance from the BTLJ editors and staff. I am absolutely liable for all remaining errors.

contact is irrelevant under patent law. No one invests in restricting contact, and information flows freely.

The obvious downside of the current regime is that sometimes, an infringement defendant will really be a true independent inventor; no copying, subtle or otherwise, took place. In these cases, as the literature shows, absolute liability imposes significant costs. These costs must be weighed against the benefits of information freely shared under the current absolute liability regime. It is hard to say how the costs and benefit net out. Under these circumstances, a recent innovation in U.S. patent law, the new “prior commercial use” defense under the America Invents Act (AIA), may prove helpful. This defense permits one who can prove early commercialization of an invention to avoid liability, even though he or she otherwise infringes. Because commercialization before a patentee’s filing date will often be associated with independent invention, this rule reduces some of the costs of the absolute liability standard. It is difficult to say whether this combination of (1) absolute liability, plus (2) an exception for early commercialization, provides the optimal set of incentives for potential patent infringers. But at least the prior commercial use rule encourages activity that has independent social value, i.e., rapid introduction to the market. Given the real benefits to the longstanding rule of absolute liability in patent law, providing an exception under these circumstances makes for a reasonable policy.

TABLE OF CONTENTS

I. INTRODUCTION.....	3
II. COPYING, DIFFUSION, AND PROBLEMS OF PROOF	10
A. INFORMATION DIFFUSION AND INADVERTENT	
PLAGIARISM	10
1. <i>Diffusion Studies</i>	11
2. <i>Fallible Memory and Inadvertent Plagiarism</i>	17
B. PROVING COPYING: CASES FROM OTHER AREAS OF IP	
LAW.....	20
1. <i>The Access Requirement in Copyright Law</i>	21
2. <i>Derivation in Patent Law</i>	22
C. FROM COPYING TO ACQUIRING	28
III. EXCESS PRECAUTION: THE COST OF REQUIRING	
PROOF OF COPYING.....	30
A. THE TORT THEORY PERSPECTIVE.....	31
B. HOW ABSOLUTE LIABILITY FEEDS THE INFORMATION	
COMMONS.....	36
IV. PRIOR COMMERCIAL USE AND OTHER DOCTRINAL	
ALTERNATIVES	38
A. PRIOR COMMERCIAL USE: THE GOOD AND THE BAD	38
B. ALTERNATIVE “MIDDLE GROUND” RULES	41
V. CONCLUSION	44

I. INTRODUCTION

Through its doctrinal structure, patent law essentially *presumes* copying on the part of any company that makes or uses technology developed after the date a patent issues. Liability follows when a product accused of infringing a patent falls within the linguistic boundaries of (at least one of) the patent's claims. It is irrelevant under current law whether the defendant actually copied the patentee's technology, let alone whether it intentionally, recklessly, negligently or inadvertently copied the patentee's technology. Put simply, patent infringement is an absolute liability regime.¹

Many commentators have decried this aspect of patent law. Absolute liability means that a patent owner can sue anyone who makes, uses, or sells the same invention covered by the patent's claims. The strong consensus on absolute liability is clear: it is a very bad idea. Most commentators agree that what patent law needs is an "independent invention" defense.² Absolute liability, they argue, should give way to a

1. Actually, as Blair and Cotter point out, this is not completely true. That is because of the patent marking statute, which precludes the awarding of damages in cases where the patentee could have marked its product with notice of a patent, but did not. There are quirks in the marking statute that mitigate this to some extent. And also, in these cases an injunction is still possible, even in the absence of notice. See Roger D. Blair & Thomas F. Cotter, *Strict Liability and Its Alternatives in Patent Law*, 17 BERKELEY TECH. L.J. 799 (2002). Also, note that this refers only to what is called direct infringement. Indirect infringement—which means that the accused party falls short of practicing every single element of a claimed invention—requires the patentee to establish that the accused infringer acted with some measure of intentionality or scienter. See *Limelight Networks, Inc. v. Akamai Tech., Inc.*, 134 S. Ct. 2111 (2014).

2. See, e.g., Stephen M. Maurer & Suzanne Scotchmer, *The Independent Invention Defence in Intellectual Property*, 69 ECONOMICA 535 (2002); Carl Shapiro, *Prior User Rights*, 96 AM. ECON. REV. 92 (2006); Oskar Liivak, *Rethinking the Concept of Exclusion in Patent Law*, 98 GEO. L.J. 1643 (2010); Samson Vermont, *Independent Invention as a Defense to Patent Infringement*, 105 MICH. L. REV. 475 (2006); see also Mark A. Lemley, *Should Patent Infringement Require Proof of Copying?*, 105 MICH. L. REV. 1525, 1525 (2007) (identifying the primary concern with the independent invention defense as a potential reduction in incentives to invent; proposing alternatives, including prior user rights; making independent invention a defense to willful infringement; and making third party independent invention a secondary consideration weighing against nonobviousness). There are, however, a few voices that favor the status quo. *But see* Lemley, *supra*, at 1529, 1535–36 (arguing that the independent invention defense may interfere with patent law's incentive structure); John F. Duffy, *Inventing Invention: A Case Study of Legal Innovation*, 86 TEX. L. REV. 1, 9 (2007) ("A narrow right that allows for independent creation and protects only the precise details of a particular embodiment of the invention is unlikely to give sufficient protection, as a practical matter, to encourage the type of investments and work that society wants to encourage."); Clarisa Long, *Information Costs in Patent and*

more fault-based standard. Only defendants who have copied from the patentee ought to be liable. Defendants who do not copy, who independently invent, should not be. Until the rule is changed, many observers presume that a good deal of patent litigation will be illegitimate: it pits an opportunistic patent owner against an innocent accused infringer, who learned nothing from the owner's patent because the accused party developed its technology completely independently. This means that in some sense, most cases of patent infringement can be described as rent-seeking, pure and simple—an attempt by the holder of a legal right to extract value from a company that is earning money on the basis of its own research.³ When we think of patents and rent-seeking these days, of course, it raises the specter of patent trolls.⁴ It is therefore

Copyright, 90 VA. L. REV. 465, 528 (2004) (arguing that “an independent creation privilege in patent law would too drastically reduce incentives to create”). From the practitioner viewpoint, see Roger Milgrim, *An Independent Invention Defense to Patent Infringement: The Academy Talking to Itself: Should Anyone Listen?*, 90 J. PAT. & TRADEMARK OFF. SOC'Y 296, 297 (2008) (suggesting that academic proposals for an independent invention defense are in need of “adult supervision”).

3. The alternative argument depends on a pure incentive story: inventors require an exclusive right to create and develop an expensive technology, and the benefits of exclusivity outweigh the costs of duplicative R&D. This incentive story typically plays out in the context of a patent “race”—when two or more firms are competing to achieve a given patentable invention. See generally Suzanne Scotchmer, *Incentives to Innovate*, in THE NEW PALGRAVE DICTIONARY OF ECON. & THE LAW 273, 275 (1998) (noting “two views on patent races: that they inefficiently duplicate costs, and that they efficiently encourage higher aggregate investment”); Jennifer F. Reinganum, *The Timing of Innovation: Research, Development, and Diffusion*, in 1 HANDBOOK OF INDUSTRIAL ORGANIZATION 849, 853–68 (Richard Schmalensee & Robert D. Willig eds., 1989) (also discussing literature on patent races). A related perspective quibbles with the notion that duplicate R&D costs are really wasted. See JEAN TIROLE, THE THEORY OF INDUSTRIAL ORGANIZATION, 400 (1988) (noting that the loser in a patent race may benefit from positive spillovers, may develop another product, and may gain experience for future races); Robert P. Merges & Richard R. Nelson, *On the Complex Economics of Patent Scope*, 90 COLUM. L. REV. 839, 870–79 (1990) (competition in the market for improvements spurs innovation, despite possible efficiency losses attributable to rivalrous invention).

4. There is a very large literature discussing the troll problem. For general orientation, see generally Robert P. Merges, *The Trouble with Trolls: Innovation, Rent-Seeking, and Patent Law Reform*, 24 BERKELEY TECH. L.J. 1583 (2009), and Mark A. Lemley & A. Douglas Melamed, *Missing the Forest for the Trolls*, 113 COLUM. L. REV. 2117, 2157 (2013). For information on the closely related phenomenon of the “secondary market” for patents (buying and selling patents, a practice that feeds patent troll formation and litigation), see Michael Risch, *Licensing Acquired Patents*, 21 GEO. MASON L. REV. 979 (2014).

not surprising to find noted scholar Mark Lemley saying: “An independent invention defense would eliminate the troll problem.”⁵

The traditional response to this charge depends on incentive theory. A single exclusive right, it is said, reserves a product market for one firm. Under certain conditions, only this level of exclusivity can ensure an adequate return for the cost and risk of developing a complex technology.⁶ Absolute liability, in this setup, is required, regardless of whether several inventors might arrive at the same invention simultaneously. Put simply, a monopoly is necessary to stimulate the investment required to develop certain technologies. This is so despite the obvious double costs of this arrangement: first, due to the wasted resources that go into duplicative effort to win the “race for the patent”; and second, as a result of high consumer prices in the monopoly market resulting from the patent.

This is a plausible story theoretically. But this “need for market exclusivity” idea seems to have lost much of its power. The reason is simple enough. The story is based on two questionable assumptions: (1) very high-cost research projects that (2) culminate in a single market-covering patent.⁷ Most contemporary inventions outside the pharmaceutical context are incremental, the result of modest discrete investments made in connection with ongoing R&D. And for this reason, in most cases today, individual patents cover but one small component, or one aspect of one component, of large, multi-component technologies. So, outside certain special cases of very high-cost research (most notably pharmaceuticals), it is difficult to make a case for exclusivity based on a single market-spanning patent.

This may well be one reason that it has become more common for patent scholars to call for proof of copying in patent law.⁸ The factual

5. Mark A. Lemley, *Should Patent Infringement Require Proof of Copying?*, 105 MICH. L. REV. 1525, 1526 (2007).

6. See Glenn C. Loury, *Market Structure and Innovation*, 93 Q.J. ECON 395, 397 (1979) (modeling patent race in which “rewards . . . become available only to the first firm that introduces an innovation”); Tom Lee & Louis L. Wilde, *Market Structure and Innovation: A Reformulation*, 94 Q.J. OF ECON. 429, 429–30 (1980) (echoing the assumption that the race has a single winner, who obtains completely exclusive rights over all other competitors, and thus, that all other competitors lose the investments made in trying to win the race).

7. *Id.*

8. Trolls, or Non-Practicing Entities (NPEs) as they are sometimes called, bring most of the lawsuits they initiate in the semiconductor and computer industries, where products tend to be covered by many patents, and thus where the “race for a single dominant patent” story has the least relevance. See, e.g., Colleen V. Chien, *Of Trolls, Davids, Goliaths, and Kings: Narratives and Evidence in the Litigation of High-Tech Patents*,

assumptions underlying the theoretical case for strong market exclusivity have dropped away.⁹ But patent law's disregard for independent invention lives on. Surely, if a single market-spanning patent is not required to call forth inventor effort and investment, then the rationale for an absolute liability regime is gone, which can only mean that it is time to do away with that regime by instituting a robust independent invention defense.

True—unless there is an alternative rationale. Is there some other plausible reason to dispense with the patentee's need to prove that the plaintiff-patentee copied its invention? Does absolute liability serve some function other than insuring complete exclusivity to the winner of a patent race?

There are two alternative rationales. First, the absolute liability rule is necessary because in some cases patentees would find it very difficult to prove copying, even though the defendant has in fact benefitted from the patentee's technological contribution. This is a tricky case to make out. It pushes against the notion that in almost every case patent infringement defendants make their inventions completely independently. Yet it assumes serious obstacles to *proving* just that. It is an argument rooted in notions of corrective justice: inventors should be compensated when they have bestowed benefits on others. But it also has a strong practical flavor: inventors may have trouble proving that others benefitted from their work. Thus, though the argument starts from considerations of fairness, it may be expressed in the language of a strong evidentiary presumption.

Second, absolute liability may actually encourage widespread dissemination of technical information between companies. This is obviously so in cases where a patentee can show immediate, direct, and full disclosure of an invention, together with explicit copying. But it is also true in cases where the communication is less complete, less direct, or harder to prove. Sometimes, inventions are copied somewhat inadvertently—not as the result of a clear-cut, well-planned effort to explicitly copy a competitor's product. Copying occurs, in other words, in all kinds of subtle ways. With absolute liability in place, there is no reason

87 N.C. L. REV. 1571, 1580–81 (2009) (footnotes omitted) (“NPEs have focused on high-tech inventions for several reasons. First, they have historically acquired their patents from distressed or bankrupt companies, principally casualties of the Internet bubble. Second, products in computer and semiconductor-related industries tend to be covered by many patents, increasing the likelihood of infringement.”).

9. On the general trend toward thinking of large numbers of patents in “portfolios” instead of single patents, see Gideon Parchomovsky & R. Polk Wagner, *Patent Portfolios*, 154 U. PA. L. REV. 1 (2005); Michael Risch, *Patent Portfolios as Securities*, 63 DUKE L.J. 89, 154 (2013).

for companies to guard against either direct, complete communications from a patentee, or these more subtle forms of copying. In both cases, precautions do not matter; if a company's product infringes one or more claims of a patent, liability follows. But if an independent invention defense were introduced, this would change. Companies would be wise to change their behavior in response. A smart company would not only limit direct contact with known patentees or patent applicants. It would also begin to screen out some of the information that comes into the hands of its researchers. Screening would reduce the company's exposure to liability, by helping to show that the company's products were developed without access to outside information. By screening out incoming technical articles or other sources of external information, companies would guard against proof that they copied ideas from the technical literature, or from others in the field generally. These efforts to cut down on access would cost society a great deal. The point, developed at greater length below, is that absolute liability leaves companies with no real reason to screen out information. It encourages information sharing, and thus more innovation.

But any defense of the status quo must first engage a devastating study by Mark Lemley and Chris Cotropia, who found that patent infringement plaintiffs virtually never show actual copying on the part of a defendant,¹⁰ even when there are legal and common sense reasons why such a showing would benefit them. The study implies that the vast majority of patent infringement suits involve acts of independent invention. It suggests that the absolute liability standard in patent law is essentially the *sole reason* patentees are able to succeed in litigation. Is there any reason to proceed with my argument given such solid counter-evidence planted firmly in the way?

Perhaps. The Lemley and Cotropia study is not quite as devastating as it might appear at first. Their key finding is that "a surprisingly small percentage of patent cases even involve allegations of copying, much less proof of copying."¹¹ The authors note, of course, that patent infringement is an absolute liability regime, but they also point out a number of reasons why patentees have an incentive to prove copying. Various patent

10. Christopher A. Cotropia & Mark A. Lemley, *Copying in Patent Law*, 87 N.C. L. REV. 1421 (2009) [hereinafter Cotropia & Lemley].

11. *Id.* at 1422; *see also id.*, at 1424 (reporting that copying is alleged in only 10.9% of the 193 complaints sampled; is found to have occurred in only 1.76% of the 1871 published opinions studied; and is especially rare outside the specialized context of pharmaceuticals and chemicals—for example, in cases involving computers and software, copying was found in less than 1% of the decisions studied).

doctrines require such a showing, for instance.¹² And then there is the importance of telling a good “jury story”: proof of copying certainly casts the defendant in a bad light, which has to be a major advantage for the plaintiff.

Despite these advantages, the dominant fact remains that proof of copying is not essential for a plaintiff to make out its case. While it is helpful, it could also be a difficult issue to prove in many cases. True, a blatant case of copying, where the defendant bought one of the plaintiff’s products on the open market, took it apart, and deliberately reverse engineered it, would likely leave a blazing paper trail. But copying happens in other ways as well. Consider a member of the defendant’s research team assigned to solve one problem in the design of a new product. He or she brings to bear past experience, intuition, knowledge of the state of the art, and intelligence about competing products. In this *mélange* a smoking gun may not appear, or may be difficult to piece together later. Indeed, given how things really work, it might be better to sometimes say that technology is “borrowed,” or “inadvertently incorporated,” rather than “copied” outright. Later in this Article, the more neutral term “acquired” is used to signify this kind of less-than-blatant borrowing or copying.

Given the enormous expense of discovery in patent cases, perhaps it makes sense for the plaintiff to choose not to follow up every hint or clue that one feature or aspect of the defendant’s product was derived somehow from work that owes its ultimate origin to the patentee. At a minimum, it makes sense to forgo discovery on this point when the law, in effect, already presumes that copying has taken place. Why spend resources establishing an element of a cause of action that the law *already provides* at the start?

From this perspective, one feature of the cases that Cotropia and Lemley criticize makes perfect sense. They note that lawyers for patentees, and sometimes courts, are prone to coyly slide from the established fact of technical claim infringement to the morally loaded language of copying, theft, and wrongdoing.¹³ On this point, Cotropia and Lemley are correct (and courts should guard more zealously against this subtle but influential

12. Proof of copying helps the patentee’s case with respect to nonobviousness and willfulness. An infringer who copies a patentee’s invention has a harder time proving that the invention was obvious, because if it were, why resort to copying? It helps with willfulness too; intentional copying is solid proof of intentionality on the part of the infringer—a key element of willful infringement. For willfulness, the incentive to prove copying is significant: willful infringement can mean up to a three-fold increase in the damages owing to the patentee.

13. Cotropia & Lemley, *supra* note 10, at 1436–37.

slide-step). Yet the fact remains: lawyers make this move because the law invites them to. The patentee's lawyers speak of copying because the law in essence presumes it. Because lawyers do not have to establish copying affirmatively, they seldom bother to try.

So the first argument towards a tentative case for absolute liability is that the chief critique of it is not as devastating as it might appear. Cotropia and Lemley necessarily draw their data from current practice, in which proof of copying is helpful if a plaintiff happens upon it, but not in any way necessary. And this may explain at least part of the reason why copying is so seldom actually proven. Another key point mentioned above and developed below, is that there is a spectrum of information acquisition. Blatant copying is at one end, but the spectrum also includes various other forms of unacknowledged (and even unintentional) borrowing. Thus, the lack of outright copying does not mean that the accused infringer did not at some point learn something valuable from the patentee.

However, marshaling a few arguments to counter a prior study is not the same as a full refutation of it. An empirically-grounded test of the Cotropia-Lemley study, from a strictly utilitarian perspective, would include at least these data: (1) The total number of cases in which patent infringement is likely to be alleged under absolute liability and liability-only-with-copying; (2) the portion of each in which true, complete independent invention occurs—those cases in which the defendant learned nothing, directly or indirectly, from the patentee's research; (3) the portion of each in which some learning or benefit occurred; (4) the social welfare losses due to all patent litigation; (5) the effect on inventors' incentives, as well as total social welfare, from presuming a defendant was influenced by or benefitted from the work of the patentee, as the law does now; and (6) the social welfare differential between (a) rent-seeking litigation, under the current absolute liability regime, brought by patentees against defendants who are true and pure independent inventors; and (b) litigation where the defendant copied or at least acquired something of value from the patentee.¹⁴

14. Mark Lemley makes an explicit connection between rent-seeking litigation and the absolute liability standard in patent law, particularly in the information technology industries. See Mark A. Lemley, *Should Patent Infringement Require Proof of Copying?*, 105 MICH. L. REV. 1525, 1526 (2007) (footnote omitted) ("In the information technology industries, it sometimes seems as though the overwhelming majority of patent suits are not brought against people who copied a technology, but against those who developed it independently.").

I have a confession, reader: I don't have these data. So this Article cannot mount a comprehensive analysis of the issue. Instead, it makes a tentative start in the direction of defending the current doctrine. In essence, this Article examines whether there is *anything* to be said for the absolute liability standard.

II. COPYING, DIFFUSION, AND PROBLEMS OF PROOF

Currently, patent law's absolute liability principle allows recovery by A when B incorporates into his product something of value claimed in a property right owned by A, whether or not A can prove that copying in fact occurred. Regarding the ultimate question of how many inventions are actually copied, this Article makes two claims. First, this Part will highlight how proof of copying is more difficult than one might suppose, and indeed that "copying" describes a spectrum of activities that includes but is not limited to explicit, intentional duplication. Second, Part III points out that, once proof of copying becomes relevant, researchers may respond with costly efforts to prevent the inflow of information from "outside" their own organization—with disastrous consequences for open communication and technical progress.

A. INFORMATION DIFFUSION AND INADVERTENT PLAGIARISM

Over the years, scholars have discovered some fascinating things about the way technological information percolates among researchers. Some of these findings suggest that new ideas may diffuse in ways that are not highly salient, and therefore, are difficult to track. Additionally, in some related research, psychological studies have documented a persistent phenomenon called inadvertent plagiarism. When faced with a task requiring creativity, people regularly produce a piece of information they believe is original, but in fact clearly comes from someone else.¹⁵ Taken

15. One could object to deploying this literature in the patent context because of a differential in the magnitude of the information involved. Psychological studies involve recalling short bits of information—words or names. But this Article is talking about researchers on an R&D project, who might in some cases inadvertently "recall" entire inventions that were actually developed by and learned from others. A full claimed invention in many fields may have a significant number of technical elements. Therefore, how likely is it that someone would mistakenly create an entire invention without recalling where it came from? The answer is that the key innovative aspect of an invention may in fact be quite small. Inventions, and the patent claims that cover them, do often include numerous elements; but often only one is truly innovative—what patent lawyers call the "point of novelty." Even though an entire invention might be quite complex, the key component might involve but a small piece of technical insight. This is

together, this research points toward a mechanism by which new technological ideas might plausibly move from one set of researchers to another. The point is that this movement can take place at a level that does not draw much explicit attention. Indeed, research in this area uses a provocative term—diffusion studies. The process is just that: diffuse—extended over space and time, lacking a distinct, identifiable communicative moment. Given this mechanism, proof of copying may be very difficult to come by. In addition, even when one party in an IP case asserts that a specific idea was communicated at a distinct moment, it can be very hard to prove that such an event actually took place. This Section first describes the evidence on technological diffusion, and then turns to cases from various areas of IP law to demonstrate just how difficult it can be to prove copying in court. The upshot is that proposals for an independent invention defense seriously underestimate the difficulty of proving copying. It should be obvious that it is not enough to document every detail of an R&D project, with the hope of showing that an entire invention was developed exclusively in-house. These records would also have to trace in detail the provenance of every aspect of the project. And to be convincing, the researchers would need to show that they did not use information from any external source in devising any of the key steps in the project. If the burden was on the accused infringer to affirmatively establish independent invention, it might be a heavy burden. If, on the other hand, the burden were on the patentee to affirmatively prove copying, this would eliminate infringement liability in many situations where an accused infringer learned something valuable from an inventor, but it is hard to prove. In either case, the point is the same: the many and subtle ways information diffuses around a technical field make the entire question of independent invention/copying much more complex than it might seem. Real-world complexity, in other words, renders problematic the idea of an independent invention/lack of copying defense. The upshot is this: practical problems of proof supply a fairly strong rationale for the current absolute liability regime.

1. *Diffusion Studies*

There are three reasons to review what is known about diffusion. First, the basic premise behind arguments for an independent invention defense in patent law is that multiple firms independently create most inventions. An account of just how often firms learn of new technical information

precisely the type of information that is derived from others but misattributed to oneself in a predictable number of cases in the psychological literature.

from various sources may help to counteract the idea that most inventions are created in “splendid isolation.” Showing the ubiquity of information sharing undermines the implicit narrative of each firm as an individual research silo, separate and distinct from other firms and from the technological community as a whole.

Second, a careful description of the R&D process will clarify the often hazy process by which technical details are learned within an organization. This may introduce some doubt into what has heretofore been a seemingly straightforward story. It is possible that researchers themselves do not always have a clear memory or understanding of the precise origin of each of the many technical inputs that go into a typical R&D project.

The third reason relates to the mechanics of the independent invention defense. If this defense were available, as argued in Part III, it would push researchers toward a more isolationist approach to R&D. The best way to prove independent invention is to show there was little input from the outside world into the R&D project. Experience with “clean room” procedures, developed to avoid allegations of copying in copyright law (particularly in the computer software industry) bear this out. But this move toward “R&D isolationism” would come at a great loss. Diffusion of information is so commonplace in technological communities that it is easy to underestimate its significance. The extensive literature on technological diffusion brings home the importance of open and liberal information flows among research specialists.

Economists, sociologists, and others have for many years been interested in the process by which information and ideas move through society.¹⁶ One branch of this subfield takes a particular interest in the diffusion of new products. Some classic studies from the late 1950s, for example, documented the spread of new types of hybrid corn through the world of agricultural research and farming.¹⁷ Another branch of the

16. For an overview of the field, see EVERETT M. ROGERS, *DIFFUSION OF INNOVATIONS* (3rd ed. 1983). For a sense of the range of phenomena that have been studied, see Henry C. Finney, *American Zen's "Japan Connection": A Critical Case Study of Zen Buddhism's Diffusion to the West*, 52 RELIG. MOVEMENTS & SOC. MOVEMENTS 379 (1991).

17. See Zvi Griliches, *Hybrid Corn: An Exploration in the Economics of Technological Change*, 25 ECONOMETRICA 501 (1957); Zvi Griliches, *Hybrid Corn and the Economics of Innovation*, 132 SCIENCE, NEW SERIES, NO. 3422, (1960). Griliches emphasized economic incentives as the driving force behind adoption decisions, and hence diffusion rates. Forty-five years later, sociologists were still debating whether social factors, such as education level and an associated willingness to try new approaches, should also have a prominent place in explaining diffusion rates. See, e.g., Jonathan Skinner & Douglas Staiger, *Technology Adoption from Hybrid Corn to Beta Blockers* (Nat'l Bureau of Econ.

literature concentrates on the ways in which new techniques and practices in science and technology flow into academia and industry.¹⁸ Typically, studies of diffusion use aggregate measures to analyze adoption rates and diffusion patterns through industries, professions, technological communities, and entire societies over time.

Some scholars in this tradition look at the time it takes for an important innovation to be copied or duplicated by others. In his classic 1961 study on diffusion, economist Edwin Mansfield studied twelve important innovations in four industries. He found highly differential patterns of diffusion across industries: “The number of years elapsing before half the firms [in an industry] had introduced an innovation varied from 0.9 to 15 [years]”¹⁹ In a similar study, Gort and Klepper looked at 46 major product innovations between 1887 and 1960—spanning the era from the phonograph to the laser.²⁰ They conclude: “the interval required for successful imitation has systematically declined over time.”²¹ Average imitation time in the later examples they studied was 4.9 years, considerably less time than a full patent term (17 years during this era). This, of course, does not establish the fact of patent infringement, as duplication may be possible by inventing around, rather than actual appropriation of the claimed invention; however, it does provide some useful evidence on the general rate of diffusion.

Later studies confirm the drop in the elapsed time for information diffusion. One study from 1981 found that, within four years of patent issuance, 60% of 48 innovations in four industries had been imitated by the patentee’s competitors—imitated as in copied or intentionally mimicked.²² Seventy percent of the innovations studied were patented.

Research, Working Paper No. 11251, 2005), *available at* <http://www.nber.org/papers/w11251> [<http://perma.cc/8UT4-2788>], published at Proceedings, Federal Reserve Bank of San Francisco, <http://ideas.repec.org/s/fip/fedfpr.html> [<http://perma.cc/TU66-NYX3>].

18. *See, e.g.*, THOMAS J. ALLEN, *MANAGING THE FLOW OF TECHNOLOGY: TECHNOLOGY TRANSFER AND THE DISSEMINATION OF TECHNOLOGICAL INFORMATION WITHIN THE R&D ORGANIZATION* (1977).

19. Edwin Mansfield, *Technical Change and the Rate of Imitation*, 29 *ECONOMETRICA* 741, 744 (1961); *see also* Edwin Mansfield, *The Speed of Response of Firms to New Techniques*, 77 *Q.J. ECON.* 290 (1963) (explaining variance in firms’ adoption rates as a function of their size and predicted profitability factors).

20. Michael Gort & Steven Klepper, *Time Paths in the Diffusion of Product Innovations*, 92 *ECON. J.* 630, 638 (1982).

21. Gort & Klepper, *supra* note 20, at 640.

22. *See* Edwin Mansfield, Mark Schwartz & Samuel Wagner, *Imitation Costs and Patents: An Empirical Study*, 91 *ECON. J.* 907, 913 (1981).

Those who imitated them claimed that, while patents raised the cost of imitation (by 11%), they were able to duplicate the patented technology without, in their judgment, committing patent infringement. These claims to have “invented around” the patents were not tested in court, however. The general findings here are consistent with a later study involving 100 firms and thirteen innovations, which found that, on average, information about the commencement of development projects related to major technological innovations had “leaked out” to at least one competitor within a year.²³

But what does literature on how information diffuses through technical communities have to do with the copying of inventions, and more importantly, the question of whether patent law ought to require proof of copying? Here, additional research is useful. Studies by economic sociologists establish that, within R&D organizations, certain people are adept at “spanning boundaries” between organizations.²⁴ These people, sometimes also called “gatekeepers,”²⁵ acquire information from other departments or projects within an organization (i.e., a lab or company), but also from other “external” organizations, such as universities, professional societies, and related companies (vendors, customers, suppliers).²⁶ Moreover, some boundary spanners are consulted at an above-average rate for information by colleagues; these “communication stars” have been found to be “more externally oriented” than non-stars, meaning that they acquire more “general technical/scientific information,” including attendance at many more professional meetings than non-stars.²⁷ In general, according to another study, “a minority of professional employees, often called communicators, account for a majority of the information brought into and diffused throughout the organization.

23. Edwin Mansfield, *How Rapidly Does New Industrial Technology Leak Out?*, 34 J. IND. ECON. 217, 217 (1985) (“[I]nformation concerning the detailed nature and operation of a new product or process generally leaks out within about a year.”).

24. Michael L. Tushman & Thomas J. Scanlan, *Characteristics and External Orientations of Boundary Spanning Individuals*, 24 ACAD. OF MGMT. J. 83 (1981).

25. Michael L. Tushman & Ralph Katz, *External Communication and Project Performance: An Investigation into the Role of Gatekeepers*, 26 MGMT. SCI. 1071, 1071, 1073 (1980) (“[G]atekeepers are able to gather and understand external information, and subsequently they are able to translate this information into terms that are meaningful and useful to their more locally oriented colleagues.”).

26. Tushman & Scanlan, *supra* note 24, at 86 (“Extra-organizational communication [in this study] was further separated into professional (universities, professional societies) and operational (suppliers, vendors, customers) areas.”).

27. Tushman & Scanlan, *supra* note 24, at 88 (text and Table 1, “Technical/Trade Meetings” column).

Communicators are well-connected to information sources outside and inside the organization.”²⁸ The study found that high communicators were major contributors to innovative ideas within the organization.²⁹

The reason these technical R&D professionals are valued is that they are good at gathering information from outside a particular lab or project group and orally communicating it to others within the group.³⁰ They help others solve problems. They are repositories of useful information gleaned from a wide network of contacts. There is no record of these informal consultations. Information flows freely without any documentation. That is why these communicators are effective: they know a lot, learned from many sources, and can translate this knowledge into terms immediately useful to their peers.

The significance of this well-documented communication pattern is clear: within technological communities, ideas flow freely across organizational boundaries. And it means that the provenance of an idea may be obscured or lost in this process. Put simply, no one might remember where a particular piece of technical information originated. All that matters is learning the relevant solution that is required and plugging this solution into a particular problem. An alternative scenario, explored in depth by Professor Eric von Hippel of MIT, is where engineers at rival companies “informally” exchange proprietary information in a professional peer network.³¹ In this case, no records are kept precisely because the engineers involved do not want to leave a “paper trail.”

28. Robert T. Keller & Winford E. Holland, *Communicators and Innovators in Research and Development Organizations*, 26 ACAD. MGMT. J. 742, 742 (1983).

29. *Id.* at 742.

30. See, e.g., Michael L. Tushman & Thomas J. Scanlan, *Boundary Spanning Individuals: Their Role in Information Transfer and Their Antecedents*, 24 ACAD. MGMT. J. 289, 294 (1981), (noting study protocol: recording reports of oral communication among R&D personnel).

31. Eric von Hippel, *Cooperation Between Rivals: Informal Know-How Trading*, 16 RES. POLY 291, 291 (1987) (Finding “a novel type of cooperative R&D: the informal trading of proprietary know-how between rival (and non-rival) firms. I have observed this behavior to be widespread in [the steel ‘minimill’ industry] . . . [but] it may be present in many industries.”). In this study, von Hippel builds on an earlier landmark paper studying instances of informal knowledge-sharing among members of technical communities. See Robert C. Allen, *Collective Invention*, 4 J. ECON. BEHAV. & ORG. 1 (1983) (defining “collective invention” as invention growing out of the interaction between technical employees of multiple companies; and stating that “[t]he essential precondition for collective invention is the free exchange of information about new techniques and plant designs among firms in an industry”). Allen found that the iron industry in Britain between 1850 and 1875 was constituted so that collective invention took place in the area of iron blast furnace design. *Id.* at 3.

Consider also a study of engineering research labs, primarily in the electronics and aerospace industries, which provides evidence that formally written-down ideas may at times be borrowed without proper credit. In this study of a 36 member R&D team, author Thomas Allen found that informal engineering reports, produced by engineers mostly in external private companies, and circulated among various research teams within a company, were a major source of engineering information.³² But these reports changed hands constantly:

[External reports] are necessarily limited in number [of copies] on the one hand and widely needed on the other, resulting in a situation in which they are passed back and forth among colleagues over the course of a project. . . . [T]here is a body of informal documentation that is in a state of constant flow within the laboratory. In this way, a single report very likely reaches a fairly large audience in a short period of time.³³

Allen kept track of all reported communications between individuals on the team he studied. Most importantly, he found discrepancies among pairs of researchers with regard to the number of times they communicated with each other.

[Person A], for instance, reported communication with [B] more often than [B] acknowledged communicating with [A]. There are a number of possible reasons for this difference. Some people have better memories; others are more careful in responding to questionnaires [such as those used in the study]. There is a slight tendency for the lower-status member of a communication pair to be more likely than the higher-status member to remember a transaction.³⁴

This study reflects two things that are highly useful to the task at hand: first, some information is “in a state of constant flow” within an organization; and second, not everyone remembers communications about specific pieces of information.

Jaffe et al. observed a similar pattern in a study of knowledge diffusion out of a NASA research lab, the Electro-Physics Branch (EPB) of the

32. Allen, *supra* note 18.

33. *Id.*

34. Allen, *supra* note 18, at 143. There is a sense in this passage that social hierarchy may play a role, not only in memory, but also in patterns of attribution. This is closely related to ethnographic and sociological studies of the conduct of science. *See generally* BRUNO LATOUR & STEVE WOOLGAR, LABORATORY LIFE (1979); DOMINIQUE VINCK, THE SOCIOLOGY OF SCIENTIFIC WORK 51 (2010) (describing research on hierarchies in science).

NASA-Lewis Research Center in Cleveland.³⁵ This study carefully tracked interactions between members of the lab and outside organizations to investigate the relationship between patent citations and information transfer, or “spillover.” This necessitated an inquiry into interactions between lab members whose patents cited the work of other organizations; the question was, did the citations indicate a high degree of interaction and communication? In the course of the study, however, the investigator discovered a number of occasions where close contact did *not* lead to a citation, even though it might have been expected to do so. The authors concluded: “It is also clear that contact can occur and not generate any citations: 18 patents were found in the general area of EPB’s research by organizations that had had significant contact with EPB but that did not cite the EPB patents.”³⁶ Thus the inflow of information from EPB to these other organizations was never documented. There was no record that EPB ideas and information were acquired by these other organizations. Diffusion without documentation—this provides the key rationale for dispensing with proof of copying.

Thus, when it comes to diffusion of ideas, the higher the rate of information flow the less likely that any single piece of information will receive formal credit, or even be recalled later. High throughput probably explains most occasions when diffusion occurs, but credit or citations are omitted. The next Section, on fallible memory and its close cousin, inadvertent plagiarism, explores one pathway through which this can occur.

2. *Fallible Memory and Inadvertent Plagiarism*

First consider the all-too-human possibility of a fallible memory.³⁷ In their study of patents cited in other patents, and the degree to which cited patents demonstrate actual information diffusion (“spillover score”), Jaffe et al. found that:

35. Adam B. Jaffe, Michael S. Fogarty & Bruce A. Banks, *Evidence from Patents and Patent Citations on the Impact of NASA and Other Federal Labs on Commercial Innovation*, 46 J. INDUS. ECON. 183 (1998).

36. *Id.* at 197.

37. Students of copyright law will note right away the similarity between the ideas discussed here and copyright cases on “subconscious copying.” See, e.g., *Bright Tunes Music Corp. v. Harrisongs Music, Ltd.*, 420 F. Supp. 177, 180–81 (S.D.N.Y. 1976) (finding that George Harrison subconsciously copied his song “My Sweet Lord” from an earlier song called “He’s So Fine”); see also Carissa L. Alden, *A Proposal to Replace the Subconscious Copying Doctrine*, 29 CARDOZO L. REV. 1729 (2008) (collecting cases; suggesting that the doctrine as applied makes it too easy for copyright holders to establish liability).

[T]he spillover score [in cited patents] is higher if the cited patent is more recent. . . . For the citations, this is consistent with more recent patents being more useful and older citations being more likely to be nonspillovers included by the lawyer or examiner. It could also reflect the possibility that the inventor's memory of actual communication is better with respect to more recent technology.³⁸

Experimental psychologists call this general phenomenon implicit memory, or cryptomnesia. The latter term is defined as “generating a word, an idea, a song, or a solution to a problem with the idea that it is either totally original, or at least original within the present context.”³⁹ Cryptomnesia relates to the findings of an extensive literature on the psychological phenomenon of “inadvertent plagiarism.” Studies with experimental subjects routinely show that people in a small but appreciable number of cases will provide information they believe is original to them, but that has in fact been derived from another source. Commonly, these studies provide an original list of information, either from a group session or via a computer. Then participants are asked to supply new pieces of information in the general category of the original information supplied (e.g., types of sports, words beginning with the letters “BE”). A small but persistent percentage of people give as new information things that were supplied in the original list. And, interestingly, the effect is amplified by a delay between the information-supply phase and the new-information phase of the study. There is little evidence that the subjects are liars; researchers consistently conclude that the subjects really *believe* they are the origin of the information.⁴⁰

38. ADAM B. JAFFE & MANUEL TRAJTENBERG, PATENTS, CITATIONS, AND INNOVATIONS: A WINDOW ON THE KNOWLEDGE ECONOMY 394 (2005).

39. Alan S. Brown & Dana R. Murphy, *Cryptomnesia: Delineating Inadvertent Plagiarism*, 15 J. EXPERIMENTAL PSYCHOL.: LEARNING, MEMORY, & COGNITION, 432, 432 (1989).

40. See, e.g., Brown and Murphy, *supra* note 39; Patricia L. Tenpenny, Maria S. Keriazakos, Gavin S. Lew & Thomas P. Phelan, *In Search of Inadvertent Plagiarism*, 111 AM. J. PSYCHOL. 529 (1998); Richard L. Marsh & Joshua D. Landau, *Item Availability in Cryptomnesia: Assessing its Role in Two Paradigms of Unconscious Plagiarism*, 21 J. EXPERIMENTAL PSYCHOL.: LEARNING, MEMORY, & COGNITION, 1568 (1995); Richard L. Marsh & Gordon H. Bower, *Eliciting Cryptomnesia: Unconscious Plagiarism in a Puzzle Task*, 19 J. EXPERIMENTAL PSYCH.: LEARNING, MEMORY, & COGNITION, 673 (1993). In a related set of experiments, researchers found that when told of a striking event experienced by another, people will sometimes incorporate that event or something similar into their own autobiography. See Alan S. Brown & Elizabeth J. Marsh, *Evoking False Beliefs About Autobiographical Experience*, 15 PSYCHONOMIC BULL. & REV. 186, 186 (2008).

One early contributor to this literature described the phenomenon at work in fields requiring creativity:

When an event consists of information about some original creation in the world of art, literature and thought, and the logical memory of the event has deteriorated to the point at which the information is no longer recognized as a memory, cryptomnesia may give rise to *unintended plagiarism*. This happens when the logical memory is activated fortuitously or by some mental scanning process so that the information appears in consciousness as a cryptomnesically unfamiliar train of thought whose originality and value is appreciated. The train of thought may then be proudly, though mistakenly, claimed as a personal creation.⁴¹

A substantial body of research shows that many pieces of technical information are transmitted over distance in various ways, usually without direct personal contact. For example, there is some very instructive research on the origins and significance of patent citations. Adam Jaffe and Manuel Trajtenberg, in their monograph on patent citations and spillovers, found that patent citations are generally an effective, albeit “noisy,” indicator of spillovers. In other words, many citations are evidence that real information changed hands. Of immediate interest is their finding, based on extensive interviewing of researchers, that only 18% of patents cited by inventors stem from interpersonal interactions with other researchers.⁴² What about the other 82%? These, possibly, are the product

41. F. Kräupl Taylor, *Cryptomnesia and Plagiarism*, 111 BRIT. J. PSYCHIATRY. 1111, 1112–13 (1965) (emphasis in original). Taylor goes on to provide several examples involving Freud and Nietzsche (discovered by Jung). *Id.* at 1113.

42. Jaffe & Trajtenberg, *supra* note 38, at 390. Interpersonal interaction is an important diffusion mechanism. *See, e.g.*, ALAN HYDE, WORKING IN SILICON VALLEY: ECONOMIC AND LEGAL ANALYSIS OF A HIGH-VELOCITY LABOR MARKET XVII (2003) (“When the workforce is mobile, people know what is happening in other shops. Even temporary employees know useful things about how other firms in the industry do things.”). More generally, Jaffe & Trajtenberg report that 28% of inventors had “high familiarity” with the patents they cite. Jaffe & Trajtenberg, *supra* note 38, at 389. Of course, patent lawyers add many patent citations during the preparation of patent applications. The point is simply that there is a lot of technical information that flows to researchers in highly diffuse and generalized ways, and that it is the exception rather than the rule that direct personal knowledge and attribution are involved, even when a patent is cited. And, as the studies on inadvertent plagiarism show, people are loathe to give credit—either because they do not think it is due, or because it will detract from the credit they themselves would receive. *See, e.g.*, James B. Gambrell, *The Impact of Private Prior Art on Inventorship, Obviousness, and Inequitable Conduct*, 12 FED. CIR. B.J. 425, 448 (2002–2003):

of the many diverse and subtle diffusion mechanisms existing in technological communities: word of mouth, formal instruction, tacit knowledge picked up but not really remembered. Lack of direct interpersonal copying does *not* mean that someone who worked on an R&D project did not borrow ideas from an earlier inventor. There is a great gap between pure, unmediated invention and outright slavish copying. Patent law seeks to fill this gap by dispensing with proof of actual copying. And this, in turn, is embodied in the doctrine of absolute liability for patent infringement.

B. PROVING COPYING: CASES FROM OTHER AREAS OF IP LAW

Intellectual property doctrines, such as subconscious copying in copyright law, demonstrate that there can be copying or borrowing, even when it is difficult to prove. This Section draws evidence for this assertion from various areas of intellectual property law where proof of copying is required, or at least relevant. The two most important areas are (1) copyright, where independent creation is a defense and therefore copying of some kind must be proven; and (2) the derivation defense in patent law, which allows someone accused of infringing a patent to prove that the patented invention was copied (or “derived,” to use the polite patent term) from another person. A quick tour through these areas of IP law shows just how difficult it can be to prove actual copying conclusively. Copyright law settles for proof of what might be called “an opportunity to copy” in most cases;⁴³ and in cases of “striking similarity,” it sometimes dispenses completely with the need to prove the copier had “access” to the original. Moreover, under patent law’s derivation defense, the cases show how hard it can be to establish copying, even with solid evidence of extensive contact

When the issue of § 102(f) prior art surfaces during litigation to enforce a patent, the problems get more complicated. . . . [T]he patentee may discover some possible § 102(f) prior art in searching through its documents or as a result of detailed discussions with the inventors named on the patent and their professional counterparts. However the presence or possibility of § 102(f) prior art becomes an issue in the litigation, it is a complication that must be addressed by the patentee and his assignee. The first step is straightforward—be candid with your opponent and up front with the court. The reflex is to Deny! Deny! And Deny! This is a bad strategy and a shortsighted policy.

43. MELVILLE NIMMER & DAVID NIMMER, 4 NIMMER ON COPYRIGHT, § 13.02[A] (footnote omitted) (“[I]t is clear that, even if evidence is unavailable to demonstrate actual viewing, proof that the defendant had the opportunity to view (when combined with probative similarity) is sufficient to permit the trier to conclude that copying as a factual matter has occurred . . . perhaps the more prevailing definition of access [is] . . . the opportunity to copy.”). *Id.*

between an original inventor and an allegedly deriving patentee. Taken together, copyright and derivation demonstrate that proof of copying is no simple matter. While this cannot constitute a sufficient reason to adopt absolute liability, it should at least give pause to those who claim the obvious superiority of requiring proof of copying in patent law. The difficulty and complexity of the copying issue is an argument favoring absolute liability—at least for anyone concerned that copying takes place in a fair number of cases where proof is elusive.⁴⁴

1. *The Access Requirement in Copyright Law*

Judges recognize that inadvertent plagiarism can be the source of duplicated ideas. Though patent case law has occasionally appreciated this phenomenon,⁴⁵ the clearest example is the “subconscious copying” doctrine, which has been around in U.S. copyright law for many years. In general, copyright law requires a copyright owner to establish proof of access and substantial similarity. If the copyright owner cannot show the accused infringer had access to the copyrighted work, there is a high likelihood that the accused infringer independently created his or her work.

Yet copyright cases show a healthy respect for the difficulty of directly proving access. Courts will often infer access on the basis of circumstantial evidence.⁴⁶ But even when a copyrighted work has been widely disseminated, it may be difficult to show that a particular defendant had actual access. One way to deal with this problem is by determining the degree of similarity between the copyrighted work and the accused infringement. If they are “strikingly similar,” some courts are willing to infer access without more.⁴⁷ In other cases, however, courts have developed

44. Indeed, one’s attitude toward this issue bears heavily on how one views absolute liability. If undetected copying is a very large concern, then it makes sense to design the legal system to prevent this serious wrong in as many cases as possible—thus, absolute liability. But if, on the other hand, one is more concerned with imposing liability on someone who has independently invented, then a few cases of actual yet unproven copying may seem well worth the cost. As is so often true in IP law, the empirical questions are well nigh intractable, and when confronted with a tough case we may resort to policy arguments based on rights-based theories of IP or fairness/distributional considerations. Cf. Robert P. Merges, JUSTIFYING INTELLECTUAL PROPERTY (2011).

45. See, e.g., *Barrett v. Hall*, 2 F. Cas. 914, 924 (No. 1047) (C.C.D. Mass. 1818) (“[A] party may innocently mistake, as to the extent of his own claims. . . . [A] party may suppose, that he has invented, what in truth has been partly suggested by another mind.”).

46. See, e.g., Mark A. Lemley, *Our Bizarre System for Proving Copyright Infringement*, 57 J. COPYRIGHT SOC’Y U.S.A. 719, 720–21 (2009–2010). See generally Pamela Samuelson, *A Fresh Look at Tests for Nonliteral Copyright Infringement*, 107 NW. U. L. REV. 1821, 1824 (2013) (describing time-honored tests for nonliteral infringement).

47. See, e.g., *Three Boys Music Corp. v. Bolton*, 212 F.3d 477, 484 (9th Cir. 2000).

the doctrine of “subconscious copying.”⁴⁸ Given a high degree of similarity, plus a showing of an “opportunity to copy,” courts infer access and conclude that the copyrighted work lodged in the infringer’s unconscious, only to emerge later at the infringer’s putative moment of creation. The doctrine has a long history in copyright law, going back at least to Judge Learned Hand in the 1920s.⁴⁹ It continues to evolve, and cause controversy.⁵⁰

Subconscious copying bears a close resemblance to the results of extensive studies on inadvertent plagiarism. The cases and the studies certainly suggest a plausible mechanism by which copying or borrowing may occur. And there is no reason why it should be any less common in technology fields covered by patents than in creative fields governed by copyright law. In this light, patent law’s absolute liability standard might be defended yet again. In cases where borrowing occurs but is hard to prove, absolute liability serves the same function as the doctrine of subconscious copying in copyright law. It establishes liability even where direct evidence of copying is not available.

2. *Derivation in Patent Law*

a) Full Derivation (35 U.S.C. § 102(f))

Obviously, all sorts of learning and information exchange occur in settings other than face-to-face meetings. Yet in virtually every reported case on the issue, claims of derivation involve a face-to-face meeting of some kind. This is in large part a function of the standard of proof in derivation cases. The person trying to invalidate a patent by asserting derivation must show both (1) a prior conception of the later-claimed invention (i.e., conception by the “derivee”), and (2) that this complete

48. See, e.g., *ABKCO Music, Inc. v. Harrisongs Music, Ltd.*, 722 F.2d 988 (2d Cir. 1983).

49. See *Fred Fisher, Inc. v. Dillingham*, 298 F. 145, 147–48 (S.D.N.Y.1924) (L. Hand, J.) (“Everything registers somewhere in our memories, and no one can tell what may evoke it Once it appears that another has in fact used the copyright as the source of his production, he has invaded the author’s rights. It is no excuse that in so doing his memory has played him a trick.”). In *Fred Fisher*, Judge Hand found that the similarities between the songs “amount[ed] to identity” and that the infringement had occurred “probably unconsciously, [based on] what he had certainly often heard only a short time before.” *Id.* at 147.

50. Carissa L. Alden, Note, *A Proposal to Replace the Subconscious Copying Doctrine*, 29 CARDOZO L. REV. 1729 (2008).

conception was fully communicated to the later patent claimant (the “deriver”).⁵¹

Proof of prior conception comes with all the rigmarole of a patent priority contest, including a heavy burden of proof and a demanding corroboration requirement.⁵² These issues are some of the reasons that patent interferences under the 1952 Patent Act are so notoriously complex; they in no small measure contributed to the scrapping of the 1952 Act’s “first to invent” standard in favor of a (modified) first-to-file system under the America Invents Act of 2011.⁵³

Once the “derivee” proves prior conception, he then faces the daunting task of showing full communication of the invention to the deriver. Many derivation cases end right here. The standard is strict: the *full* invention must be communicated to prove derivation.⁵⁴ Thus courts decline to find derivation when the proven details of the alleged derivee’s communication diverge from the claims of the alleged deriver’s patent.⁵⁵

More generally, the § 102(f) cases show many instances where a prior inventor taught or influenced later researchers in ways that fall short of actual legal derivation. For example, in one case a Florida State University research lab was shown to have developed a method for making a class of chemical compounds. Another research group, working at a company called American Biosciences, later patented similar compounds using the

51. *Creative Compounds, LLC v. Starmark Lab.*, 651 F.3d 1303, 1313 (Fed. Cir. 2011) (“In order to establish derivation, [a challenger to the validity of a patent] was required to ‘prove both prior conception of the invention by another and communication of that conception to the patentee.’ *Eaton Corp. v. Rockwell Int’l Corp.*, 323 F.3d 1332, 1334 (Fed. Cir. 2003).”).

52. *Egnot v. Looker*, 387 F.2d 680 (C.C.P.A. 1967); *Williams v. Clemons*, 19 F.2d 798 (D.C. Cir. 1927) (applying the strict corroboration requirement from prior invention cases such as *The Barbed Wire Patent*, 143 U.S. 275, 285 (1892) to derivation cases); *see also Price v. Symsek*, 988 F.2d 1187 (Fed. Cir. 1993) (an inventor’s testimony, standing alone, cannot support a claim of derivation under § 102(f)).

53. *See* Robert P. Merges & John F. Duffy, *PATENT LAW AND POLICY: CASES AND MATERIALS* at 341–44 (6th ed. 2012).

54. *See, e.g., Hedgewick v. Akers*, 497 F.2d 905, 908 (C.C.P.A. 1974) (“The issue of derivation is one of fact, and the party asserting derivation has the burden of proof. . . . Derivation is shown by a prior, complete conception of the claimed subject matter and *communication of the complete conception to the party charged with derivation.*”) (emphasis added); *see also International Rectifier Corp. v. IXYS Corp.*, 361 F.3d 1363, 1376–77 (Fed. Cir. 2004) (no derivation because no proof of prior conception of invention identical to the one claimed).

55. *Pentech Intern., Inc. v. Hayduchok*, 1990 WL 180579, at *8 (S.D.N.Y. Nov. 12, 1990) (Leval, J.) (“Because I find that plaintiff has not proven prior invention of the patented technology, the derivation challenge fails. There has not been an adequate showing that the Reinol formulas were identical to those . . . in the patent.”).

Florida State method.⁵⁶ One inventor on the American Bioscience research team had formerly worked in the Florida State lab that developed the chemical method. Other inventors on the team attended a conference at which the head of the Florida State lab presented. Even so, the court held that there was no improper derivation, and no need to add any Florida State inventor to the American Bioscience patent. One reason given was that no one in the Florida State lab directly communicated the specific compounds claimed in the American Biosciences patent.⁵⁷ Under very similar facts, courts have refused to find co-inventorship in addition to rejecting charges of derivation.⁵⁸

Thus, the stringency of proof required for derivation limits this doctrine to a small number of cases. The patent challenger asserting derivation must demonstrate immediate, direct, usually face-to-face communication. Without something more—without a liberal rule covering other sorts of influences and communications—many instances of learning from another would never be detected or proven in the patent system. The absolute liability standard again serves as a response to the unfairness that would result in its absence.

56. Bd. of Educ. ex rel. Florida State Univ. v. Am. Bioscience, Inc., 333 F.3d 1330 (Fed. Cir. 2003) [hereinafter Bd. Of Educ.]

57. Okwedy v. Molinari, 333 F.3d 339, 342 (2d Cir. 2003).

58. *See, e.g.*, Cook Biotech Inc. v. ACell, Inc., 460 F.3d 1365 (Fed. Cir. 2006) (Evidence that alleged co-inventor discussed technique with named inventor on patent was not sufficient to create a genuine issue of material fact “as to whether [the alleged co-inventor] ‘contributed to the conception of the claimed invention . . .’ by sharing his knowledge of [the claimed technique], some of which is reflected in [a document authored by the alleged co-inventor].”). Professor Chisum discusses cases where a later inventor builds on the published work of an earlier inventor, but there is no coordination between the two; there can be no joint invention under the Patent Act in such cases:

[T]here can be joint invention of subject matter Y when (1) inventor A works on a problem up to the stage X; (2) A turns the partial solution over to inventor B in a remote but nonpublic manner; and (3) B uses X to perfect Y. However, there can be no such joint invention when A develops X and then publishes or otherwise makes it known and B independently uses X to develop Y.

Donald Chisum, CHISUM ON PATENTS § 2.02[2][f] (citing B.J. Serv. Co. v. Halliburton Energy Serv., Inc., 338 F.3d 1368 (Fed. Cir. 2003) (later inventor claimed method of using material invented by earlier inventor; no evidence of any collaboration between them; so no joint invention). It follows that granting co-inventorship status is not a viable way to capture the influence of an earlier contributor when the later contributor is influenced, or learns from, the earlier contributor, but does not, strictly speaking, copy from her. Where appropriate, patent infringement liability is the only way to capture this type of contribution.

b) The Tricky Case of Partial Derivation

Requiring proof of copying means requiring proof that an accused infringer intentionally duplicated someone else's entire invention.⁵⁹ But what if an accused infringer copies less than an entire invention? That cannot comfortably be called copying—even if what is borrowed is the key technological insight that adds value to the patented invention and the accused infringer's product. One rationale for dispensing with proof of copying, then, is to provide compensation when what has been copied is less than an entire invention.

Derivation law again provides some insight. To invalidate a patent under § 102(f), one generally must show that the patentee was told about the *complete* invention he or she later claimed. However, courts have found that § 102(f) prior art is relevant under the nonobviousness provision, § 103. Consequently, a partial deriver will not receive a patent if his or her claim is obvious in light of the combination of the derivee's § 102(f) disclosure and other prior art.⁶⁰ But a partial deriver *will* receive a patent if he adds enough to make his claim nonobvious in light of what was disclosed by the derivee. In this scenario, where the deriver adds a nonobvious contribution to information acquired from the derivee, absolute liability protects the valuable information disclosed by the derivee. The fact that the deriver in this situation added something valuable does not detract from the point that he or she learned something of significant value from the derivee. The deriver might even obtain a patent for his or her valuable variant, but this patent does not protect the deriver from liability for infringing the derivee's patent. Such a “blocking patent” situation leaves both contributors with a property right. But absolute liability protects the contribution of the derivee.

Infringement doctrine protects the derivee when § 102(f) does not. That is because infringement doctrine dispenses with the need to show complete copying, while § 102(f) requires it. To see how separate

59. It has been observed that the America Invents Act of 2011, in eliminating § 102(f) on derivation, runs the risk of requiring the PTO to issue many patents on inventions that are at least partially derived from others. See Josh D. Sarnoff, *Derivation and Prior Art Problems with the New Patent Act*, 2011 PATENTLY-O PAT. L.J. 12, 16 n. 13 (“[A] non-obvious variant of a derived invention may or may not be severable from the original invention” . . . [And] a patented non-obvious variant may not infringe the originator's invention, but if it does, it may become a blocking patent (if the original invention issues as a patent)”). Once again, absolute infringement liability allows us to reach through the act of copying and impose liability where the derivee learned something substantial and important.

60. See *Oddzon Prod., Inc. v. Just Toys, Inc.*, 122 F.3d 1396 (Fed. Cir. 1997).

patenting of an invention protects an inventor when invalidity of the deriver's patent does not, it is important to understand that there are really two disparate types of partial derivation. In one, the deriver appropriates the entirety of an earlier inventor's contribution—lifts it whole hog in other words—and then, in addition to the entire prior earlier invention, adds other elements or features. Infringement law easily addresses this “plus” type of partial derivation. As long as the deriver copied all of a prior invention, and the deriver's marketed product includes that prior invention in its entirety, what other features the deriver may have added are irrelevant. A court will find patent infringement.⁶¹ So in cases of “plus” type partial derivation, infringement liability will follow regardless of whether patent law requires proof of copying. A deriver may also be able to prove that the derivee's patent is invalid under § 102(f) in such a case. The key will be whether the deriver can show that it communicated the entire invention to the derivee, and that the derivee copied it in making its own invention.

This is not true in the second type of cases, however. In these cases § 102(f) is no help to the derivee; the only hope is for the derivee to have its own patent. In these cases the deriver copies only a portion of a derivee's invention. This “fractional” type of derivation will not result in invalidity under § 102(f).⁶² But patent infringement under absolute

61. Technically, liability in the case of a “plus” type infringement depends on how the derivee/patentee drafts her claims. For “plus” type infringement to give rise to liability, the patentee must claim her invention in a way that contemplates liability even when the infringer adds components or elements beyond what was derived from the derivee. Such claims are said to be “open” claims, and they usually include the word “comprising” in them. See Merges & Duffy, *supra* note 53, at 28. Because open claims are broader, they are preferred by patent drafters. And empirically, they are much more common than “closed” claims, which usually include the limiting term of art “consisting of.” A rough estimate of the ratio of open (“comprised”) claims to closed (“consisting of”) claims can be made based on patents issued in 2014. In that year, there were 294,427 patents issued that included “comprising” in the claims, and only 5,060 that included the words “consisting of.” These data were obtained using the www.uspto.gov patent search website. Note that the “consisting of” patent count was constructed of (1) all patents with “consisting of” in the claims, minus (2) all patents that included the phrase “group consisting of.” This latter claim format, called a Markush group by patent drafters, is fairly common, so including patents with this language in the claims badly overstates the number of patents with closed claims. Having said that, it should be apparent that there will be a small number of patents that include “consisting of” in a limiting sense, together with claims that include the Markush terminology “group consisting of.” Put simply, the method used for the count reported here may slightly *undercount* the number of patents that include “consisting of” in their claim language.

62. See, e.g., *Bd. of Educ.*, *supra* note 56 (alleged co-inventors, and hence derivees, would have been properly named as co-inventors on deriver's original broad claims; but

liability might apply in such a case. If the deriver is found to be using something that meets all the elements of the derivee's patent claim, infringement will follow—even if the derivee cannot prove that every element in its claim was communicated to the deriver. In other words, it does not matter what portion of an invention is copied. Absolute liability means that the infringer will be liable when he or she adds one or more elements to the portion of an invention derived from the patentee. Absolute liability in this sense fills in the gaps between the derived portion of an invention and the full version claimed in the deriver's patent.

To summarize: these cases of patent *validity* under § 102(f) illustrate several points about liability for patent *infringement*. In every case of “partial derivation,” the party arguing invalidity under § 102(f) would be unable to invalidate the deriver's patent. This result stems from the rule requiring the derived invention as claimed to incorporate the complete invention as disclosed by the derivee. But when we assume that the derivee in these cases also has a patent, in addition to the one held by the deriver (and whose validity is in doubt under § 102(f)), we see how absolute liability helps the derivee. Derivation doctrine precludes a derivee from invalidating a deriver's patent if the deriver's claims cover something other than the complete invention communicated by the derivee, or a nonobvious variant of that communicated invention. In this situation, however, because patent law does not require proof of copying, the derivee with a patent covering his or her invention can successfully assert it against the deriver. Disclosure of less than the full invention is irrelevant to infringement liability. Indeed, there is no need to prove any disclosure at all—that is the purpose of absolute liability. Again, by dispensing with proof of copying, the law will capture cases where someone acquired valuable information from someone else, even though the details of the borrowing fall short of complete “copying.”

Those who want patent law to require proof of copying for infringement liability might argue that infringement is no less serious an

they did not need to be added as co-inventors on deriver's subsequent narrower claims, despite the fact that evidence showed the deriver learned general techniques for making claimed compounds from the derivee). *Cf. Alexander v. Williams*, 342 F.2d 466 (C.C.P.A. 1965) (Invention had two elements, X and Y; inventor A communicated element X to inventor B, and B thereupon conceived of element Y. The court held that B had not established priority in an interference (priority contest) between A and B, because of A's prior communication of element X). Note that in *Alexander*, if inventor B made, used or sold a product that included all the elements of one of A's patent claims, B would infringe, despite the fact that B only partially derived its product from A's prior invention.

offense than derivation—and therefore that the stringency of proof should be the same in both cases. If the deriver is a “thief,” then so is the infringer. Under this argument, the “thief” label should not be attached lightly; it should require equal levels of proof for both derivation and infringement. By placing one’s name on a patent, one does not deserve worse treatment than when incorporating a patented invention in a commercial product. But in any event, we can use the differences in the two areas of law to make an interesting point. Derivation requires stringent proof—and in the process, situations where researcher A clearly teaches, influences, and contributes to B’s research do not always amount to derivation under § 102(f). But the same levels of influence, teaching, and contribution may permit A, if he or she has a patent, to successfully sue B for patent infringement. Patent law in effect makes it easy for a patentee to capture instances of less-than-complete copying. But the law makes it hard for a person to invalidate another’s patent under the same circumstances. The derivation cases thus show, first, that proof of copying is difficult. And second, they may help to show why, for purposes of infringement, the law dispenses with proof of copying altogether. Even where B incorporates less than A’s full invention, the law may want to compensate A for the portion of its invention that B did acquire. It does this by dispensing with the need for proof of full acquisition (copying). The law presumes, in effect, that if B is selling a commercial product that is covered by one or more claims in A’s patent, B has learned enough from A that patent infringement liability ought to follow. Both the difficulty of proving full copying (the partial derivation cases), and the plausible fairness of requiring compensation for less-than-complete acquisition, argue in favor of this result.

C. FROM COPYING TO ACQUIRING

A few times now, this Article has argued that the transfer of technological information from one researcher to another occurs along a spectrum of related acts. There is deliberate copying of a complete invention. Then there is “copying plus,” i.e., deliberate copying plus new contributions from the copyist. But then there are also less conscious, less blatant ways that a prior inventor can communicate valuable information to other researchers. The field of diffusion studies names them well: they are mechanisms by which an original idea is spread around a group of interested people. When one of these people learns from the prior inventor, or borrows from what is taught, his or her act is usually not referred to as copying. Diffusion studies bear this out. They analyze the flow of information through a technical community, rather than discrete

acts of learning, duplication, or copying. Implicitly, the field understands that ideas percolate and spread through a wide variety of mechanisms.

In a way, the absolute liability standard in patent law reflects this way of thinking. It dispenses with the need to prove actual copying. And in so doing, it leaves room for all manner of information transmission. Patent law tries to capture a wide spectrum of ways an inventor might teach or influence others in a technical community. It eschews reliance on one discrete endpoint in the spectrum of influence (i.e., actual copying) by crediting an inventor when his or her valuable ideas have influenced or taught others in ways that are indirect, subtle, and hard to prove. Patent law, in other words, attaches liability for various types of information acquisition, and not just for direct copying.

There are many reasons to avoid outright theft of another's invention: trade secret protection; derivation proceedings in patent law; and concern for willful infringement (with the potential for treble damages). But it takes a great deal of solid evidence to establish any of these bases of liability. There are several species of idea acquisition that fall well short of the high standards required in these areas, and these species ought to give rise to legal liability because the acquisition involved is nevertheless significant. These types of borrowings do not trigger legal liability under derivation, trade secrecy, and the like. They are, in effect, the unique species of acquisition that are exclusively protected against by patent law's general standard of absolute liability.

Several distinct types of acquisition compose this residual category. One is borrowing that falls short of explicit, intentional copying. This runs the gamut from partial, unacknowledged borrowing; to inadvertent borrowing; to completely subconscious (and one might even say unintended) acquisition. The studies described earlier capture these categories well. For each, proof of deliberate copying will not be possible because no deliberate copying occurred in the first place.

The second type involves partial acquisition—borrowing of less than a full, coherent inventive concept. As described earlier, patent law's derivation defense does not reach this activity. But absolute liability does. By dispensing with proof of copying, it covers the case where a person acquires key ideas from someone who later obtains a patent.

To generalize, then, absolute liability ends up covering residual categories of information acquisition. It makes borrowers liable when they have borrowed, but (1) the patentee cannot prove borrowing; (2) the borrowing was inadvertent; or (3) the borrowing was partial, not complete. Because acquisition in these instances should arguably still give rise to liability, absolute liability steps in. It covers instances where there has been

some significant degree of idea transmission, but the law would not otherwise impose liability.

Implicit in this formulation is a preference for the law to favor the information-supplier in these scenarios. But why should we be so solicitous of this person, at the expense of others who merely “learn” or “acquire information” from him or her? Admittedly, this is perhaps a weak point of this Article’s argument. The assumption here is that, where possible, we want to credit (and compensate) originators of information, even when proof of learning or acquisition is difficult to come by. This argument assumes a bias in favor of those who originate information, and upon whose work others build. To make a full case in favor of absolute liability would require extensive proof that this bias in favor of idea originators is warranted. This might prove difficult, as the free flow of information is a point on which much of the argument here is built. If absolute liability encourages the free flow of information, then why not eliminate liability completely in more cases, and thereby encourage information flow even more? It is enough, in keeping with the spirit of this Article, to put forth a set of arguments in favor of absolute liability. The point is merely to put something positive on the other side of a scale that has been tipping heavily in favor of eliminating absolute liability in recent years. The full case would be hard; so we begin with an easier task—saying a few kind words in favor of absolute liability. It is a key assumption in this argument that we want to favor idea originators. Perhaps further empirical research will illuminate the issue. In the meantime, it is best to proceed on the basis that rewarding idea originators is a good idea, and note that absolute liability is a good way of achieving that goal.

III. EXCESS PRECAUTION: THE COST OF REQUIRING PROOF OF COPYING

The first argument was that copying is hard to prove, and that it is indeed more of a spectrum of related behaviors than a single discrete event. The second argument is that if the law requires proof of copying, many firms may well invest in elaborate systems to *disprove* that copying has taken place—and that these systems are bad for society. When IP owners must prove copying, people who are likely to be targeted for lawsuits will take steps to keep “outside” information from entering their organizations. This cuts down on the flow of information across

organizations, which in turn suppresses the rate at which information spreads in a field or industry.⁶³

Paradoxically perhaps, a strict liability standard actually encourages communication. Under strict liability, who communicated what to whom, and when, are all irrelevant. If the volume of communication is irrelevant to one's chances of liability, then there is no patent-related disincentive to communicate. Talking and sharing freely have no identifiable cost for purposes of patent law. But there are benefits. So people and companies tend to share.

A. THE TORT THEORY PERSPECTIVE

One way to capture this thought is to describe the copying rule in patent law in terms familiar to students of tort law. To do so, this Article treats the accessing of technical information produced by third parties as a potentially risky activity. It is commonplace for a researcher from Company A to read a technical paper written by researchers at Company B, and incorporate the information from B's paper into a product later sold by A. If B patents the information published in its technical paper, A may find itself liable for patent infringement. From the point of view of potential liability, A's reading of B's technical paper creates a risk of harm to B. Of course, B may choose not to patent. Or A's product might benefit from B's information, yet not meet all the elements of any claim in B's patent. Which means that in reality, A will have to analyze the situation taking into account both A's chances of infringement and B's "propensity to patent."

Once things are framed this way, we can look to tort theory for guidance on the best way to handle this risk. Before doing so, I have two quick points to make. First, I want to recognize that the "risk" discussed here is different from the types of risks that are usually associated with torts. Tort law typically concerns physical risks, or at least risks to interests that seem quite basic. Car accidents are the classic example. When

63. One response to these costs is simple and drastic: weaken IP rights, prevent them from being applied in some fields, or perhaps do away with them altogether. *See, e.g.,* KAL RAUSTIALA & CHRISTOPHER SPRIGMAN, *THE KNOCKOFF ECONOMY* (2012) (describing industries that thrive despite the absence of effective IP rights); MICHELE BOLDRIN & DAVID K. LEVINE, *AGAINST INTELLECTUAL MONOPOLY* (2008) (general case against IP rights); *See also* Robert P. Merges, *Economics of Intellectual Property Law*, in *OXFORD HANDBOOK OF LAW AND ECONOMICS* (Francesco Parisi, ed., forthcoming), <http://ssrn.com/abstract=2412251> (reviewing positive economic case for IP rights, and noting limitations of research showing that creativity flourishes in some industries despite the absence of IP rights).

discussing optimal tort rules, it is plain to everyone that cost-effective minimization of the risk of physical injury is an important, or even essential, interest that society ought to care about a great deal. Patent infringement is entirely different. It does not involve physical harm. Moreover, it does not seem to involve an interest that is nearly as important or essential as those at stake in many tort cases. Indeed, it might be argued that the “harm” and “risk” of patent infringement is more the product of a legal policy than an affront to a central and personal interest such as physical integrity.

The answer to this objection is twofold. It requires that we recognize the importance of innovation to economic well-being, and of economic well-being to other important social values such as stability and opportunities for self-advancement. The interests involved, in other words, *are* important ones. And it requires an acknowledgement that though tort law canonically deals with physical injuries, it also embraces a wide spectrum of economic injuries. Perhaps some of the intuitive force of tort law is weakened when it concerns injuries and risks that are purely economic. But because tort law provides powerful tools for thinking about risks and harms, it is worth moving forward with an analysis of patent infringement from a tort-centric perspective. Although the interest at stake in patent infringement may seem more “socially constructed,” and somehow less “essential,” it is nevertheless an important interest. And it is therefore worth looking at how harms to that interest are, and should, be handled by the law.

Another fundamental objection to the basic approach used here has to do with the nature of the “harmful activity.” As previously stated, the risk or harm we are talking about occurs when A uses B’s technical information. This often begins with A reading or otherwise learning about technical research that B has performed. Some readers may balk at a framework that takes the acts of reading or learning as a “harm” or “injury.” It may seem wrongheaded, offensive, or even vaguely unconstitutional to treat *reading* as a potentially harmful act. While this view deserves some sympathy, consider also two important points. One is that reading alone never creates infringement liability; patent infringement occurs when an infringer makes, uses, or sells a product incorporating the patented invention. So it is the *use* of the acquired information that triggers legal liability. And second, there are other areas of law where simply accessing information is seen as a harm. Trade secret misappropriation, insider trading, and access via computer hacking are examples. Further afield are national security-related offenses. The point is that despite the general disposition of society and our legal system, which

broadly and generally favors free and easy access to as much information as possible, there are situations where reading and learning about a particular piece of information can lead to legal liability.

Of course, even under absolute liability, one's chances of liability for infringement increase when one receives information from a source that will later obtain a patent. So it might be asked: why don't people currently screen out information they receive from sources that are likely to file for patents? Why don't they keep out information from "patent likely" sources, and allow in only information from "patent unlikely" sources? The response to this starts with two basic facts. First, the probability that infringement liability will follow directly from receiving and using any *particular* piece of information is quite low. The cost of screening information by source—of differentiating between information transmitters that are at high likelihood to later obtain patents versus those at low likelihood—would be quite high. And, crucially, keeping out information from a source that is likely to patent only *reduces* the chance of legal liability. If in-house researchers learn the same information from another source, or create it independently themselves, the company could still be liable. Screening outside information under the current rules eliminates one potential source of legal liability. But it cannot eliminate the threat of liability altogether because under the current rule, liability is independent of any particular source of information. Eliminating a particular source of information might help lower the risk of liability—but only to the extent that this particular source is unique, and only if in-house researchers will not duplicate the information from the source or receive it from somewhere else.

But a move away from absolute liability would significantly increase the payoff from screening. It would make it much harder to prove liability when a company systematically weeds out all information from external sources. Screening under today's absolute liability standard merely reduces the potential incidences of idea acquisition. It does not affect liability once information held by a particular outside source is either acquired elsewhere or is recreated in-house. This changes under an independent invention defense. The ability to escape liability by showing independent invention, in effect, makes in-house research completely safe from any risk of liability—assuming one can prove that no external information ever entered the research process.⁶⁴ Liability follows only after proof that

64. To be precise, independent invention privileges (1) truly in-house research, as well as (2) externally acquired information for which there is no proof of access. The

invention came from a particular source—an external researcher who later patented the information. Eliminate proof of access to this source, and you eliminate liability. Put another way, screening under independent invention does not just eliminate one source of duplicated information, it eliminates all risk of liability. It privileges in-house research, which happens to duplicate external information, by making it immune from legal liability. This creates a much bigger incentive to eliminate outside information.⁶⁵

evidentiary burden of proving access to external information effectively renders impossible-to-prove external access the equivalent of in-house research.

65. To be specific: information from outside a firm can do one of three things. It can (1) bring general benefits, without increasing the risk that the recipient will infringe any patents owned by the sender of the information; (2) it can cause infringement, when the recipient copies it and would not have independently invented it; or (3) it can increase the risk of liability when the recipient does *not* copy it, by undermining the recipient's ability to prove that it independently invented. Under either legal rule—absolute liability or proof of copying/independent invention—the decision whether to screen information from outside the firm involves a tradeoff between (1) and (2). Because under the existing (absolute liability) rule, screening appears to be very rare, we can conclude that firms find (1) outweighs (2). My argument thus concerns the addition of (3) to the equation. Screening under a proof of copying/independent invention rule brings an additional benefit over the current situation. It eliminates liability in some cases where it cannot be eliminated under absolute liability. The magnitude of this marginal effect will determine whether my concern about extra screening is valid. The social welfare effects depend on two factors. First, the firm faced with the decision whether to screen or not must do its homework correctly. A crucial issue is how often external idea acquisition will lead to legal liability. If a firm overestimates this (which could well happen due to the high salience of very large patent damages in rare cases), it might decide to screen when that is not in fact efficient. Second, the firm might underestimate the lost value of external information. Firms so frequently disregard valuable external information that the phenomenon has a name: the “not invented here syndrome.” See Ralph Katz & Thomas J. Allen, *Investigating the Not-Invented-Here (NIH) Syndrome: A Look at Performance, Tenure and Communication Patterns of 50 R&D Project Groups*, 12 RES. & DEV. MGMT. 7 (1982). If the information that is screened out to reduce liability would have yielded large firm-level benefits (in the form of less duplication of effort, and the stimulation of improvements and variants that would have resulted if the in-house team had access to the external information), then screening may involve difficult-to-estimate, but very real costs. As a separate matter, it may be costly from a social point of view. Screening may reduce the incidences of external copying below the socially optimal level. This will occur when the screening firm does not bear all the costs of screening (e.g., its in-house team may miss out on the chance to build on external information that would be hard for the firm itself to capitalize on, but that would benefit society). It is exceedingly difficult to estimate these “lost potential social spillover costs” that occur when an in-house team is denied access to external information. All that can be said is that under current conditions, we observe that there is a very large volume of information sharing among researchers; and there are very large social spillovers from much organized research activity. We might well be loath to disturb the legal rules that produce this favorable equilibrium.

This shift *might* be enough to change the behavior of firms—to move them to an aggressive use of information screening. It would still be true that the probability of infringement is low for each piece of information received, but the ability to eliminate liability in some cases by screening might make it a good idea to screen. Once again, efforts to reduce copying of information might make sense from the perspective of each individual firm, but would entail high social costs. Paradoxically, and against the weight of scholarship on this topic, requiring proof of copying might make duplicative effort more common, not less common.⁶⁶

The logic is simple. People will invest in precautions against copying so long as the expected payoff from these investments exceeds the (probability-weighted) expected loss from patent infringement liability.⁶⁷ They will ask only: is the potential cost of infringement payouts to prospective patentees greater or lesser than the cost of preventing copying? This prevention cost would include both the immediate costs of setting up and running a screening system, as well as the loss of value from not being able to use information that comes from patent-likely sources—with this latter cost itself being comprised of two components: (1) invention opportunities lost or forgone because of the absence of a key piece of information that would have come from outside, had it not been screened out; and (2) the added cost of recreating information in-house that would have been obtained for free from external sources, had screening not been imposed. From the point of view of a single research entity, if the sum of these prevention costs is lower than potential infringement payouts, then it should choose preventive screening. Given high damages awards in patent cases, this is entirely possible.⁶⁸

66. The alert reader will note that here we are equating the requirement to prove copying with an independent invention defense. Of course, there is a subtle but perhaps important distinction between the two. Proof of copying puts the burden on the patentee to prove that the infringer copied. Independent invention puts the burden on the infringer to prove that it in effect did *not* copy—that it independently invented. In either case, the accused infringer would benefit from being able to prove that it did not have access to—and so could not have copied—the patentee’s invention. So in this sense they are equivalent.

67. This is a simple application of the Learned Hand negligence rule first set out in *U.S. v. Carroll Towing*, 159 F.2d 169, 172 (2d Cir. 1947) (L. Hand, J.). See generally Richard A. Posner, *Instrumental and Noninstrumental Theories of Tort Law*, 88 IND. L.J. 469, 469 (2013) (describing Hand formula).

68. Note that the torts analogy suggests a completely different approach to the copying question. As with other harms, the law seeks to affect the ex ante incentives of decision makers as a way to encourage just the “right” amount of harm. We typically care about the calculus of the decision maker whose activities create a risk of harm and who must choose how much to invest in precautions beforehand. So arguably, the real

An individual research unit may well get it wrong when making the decision about investing in precautions. That is because not all the variables mentioned in the prior paragraph are equally salient to the average decision maker. Past patent infringement awards, for example, may stick out in a decision maker's mind; press accounts of whopping damage awards are fairly common, and they rarely mention that the reported damages are far in excess of averages or medians. At the same time, cost component (1) above—the cost of foregone invention opportunities that follows from screening out potentially valuable information—is very hard to measure, and may well fall victim to the well-known propensity of research units to undervalue outside information.⁶⁹ Likewise, it may be difficult for a firm to figure out which pieces of important information that came from in-house sources would have been instead received from outside in the absence of screening. This all adds up to a single point: the benefits of preventive screening may be quite visible or salient, but the costs may be hard to fathom. And that implies excessive screening.

B. HOW ABSOLUTE LIABILITY FEEDS THE INFORMATION COMMONS

When looking carefully at the likely effects of requiring proof of copying, the hidden advantages of the status quo become apparent. The unseen truth about absolute liability is that it makes the *source* of information irrelevant. This makes it unnecessary for researchers to segregate the information they receive. Any researcher who invents something may be liable for patent infringement. It could happen if the researcher directly copies information. It could happen if he or she

requirement in patent law ought to be not copying per se, but *negligent* copying: copying that would have been avoided by a researcher exercising an ordinarily prudent standard of care. On this view, some copying would be expected and permissible; the only copying that would lead to liability would be negligent or inefficient copying—copying that took place because a researcher chose not to implement prudent screening in a given case. This might be hard for people in the patent system to accept; letting a copyist go free might seem outrageous. On the other hand, it might be argued that a copyist charged with infringement would be very unlikely to escape liability under the negligence standard, because—assuming a high correlation between detection of infringement and very serious economic harm—copyists hauled into court for infringement would rarely be able to show they took adequate precautions under the circumstances. An alternative theory, of course, is that copying an invention is a moral wrong and cannot be excused under an economic calculus. Cf. Posner, *supra* note 67, at 469 (explaining non-instrumental views of torts).

69. See discussion and sources cited *supra* note 65; Michael J. Meurer, *Inventors, Entrepreneurs, and Intellectual Property Law*, 45 HOUS. L. REV. 1201, 1215 n.55 (2008) (citing “not invented here” syndrome described in a research source).

partially copies. It could happen if he or she inadvertently copies. And it could happen if he or she never copies at all—if an invention springs strictly and solely from his or her own inspiration. Absolute liability makes it completely irrelevant where the inspiration for an invention comes from. Put another way, there is nothing anyone can do about patent infringement liability. It just is.

Since there is nothing anyone can do about infringement liability, when it comes to the sources of technical information, no one does anything. And therein lies the great advantage. It makes very little sense to screen technical information; doing so will not reduce the chances of liability enough to make it worthwhile. (We know that for a fact because no one does it now.) What this means is that researchers can gather information from any and all sources. They can acquire, store up, and use information without regard to where it comes from or whether it will one day find its way into someone else's patent. Absolute liability provides a sort of umbrella of legal risk. Under this umbrella, which covers all research activity, it makes no sense to try to avoid legal liability. So no one does. The result is that information is shared and acquired rather freely. Indeed, because of the potential for a disclosing party to obtain a patent even after the information is disclosed, it might be said that the patent regime provides an almost ideal set of incentives to disclose technical information. The strange feature of absolute liability, then, is that because it is indiscriminate in fixing liability, it permits researchers to be indiscriminate in obtaining information from any and all sources. Who would have thought that?—absolute liability means that patent infringement can come out of nowhere. And that frees researchers to acquire information from everywhere.

But here is another thought: in mixed technological communities—those where some members abjure IP rights, and others systematically claim them—absolute liability might be a plus as well. This is because absolute liability removes the need to selectively screen contributions from the two types of community members, those committed to “open sharing” and those who believe in obtaining patents. On the assumption that technical information from a “pro-patent” member would more commonly lead to a later charge of infringement, community members might well erect barriers against information from pro-patent members. Not only would this reduce the total volume of technical information available to a community member, it would also be costly in and of itself. Technically trained researchers would have to be put in place to screen technical articles, conference presentations, and other sources of information. Each item would have to be labeled by source: pro-patent or non-patent. Only

information coming from non-patent sources would be allowed through. This screening would be necessary because of the likelihood that a pro-patent source communicating technical information today will file for a patent on that information. So if the recipient incorporated the ideas in a communication into one of its own products, it would be potentially liable for patent infringement when a patent issued at some later time to the company from which the information originated. This type of screening would be expensive.

IV. PRIOR COMMERCIAL USE AND OTHER DOCTRINAL ALTERNATIVES

In light of this research, two related issues deserve consideration. One is the new prior commercial use right, § 273 of the America Invents Act of 2011 (AIA). The other is a set of scholarly suggestions for an independent invention defense.

A. PRIOR COMMERCIAL USE: THE GOOD AND THE BAD

Under AIA § 273, a person who uses certain inventions commercially more than one year before another applies for a patent may continue to use the invention despite issuance of a patent to the other person.⁷⁰ This “prior commercial use” (PCU) defense is quite limited, however. First, of course, it only applies to patents issued after September 2011, when the AIA was passed. Second, the commercial use must be continuous during

70. The statute has an alternative one-year limit. The prior commercial use defense fails if the prior user cannot establish his or her use more than a year before the applicant filed. In addition, the defense fails if the prior user cannot prove use more than a year before the patent applicant first discloses his or her invention publicly, prior to filing. 35 U.S.C. § 273(a)(2)(A), (B) (2012). The AIA’s legislative history describes the purpose behind the prior commercial user defense:

Many countries include a more expansive prior-user rights regime within their first-to-file system. In the United States, this is particularly important to high-tech businesses that prefer not to patent every process or method that is part of their commercial operations. . . . This narrow expansion of prior-user rights balances the interests of patent holders, including universities, against the legitimate concerns of businesses that want to avoid infringement suits relating to processes that they developed and used prior to another party acquiring related patents.

AMERICA INVENTS ACT, H.R. REP. NO. 112-98, pt. 1, at 44 (2011). The reference to universities indicates the defense is not available for patents growing out of grant-funded university research. *See* 35 U.S.C. § 273(d)(5)(A) (2012); On parallel “public disclosure” provisions in the grace period under the AIA, *see* Robert P. Merges, *Priority and Novelty Under the AIA*, 27 BERKELEY TECH. L.J. 1023 (2012).

the infringement period, which creates a risk that the defense will be limited based on abandonment. Third, there is a heavy burden of proof on the defendant—“clear and convincing evidence.” Fourth, there is punishment for a defendant who “unreasonably” pleads the defense—payment of the plaintiff’s attorney fees. Fifth, it applies only to process inventions and products “used in a manufacturing or other commercial process.”⁷¹

There is absolutely no indication that the contours of § 273 were designed in light of the diffusion research described in this Article. Nevertheless, the many limitations of this new provision make at least some sense when viewed from this perspective. For one thing, the timing built into § 273 makes it possible for the originator of an idea to file for a patent within one year of first publicly disclosing it, with the certain knowledge that anyone who borrows the idea will have to honor the inventor’s patent. By the same token, anyone who files within a year of public disclosure, when met with evidence that someone else was indeed using the same idea more than a year before the inventor’s filing date, can be confident that it is truly a case of independent invention. Moreover, the burden of proof required in the new provision eliminates the need for the inventor/patentee to prove copying. Given the general nature of diffusion, and the potential for inadvertent plagiarism, this may be a good thing. Idea duplication can be very subtle indeed, and it makes sense to put the burden on the party asserting independent invention to show affirmatively that all components of their research originated with them, or at any rate did not originate with the inventor/patentee. On the other hand, for reasons discussed in the next Section, the proof requirement entails some costs as well.

One reason to favor the AIA’s PCU defense is that this defense furthers a valuable policy goal beyond simply defending independent invention. The key is its emphasis on *use*. To qualify, one must not only invent earlier than a patentee who asserts a patent, one must also *use* the invention in a positive and constructive way. In light of what we know about the incidence of copying, the defense makes sense. First, the conservative timing requirements make it less likely that one who asserts the defense will have actually learned something crucial, at an earlier date, from the researcher who later obtains a patent. Second, even where learning takes place, the defense recognizes the value of rapid implementation. One who learns from another researcher—but also

71. 35 U.S.C. § 273(a) (2012).

applies the learned information quickly in a commercial manner—makes an independent contribution to society. The defense recognizes this contribution, even when one learns from a researcher that later obtains a patent.⁷² It could even be argued that the emphasis on commercial use in § 273 harkens back to an earlier era in U.S. patent law when the courts favored active implementation over the mere pursuit of legal rights.⁷³

As previously suggested, PCU is a rather limited defense under the AIA. The diffusion research emphasized earlier can be read to support this, implying that remote researchers may subtly influence the ideas of others, in ways that are difficult to trace and document. Placing the burden of proving PCU on the defendant can be seen as a reflection of this. It might be defended this way: we place the burden of proving independent invention on the accused infringer because the evidence of independent invention is close at hand and easy for that party to pull together. Proof of copying, on the other hand, would be far more difficult for the patentee to produce. The evidence may be buried deep within the files and records of the accused infringer, making it hard for the patentee to reconstruct, through discovered documents and testimony, the chain of

72. For an argument that the prior commercial user defense represents a boon to domestic U.S. manufacturers, see Martin Gomez, *Manufacturing, Please Come Home: How AIA's Prior User Right Could be the American Economy's Savior*, 13 U.C. DAVIS BUS. L.J. 61 (2012). It should be noted that while this student author may be right, experience with overseas patent systems has shown that prior user rights are seldom used as an effective defense, at least in reported infringement cases. See, e.g., Pierre Jean Hubert, *The Prior User Right of H.R. 400: A Careful Balancing of Competing Interests*, 14 SANTA CLARA COMPUTER & HIGH TECH. L.J. 189, 213 (1998) (“[T]he limited data available relating to operation of the prior user right in foreign countries suggests the incidence of prior user right problems which would arise in practice in the United States would be very small.”); Keith M. Kupferschmid, *Prior User Rights: The Inventor's Lottery Ticket*, 21 AIPLA Q.J. 213, 223–26 (1993) (“[P]rior user right litigation is minimal in countries presently having the right[.] [I]t is safe to conclude that there should be an extremely small number of prior user rights cases in the United States.”).

73. The now-discredited “paper patent” doctrine is an example of this. See Robert P. Merges, *From “Paper Patents” to The Paper Bag Case: Economic Change and Patent Doctrine, 1870–1910*, (Working Paper, April, 2013, on file with BTLJ) (arguing that nineteenth-century patent doctrine is a good example of general nineteenth-century thinking which encouraged the “release of entrepreneurial energy”, a phrase made famous by the legal historian J. Willard Hurst); see also J.W. HURST, LAW AND THE CONDITIONS OF FREEDOM IN THE NINETEENTH-CENTURY UNITED STATES, at 3–32 (1956). Hurst elaborates this theme in his monumental book, J.W. HURST, LAW AND ECONOMIC GROWTH: THE LEGAL HISTORY OF THE LUMBER INDUSTRY IN WISCONSIN, 1836–1915, at 358 (1964) (“Nineteenth century public policy in the United States generally favored action and the venture of capital in production.”).

events by which the patentee's invention made its way into the infringer's product design.

First, the good news. The emphasis in the AIA is on use, not invention. Whether the accused infringer learned of an idea or a new way of doing something from the patentee does not matter. All that matters is use—a much more tractable issue of proof. It might be argued that this emphasis on use also returns the patent system to an emphasis on implementation that has been lost or disregarded in recent years. In the nineteenth century, the “paper patent” doctrine and other rules disfavored patents that were never actually implemented or put into practice.⁷⁴ Arguably, a rule centered on proof of actual commercial use represents a partial return to the spirit of these nineteenth century rules. What matters, again, is practical use, and not just clever claim drafting and timely filing.

However, the PCU defense may not, in the end, make anyone very happy. For patentees, the fact that liability hinges on use may seem unfair. An infringer can indeed learn of an intriguing idea, and as long as it acts well before the inventor takes action (by filing or disclosing), the infringer is off the hook. For infringers, the stringent requirements of the PCU may prove quite burdensome. The records and evidence mentioned earlier may be difficult to assemble. In addition infringers may find it difficult to meet the stringent burden of proof. And above all, the defense applies only to process inventions and products “used in a manufacturing or other commercial process.”⁷⁵ It could be that the PCU in the U.S. will be doomed to the same fate it has experienced in other jurisdictions—a good defense in theory, but one that rarely proves effective in practice.

B. ALTERNATIVE “MIDDLE GROUND” RULES

Thus far, this Article attempts to make a case for the absolute liability rule in patent law. Even so, at several points we have seen that this rule sometimes produces unfair outcomes. This Section briefly considers some in-between rules that have some of the positive features of the AIA's PCU defense, but that still fall short of a full and true independent invention defense.

First, consider the independent invention defense proposed by legal scholar Samson Vermont. Vermont's defense, styled a “reinvention defense,” would attach prior to the time when a patented invention was widely publicized. In other words, actual or constructive notice of the

74. See *Merges*, *supra* note 73.

75. 35 U.S.C. § 273(a) (2012).

existence of the patent would cut off the possibility of an independent invention defense. This constructive notice rule would obviously change the current rule of absolute liability, but only in cases where a reinventor had actual or constructive notice of a patented invention. As Vermont explains it:

Publication that would likely satisfy the standard for purposes of constructive notice includes English-language publication in an issued patent, a published patent application, publication in a mainstream scientific journal, or publication via presentation at a conference open to the relevant public. Note that, even with the stricter standard, an unavoidable evil of letting constructive notice shut the reinvention window is that legitimate reinventors who look for but never see the first inventor's good faith publication will nonetheless lose the defense if they fail to complete reinvention before the date of that publication.⁷⁶

Vermont's broad dissemination requirement is close to the PCU defense of § 273 in several respects. Although the PCU defense relies less on concepts of notice, intending instead to reward the application or commercialization of technology, a technology that has been commercialized is more likely to come to the attention of researchers and competitors. Therefore, commercialization will often (though not always) be correlated with dissemination. Even when it is not, dissemination and commercialization each serve a positive purpose—which means that both the Vermont proposal and the PCU alter absolute liability in ways that enhance social welfare. They are, as a result, similar at least in the broad sense of deviating from absolute liability only when doing so promotes an important policy.

A similar proposal by Roger Blair and Tom Cotter calls for infringement liability only for what they call “idle patents.”⁷⁷ Their primary motive is to reduce the patent search costs of a person or company that wants to market a new product. In this sense, their proposal is essentially a tort setup—patent infringement is a harm, and they seek to minimize the social welfare costs stemming from that harm. For patents in active use, society benefits by active deployment of the underlying technology. But no such benefit accrues in the case of “idle patents.” So for

76. Samson Vermont, *Independent Invention as a Defense to Patent Infringement*, 105 MICH. L. REV. 475, 487 (2006).

77. See Roger D. Blair & Thomas F. Cotter, *Strict Liability and Its Alternatives in Patent Law*, 17 BERKELEY TECH. L.J. 799 (2002).

“idle patents,” liability would require proof of actual notice to the infringer.

Both proposals have merit. Vermont’s reinvention defense protects infringers from incurring liability when a patent is very difficult to discover in advance. It would also, if adopted, incentivize patent owners to widely disseminate information about their inventions. The Blair and Cotter proposal shares an important feature of the PCU defense of the AIA: an emphasis on applying technology rather than merely stockpiling patents. Varying liability standards according to whether patented technology is “in use” or “idle” would reward active deployment, much as rewarding “commercial use” with a defense against infringement.

In many ways these proposals would bring patent law closer to copyright with respect to proof of copying. Recall that the “access” requirement in proving copyright infringement often boils down to proof of an “opportunity to copy.” The wide dissemination aspect of the Vermont proposal is quite similar; obviously an invention that is widely disseminated provides a greater opportunity to copy. Likewise, when Blair and Cotter speak of rewarding active deployment of technology, they in effect make it much more likely that a prospective infringer can discover the existence of a patent—on the theory that competitors are more likely to investigate the patent status of a technology when that technology has been put into practice.

Taken together, these two proposals present something of a middle ground as regards the patent liability standard. They are well short of requiring direct proof of copying, of course. But they also go well beyond absolute liability. They call for the patentee to establish facts that show it was quite possible the infringer learned or could have learned of the patented invention from the patent owner. I would call this a “plausible mechanism” requirement. Under it, the patentee would have to show not only that the infringer made, used, or sold an invention falling within one or more of the patentee’s claims, the patentee would have to establish a “plausible mechanism” through which the invention *might* have been transmitted from the patentee to the infringer. In the case of a very obscure invention, one that had not been widely disseminated or deployed, proving a plausible mechanism would be difficult. Proof of actual, direct copying would of course suffice. Short of this, perhaps some idiosyncratic facts could be established—a plausible chain of communication, for example, extending from the patentee to the infringer. In the absence of any such facts, however, there would be no liability for patent infringement.

V. CONCLUSION

Absolute liability dispenses with the need to prove often-complex facts. A right holder proves a violation; liability follows. The best defense of patent law's absolute liability rule—which eliminates any opportunity for an infringer to argue independent invention—makes a virtue of this stripped-down liability standard.

In some ways, what this Article has been arguing parallels the famous Fuller and Perdue explanation of the importance of reliance in contract law. For them, you might recall, the doctrine of consideration embodies a deep commitment to the protection of the “reliance interest” among contracting parties. Reliance is so important, they said, that consideration doctrine had evolved to eliminate the need for a party to actually prove it. In their words, judges had decided that in contract law: “To encourage reliance we must . . . dispense with its proof.”⁷⁸ The equivalent I am suggesting would say instead: To encourage disclosure and diffusion in general, patent law, for purposes of establishing liability, dispenses with its proof. By making proof of disclosure irrelevant, patent law eliminates liability-proofing strategies that might well isolate researchers much more than is good for them, or for us.

If the influence of an earlier inventor could be easily ruled out, and if an independent invention defense caused few distortions in the way research is conducted, an independent invention defense would be a fine thing. But neither point has been established. The many and subtle ways that earlier inventors can influence later ones means that absolute liability will, in important cases, lead to the fair outcome. And the ability to show independent invention makes it more desirable to screen out external sources of information. The likely result—research groups working on isolated “islands”—would significantly undermine the free flow of information amongst members of technical communities. The irony is that an exacting standard of liability means there is little reason to attempt reducing liability by increasing the degree of isolation. Researchers are thus in a sense united under the umbrella of absolute liability. And the sharing of information is encouraged because there are good reasons to learn from one another, but limited payoffs from increasing isolation. The whole thing sounds counterintuitive, paradoxical even. Yet it is true. Strong liability reduces the barriers to information flow among researchers. Before digging into it, who would have thought that?

78. Lon L. Fuller & William R. Perdue, Jr., *The Reliance Interest in Contract Damages: I*, 46 YALE L.J. 52, 62 (1936).

STRATEGIC DECISION MAKING IN DUAL PTAB AND DISTRICT COURT PROCEEDINGS

Saurabh Vishnubhakat, Arti K. Rai & Jay P. Kesan[†]

ABSTRACT

The post-grant review proceedings set up at the U.S. Patent and Trademark Office's Patent and Trial Appeal Board by the America Invents Act of 2011 have transformed the relationship between Article III patent litigation and the administrative state. Not surprisingly, such dramatic change has itself yielded additional litigation possibilities: *Cuozzo Speed Technologies v. Lee*, a case addressing divergence between the manner in which the PTAB and Article III courts construe patent claims, will soon be argued at the U.S. Supreme Court.

Of the three major new PTAB proceedings, two have proven to be popular as well as controversial: *inter partes* review and covered business method review. Yet scholarly analysis of litigant behavior in these proceedings has been limited thus far to descriptive data summaries or specific policy perspectives on these types of post-grant challenges, such as their impact on the well-rehearsed patent troll debate. In this article, we present what is to our knowledge the first comprehensive empirical and analytical study of how litigants use these *inter partes* review and covered business method review proceedings relative to Article III litigation.

A major normative argument for administrative ex post review is that it should be an efficient, accessible, and accurate substitute for Article III litigation over patent validity. We assess the substitution hypothesis, using individual patents as our general unit of analysis as well as investigating patent-petitioner pairs and similar details in greater

DOI: <http://dx.doi.org/10.15779/Z38XG2G>

© Saurabh Vishnubhakat, Arti K. Rai & Jay P. Kesan.

[†] Saurabh Vishnubhakat is an Associate Professor of Law at the Texas A&M University School of Law, a Fellow at Duke Law Center for Innovation Policy, and a former Expert Advisor to the Chief Economist of the United States Patent and Trademark Office. Arti K. Rai is the Elvin R. Latty Professor of Law at Duke Law School, Faculty Co-Director of the Duke Law Center for Innovation Policy, and the former Administrator for Policy and External Affairs of the USPTO. Jay P. Kesan is a Professor of Law and the H. Ross & Helen Workman Research Scholar at the University of Illinois College of Law and was an inaugural Thomas Alva Edison Distinguished Scholar at the USPTO. The arguments in this writing are the authors' and should not be imputed to the USPTO or to any other organization. For thoughtful questions and suggestions, we are grateful to workshop participants and commentators at the University of Pennsylvania Law School, the University of San Diego School of Law 6th Annual Patent Law Conference, the University of Illinois College of Law 4th Annual Empirical Patent Law Conference, and the Duke Law Innovation Policy Roundtable on the PTAB's Post-Grant Review Proceedings.

depth. Our data indicate that the “standard model” of explicit substitution—wherein a district court defendant subsequently brings an administrative challenge to patent validity—occurs for the majority (70%) of petitioners who bring *inter partes* review challenges. An important implication of this effect is that the PTAB should use a claim construction standard that mirrors that of the district court, so that substituting administrative process for judicial process does not lead to substantively different outcomes.

Notably, however, standard substitution is not the only use of the PTAB: particularly in the area of *inter partes* reviews, we also see a surprising percentage of cases (about 30%) where the petitioner is not the target of a prior suit on the same patent. The frequency of these nonstandard petitioners, as well as their tendency to join the same petitions as an entity that has been sued, varies by technology. Our data on nonstandard petitioners provide some insight into the extent to which patent challengers are engaging in collective action to contest the validity of patents. Depending on the details of how nonstandard petitioning and collective action are being deployed, this activity could provide a social benefit or constitute a form of harassment.

TABLE OF CONTENTS

I.	INTRODUCTION	47
II.	EX POST REVIEW OF PATENT VALIDITY	50
	A. MOTIVATIONS FOR (AND CONCERNS REGARDING) ADMINISTRATIVE REVIEW	51
	B. OPPOSITION MECHANISMS PRIOR TO THE AIA	55
	1. <i>Ex Parte</i> Reexamination	55
	2. <i>Inter partes</i> Reexamination	57
	C. ADMINISTRATIVE OPPOSITION UNDER THE AIA	58
	1. <i>Inter Partes</i> Review	59
	2. <i>Covered Business Method</i> Review	60
	3. <i>Intersection with the Courts</i>	62
III.	AGENCY OR COURT: STRATEGIC CHOICES	64
	A. LITIGANT BEHAVIOR	65
	1. <i>IPR and CBM</i> Petitions: <i>Descriptive Statistics</i>	66
	2. <i>Article III</i> Litigation	69
	B. AGENCY DECISIONS	78
	C. COURT DECISIONS	80
IV.	DISCUSSION	81
V.	CONCLUSION	83
	APPENDIX A: DATA	84
	A. PTAB DATA	84
	B. DISTRICT COURT DATA	86
	APPENDIX B: FIGURES	87
	APPENDIX C: TABLES	113

I. INTRODUCTION

This is the first paper in a multipart project studying the new post-grant review proceedings set up at the U.S. Patent and Trademark Office (USPTO) Patent and Trial Appeal Board (PTAB) by the America Invents Act of 2011 (AIA).¹ These new administrative trial-type proceedings represent a significant change in the relationship between the system of patent litigation in Article III courts and the administrative state.

Although PTAB proceedings have proved to be quite popular, scholarly analysis of litigant behavior has thus far been limited to descriptive data summaries or specific policy perspectives on post-grant challenges, such as their impact on the well-rehearsed patent troll debate.² This Article is the first comprehensive empirical and analytical study of how litigants use these administrative procedures relative to Article III litigation. In addition to assessing the behavior of litigants, we analyze the behavior of both the PTAB and the courts.

Under the AIA, defendants, potential defendants, and third parties now confront the question of whether and when to challenge the validity of patents by filing one or more petitions for *inter partes* review (IPR) or, if applicable, petitions for covered business method (CBM) review. IPR petitions are filed against individual patents (and claims thereof), but multiple petitions against a patent may be filed by the same or different parties, and a single petition may be filed or joined by multiple parties. Similarly, CBM petitions are filed against individual patents and claims that are directed to eligible business method-related inventions.³

Meanwhile, patent owners still face the question of which patents to assert, when and where to assert them, and against whom to assert them. The AIA's anti-joinder provision for Article III litigation arguably increases complexity by substantially reducing owners' ability to sue multiple defendants in a single case.⁴ Thus, patent owners wishing to sue multiple defendants on a given patent generally have to sue them individually. More importantly for our purposes, the rise of the PTAB

1. Leahy-Smith America Invents Act, Pub. L. No. 112-29, 125 Stat. 284.

2. *E.g.*, Brian J. Love & Shawn Ambwani, *Inter Partes Review: An Early Look at the Numbers*, 81 U. CHI. L. REV. DIALOGUE 93 (2014).

3. In ongoing work, discussed in summary below, we are looking in detail at patents that are the subject of more than one petition. We are dividing these patents into two categories: those that are challenged by the same petitioner multiple times, and those that are challenged by different petitioners. We are further subdividing the two categories by claims and grounds.

4. 35 U.S.C. § 299. *See generally* David O. Taylor, *Patent Misjoinder*, 88 N.Y.U. L. REV. 652 (2013) (discussing rationale for anti-joinder provision).

forces patent owners to factor in the strong possibility of retaliatory or even preemptive patent validity challenges at the PTAB. As a result, two complex frameworks of resolving patent disputes now coexist: ordinary infringement litigation and declaratory judgment actions in Article III courts, along with administrative invalidation actions in the PTAB.

Multiple proceedings with many potential parties offer a number of strategic possibilities. Two examples of ongoing litigation involving certain highly asserted and highly petitioned patents provide an illustration of the complexities and the correspondingly complicated strategic questions. Although these cases are hardly representative, they do provide clear examples of the multiple, perhaps even combinatorial, strategic possibilities.

In a set of seven cases filed between July 1 and July 9, 2013, Zond, a plasma discharge technology developer, asserted a suite of patents in Massachusetts district court against nine defendants.⁵ Intel, one of the defendants, responded by filing IPR petitions on all of the asserted patents.⁶ In April 2014, Intel persuaded the Massachusetts district court to grant a stay of the litigation.⁷ Within two months of the court granting a stay to Intel, all but one of the defendants had filed IPR petitions on the same claims and the same grounds.⁸ All of the petitioning defendants received stays, and the PTAB joined them to the Intel petitions. Although Intel ultimately settled, PTAB review of the challenged patents continues, albeit with a new lead petitioner.⁹

In another set of cases, e-Watch sued eleven firms on two digital signal transmission patents in the Eastern District of Texas.¹⁰ A third-

5. Zond, Inc. v. Gillette Co., No. 1-13-cv-11567 (D. Mass., July 1, 2013); Zond, LLC v. Advanced Micro Devices, Inc., No. 1-13-cv-11577 (D. Mass., July 2, 2013); Zond, LLC v. Intel Corp., No. 1-13-cv-11570 (D. Mass., July 2, 2013); Zond, Inc. v. SK Hynix Inc., No. 1-13-cv-11591 (D. Mass., July 3, 2013); Zond, Inc. v. Toshiba America Elec. Components, Inc., No. 1-13-cv-11581 (D. Mass., July 3, 2013); Zond, Inc. v. Renesas Elecs. Corp., No. 1-13-cv-11625 (D. Mass., July 8, 2013); Zond, Inc. v. Fujitsu Ltd., No. 1-13-cv-11634 (D. Mass., July 9, 2013).

6. The 27 *inter partes* review petitions filed by Intel are listed in Table 1 of Appendix C.

7. Order Granting Motion to Stay Pending Inter Partes Review, Case No. 1-13-cv-11570, Paper No. 120 (D. Mass. Apr. 18, 2014).

8. The 90 *inter partes* review petitions filed by defendants are listed in Table 2 of Appendix C.

9. Joint motions to terminate proceedings, all filed simultaneously on Sept. 12, 2014, settled the Intel-initiated IPR petitions on Zond's patents. The settlement agreement between Intel and Zond that governs the termination of all these proceedings is confidential.

10. e-Watch, Inc. v. LG Elecs., Inc., No. 2-13-cv-01064 (E.D. Tex., Dec. 9, 2013); e-Watch, Inc. v. Samsung Elecs. Co., No. 2-13-cv-01062 (E.D. Tex., Dec. 9, 2013); e-

party firm filed the first PTAB petition related to those patents.¹¹ Subsequently, HTC, a defendant, instituted a petition, and the institution of the HTC petition triggered other petitions.¹²

A major normative argument for administrative ex post review is that it should be an efficient, accessible, and accurate substitute for Article III litigation over patent validity.¹³ In this paper, we assess the substitution hypothesis, using as our unit of analysis the individual patent. Our data indicate that the “standard model” of substitution—wherein a district court defendant subsequently brings an administrative challenge to patent validity—is indeed occurring. In addition to providing overall data, we analyze standard substitution by technology area and district court.

Notably, however, standard substitution is not the only use of the PTAB: particularly in the area of IPRs, we also see a surprising percentage of cases (about 30 percent) where the petitioner is *not* the target of a prior suit on the same patent. The frequency of these nonstandard petitioners, as well as their tendency to join the same petitions as an entity that has been sued, varies by technology. Our data on nonstandard petitioners thus provide some insight into the extent patent challengers are engaging in collective action to challenge patents.

Depending on the details of how nonstandard petitioning and collective action are being deployed, this activity could provide a social benefit or constitute a form of harassment. As we discuss in Part II, many commentators have noted that challenging an invalid patent, particularly in expensive Article III litigation, represents a collective action problem.

Watch, Inc. v. Apple Inc., No. 2-13-cv-01061 (E.D. Tex., Dec. 9, 2013); e-Watch, Inc. v. HTC Corp., No. 2-13-cv-01063 (E.D. Tex., Dec. 9, 2013); e-Watch Inc. v. BlackBerry Ltd., No. 2-13-cv-01078 (E.D. Tex., Dec. 10, 2013); e-Watch Inc. v. Sharp Corp., No. 2-13-cv-01074 (E.D. Tex., Dec. 10, 2013); e-Watch Inc. v. ZTE Corp., No. 2-13-cv-01071 (E.D. Tex., Dec. 10, 2013); e-Watch Inc. v. Sony Corp., No. 2-13-cv-01073 (E.D. Tex., Dec. 10, 2013); e-Watch Inc. v. Nokia Corp., No. 2-13-cv-01075 (E.D. Tex., Dec. 10, 2013); e-Watch Inc. v. Huawei Tech. Co., No. 2-13-cv-01076 (E.D. Tex., Dec. 10, 2013); e-Watch Inc. v. Kyocera Commc’ns, Inc., No. 2-13-cv-01077 (E.D. Tex., Dec. 10, 2013).

11. Petition for Inter Partes Review by Iron Dome LLC, No. IPR2014-00439 (P.T.A.B. Feb. 18, 2014).

12. The twelve *inter partes review* petitions filed are listed in Table 3 of Appendix C.

13. Others (including one of us) have argued that to the extent the procedures set up by the AIA resemble formal adjudications, they could serve as a vehicle not simply for error correction but also for legal and policy development. See, e.g., Arti K. Rai, *Patent Validity Across the Executive Branch: Ex Ante Foundations for Policy Development*, 61 DUKE L.J. 1237 (2012); Melissa F. Wasserman, *The Changing Guard of Patent Law: Chevron Deference for the PTO*, 54 WM. & MARY L. REV. 1959 (2013). In this Article, however, we focus on error correction.

Administrative alternatives may ease the collective action problem, but they may also provide opportunities for harassing patent owners.¹⁴ As another indicator of potential harassment and delay, we also look at the frequency of serial petitioning on a given patent.

Of course, substitution of any sort (as contrasted with duplication) can occur only if administrative review is accurate and efficient, and courts generally stay any related Article III litigation pending administrative review. In the case of declaratory judgment (DJ) litigation, the AIA both bars a DJ litigant from bringing a subsequent administrative review and provides for automatic stays of any subsequent DJ actions.¹⁵ So the issue of duplication primarily arises in the context of infringement litigation brought by the patent owner. Although a full answer to the duplication issue awaits further decision making in cases currently before the PTAB and the courts, we provide some initial data on the question.

In this Article, Part II discusses the normative arguments for and against administrative ex post validity review as a substitute for judicial review. It reviews these arguments as they developed in earlier incarnations of administrative review and as they developed in the far more robust AIA proceedings. Part III provides the large-scale empirical data we have gathered. It discusses various indicia of a general substitution effect in the context of particular technologies and in particular district courts. We also discuss the phenomenon of nonstandard petitioners and the collective action in which they sometimes engage. Additionally, Part III presents data regarding multiple IPR petitions filed against the same patent. Based on these data, Part III examines agency and court decision-making in the face of strategic behavior by the parties before them. Part IV discusses our major findings, suggests directions for further research, and outlines our ongoing agenda to advance these research goals.

I. EX POST REVIEW OF PATENT VALIDITY

This Part discusses the normative arguments that have motivated administrative review of patent validity, particularly as a substitute for litigation in the federal courts. Against the backdrop of this normative

14. *E.g.*, Gregory Dolin, *Dubious Patent Reform*, 56 B.C. L. REV. 881 (2015); Jay P. Kesau, *Carrots and Sticks to Create a Better Patent System*, 17 BERKELEY TECH L.J. 145, 165 (2002) (discussing the dangers of delay and harassment in post-issuance patent office proceedings); Raymond A. Mercado, *Ensuring the Integrity of Administrative Challenges to Patents: Lessons from Reexamination*, 14 COLUM. SCI. & TECH. L. REV. 558 (2013).

15. Perhaps not surprisingly, since patents became available for PTAB review, DJ actions have fallen both in absolute terms and as a percentage of case filings.

framing, we then evaluate *ex parte* and *inter partes* reexamination (the latter now defunct) as well as the new *ex post* review procedures introduced by the AIA.

A. MOTIVATIONS FOR (AND CONCERNS REGARDING)
ADMINISTRATIVE REVIEW

The initial patent examination process will inevitably produce some improper patent grants. To the extent improperly granted patents impose unnecessary costs and call into question the credibility of the patent system,¹⁶ these improper grants ought to be corrected.¹⁷ The importance of error correction remains a dominant theme in *ex post* patent review, especially in evaluating the success of the AIA.¹⁸ Perhaps even more important, however, is the recurring theme of institutional design: the USPTO's examination errors should not merely be corrected, but should be corrected outside the federal courts.

Several interrelated arguments counsel in favor of administrative review. Most obviously, Article III litigation is quite costly. The biennial economic survey of the American Intellectual Property Law Association indicates that even for the lowest-stakes category of patent lawsuits (in which less than \$1 million was at risk), median litigation costs have risen from \$650,000 in 2005 to \$700,000 in 2013.¹⁹ And for the highest-stakes

16. See generally Mark D. Janis, *Rethinking Reexamination: Toward a Viable Administrative Revocation System for U.S. Patent Law*, 11 HARV. J.L. & TECH. 1, 7–36 (1997) (considering USPTO's role in patent revocation via administrative reexamination). For purposes of this article, we need not engage the robust academic debate over the level of error the initial examination process should tolerate.

17. *Ex post* review as a means for correcting USPTO examination errors has been a consistent theme in institutional discussions of patent quality. See, e.g., *In re Swanson*, 540 F.3d 1368, 1375 (Fed. Cir. 2008) (“Congress intended reexaminations to provide an important ‘quality check’ on patents that would allow the government to remove defective and erroneously granted patents.”); *In re Recreative Techs. Corp.*, 83 F.3d 1394, 1396–97 (Fed. Cir. 1996) (“The reexamination statute’s purpose is to correct errors made by the government . . . and if need be to remove patents that never should have been granted.”); *Patlex Corp. v. Mossinghoff*, 758 F.2d 594, 603 (Fed. Cir. 1985) (“The legislative history of the reexamination statute makes clear that its purpose is to cure defects in administrative agency action with regard to particular patents and to remedy perceived shortcomings in the system by which patents are issued.”).

18. The degree to which the AIA *ex post* administrative review procedures are actually capturing and correcting *ex ante* examination errors is the focus of related large-scale empirical research relying on much of the same data as the present project. See Saurabh Vishnubhakat, David L. Schwartz & Alan C. Marco, *What Ex Post Review Has Revealed About Patents* (forthcoming).

19. AM. INTELL. PROP. LAW ASS'N, REPORT OF THE ECONOMIC SURVEY 34 (2013) [hereinafter AIPLA SURVEY 2013].

lawsuits (in which more than \$25 million was at risk), median litigation costs rose over the same time period from \$4.5 million to \$5.5 million.²⁰

The high cost of litigation would be less problematic if these great expenditures yielded great accuracy in judicial outcomes. As standard economic accounts of procedure note, the goal of procedure is the minimization of litigation costs and error costs.²¹ But decisions reached in Article III litigation may not be particularly accurate.²² Because patent law often uses science-based proxies such as “ordinary skill in the art” to tackle relevant legal and policy goals, the subject matter of patent law can be highly complex as a scientific matter.²³ And even if a case is not highly complex as a scientific matter, the manner in which factual findings interact with law and policy can be complex.²⁴ With the possible exception of Federal Circuit judges, judges in the federal courts tend to be generalists who may not be equipped to tackle complex questions at the intersection

20. *Id.*

21. See generally Louis Kaplow & Steven Shavell, *Accuracy in the Determination of Liability*, 37 J.L. & ECON. 1 (1994) (modeling the relationship between the design of legal rules and the likelihood of reaching accurate outcomes); Louis Kaplow & Steve Shavell, *Accuracy in the Assessment of Damages*, 39 J.L. & ECON. 191 (1996) (modeling the relationship between the design of legal rules and the likelihood of imposing accurate monetary sanctions).

22. E.g., David L. Schwartz, *Practice Makes Perfect? An Empirical Study of Claim Construction Reversal Rates in Patent Cases*, 107 MICH. L. REV. 223 (2008) (finding high reversal rate for district court claim construction). See generally Anup Malani & Jonathan S. Masur, *Raising the Stakes in Patent Cases*, 101 GEO. L.J. 637, 659 (2013) (offering a brief survey of scholarly proposals to improve judicial accuracy in patent adjudication).

23. For example, a patent may be challenged as being invalid because the invention that it claims was obvious under 35 U.S.C. § 103 in light of the prior technical knowledge available to those in the field at the time of invention. Whether a claimed invention is obvious is a question of law reviewed de novo on appeal. But the legal determination is based on predicate findings of fact regarding the prior art and the level of skill in the art. These findings are supposed to be reviewed deferentially on appeal. See, e.g., *In re Kubin*, 561 F.3d 1351, 1355 (Fed. Cir. 2009) (concerning review from PTO patent denial); see also *Graham v. John Deere*, 383 U.S. 1, 17 (1966) (explaining factual findings made by district court are subject to clearly erroneous review). Meanwhile, the USPTO’s factual findings in granting a patent are presumed correct and must be rebutted by clear and convincing evidence in the courts. *Microsoft Corp. v. i4i Ltd. P’ship*, 131 S. Ct. 2238 (2011). See generally Arti K. Rai, *Engaging Facts and Policy: A Multi-Institutional Approach to Patent System Reform*, 103 COLUM. L. REV. 1035, 1068–75 (2003) (discussing the technical complexity often involved in applying patent law’s fact-based standards).

24. For example, the practical probative value of factual findings toward obviousness analysis may vary by the inherent unpredictability of the given technology: whereas mechanical inventions operate in relatively predictable and well-understood ways, small technical changes may lead to dramatic and unexpected results in biochemistry. *Pfizer, Inc. v. Apotex, Inc.*, 480 F.3d 1348, 1364 (Fed. Cir. 2007).

of law, science, and policy.²⁵ Moreover, district courts have to contend with juries, which may be even less equipped than federal judges to address complex questions of law and science.²⁶ In contrast, administrative patent judges have long been required to be “persons of competent legal knowledge and scientific ability.”²⁷

Empirical research bears out concerns about the capacity of judges to resolve patent disputes. The generalist background of most district judges has led many plaintiffs to seek out specific districts,²⁸ with the natural result being a certain amount of de facto specialization.²⁹ A few districts see a disproportionate number of patent cases, and some have reputations as “rocket docket” for resolving them quickly.³⁰ Empirical evidence suggests that, among the subset of judges who preside over patent cases regularly, increased experience may produce more efficient and accurate case outcomes.³¹ Yet this private ordering toward certain districts only underscores the overall lack of expertise among district court judges.³² Moreover, some commentators have argued that aggressive attempts to

25. Peter Lee, *Patent Law and the Two Cultures*, 120 YALE L.J. 2, 4–6 (2010).

26. See Mark Lemley, *Why Do Juries Decide If Patent Are Valid?*, 99 VA. L. REV. 1673, 1705 (2013) (noting that as far back as the 1950s, modern technology was already “judged too complex for a jury to understand, so it made no sense to give them the patent questions” where avoidable).

27. 35 U.S.C. § 6 (requiring that “administrative patent judges shall be persons of competent legal knowledge and scientific ability”).

28. Commentators have long discussed forum shopping at both the appellate and district court level. *E.g.*, Scott Atkinson, Alan C. Marco & John H. Turner, *The Economics of a Centralized Judiciary: Uniformity, Forum Shopping, and the Federal Circuit*, 52 J.L. & ECON. 411 (2009) (forum shopping prior to the creation of the Federal Circuit); Kimberly A. Moore, *Forum Shopping in Patent Cases: Does Geographic Choice Affect Innovation?*, 79 N.C. L. REV. 889 (2001) (early discussion of forum shopping at district court level).

29. Jay P. Kesan & Gwendolyn G. Ball, *Judicial Experience and the Efficiency and Accuracy of Patent Adjudication: An Empirical Analysis of the Case for a Specialized Patent Trial Court*, 24 HARV. J.L. & TECH. 393, 447 (2011) (showing in Table III that the distribution of how many cases judges hear is highly skewed such that most judges hear fewer than ten patent cases each whereas roughly the top fifth of high-volume judges hear over three-fifths of all patent cases).

30. Saurabh Vishnubhakat, *Reconceiving the Patent Rocket Docket: An Empirical Study of Infringement Litigation 1985–2010*, 11 J. MARSHALL REV. INTELL. PROP. L. 58 (2011).

31. Kesan & Ball, *supra* note 29, at 423–43.

32. This argument is particularly compelling when offered by judges themselves. *See, e.g.*, Judge James F. Holderman, *Judicial Patent Specialization: A View from the Trial Bench*, 2002 U. ILL. J.L. TECH. & POL’Y 425, 430–31 (2002); Judge James F. Holderman & Halley Guren, *The Patent Litigation Predicament in the United States*, 2007 U. ILL. J.L. TECH. & POL’Y 1, 5–6 (2007); The Hon. Ed Kinkeade, *Point-Counterpoint: Two Judges’ Perspectives on Trial by Jury*, 12 TEX. WESLEYAN L. REV. 497, 498 (2006).

specialize in patent disputes by judges whose districts are found outside traditional technology centers lead to overly plaintiff-friendly procedures rather than accurate adjudication.³³

Another reason to favor low-cost administrative review, rather than high-cost Article III review, is that patent plaintiffs and defendants have asymmetric incentives. Supreme Court case law builds into the patent doctrine asymmetric incentives to litigate. Under the law's estoppel provisions, a challenger who successfully invalidates a patent provides a public good—the challenger benefits not only itself but also all other potential challengers.³⁴ By contrast, the challenger who loses is uniquely estopped from challenging the patent again.³⁵

Although the public-good-type incentive may exist in the administrative context as well³⁶ (and, indeed, exists in both pre-AIA and post-AIA *inter partes* proceedings as a consequence of the congressional decision to implement relatively strong statutory estoppel provisions), the significantly lower cost of the administrative proceeding presumably reduces its scale. In other words, although a challenger may still be reluctant to provide a public good, a public good that costs a few hundred thousand dollars is quite different from one that costs several million dollars. Moreover, the absence of a standing requirement in IPR

33. See, e.g., J. Jonas Anderson, *Court Competition for Patent Cases*, 163 U. PA. L. REV. 631 (2015); Jeanne C. Fromer, *Patentography*, 85 N.Y.U. L. REV. 1444 (2010); Daniel Klerman & Greg Reilly, *Forum Selling*, 88 S. CAL. L. REV. (forthcoming 2016), <http://www.ssrn.com/abstract=2538857> [<http://perma.cc/528U-TJS8>].

34. See *Blonder-Tongue Labs., Inc. v. Univ. of Ill. Found.*, 402 U.S. 313, 350 (1971) (stating a patent invalidity finding creates nonmutual defensive collateral estoppel, so that a patent that is invalid as against one party is invalid as against the world); see also Joseph Farrell & Robert P. Merges, *Incentives to Challenge and Defend Patents: Why Litigation Won't Reliably Fix Patent Office Errors and Why Administrative Patent Review Might Help*, 19 BERKELEY TECH. L.J. 943 (2004) (noting that, in addition to public good problem, disincentives to challenge may be created in certain situations involving oligopolistic competition between licensee that pay sales-based royalties to the patentee).

35. Interestingly, some practitioners have suggested that the pre-AIA tendency of so-called patent trolls to sue multiple defendants in one suit might have facilitated some collective action through informal or formal joint defense agreements. See Daniel Bream & Lee Cheng, *Benefits of a Coordinated Joint Defense in Patent Cases*, LAW360 (Nov. 27, 2012). Whether or not that was the case, the AIA's anti-joinder provision may limit this potential nudge towards collective action. *Id.*

36. See John R. Thomas, *Collusion and Collective Action in the Patent System: A Proposal for Patent Bounties*, 2001 U. ILL. L. REV. 305 (2001). But see Stuart M. Benjamin & Arti K. Rai, *Who's Afraid of the APA? What the Patent System Can Learn From Administrative Law*, 95 GEO. L.J. 269, 323–27 (2007) (noting that administrative review that relied on *Chevron* deference by the courts rather than estoppel against the patent challenger could substantially reduce collective action problems).

proceedings creates possibilities for additional entities, including entities that represent groups of potential defendants in a given industry, to challenge patents.

For these reasons, a less costly, more expert, and more widely accessible institution has long been thought desirable. On the other hand, even advocates of an administrative mechanism have noted the potential for harassment of patent owners that might arise in such administrative review.³⁷ Harassment potential exists as the obvious flip side of access created by low cost and the absence of an Article III standing requirement.³⁸ Moreover, to the extent that courts do not believe that administrative review will in fact be accurate and efficient, and thus do not stay any related Article III litigation, such review may create costly duplication rather than efficiency.

B. OPPOSITION MECHANISMS PRIOR TO THE AIA

In 1980 Congress created a mechanism for USPTO *ex parte* reexamination of patent validity,³⁹ and in 1999 Congress created a mechanism for *inter partes* reexamination.⁴⁰ These procedures have realized their error-correction and efficiency goals to varying degrees and have interacted in important ways with federal court litigation.

1. Ex Parte Reexamination

As the Federal Circuit recognized in a contemporaneous opinion, the *ex parte* reexamination system was an effort to reap

three principal benefits. First, the new procedure could settle validity disputes *more quickly and less expensively* than the often protracted litigation involved in such cases. Second, the procedure would allow courts to refer patent validity questions to the *expertise of the Patent Office*. Third, reexamination would

37. *E.g.*, Kesan, *supra* note 14 (noting the potential for delay and harassment in patent office proceedings); Jonathan Masur, *Patent Inflation*, 121 YALE L.J. 470, 522 (2011) (recognizing that “inter partes review could potentially be abused by parties interested only in delaying and harassing competitors”); Joe Matal, *A Guide to the Legislative History of the America Invents Act: Part II of II*, 21 FED. CIR. B.J. 539, 550 (2012) (noting that the AIA’s own statutory text directs the USPTO to penalize abuses of administrative validity challenge proceedings “such as to harass or to cause unnecessary delay or an unnecessary increase in the cost of the proceeding”).

38. That said, the challenger who loses at the administrative level may have to meet Article III standing requirements in order to appeal. *See* *Consumer Watchdog v. Wis. Alumni Research Found.*, 753 F.3d 1258 (Fed. Cir. 2014).

39. Bayh-Dole Act, Pub. L. No. 96-517, 94 Stat. 3015 (1980).

40. American Inventors Protection Act, Pub. L. No. 106-113, 113 Stat. 1501 (1999).

reinforce investor confidence in the certainty of patent rights by affording the USPTO a *broader opportunity to review* doubtful patents.⁴¹

The parameters of this reexamination procedure reflect its twin goals: to correct those USPTO examination errors that improperly allowed patents to issue, and to do so more cheaply, accurately, and accessibly than the federal courts could.

Cost-wise, the USPTO's *ex parte* reexamination fee has grown from \$1,500 in the early 1980s⁴² to \$12,000 at present,⁴³ and attorney costs have risen to approximately \$20,000 at the mean and \$15,000 at the median.⁴⁴ Even today, the expense of *ex parte* reexamination tends to be below \$35,000—some twentyfold less costly than the lowest-stakes category of litigation.⁴⁵ Moreover, when the USPTO decides to deny a request for *ex parte* reexamination, the agency refunds most of the fee to the requester, further lowering the financial hurdle, and risk, to a patent validity challenge.⁴⁶ Consistent with its mandate to correct examination errors, *ex parte* reexamination requires a “substantial new question of patentability” as to one or more of the challenged patent claims, and this standard may be met by reargument of information that was previously before the patent examiner.⁴⁷ Access to reexamination is also unconstrained by traditional Article III standing requirements. Anyone at any time may seek reexamination of a patent, including the patent owner and the USPTO itself,⁴⁸ and courts narrowly construe agreements not to challenge patent validity, in favor of access to reexamination.⁴⁹

However, for all its intended benefits, the reexamination procedure introduced in 1980 has long been criticized for its *ex parte* nature, which

41. See *Patlex Corp. v. Mossinghoff*, 758 F.2d 594, 602 (Fed. Cir. 1985) (internal quotations and citations omitted) (emphasis added).

42. Revision of Patent Fees, 50 Fed. Reg. 31818-01 (Aug. 6, 1985), codified in various parts of 37 C.F.R. Part 1.

43. 37 C.F.R. § 1.20(c)(1).

44. See AIPLA SURVEY 2013, *supra* note 19, at I-112 (tabulating attorney costs reported for *ex parte* reexamination); AM. INTELL. PROP. LAW ASS'N, REPORT OF THE ECONOMIC SURVEY I-136 (2011); AM. INTELL. PROP. LAW ASS'N, REPORT OF THE ECONOMIC SURVEY I-114 (2009); AM. INTELL. PROP. LAW ASS'N, REPORT OF THE ECONOMIC SURVEY I-76 (2007).

45. Janis, *supra* note 16.

46. See 35 U.S.C. § 303(c); 37 C.F.R. § 1.26(c).

47. 35 U.S.C. § 303(a).

48. 35 U.S.C. §§ 302, 303(a).

49. See, e.g., *Joy Mfg. Co. v. Nat'l Mine Serv. Co., Inc.*, 810 F.2d 1127 (Fed. Cir. 1987).

excludes any third-party participation beyond filing the initial request.⁵⁰ According to the PTO's statistics,⁵¹ 29 percent of the *ex parte* reexaminations filed between July 1981 and September 2014 were filed by the patent owner itself, presumably as a potential mechanism for strengthening the patent.⁵²

2. Inter Partes Reexamination

To improve public participation in the administrative review of patent validity, in 1999 Congress created a new procedure: *inter partes* reexamination.⁵³ Designed to coexist with the old *ex parte* procedure, *inter partes* reexamination conferred significant rights upon third-party requestors to participate in the USPTO's review of patent validity. A requestor could comment on every substantive response by the patent owner to an examiner action and could appeal the examiner's decision to the USPTO's administrative review board.

However, *inter partes* reexamination also posed significant barriers. One was a strong estoppel provision, barring the challenger from raising in Article III litigation any issues it raised or could have raised during the *inter partes* reexamination.⁵⁴ Even more significant was the prolonged duration of reexamination. The reexaminations themselves took an average of 39.5 months, and then had to be appealed to the Board of

50. Janis, *supra* note 16, at 6 n.12 (citing Shannon M. Casey, *The Patent Reexamination Reform Act of 1994: A New Era of Third Party Participation*, 2 J. INTELL. PROP. L. 559 (1995)); Marvin Motsenbocker, *Proposal to Change the Patent Reexamination Statute to Eliminate Unnecessary Litigation*, 27 J. MARSHALL L. REV. 887, 898 (1994); Gregor N. Neff, *Patent Reexamination—Valuable, But Flawed: Recommendations for Change*, 68 J. PAT. & TRADEMARK OFF. SOC'Y 575 (1986).

51. See USPTO, *Ex parte* Reexamination Filing Data—September 30, 2014, <http://www.uspto.gov/learning-and-resources/statistics/reexamination-information> [<http://perma.cc/WP77-V8T8>].

52. Certain Federal Circuit cases have indicated that patents that survive reexamination should be viewed even more deferentially by the courts than ordinary patents. *E.g.*, *Gould v. Control Laser Corp.*, 705 F.2d 1340, 1342 (Fed. Cir. 1983) (explaining that reexamination ought to “facilitate trial of [the reexamined] issue by providing the district court with the *expert view of the PTO* (when a claim survives the reexamination proceeding)”) (emphasis added).

53. See generally Kenneth L. Cage & Lawrence T. Cullen, *An Overview of Inter Partes Reexamination Procedures*, 85 J. PAT. & TRADEMARK OFF. SOC'Y 931 (2003) (explaining the structural details and concerns motivating the 1999 *inter partes* reexamination procedures).

54. 35 U.S.C. § 317(b) (2000); see, e.g., M. Patricia Thayer et al., *Examining Reexamination: Not Yet an Antidote to Litigation*, 5 SEDONA CONF. J. 23, 24 (2004) (noting that estoppel makes “*inter partes* reexamination something of a double-or-nothing gamble”).

Patent Appeals and Interferences.⁵⁵ From its creation in 1999 through 2012, when it was subsumed under the new administrative review system established by the AIA, *inter partes* reexamination was never widely used as a means for challenging the validity of patents.⁵⁶

C. ADMINISTRATIVE OPPOSITION UNDER THE AIA

The America Invents Act of 2011⁵⁷ (AIA), which created four new procedures for reevaluating the validity of patents, significantly strengthened the U.S. system for administrative review of patent validity. One procedure, post-grant review, is just beginning its operation as it only applies to patents that issue from applications filed under the AIA's new first-inventor-to-file regime.⁵⁸ Another procedure, supplemental examination, allows patent owners themselves to provide new information that helps fortify the validity of their patents.⁵⁹ Thus far, the opportunity for robust public participation in challenging pre-existing patents has arisen in the two remaining procedures: *inter partes* review (IPR) and the transitional program for covered business method (CBM) review.

Before turning to the specifics of these proceedings, we note the AIA's desire to promote administrative review is probably clearest in the situation where the petitioner is a prior district court defendant. In the case of *inter partes* review, this "standard model" is endorsed, at least implicitly, by several features we discuss in detail below. One feature is the requirement that IPR review occur within one year of a prior district court lawsuit. The congressional decision to include a strong estoppel provision, and thus potentially set up in the administrative context the Article III collective action problem for challengers,⁶⁰ may also reflect congressional embrace of the standard model.

55. This average included some outlier cases, but the median was a lengthy 34.1 months. See *Inter Partes Reexamination Filing Data*, USPTO, <http://www.uspto.gov/learning-and-resources/statistics/reexamination-information> [<http://perma.cc/82NT-8L67>].

56. From November 29, 1999, through the abolition of *inter partes* reexamination effective September 16, 2012, fewer than 2,000 requests were filed, and in most years the usage of *inter partes* reexamination represented only a fraction of *ex parte* reexamination. See *Reexamination Statistics*, USPTO, <http://www.uspto.gov/learning-and-resources/statistics/reexamination-information> [<http://perma.cc/82NT-8L67>].

57. Leahy-Smith America Invents Act, Pub. L. No. 112-29, 125 Stat. 284.

58. As of October 31, 2015, only thirteen post-grant review petitions have been filed in the USPTO. See USPTO, PATENT REVIEW PROCESSING SYSTEM (Oct. 31, 2015) at 2, <http://www.uspto.gov/sites/default/files/documents/2015-10-31%20PTAB.pdf> [<http://perma.cc/V2N8-EJT9>].

59. 35 U.S.C. § 257.

60. See *supra* text accompanying notes 37–41.

In the case of CBM review, Congress embraced the standard model even more fully. CBM review explicitly requires the petitioner to be “charged with infringement,” language the PTO has interpreted as requiring the petitioner to prove standing necessary to bring a declaratory judgment action in district court.⁶¹

That said, in the case of IPRs, the statutory language certainly does not preclude petitioners that are outside the standard model. IPR proceedings have no standing requirement, and the AIA also provides for potential collective action by allowing joinder to existing petitions.

Thus far, the new AIA proceedings do appear substantially cheaper than district court litigation. According to the 2015 AIPLA Economic Survey, the median cost of an IPR through a PTAB hearing was \$275,000 and through appeal was \$350,000.⁶² Although the AIPLA survey does not differentiate between IPRs based on amount of money at risk, these figures are substantially lower than the median cost of district court litigation even for the lowest stakes cases.⁶³

1. *Inter Partes Review*

IPR challenges are available to anyone, other than the patent owner,⁶⁴ who has not previously sought to invalidate the patent through a civil action⁶⁵ and who has not been sued more than one year earlier for infringing the patent in question.⁶⁶ An IPR petition may not be filed anonymously: the petitioner must disclose all its real parties in interest.⁶⁷ For any patent that issued under the old first-to-invent regime, an IPR petitioner may file a challenge immediately.⁶⁸

For a patent that issues under the new first-inventor-to-file regime, an IPR petitioner may file a challenge only after nine months from the patent’s date of grant or after the termination of any post-grant review that

61. 37 C.F.R. § 42.302(a) (“Charged with infringement means a real and substantial controversy regarding infringement of a covered business method patent exists such that the petitioner would have standing to bring a declaratory judgment action in Federal court.”).

62. AM. INTELL. PROP. LAW ASS’N, REPORT OF THE ECONOMIC SURVEY 38 (2015).

63. *See supra* Section I.A.

64. 35 U.S.C. § 311(a).

65. *See* 35 U.S.C. §§ 315(a)(1), (3) (providing that a counterclaim challenging the validity of a patent claim in an infringement action is not a civil action).

66. 35 U.S.C. § 315(b).

67. This disclosure is necessary because the constraints on who can petition also apply to all legal privies and real parties in interest of the would-be petitioner. *See* 35 U.S.C. §§ 315(a)(1), (b).

68. 35 U.S.C. § 311(c).

has been instituted as to the patent, whichever is later.⁶⁹ An IPR may challenge patent claims only on the grounds that they fail to satisfy the novelty requirement⁷⁰ or the nonobviousness requirement,⁷¹ and may only argue on the basis of prior patents or printed publications.⁷² To decide that an IPR petition warrants institution of an IPR proceeding, the USPTO must find a “reasonable likelihood that the petitioner would prevail with respect to at least [one] of the claims challenged in the petition.”⁷³

2. *Covered Business Method Review*

CBM challenges are available only to parties who have previously been sued for infringing, or charged with infringing, the patent in question.⁷⁴ Like IPR petitions, CBM review petitions may not be filed anonymously—they must disclose real parties in interest.⁷⁵ Through rulemaking, the PTO has interpreted the statutory “charged with infringement” language to mean “a real and substantial controversy regarding infringement . . . exists such that the petitioner would have standing to bring a declaratory judgment action in Federal Court.”⁷⁶ Congressional creation of a standing requirement in a CBM review proceeding contrasts with the absence of such a requirement in an IPR proceeding. In general, to the extent that IPR and CBM review provisions differ (and we detail other differences below), these differences may arise because CBM review was inserted into the AIA relatively late in the day, and the members of Congress most responsible for the review saw CBM patents as categorically suspect.⁷⁷

69. *Id.*

70. 35 U.S.C. § 102.

71. 35 U.S.C. § 103.

72. 35 U.S.C. § 311(b).

73. 35 U.S.C. § 314(a).

74. Leahy-Smith America Invents Act, Pub. L. No. 112-29, § 18(a)(1)(B), 125 Stat. 330 (stating the prior infringement suit may be one that targeted the CBM petitioner itself or its privies or real parties in interest).

75. *See id.*

76. 37 C.F.R. § 42.302(a).

77. Senators Schumer and Kyl, who proposed a version of the CBM provision as part of a floor managers’ amendment on March 1, 2011, were highly suspicious of all business method patents. In his March 2011 Senate floor testimony, Senator Schumer described business method patents as “the bane of the patent world” and castigated the decision the Federal Circuit in *State Street Bank* to allow such patents. 157 CONG REC. S1363 (Mar. 8, 2011) (statement of Sen. Schumer). Among many Senators on the Republican side, positions were equally strong. The Senate Republican Policy Committee’s summary of § 18, introduced into the Congressional Record by Senator Kyl, stated (somewhat inaccurately):

For any eligible business method patent⁷⁸ that issued under the old first-to-invent regime, a CBM petitioner may file a challenge at any time after the procedure was established on September 16, 2012.⁷⁹ For an eligible business method patent that issues under the new first-inventor-to-file regime, a CBM petitioner may file a challenge only after nine months from the patent's date of grant or after the termination of any post-grant review that has been instituted as to the patent, whichever is later.⁸⁰ A CBM petition may challenge patent claims on essentially the complete range of patentability criteria, including subject-matter eligibility,⁸¹ novelty,⁸² nonobviousness,⁸³ utility,⁸⁴ single invention,⁸⁵ enablement,⁸⁶ written description,⁸⁷ definiteness,⁸⁸ and others.⁸⁹ To decide

Recent court decisions, culminating in last year Supreme Court decision in *Bilski v. Kappos*, have sharply pulled back on the patenting of business methods, emphasizing that these "inventions" are too abstract to be patentable. In the intervening years, however, PTO was obliged to issue a large number of business-method patents, many or possibly all of which are no longer valid. The Schumer proceeding offers a relatively cheap alternative to civil litigation for challenging these patents, and will reduce the burden on the courts of dealing with the backwash of invalid business-method patents.

157 CONG. REC. S1367 (Mar. 8, 2011) (statement of Sen. Kyl).

78. The AIA defines an eligible "business method" patent as: "a patent that claims a method or corresponding apparatus for performing data processing or other operations used in the practice, administration, or management of a financial product or service, except that the term does not include patents for technological inventions." Leahy-Smith America Invents Act, Pub. L. No. 112-29, § 18(d)(1), 125 Stat. 331. USPTO regulations further define a "technological invention" based on "whether the claimed subject matter as a whole recites a technological feature that is novel and unobvious over the prior art; and solves a technical problem using a technical solution." 37 C.F.R. § 42.301(b).

79. CBM challenges became available one year from the enactment of the AIA, which was signed into law on September 16, 2011. Leahy-Smith America Invents Act, Pub. L. No. 112-29, § 18(a)(1), 125 Stat. 329.

80. *See* Leahy-Smith America Invents Act, Pub. L. No. 112-29, § 18(a)(1), 125 Stat. 329 (incorporating into CBM review the same standards that apply to post-grant review proceedings as codified in 35 U.S.C. §§ 321–329).

81. 35 U.S.C. § 101.

82. 35 U.S.C. § 102.

83. 35 U.S.C. § 103.

84. 35 U.S.C. § 101.

85. *See* 35 U.S.C. § 101 (allowing an inventor or discoverer to "obtain a patent") (emphasis added).

86. 35 U.S.C. § 112(a).

87. 35 U.S.C. § 112(a).

88. 35 U.S.C. § 112(b).

89. *See* 35 U.S.C. § 321(b) (allowing challenges on any invalidity defense available under 35 U.S.C. § 282(b)(2) and § 282(b)(3), and under 35 U.S.C. § 251). Additional

that a CBM petition warrants institution of a CBM review, the USPTO must find that “it is more likely than not that at least [one] of the claims challenged in the petition is unpatentable,”⁹⁰ or that “the petition raises a novel or unsettled legal question that is important to other patents or patent applications.”⁹¹

3. *Intersection with the Courts*

Both IPR and CBM review proceedings generate estoppel effects, though not in the same way. The estoppel generated by IPR is quite strong. An IPR resulting in a final written decision precludes the petitioner⁹² from asserting any claim in either the USPTO, the federal courts, or the International Trade Commission (ITC) that the petitioner raised, or could have raised, in the IPR proceeding.⁹³

By contrast, a CBM review that results in a final written decision creates full estoppel within the USPTO only—it precludes the petitioner⁹⁴ from asserting any claim in the USPTO that the petitioner raised, or could have raised, in the CBM review proceeding.⁹⁵ With respect to the courts, CBM petitioners are not estopped from invoking those invalidity grounds that they raised at the USPTO.⁹⁶

Both IPR and CBM review proceedings trigger automatic stays of co-pending declaratory judgment litigation. Just as a would-be petitioner cannot challenge a patent in an IPR if it has previously challenged that patent in a civil action,⁹⁷ if a petitioner files such a civil action *after* the IPR petition, then that civil action is automatically stayed.⁹⁸ The stay may be lifted only if the patent owner requests it, if the patent owner claims or counterclaims infringement against the petitioner, or if the petitioner

constraints also apply to novelty- or nonobviousness-based challenges based on pre-AIA § 102 or § 103. *See* Leahy-Smith America Invents Act, Pub. L. No. 112-29, § 18(a)(1)(C), 125 Stat. 330.

90. 35 U.S.C. § 324(a).

91. 35 U.S.C. § 324(b).

92. The same estoppel effects bind the petitioner’s privies and real parties in interest.

93. 35 U.S.C. § 315(e).

94. The same estoppel effects bind the petitioner’s privies and real parties in interest.

95. 35 U.S.C. § 325(e)(1). In general, the provisions of post-grant review, which apply to patents filed under the first-inventor-to-file system, also apply to CBM review, unless Section 18 of the AIA otherwise specifies.

96. 35 U.S.C. § 325(e)(2) generally estops the relitigation in court of arguments that were raised or could have been raised in the USPTO, but AIA § 18 provides that § 325(e)(2) does not apply to CBM proceedings—meaning that patent validity challengers are free to raise those arguments again in the courts.

97. *See* 35 U.S.C. §§ 315(a)(1), (3) and *supra* text accompanying note 65.

98. 35 U.S.C. § 315(a)(2).

dismisses its civil action.⁹⁹ Likewise, if a petitioner files a civil action challenging the patent after filing a CBM petition on the same patent, then that civil action must automatically be stayed.¹⁰⁰ As with IPR, an automatic CBM stay may be lifted only if the patent owner requests it, if the patent owner claims or counterclaims infringement against the petitioner, or if the petitioner dismisses its civil action.¹⁰¹

Meanwhile, courts still have the discretion to stay existing infringement litigation brought by a patent owner pending the outcome of an IPR or CBM review proceeding. For IPRs, where the AIA does not specify the standard for such stays, prior standards pertaining to *ex parte* and *inter partes* reexamination remain valuable, though not conclusive, precedent.¹⁰² Drawing on the reexamination case law, courts continue to consider the familiar three factors in deciding whether to issue stays: the potential for prejudice or tactical disadvantage; the timing of the desired stay relative to that of the administrative proceeding itself; and the likelihood that resolution of the administrative proceeding may simplify the pending litigation.¹⁰³

Notably, the AIA specifies a four-factor test for CBM-related stays. This four-factor test encompasses three factors courts previously used in determining contested motions for stay under the old reexamination system and adds a fourth factor—“whether a stay, or the denial thereof, would reduce the burden of litigation on the parties and on the court.”¹⁰⁴ Moreover, for purposes of “ensur[ing] consistent application of established precedent,” the AIA provides for immediate interlocutory appeal of the district court’s decision regarding stays.¹⁰⁵ It also states the Federal Circuit’s standard of review on appeal from a district court decision “may be de novo.”¹⁰⁶ Using this standard, the Federal Circuit has held that

99. *Id.*

100. 35 U.S.C. § 325(a)(2).

101. *Id.*

102. See Matthew R. Frontz, *Staying Litigation Pending Inter Partes Review and Effects on Patent Litigation*, 24 FED. CIR. B.J. 469, 469 (2015) (“The courts have precedent in reviewing motions to stay litigation pending the reexamination procedure; however, the newly enacted statutory limitations have made this issue ripe for judicial review.”).

103. *Id.* at 473 (citing *Universal Elecs., Inc. v. Universal Remote Control, Inc.*, 943 F. Supp. 2d 1028, 1030–33 (C.D. Cal. 2013)); see also 35 U.S.C. § 311(b) and accompanying text.

104. Leahy-Smith America Invents Act, Pub. L. No. 112-29, § 18(b)(1), 125 Stat. 284.

105. Leahy-Smith America Invents Act, Pub. L. No. 112-29, § 18(b)(2), 125 Stat. 331.

106. *Id.*

district courts have limited discretion to deny CBM-related stays when all claims asserted in litigation are also under CBM review.¹⁰⁷

In general, the AIA's legislative history indicates Congress wanted both IPRs and CBM reviews to serve as a substitute for Article III litigation over patent validity. However, for those defendants who are charged with infringement of a patent that falls within the "covered business method" designation, the broader number of grounds available for challenge, less onerous estoppel provisions, and a codified stay provision likely make CBM review even more attractive than IPR.

II. AGENCY OR COURT: STRATEGIC CHOICES

To describe more fully how these doctrinal frameworks operate in practice, we offer here the largest-scale empirical study to date of ex post administrative scrutiny of patent validity. Our analysis is based on a new dataset of all IPR and CBM petitions filed in the USPTO since the creation of these procedures under the AIA, as well as data on Article III patent cases filed contemporaneously with IPR and CBM petitions, and on requests for litigation stays pending the outcome of administrative challenges to patent validity. Our findings provide a comprehensive view of ex post administrative review that assimilates the more localized findings of prior empirical studies.¹⁰⁸ We use the individual patent as our unit of analysis and, unless otherwise specified, our time period is from September 16, 2012 through June 30, 2015.

Our analysis can be replicated using data from the DocketNavigator service, which provides free and low-cost access to coded metadata about patent cases in the U.S. federal courts as well as the PTAB.¹⁰⁹ Like LexMachina¹¹⁰ and other widely used patent litigation data services, DocketNavigator obtains its underlying litigation data from the federal judiciary's Public Access to Court Electronic Records (PACER) service,¹¹¹

107. *VirtualAgility Inc. v. Salesforce.com, Inc.*, 759 F.3d 1307, 1309–10, 1320 (Fed. Cir. 2014).

108. *E.g.*, Love & Ambwani, *supra* note 2.

109. DOCKETNAVIGATOR, <http://home.docketnavigator.com/ourstory> [<http://perma.cc/B4AP-SB4M>].

110. LEXMACHINA, <https://lexmachina.com/what-we-do/how-it-works> [<http://perma.cc/WA5J-UEDV>].

111. PACER, <http://www.pacer.gov> [<http://perma.cc/YP39-UJZ3>]; *see* Judy L. Heier, *Researching Patent Litigation Made Easy*, RECORDER (May 13, 2013), <http://home.docketnavigator.com/wp-content/uploads/2013/08/The-Recorder-Article.pdf> [<http://perma.cc/3ERK-XS3T>] (stating that DocketNavigator obtains litigation data from PACER).

which is the principal data source of many innovation studies.¹¹² Neither PACER nor the commercial services that rely on it permit researchers to disclose significant portions of their database. Accordingly, we describe the DocketNavigator data we used with the understanding that other researchers can readily access it to replicate our study.¹¹³

A. LITIGANT BEHAVIOR

Like the administrative *ex post* validity challenge mechanisms that preceded the AIA, the IPR and CBM review procedures were established to provide more affordable, more expert, and more accessible adjudication than litigation. However, what would-be patent challengers regard as barriers¹¹⁴ to contesting validity, are safeguards from the perspective of patent owners. We are quite interested, therefore, in discovering whether and under what circumstances IPR and CBM reviews are serving as defensive tools for defendants previously charged in district court with infringement; as tools for preemptive attacks upon patent owners; as mechanisms for harassment and abuse; or as a mix of these functions.

In general, we show that most patents challenged in the PTAB are also challenged in Article III litigation. However, there is no clear relationship between the number of times a patent is challenged in the PTAB and the numbers of times it is asserted in district court. Additionally, while Chemical patents are disproportionately likely to be the subject of a PTAB-only challenge, Computers and Communications (CCM) patents are disproportionately *unlikely* to be challenged only in the PTAB.

We also studied behavior at the level of the individual petitioner. For both CBM reviews and IPRs, the standard substitution model describes

112. *E.g.*, John R. Allison, Mark A. Lemley & David L. Schwartz, *Understanding the Realities of Modern Patent Litigation*, 92 TEX. L. REV. 1769, 1772 (2014) (identifying Lex Machina, which obtains and cleans original PACER information, as the data source); Christopher A. Cotropia & Mark A. Lemley, *Copying in Patent Law*, 87 N.C. L. REV. 1421, 1440–41 (2009) (identifying PACER as the data source); Jay P. Kesan & Gwendolyn G. Ball, *How Are Patent Cases Resolved? An Empirical Examination of the Adjudication and Settlement of Patent Disputes*, 84 WASH. U. L. REV. 237, 266 (2006) (identifying PACER as the data source).

113. *See infra* Appendix A.

114. Such barriers include, for example, potential estoppel in the federal courts from initiating an administrative validity challenge in the USPTO. *Supra* Section I.C.3. More generally, as discussed in detail in Part II, patent challengers face a significant collective action problem. *See* Thomas, *supra* note 36, at 333 (noting that third parties to a successful validity challenge “may readily free ride from the efforts of the former patentee and the opponent, employing the teachings of the invalidated patent to practice the invention without compensation to anyone”).

the majority of cases. Notably, however, in the context of IPRs, the percentage of petitioners who fall outside the standard model because they have not *themselves* previously been sued on the patent in question is surprisingly substantial, on the order of 30 percent. This percentage is particularly high with respect to Drugs and Medical patents. Also notable is the extent to which petitioners that have not previously been sued join the same petitions as those that have been sued. In the case of Drugs and Medical patents, for example, petitioners that have not previously been defendants disproportionately appear to be engaged in collective action with those that are defendants.

1. *IPR and CBM Petitions: Descriptive Statistics*

a) IPR Petitions

Through the end of June 2015, petitioners have filed 3,157 petitions for *inter partes* review. As Figure 1 shows, these filings began slowly in September 2012, when the IPR procedure became available, and have risen from twenty petitions per month to roughly 140 petitions per month.

These petitions have been distributed unevenly across technology areas. The National Bureau of Economic Research (NBER) categorizes patents into six different technology areas: (1) Chemical (excluding Drugs); (2) Computers and Communications (CCM); (3) Drugs and Medical; (4) Electrical and Electronics; (5) Mechanical; and (6) Others.¹¹⁶ As Figure 2 shows, IPR petitions disaggregated by NBER's six-part category scheme have predominantly challenged CCM-related patents, which account for just over half (50.4%) of all IPR petitions. Figure 3 confirms this trend has persisted from the start, with cumulative filings in CCM-related IPR petitions rising considerably faster than those in all other technology areas.

Although IPR petitions may challenge patent claims as to either novelty or nonobviousness, nonobviousness challenges predominate across all major technology areas. As Figure 4 shows, nearly all IPR petitions include a nonobviousness challenge, whereas the proportion of IPR petitions that include a novelty challenge varies considerably by technology. The preference for including nonobviousness as a basis for

115. Figures are presented in Appendix B

116. See generally Bronwyn H. Hall, Adam B. Jaffe & Manuel Trajtenberg, *The NBER Patent Citations Data File: Lessons, Insights and Methodological Tools* 13 (Nat'l Bureau of Econ. Research, Working Paper No. 8498, 2001), <http://www.nber.org/patents> [<http://perma.cc/NY76-VHVV>] (articulating and defining the NBER classification system and its concordance with the U.S. Patent Classification system).

challenge is not surprising. While a novelty-based challenge must rest on a single reference, a nonobviousness-based challenge can presumably take advantage of the ability of PTAB judges to engage in complex reasoning that combines multiple references.¹¹⁷

b) CBM Petitions

Compared to IPR petitions, usage of the CBM procedure has been considerably smaller in scale. Through the end of May 2015, petitioners have filed 362 petitions for CBM review. As Figure 5 shows, these filings have averaged between ten and fifteen petitions per month. Moreover, because CBM proceedings are oriented by definition toward business-method-related technologies such as information and communications, it is unsurprising that an overwhelming majority (82.2%) of CBM petitions challenge Computers and Communications-related patents. Mechanical-related patents make up another 15.9% of CBM petitions, and only a negligible share of CBM petitions fall in any other category. Figure 6 illustrates these trends.

Unlike IPR petitions, CBM petitions may challenge patent claims on a fuller range of patentability requirements: in addition to novelty and nonobviousness, subject-matter eligibility, enablement, written description, and indefiniteness are available grounds. Across this range of options, however, petitioners have focused their attention primarily on subject-matter eligibility and nonobviousness. As Figure 7 shows, 68.6% of CBM petitions challenged the subject-matter eligibility of the patent in dispute, and 71.1% challenged the nonobviousness of the patent. Just under half (48.3%) challenged novelty. By contrast, challenges as to enablement, written description, and indefiniteness each arose in fewer than 20% of petitions.

As with IPR petitions, the relative preference for nonobviousness challenges over novelty challenges in CBM petitions is rational given the greater availability of combining prior art references in evaluating nonobviousness. In addition, the strong preference for subject-matter eligibility challenges is consistent with the widespread view among critics of business method patents that such patents are not just narrowly

117. John Schroeder, *First Ever Inter Partes Review Decision Finds Claims Not Patentable*, LEXOLOGY (Nov. 25, 2013), <http://www.lexology.com/library/detail.aspx?g=d699d660-d5da-4953-af0f-a88e3d3152d2> [perma.cc/CW4C-DGK6] (noting “the general consensus that *inter partes* review may yield better results [than juries in district court litigation] when relying on complex invalidity arguments hinging on a combination of prior art references”).

problematic for inadequate disclosure in the patent specification or lack clarity in the claims—problems that are more the purview of enablement, written description, and indefiniteness—but instead are outside the scope of what should be eligible for patent protection in the first place.¹¹⁸

Beyond these basic PTAB filing trends, we find that a number of patents have been targets of serial challenges spread across both multiple petitions and multiple challengers in IPR petitions. Patents in the Chemical, CCM, and Electrical areas are particularly prone to multiple petitions. As Figure 8 shows, a majority of patents in each of these fields were the subject of multiple IPR petitions: 60.6% of Chemical patents, 50.9% of CCM patents, and 58.4% of Electrical patents. Figure 9 shows how these serial challenges are distributed within technology categories, notably that the highest volume of serial challenges is in the CCM area. We are currently studying the precise nature of these serial challenges (for example, whether they are being brought by the same petitioner) to determine whether they could represent harassment and therefore are problematic from a policy perspective. The frequency of serial challenge to a patent may also be related to the number of defendants against whom the patent is asserted in court.

We turn next to the general question of the relationship between patent challenges at the PTAB and patent litigation in the district courts.

118. This view was held by Senators Schumer and Kyl, who proposed a version of the CBM provision as part of a floor managers' amendment on March 1, 2011. In his March 2011 Senate floor testimony, Senator Schumer described business method patents as "the bane of the patent world" and castigated the decision the Federal Circuit in *State Street Bank* to allow such patents. 157 CONG. REC. S1363 (March 8, 2011) (statement of Senator Schumer). Among many Senators on the Republican side, positions were equally strong. The Senate Republican Policy Committee's summary of § 18, introduced into the Congressional Record by Senator Kyl, stated (somewhat inaccurately):

Recent court decisions, culminating in last year Supreme Court decision in *Bilski v. Kappos*, have sharply pulled back on the patenting of business methods, emphasizing that these "inventions" are too abstract to be patentable. In the intervening years, however, PTO was obliged to issue a large number of business-method patents, many or possibly all of which are no longer valid. The Schumer proceeding offers a relatively cheap alternative to civil litigation for challenging these patents, and will reduce the burden on the courts of dealing with the backwash of invalid business-method patents.

157 CONG. REC. S1367 (Mar. 8, 2011) (statement of Sen. Kyl); see also *eBay Inc. v. MercExchange, L.L.C.*, 547 U.S. 388, 396–97 (2006) (Kennedy, J., concurring) (emphasizing the "nature of the patent" as a newly relevant consideration in enforcement and accusing business method patents in particular of "potential vagueness and suspect validity").

2. *Article III Litigation*

Contemporaneous with petitions for IPR and CBM review in the USPTO, patent litigation in the federal courts has continued apace. To investigate the interaction between these two fora, we collected data on all 24,162 patent cases filed between September 16, 2011, and June 30, 2015, in the federal district courts.¹¹⁹

Many of these cases involved multiple patents-in-suit, and we observed a total of 47,764 patent assertions across these cases,¹²⁰ or an average of 1.98 assertions per patent case. Figure 10 shows the trend in patent cases over this period rising from 150 case filings per month in September 2011 to an average of over 500 case filings per month by June 2015. These petitions have also been distributed unevenly across technology areas. Figure 11 shows that patent cases have predominantly involved CCM-related patents, which far outpace all other technology areas, and that this trend has persisted from the start, with cumulative filings in CCM-related patent cases rising considerably faster than in all other technology areas.¹²¹

During this time, a total of 14,218 patents were either challenged in an IPR or CBM petition, asserted in litigation, or both. A subset of 13,557 patents were involved in litigation alone; 298 patents were involved in a USPTO proceeding alone; and 1,968 patents were involved in both. Accordingly, about 12.7% of litigated patents are also being challenged in the PTAB,¹²² and about 86.8% of IPR- or CBM-challenged patents are also being litigated in the federal courts.¹²³

119. We chose September 16, 2011 as our starting date for district court litigation because it represents the first date on which patents asserted in litigation could become the subject of a PTAB filing. Consistent with our interest in examining the interaction between assertion by patent owners and PTAB petitions, we excluded declaratory judgment actions. In any event, as discussed in the text, the AIA essentially makes declaratory judgment actions unavailable to those who file PTAB petitions. *See supra* Part I.

120. Though the data that we collected include cases where design and plant patents were asserted (either exclusively or together with utility patents), we focus our analysis on utility patents.

121. Because district court cases can (and frequently do) involve multiple-patents in a single suit—unlike IPR or CBM petitions, which are necessarily limited to a single patent—we calculate technology trends by aggregating a technology's relative share among the patents that were asserted in each case. For example, a patent case involving three CCM patents and two Electrical patents would have been counted as 0.6 of a CCM case and 0.4 of an Electrical case.

122. This is calculated as $1968 / (1968 + 13557) = 12.7\%$.

123. This is calculated as $1968 / (1968 + 298) = 86.8\%$.

These measures suggest validity challenges in the USPTO are, indeed, connected with the threat or fact of infringement litigation, for a large majority of challenged patents are also asserted in court. Our data indicate that patents challenged in the PTAB are, on average, also asserted at least three times in court. As Figure 12 indicates, however, this average reflects considerable variation (as shown by the error bars representing one standard deviation of the mean). At least when the group of patents involved in IPR and CBM proceedings is considered as a whole—that is, without disaggregation by technology and district court—the relationship between the number of IPR or CBM petitions that were filed on a patent and the number of times that the patent was asserted in district court is not monotonic. Finally, of course, most patents challenged in district court are not asserted at the PTAB.

To further investigate the relationship between PTAB challenges and Article III assertions, we evaluated a series of measures constructed from the underlying data.

a) IPR and CBM Reviews with Litigation in the Offing

In light of the intended uses of IPR and CBM review as substitutes for federal court litigation, notably, in a number of cases, a given patent was challenged in a PTAB petition *before* that patent was asserted in litigation. This is a relatively rare occurrence. As of June 30, 2015, 1,968 patents have been the subjects of both a PTAB challenge (either in IPR or CBM) and of district court litigation. Only fifty-eight of these patents (3.01%) were challenged in district court litigation simultaneously with or after the first PTAB challenge, rather than before.

Their small number notwithstanding, these cases arguably represent a challenge to the standard model of a PTAB challenge as a substitute for ongoing litigation. However, a relatively small variation to that standard model could encompass the case where litigation was actually imminent. In other words, in these circumstances the filing of a petition in the PTAB was similar to a declaratory judgment action. That is, indeed, what we find. Of the fifty-eight patents that were challenged in the PTAB before any litigation, forty-seven patents (81.0%) were challenged by petitioners who were subsequently named as defendants in federal court litigation over the same patents.

b) IPR and CBM Reviews with No Related Litigation

Another phenomenon that must be reconciled with the standard model is that some patents are challenged in the PTAB but have not been observed in litigation at all, either before or after the petition for IPR or

CBM review. Though a PTAB validity challenge is a reasonable substitute for litigation that has already begun or is imminent, it may be a potentially counterproductive approach for anyone else:¹²⁴ particularly in the case of an IPR (where, as contrasted with the CBM review, the petitioner does not have to be charged in any way with infringement), such a challenger might simply raise unwanted attention to its potentially infringing activities. Indeed, IPRs or CBM reviews with no related litigation are a somewhat rare occurrence. As of June 30, 2015, only 298 patents (13.2%) have been challenged in the PTAB with no related litigation observed in the federal courts. But even the existence of such a subset might be considered peculiar.

There are several potential reasons for this unexpected subset. One is statistical censoring: the PTAB challenges are simply so recent that the patent owner has not filed responsive litigation yet, but may do so in the relatively near term. Censoring, however, does not appear to explain the subset fully. Petitions on such “PTAB-only” patents have been filed from the earliest days of IPR and CBM review in September 2012. As of June 30, 2015, 332 petitions have been filed on patents with no related litigation observed and 163 (49.1%) of those petitions were filed more than one year ago from the present or in or before June 2014. In other words, many of the patent owners have had ample time to bring infringement actions against the petitioners who filed for IPR or CBM review and have not yet done so. So it is still possible, but increasingly unlikely, that a patent owner who has not asserted a patent against an IPR or CBM challenger will do so now.

A second possible reason for this phenomenon is statistical selection, including technology-specific selection: where a PTAB validity challenge is sufficiently strong, and a patent owner’s countervailing infringement claim against the PTAB challenger is sufficiently weak, an invalidity challenge might arise without any corresponding infringement assertion. This kind of selection effect, however, would require that both parties have information *ex ante* about the relative merits of each other’s case, i.e., about the boundaries and legal viability of the patent in dispute, that is

124. For example, the filing fees for IPR are \$9,000 at the petition stage and \$14,000 at the post-institution stage. 37 C.F.R. § 42.15(a). The filing fees for CBM review are even higher: \$12,000 at the petition stage and \$18,000 at the post-institution stage. 37 C.F.R. § 42.15(b). Contemporaneous estimates of average attorney costs were over \$130,000. Olga Berson, *Challenging Patent Validity Under the AIA: Strategic and Tactical Considerations When Deciding Whether to Pursue Ex parte Reexamination or Inter Partes Review As Part of the Overall Litigation Strategy*, 2012 WL 6636452, *12 (2012).

both adequate and roughly symmetric. Such ex ante clarity may be possible for Chemical and Drugs and Medical patents, where technical nomenclature is standardized and the boundaries of the invention are amenable to delineation.¹²⁵ Ex ante clarity may even be possible for Electrical and Mechanical patents if the patent discloses sufficiently detailed structural information. However, patents on CCM inventions that are claimed in functional terms would be much less likely to provide enough ex ante clarity that a PTAB challenge would be so plainly strong, and a retaliatory infringement suit so plainly weak, as to produce an IPR or CBM review with no litigation in response.

Additionally, in at least some technology areas, the number of patents that are clearly “important” as a matter of potential litigation risk may be relatively small and easy to identify. Particularly in the context of IPRs (which can be filed even without any assertion of infringement on the part of the patentee), the high volume of CCM-related patents may make it unclear which patents are most important.

The data are consistent with technology-specific selection effects across the three subsets of (1) patents that were only challenged in the PTAB, (2) patents that were only asserted in litigation, and (3) patents that were both challenged in the PTAB and asserted in litigation as summarized in Figure 13. Comparing PTAB-only patents with district court-only patents, the technology distributions were mostly similar. In both subsets, CCM patents accounted for about a third (32.8% and 37.1%, respectively); Drugs and Medical patents about a fifth (20.6% and 19.7%, respectively); Electrical patents a little less than a seventh (13.9% and 11.3%, respectively); and Mechanical patents a little more than a tenth (11.5% and 10.1%, respectively).¹²⁶ Only Chemical patents occupied a significantly greater share of PTAB-only patents (12.5%) than of district court-only patents (4.9%).¹²⁷

The most notable difference was for patents that were both challenged in the PTAB and asserted in district court. A majority of these PTAB-and-district-court patents (54.7%) were in the CCM technology area, as

125. Peter S. Menell & Michael J. Meurer, *Notice Failure and Notice Externalities*, 5 J. LEGAL ANALYSIS 1, 36 (2013). Indeed, in the case of certain drugs (so-called small molecule drugs), patents asserted to cover the drug are specifically on the FDA “Orange Book.”

126. These differences were not statistically significant ($p > 0.05$ using a two-tailed test of proportions).

127. Conversely, “Other” patents occupied a greater share of district court-only patents (16.9%) than of PTAB-only patents (8.8%).

compared with 32.8% of PTAB-only patents.¹²⁸ This underrepresentation of CCM patents in the PTAB-only group is consistent with the expected lower likelihood that CCM patents offer enough ex ante clarity and evidence of importance to produce PTAB challenges in situations where there is no federal court litigation.

Having considered the special cases of PTAB validity challenges that either precede a district court litigation or have no related litigation at all, we now turn to the standard model of PTAB validity challenge as a direct response by a defendant in prior infringement litigation.

c) CBM and IPR Challenges As Direct Self Interest

As we have discussed, a defendant that challenges a patent's validity in the USPTO *after* the patent has been asserted in litigation is the standard use of CBM and IPR petitions. The USPTO's expertise substitutes for the generalist orientation of the courts. We find that, overall, CBM and IPR petitions are in fact predominantly assertions of the petitioners' own direct interests with respect to infringement liability on the particular patent being challenged.

In the majority of cases, petitioners for CBM review have previously been defendants in federal court litigation where the same patent was asserted. Two related measures support this finding. One is the share of CBM *petitioners* (77.9%) who have previously been defendants in district court litigations involving the patents they later challenge in CBM review. The other is the share of CBM *petitions* (82.7%) in which at least one petitioner was previously a defendant as to the patent now being challenged. These results are perhaps unsurprising, as CBM petitions can only be brought by those sued for, or charged with, infringement. Additionally, though it is not particularly meaningful to speak of technology differences among CBM petitions,¹²⁹ Figures 14a and 14b show that the finding also persists for each NBER technology category.

Similarly, in the case of IPRs, the majority (70%) of IPR petitioners have previously been defendants in district court litigations involving the patents they now challenge. The remaining 30% of cases in which petitioners are not prior defendants do, however, represent an interesting

128. This difference was highly statistically significant ($p < 0.0001$ using a two-tailed test of proportions).

129. This is because the availability of CBM review is defined, and limited, by technology, and as a result, CCM patents have accounted for 82.2% of all CBM Petitions, with 15.9% coming from Mechanical patents and 1.9% from Other patents. *See infra* Figure 6.

puzzle, particularly if one looks across technologies, and also at the percentage of petitions in which at least one petitioner was previously a defendant. We turn next to this puzzle.

d) IPR Challenges by Entities That Were Not Prior Defendants

As Figure 15a shows, the percentage of IPR petitioners who were not prior defendants varies substantially across technologies. Notably, because only about 48% of petitioners in the Drugs and Medical area have previously been sued, over half of all petitioners in this technology are non-standard. In some cases, generic firms may be filing even prior to being sued in order to clear the path toward eventual entry into the market. In other cases, we know from reading IPR petitions to identify petitioners that third parties have been active. One active third party is J. Kyle Bass, the principal of Hayman Capital Management and of the Coalition for Affordable Drugs, who has filed at least twenty-eight petitions.¹³⁰ Another is Erich Spangenberg, the chief executive of the IP Navigation Group and of nXn Partners, who has also filed at least twenty-eight petitions.¹³¹ Both Mr. Bass and Mr. Spangenberg have thus far focused their validity challenges entirely on Drugs and Medical-related patents.

Figures 15a and 15b also reveal substantial disparities in certain technology areas between the share of *petitioners* who were previously sued and the share of IPR *petitions* with at least one petitioner who was previously a defendant on the challenged patent. Specifically, the petitioner vs. petition disparities are quite substantial in the categories of Drugs and Medical (48.5% vs. 70.8%), Mechanical (53.1% vs. 70.2%), and Other (65.5% vs. 82.6%). The disparities reveal that, in each of these technology areas, petitioners who are not prior defendants are joining petitions filed by prior defendants.

Arguably, this collective action is socially beneficial, as it directly addresses the general collective action problem in challenging invalid patents.¹³² However, to the extent collective action takes the form of serial petitions that are joined later to the petition of a prior defendant, it could be seen as harassment and delay. Currently, our data do not allow us to

130. Joseph Walker & Rob Copeland, *New Hedge Fund Strategy: Dispute the Patent, Short the Stock*, WALL ST. J., Apr. 7, 2015, <http://www.wsj.com/articles/hedge-fund-manager-kyle-bass-challenges-jazz-pharmaceuticals-patent-1428417408> [<http://perma.cc/X26M-53QM>].

131. David Segal, *Has Patent, Will Sue: An Alert to Corporate America*, N.Y. TIMES, July 13, 2013, <http://www.nytimes.com/2013/07/14/business/has-patent-will-sue-an-alert-to-corporate-america.html> [<http://perma.cc/R2X6-8D49>].

132. See *supra* text accompanying notes 36–38.

determine exactly *when* nondefendant petitioners are joining the petitions of defendants. PTO regulations do require, however, that a joinder request be filed no later than one month after the institution date of any *inter partes* review for which joinder is requested.¹³³ In ongoing research, we are parsing the joinder data more finely to look at timing and how the regulations are being applied.

In this regard, it bears mention that fostering collective action is the explicit mission of organizations such as Unified Patents, which files patent validity challenges on behalf of its member companies in order to reduce their patent litigation risk.¹³⁴ We expect that, in order to be effective, such member-based organizations would likely file significant numbers of IPR petitions and focus their efforts largely on a single technology area. Unified Patents, for example, has filed at least twenty-four petitions of which seventeen (71%) are against CCM-related patents.

e) Timing Between the Courts and the USPTO

Closely related to the “non-standard” petitioner issue is the question of time lag between Article III assertion and PTAB challenge. Unless the petition includes a request for joinder, a petitioner cannot file an IPR challenge more than a year after it has been sued for infringing a particular patent.¹³⁵ As a result, administrative validity challenges filed more than one year after the last federal court lawsuit prior to a petition are likely to reflect either non-standard petitioners and/or petitioners seeking joinder to earlier petitions.

To investigate these issues further, we measured the time lag between the first IPR petition on a given patent and the federal court litigation on that patent filed *most recently prior* to the first IPR petition. (By definition, the first IPR petition cannot request joinder.) As an additional frame of reference for these results, we calculated the lag between the first IPR petition on a given patent and the *earliest* observed federal court litigation on that patent. The latter measure takes a broad view of how court-agency lags are distributed and is likely to contain a small, but non-trivial number of instances where the lag is greater than one year. The reason is that, for repeatedly-asserted patents, the first defendant sued need not be the one that mounts a validity challenge in the USPTO.

133. 37 C.F.R. § 42.122(b).

134. UNIFIED PATENTS INC., <http://www.unifiedpatents.com/faq> [<http://perma.cc/K4XC-4Y23>].

135. *See* 35 U.S.C. § 315(b).

As Figure 17 shows, quite a few patents fit this latter profile: nearly a quarter of the distribution (23.4%) exceeds the one-year lag from the earliest observed federal court litigation on a given patent, reaching upwards of three years for some patents. Notably, a small share of patents, roughly 3.3%, shows a negative lag indicating the first IPR petition against the patent *preceded* the first federal court assertion of the patent.¹³⁶ For these patents, administrative validity challenges are not defensive in the traditional sense, as no offensive litigation has yet been observed; rather, they are, at most, preemptive. Most IPR petitions, however, fall within the zero-to-one-year range, distributed symmetrically about a median lag of six months, with a modal spike at the one-year deadline.

Meanwhile, measuring from the *last* pre-IPR federal court lawsuit to the first IPR petition is likely to capture not only non-standard petitioners but also cases where earlier lawsuits against others have revealed useful information about the patent owner's enforcement strategy so that less time is needed to decide whether and how to prepare an IPR challenge. This is, in fact, what the data reveal in Figure 18. The majority of cases fall again within the zero-to-one-year range, but with a median lag roughly four months less than in Figure 17. A far smaller share of the distribution (11.4%) exceeds one year—presumably this 11.4% comprises of non-standard petitioners only. As before, a modal spike near and at a one-year lag indicates that litigants wait for the statutory deadline.

These direct and indirect measures suggest that challenges to patent validity through *inter partes* review are primarily—though not exclusively—a defensive response to existing litigation. In most cases, a prior defendant files an administrative challenge. Other entities, acting on this revealed information, may also respond with petitions for validity review.

We now turn to another aspect of strategic behavior in patent litigation that has previously presented policy concerns: the tendency of patent cases to be filed disproportionately in a few judicial districts, so much so that these districts are now widely identified with patent litigation.

f) District-Specific Effects

Skewed distribution of patent litigation toward particular high-volume judicial districts and litigant forum-shopping, which not only results from this skew but also contributes to it, are well documented.¹³⁷ It is likely,

136. As we have discussed, these preemptively-challenged patents may reflect litigation in the offing or else no related litigation. *See supra* Sections III.A.2.a–b.

137. *See generally* notes 28–33 and accompanying text.

then, that such leading patent courts should send commensurately greater numbers of patents into PTAB validity challenges as well. Yet in this regard, the data show a surprising effect. Of the eight leading district courts—which together account for nearly 70% of litigated patents during the observed time period—the top three courts were *over*represented in sending patents into PTAB validity challenges, and the remaining five were *under*represented.

Figure 16 depicts the fraction of *all litigated patents* that were litigated at least once in a given court and the fraction of all *IPR-challenged patents* that were litigated at least once in the same court, across the top eight districts for patent litigation. The latter fraction was significantly higher than the former for the District of Delaware (41.1% vs. 34.4%), the Eastern District of Texas (41.4% vs. 28.5%), and the Northern District of California (21.6% vs. 15.2%),¹³⁸ indicating that patents litigated in those districts were unusually likely to be challenged in *inter partes* review. The effect was reversed for the other high-volume patent districts, including the Central District of California (14.1% vs. 16.0%), the District of New Jersey (10.0% vs. 13.0%), and the Northern District of Illinois (4.8% vs. 9.6%).¹³⁹

The great disparity we see in the Eastern District of Texas is unsurprising—the court’s strong pro-patentee reputation¹⁴⁰ would be expected to drive defendants to a more strategically favorable forum. This effect is likely in spite of the apparently low likelihood of defendants either filing or being granted stays in the Eastern District of Texas.¹⁴¹ In the cases of the District of Delaware and the Northern District of California, the reasons for disproportionately high IPR filings are less clear. Defendants in those districts may be encouraged, however, by the high rate of stay grants in these districts.¹⁴²

138. These differences were highly significant ($p < 0.0001$ using a two-tailed test of proportions).

139. These differences were all significant as well ($p < 0.05$ using a two-tailed test of proportions).

140. *See, e.g.*, Vishnubhakat, *supra* note 30, at 65 (discussing the reputation of the Eastern District of Texas for producing pro-patentee outcomes).

141. *PTAB Stay Stats: 2012 to May 31, 2015*, WINSTON & STRAWN LLP, <http://www.winston.com/en/thought-leadership/winston-publishes-stats-on-ptab-stays.html> [<https://perma.cc/3W7H-Y3Q4>].

142. *Id.*

B. AGENCY DECISIONS

When petitioned, the PTAB must decide whether to institute an IPR or CBM review on the grounds petitioned. If it decides to institute a review, the PTAB must then adjudicate the case on its merits. Decisions on institution and on the merits are interdependent in that the legal standard for instituting an IPR is whether the petitioner is reasonably likely to succeed as to at least one claim, and the legal standard for instituting a CBM review is whether the petitioner is more likely than not to prevail as to at least one claim.¹⁴³ Therefore, the rates of institution are particularly important because the very fact of institution is, by statutory design, a credible signal about the ultimate outcome of the validity challenge.

In the case of IPR, an early study on whether the USPTO had decided to institute or not (823 IPR petitions at the time) found that 84.0% were granted as to at least one challenged claim.¹⁴⁴ Further-updated data confirm this point estimate but reveal a slow and consistent decline in the institution rate. Figure 19 compares over time (1) the running total number of IPR petition filings, (2) the running total number of institution decisions, and (3) the running total number of institution decisions granting at least one challenged claim. Calculating the institution rate as (2) divided by (3) over time, Figure 20 shows that the rate has been declining and is currently 74.8%.

The earlier study also found that 74.0% of at-least-partially instituted petitions were fully instituted. Our data conflict on this point. We find that 51.4% of at-least-partially instituted petitions were fully instituted¹⁴⁵ and, more relevantly, that 38.4% of petitions that received an institution decision were fully instituted. These trends are summarized in Figure 21.

In addition to general institution rates, we also disaggregate institution rates by technology area and by the grounds on which patent validity was

143. See 35 U.S.C. § 324(b), *supra* note 91; LEXMACHINA, *supra* note 110, and accompanying text. A CBM review may also be instituted for policy reasons. See sources cited *supra* note 111 and accompanying text.

144. Love & Ambwani, *supra* note 2, at 100.

145. In order to compare rates of partial and full institution, we keep consistent the denominator of petitions that have received an institution decision whereas the Love-Ambwani study reports full institution as a fraction of partial institution. By that measure, too, our findings differ: we find that 41.2% of at-least-partially instituted petitions were fully instituted, not 74.0%. This discrepancy controls for the amount of data available at the time of the earlier study—roughly the end of March 2014, when 843 institution decisions had been made and 699 petitions had been at-least-partially instituted.

challenged. Figure 22 shows the rates at which institutions are granted and denied across technologies for petitions arguing a lack of novelty. Petitions on Drugs and Medical-related patents have a 59.9% likelihood of being denied,¹⁴⁶ and in all other technologies, petitions are as likely as not to be instituted ($p > 0.05$). Figure 23 shows the rates at which institutions are granted and denied across technologies for petitions arguing a lack of nonobviousness. Perhaps not surprisingly, given the ability of expert judges to combine multiple references, nonobviousness petitions are more likely than not to be instituted across all technology areas. Nonobviousness challenges to Chemical patents are particularly likely to be granted, with an institution rate of 68.5%.¹⁴⁷

Meanwhile, for CBM petitions, comparing technology categories is not particularly meaningful, as the definition of covered business method patents in practice overlaps substantially with CCM-related patents. Instead, because CBM review allows the full range of legal grounds on which to challenge validity¹⁴⁸ and because petitioners themselves have availed themselves of these grounds to varying degrees,¹⁴⁹ comparing the rates at which CBM petitions have been instituted with respect to each of these grounds is more meaningful.

Figure 7 previously showed that subject-matter eligibility under § 101, novelty under § 102, and nonobviousness under § 103 were the major grounds on which CBM petitions have been filed whereas the enablement, written description, and definiteness requirements of § 112 have been employed relatively infrequently. Because CBM review arose out of categorical resistance to business methods as patent-eligible subject matter, and inception of CBM review coincided with Supreme Court decisions substantially strengthening patent eligibility requirements, we expected that subject-matter challenges would be the most fertile ground for decisions to institute CBM petitions. We expected that the remaining grounds would be likely to garner fewer PTAB institutions, though in the particular case of nonobviousness, the higher standard imposed by the Supreme Court's 2007 decision in *KSR Int'l Co. v. Teleflex Inc.*¹⁵⁰ might have an impact.

146. The differences between respective likelihoods of grant and denial are highly significant ($p < 0.001$ using a two-tailed test of proportions).

147. The differences between respective likelihoods of grant and denial are significant ($p < 0.05$) for Mechanical-related petitions, and highly significant for all other technologies ($p < 0.005$). Comparisons use a two-tailed test of proportions.

148. See 35 U.S.C. § 315(a)(2); *Frontz*, *supra* note 102; *supra* note 107.

149. See *infra* Figure 7.

150. 550 U.S. 398 (2007).

Figure 24 confirms our hypothesis that subject matter eligibility would dominate the CBM procedure. Subject matter eligibility-based CBM petitions are overwhelmingly instituted, at a rate of 70.9%.¹⁵¹ For all other grounds, decisions *not to institute* predominate by large margins: challenges based on novelty were denied at a rate of 59.3%; nonobviousness, 56.9%; enablement, 100%; written description, 71.7%; and definiteness, 64.7%.¹⁵²

C. COURT DECISIONS

While the USPTO evaluates and decides invalidity petitions, federal courts must decide how to manage ongoing patent infringement litigation on which these validity challenges can have considerable impact. The most frequent decision for courts is when to issue a stay. The ability of defendants to obtain litigation stays pending the outcome of validity challenges is a powerful strategic consideration in managing both the immediate cost of litigation and the eventual threat of liability. Conversely, the tendency of courts to grant such stays is a powerful strategic consideration for patent owners to enforce their rights effectively and deflect potential harassment and abuse by challengers.

Table 1. Results of Motions to Stay Pending *Inter Partes Review*

<i>Inter Partes Review</i>	Fully Denied	Denied without prejudice	Denied in part granted in part	Granted
Motion to Stay Pending <i>Inter Partes Review</i>	67	47	22	113
Renewed Motion to Stay Pending <i>Inter Partes Review</i>	2	0	2	11
Stipulated/Agreed Motion to Stay Pending <i>Inter Partes Review</i>	0	2	2	1
Sua Sponte Motion to Stay Pending <i>Inter Partes Review</i>	0	0	0	1
Subtotal	69	49	26	126
Share	25.6%	18.2%	9.63%	46.7%

151. The difference between likelihoods of grant and denial is highly significant ($p < 0.0001$ using a two-tailed test of proportions).

152. The differences between likelihoods of grant and denial were all significant ($p < 0.05$) and in many cases highly significant ($p < 0.005$) using a two-tailed test of proportions.

Table 2. Results of Motions to Stay Pending Covered Business Method Review

CBM Review	Denied	Denied without prejudice	Denied in part granted in part	Granted
Motion to Stay Pending CBM Review	12	7	9	26
Renewed Motion to Stay Pending CBM Review	1	0	0	7
Sua Sponte Motion to Stay Pending CBM Review	0	0	0	1
Subtotal	13	7	9	34
Share	20.6%	11.1%	14.3%	54.0%

Tables 1 and 2 provide basic statistics regarding motions for stays pending IPR and CBM proceedings, as well as federal court adjudications of such motions. As the statistics indicate, full denials of motions to stay (as contrasted to the combined total of “denials without prejudice,” partial grants, and grants) are relatively rare, particularly in the context of CBM reviews.

III. DISCUSSION

Our analysis yields several “top-line” findings regarding strategic choices by parties in PTAB proceedings. Most patents challenged at the PTAB are also in Article III litigation—PTAB petitions on patents that are not being litigated in an Article III court are relatively rare. Moreover, if there is no Article III litigation, CCM patents are *particularly* unlikely to be challenged at the PTAB. In this area of technology, district court assertion may be necessary to force parties to overcome several technology-specific barriers to a petition. These barriers may include an absence of clarity regarding the merits of a validity challenge created by lack of boundary notice, as well as informational hurdles created by the sheer volume of CCM patents.

Just as Article III litigation disproportionately accompanies PTAB petitions on CCM patents, IPR petitions in the CCM field appear to be brought largely by the same entities that are defendants in Article III litigation. Both the share of CCM petitions involving at least one prior Article III defendant (81.5%) and the share of CCM petitioners who are themselves prior defendants (76.3%) are quite high. This result suggests that non-standard petitioners are, at least thus far, playing a relatively modest role in IPR petitions brought against CCM patents. Thus, to the

extent we see a substantial amount of serial petitioning in the CCM area, this is being generated by prior defendants.

The most significant role for non-standard petitioners is in the Drugs and Medical area. For Drugs and Medical-related challenges, previously sued defendants make up only a minority of petitioners (48.5%). Non-standard petitioners also appear to be engaging in significant collective action with standard petitioners. A substantial majority (70.8%) of petitions in this area contains at least one petitioner who has previously been sued. Litigation defendants in the Drugs and Medical field are clearly bringing aboard entities that have not yet been sued. In order to address policy implications (e.g., whether it is socially beneficial collective action or possible harassment), we are currently investigating the important policy question of precisely *when* these other entities are getting on board.

In addition to technology-specific effects, we see district-specific effects. To a statistically significant degree, patents litigated in the “top three” district courts—the Eastern District of Texas, the District of Delaware, and the Northern District of California—are more likely to be the subject of an IPR than patents litigated in other districts. The statistically and numerically significant results for the Eastern District of Texas are unsurprising. Whether or not judges in the Eastern District grant stays for ongoing litigation (and the available data suggest defendants are less likely to seek or be granted stays than in other districts), the Eastern District’s “pro-plaintiff” reputation makes filing a PTAB petition an obvious choice for any defendant. In the case of Delaware and the Northern District of California, the reasons for disproportionately high IPR filings are less clear. Defendants in those districts may be encouraged, however, by the high rate of stay grants in these districts.

Agency decision-making also exhibits some interesting patterns. Perhaps because high early rates of institution spurred petitioners to challenge somewhat stronger patents, the overall institution rate has decreased over time. Agency decision-making also exhibits differential patterns across technology: specifically, IPR institution rates are significantly higher for CCM patents than for Drug and Medical patents. Meanwhile, nonobviousness represents a stronger ground for securing a favorable institution decision on an IPR than novelty. As for CBM reviews, § 101 is clearly the best route for challengers.

In current ongoing work, we are investigating both more intensively and more formally the interrelated questions of collective action and potential harassment. Specifically, we are investigating the precise nature

and timing of the collective action undertaken both by petitioners that are prior defendants and those that are not prior defendants. We are also interested in whether non-defendant petitioners do in fact become defendants at a later point in time. Additionally, we are developing regression models that assess, conditional on assertion in litigation, what factors influence the likelihood and frequency of a patent being challenged at the PTAB.

IV. CONCLUSION

Our data indicate that the standard model of prior district court defendants bringing PTAB petitions on the patents for which they have been sued explains the majority of PTAB cases. PTAB-only cases are relatively rare, and particularly rare in the case of Computers and Communication (CCM) patents.

That said, a significant minority of IPRs are being brought by entities that are not prior defendants in lawsuits over the patents that they are now challenging. In ongoing research, we are examining the precise role of these non-standard petitioners to examine whether they are engaging in beneficial collective action or in non-beneficial harassment.

Finally, at least thus far, the relative reluctance of the Eastern District of Texas to grant stays does not appear to have impeded entities' disproportionate desire to seek IPRs for patents asserted in the Eastern District. Perhaps more surprisingly, patents asserted in the Northern District of California and in the District of Delaware also see a disproportionate number of IPR petitions.

APPENDIX A: DATA

Our analysis can be replicated using data from DocketNavigator, which provides free and low-cost access to coded metadata about patent cases in the U.S. federal courts and the USPTO Patent Trial and Appeal Board.¹⁵³ In this Appendix, we describe the DocketNavigator data with the understanding that other researchers can readily access it to replicate our study.

A. PTAB DATA

DocketNavigator’s search interface allows minimal queries that can yield large result sets. Thus, to obtain all case information on all petitions filed in the PTAB, we used only one search term: “Patent Trial and Appeal Board (PTAB)” for the “Court/Agency” field. Because the total number of PTAB cases in the DocketNavigator database recently passed 4,000 and search results are displayed one hundred at a time, the results are distributed across forty pages. Detailed party information about cases is bulk-downloadable on a page-by-page basis, i.e., each download contains detailed party information about the cases displayed on the given page of results. Similarly, detailed information about the patents involved in the cases is bulk-downloadable on a page-by-page basis as well. Both sets of case information include multiple variables:

Party Information	Patent Information
Case name	Case name
Court abbreviation	Court abbreviation
Case number	Case number
Case filing date	Case filing date
Party name	Patent
Party roles	Patent title
Firm name	Parties
Attorney name	USPTO class codes
	Cooperative patent class codes

Importantly, case-identifying variables appear in both sets of downloads, allowing them to be merged. To construct our data set, we downloaded this detailed party information as well as patent information for all PTAB petitions and merged them by PTAB case number. The

153. DOCKETNAVIGATOR, <http://www.docketnavigator.com> [<http://perma.cc/QLY4-LJT7>].

merged results yielded a comprehensive set of filing, party, and patent information for each IPR and CBM review petition at the PTAB.

In addition to petition data, DocketNavigator provides searchable data on PTAB institution decisions. As the earliest PTAB institution decisions came in December 2012, a date-based search for decisions issued on or after November 1, 2012 (or any similarly early date), returns a set of all decisions. As with petition data, these results are accessible one hundred at a time on a page-by-page basis. To this end, the “Print Friendly” feature in the search result interface generates a simple formatted table to copy directly into spreadsheet software. The institution decision data contain the following variables:

Institution Decision Information
Patent number
Case name
Case number
Substantive ground for petition
Institution decision on that ground
Relevant patent claims to which the decision pertains
Order filing date

Finally, DocketNavigator provides searchable data on final determinations by the PTAB. Searching for “Patent Trial and Appeal Board (PTAB)” in the “Court/Agency” field returns a set of all determinations. These results, too, are accessible one hundred at a time on a page-by-page basis and available in a simple formatted table through the “Print Friendly” feature in the search result interface. The final determinations data contain the following variables:

Final Determination Information
Patent number
Case name
Case filing date
Determination
Judge
Order filing date

Because this institution decision data and final determination data also contain case-identifying variables that overlap with the filing, party, and patent data, we readily merged this additional information into our data set as well.

B. DISTRICT COURT DATA

We obtained case information on patent litigations filed in the U.S. district courts with a similarly minimal search query in DocketNavigator's primary search interface: "U.S. District Courts (and all districts)" for the "Court/Agency" field. The total number of patent cases in the DocketNavigator database exceeded 55,000 results. Because IPR petitions are generally time-barred one year from the date when a would-be petitioner has been sued on the same patent in U.S. district court,¹⁵⁴ we determined that a reasonably complete set of federal patent litigation would not need to extend more than one year before the IPR mechanism became available. Therefore, we narrowed our search to cases, other than declaratory judgment cases, filed on or after September 16, 2011, one year prior to the enactment of IPR and CBM review mechanisms in the PTAB.

As with PTAB cases, detailed party and patent information about U.S. district court patent cases is bulk-downloadable on a page-by-page basis. Both sets of case information include multiple variables:

Party Information	Patent Information
Case name	Case name
Court abbreviation	Court abbreviation
Case number	Case number
Case filing date	Case filing date
Party name	Patent
Party roles	Patent title
Firm name	Parties
Attorney name	USPTO class codes
	Cooperative patent class codes

To construct our data set, we downloaded this detailed party and patent information for all relevant patent lawsuits and merged them on the case number, producing a comprehensive set of filing, party, and patent information on each patent lawsuit in the U.S. district courts.

154. 35 U.S.C. § 315(b).

APPENDIX B: FIGURES

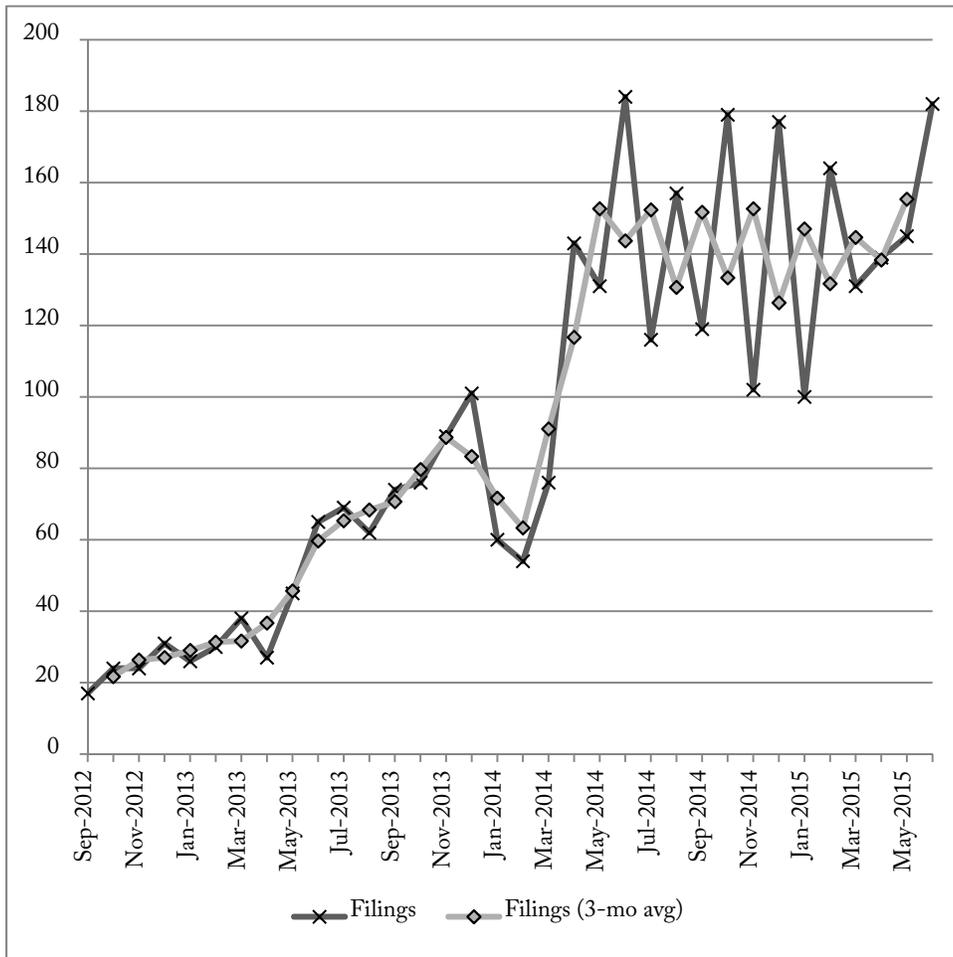


Figure 1: IPR Petition Filings by Month

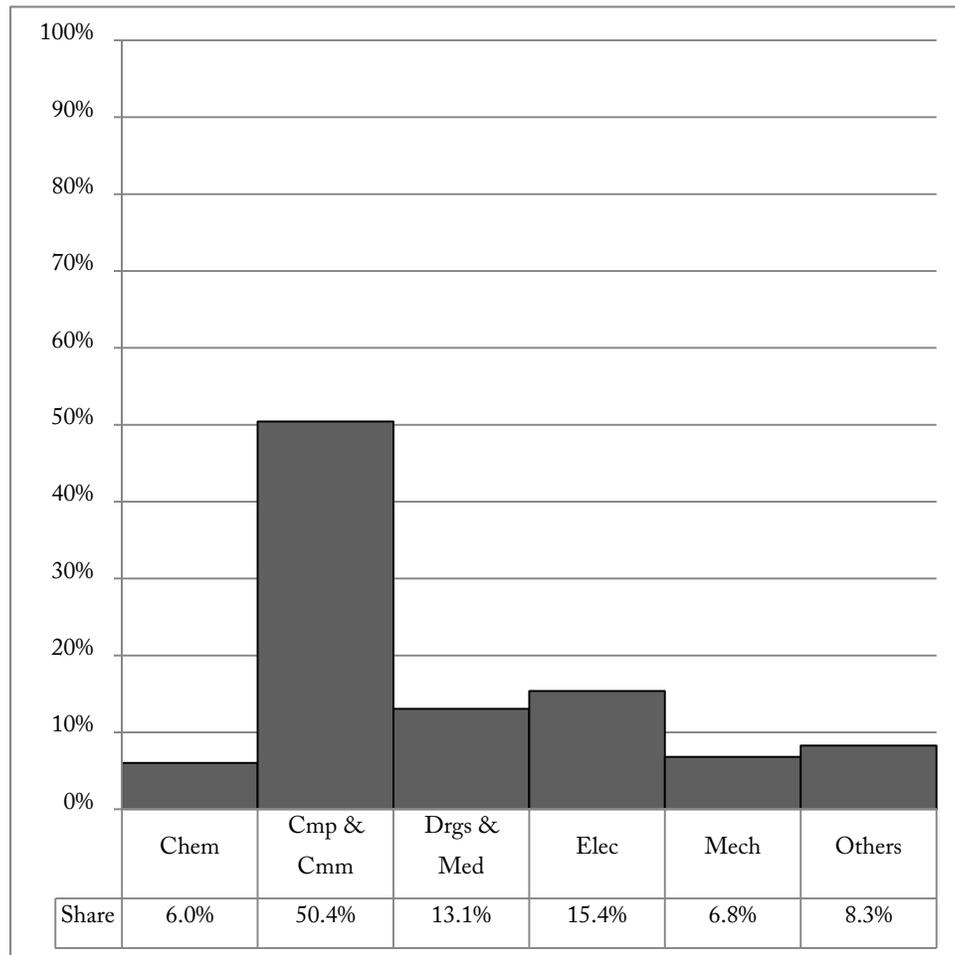


Figure 2: IPR Petition Filings Across Technology

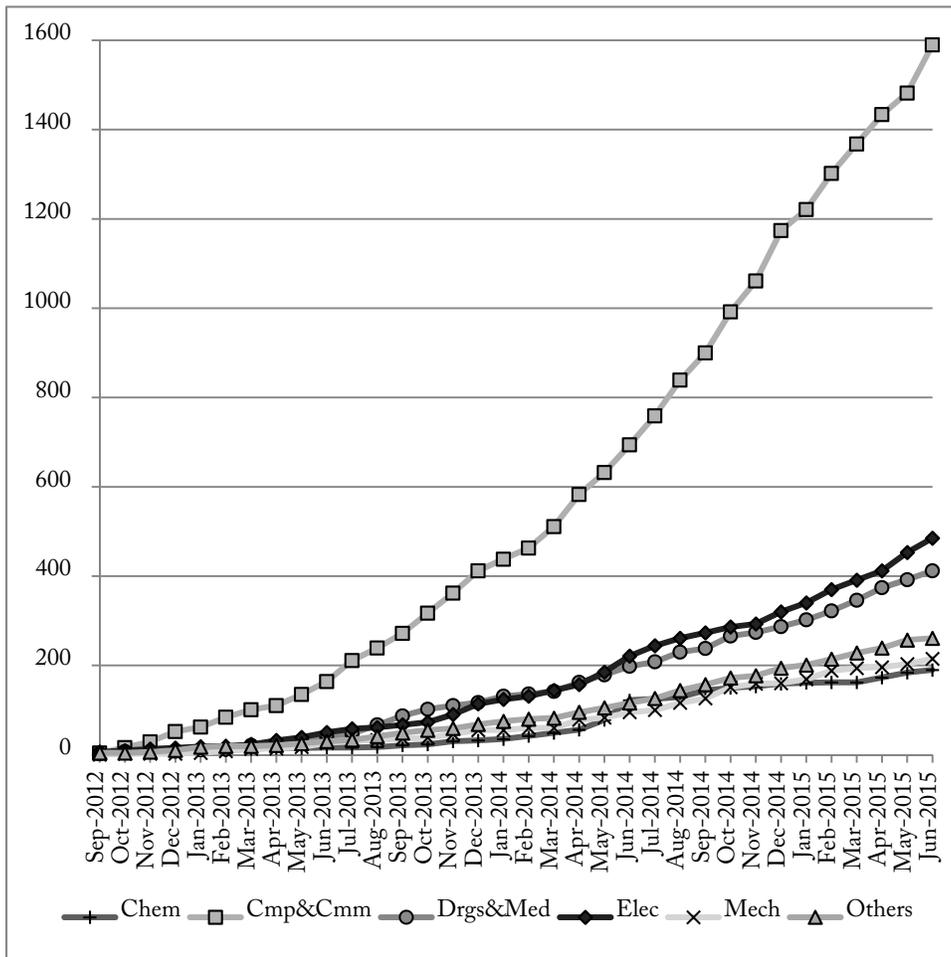


Figure 3: Cumulative IPR Petition Filings Across Technology

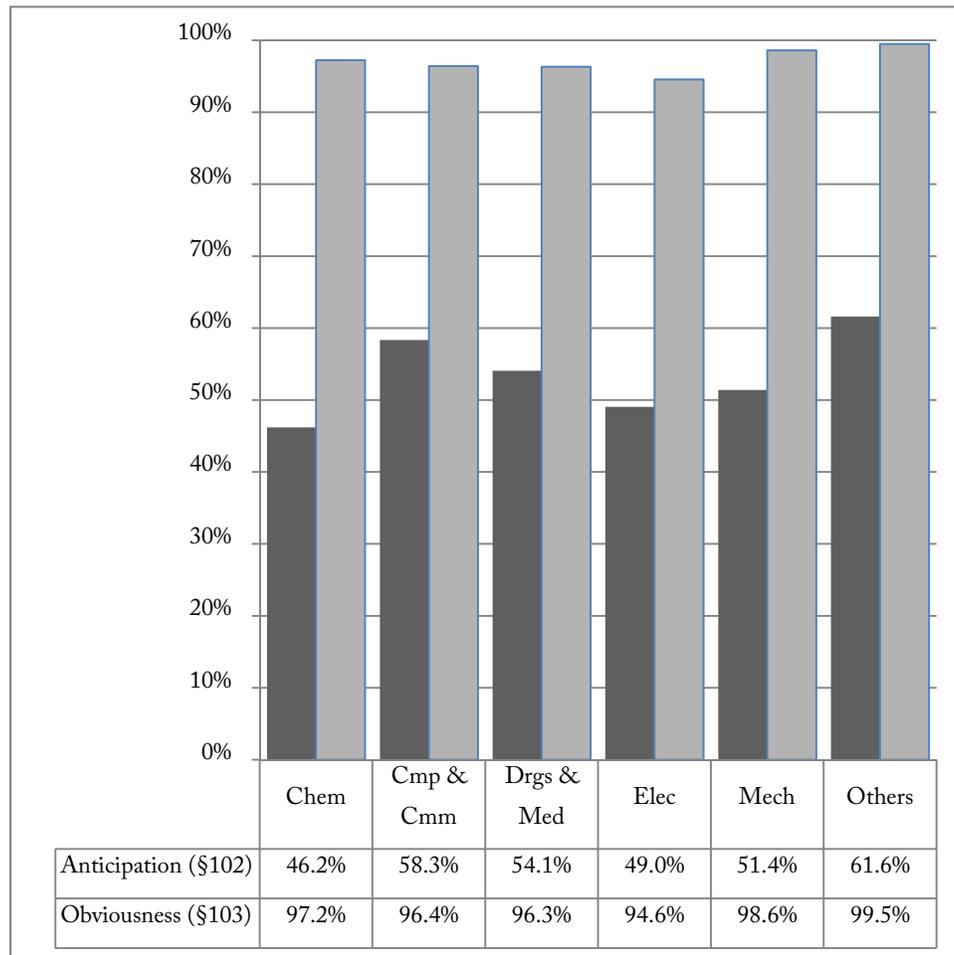


Figure 4: Proportions of IPR Petitions Containing Each Grounds for Challenge, Across Technology Area

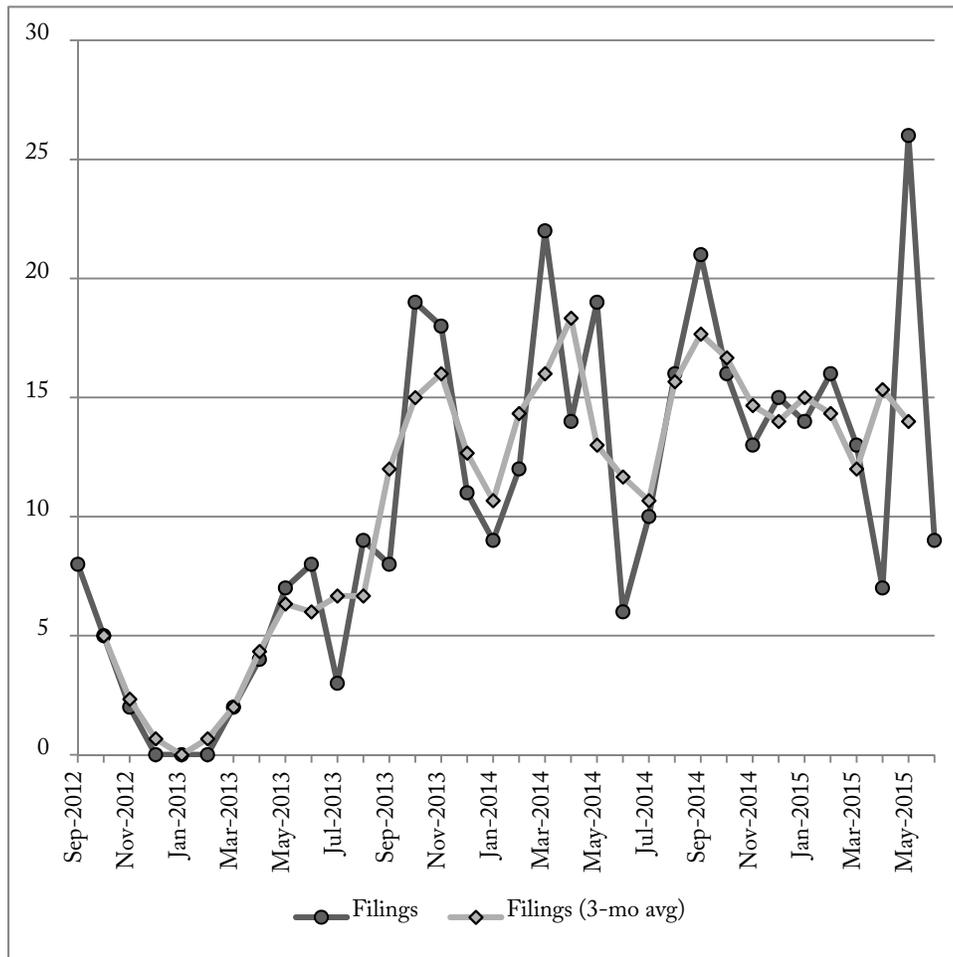


Figure 5: CBM Petition Filings by Month

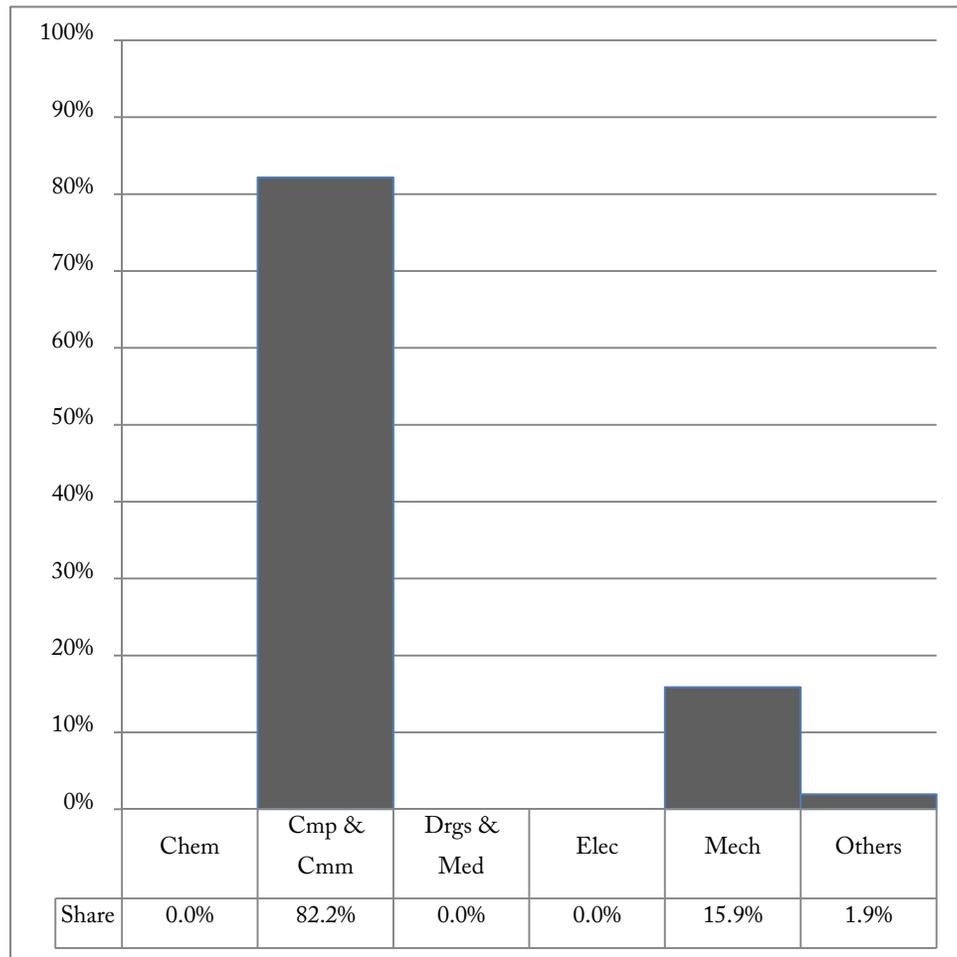


Figure 6: CBM Petition Filings Across Technology

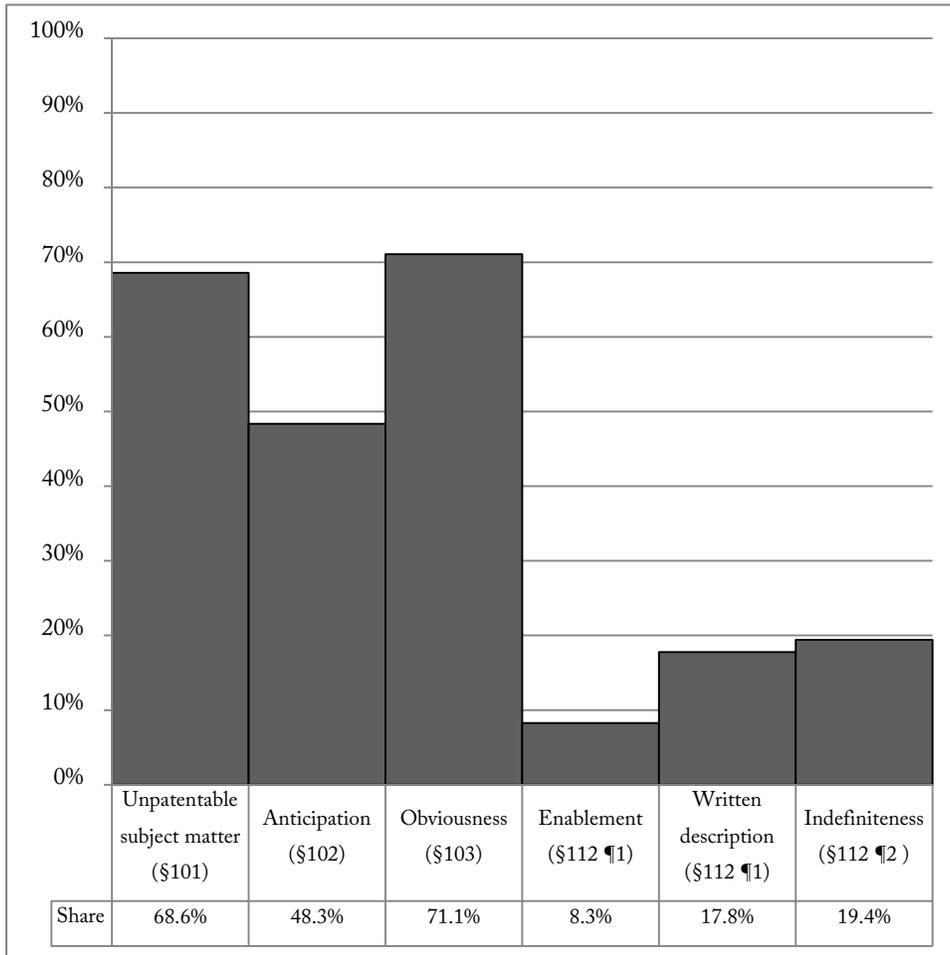


Figure 7: Proportions of CBM Petitions Containing Each Grounds for Challenge

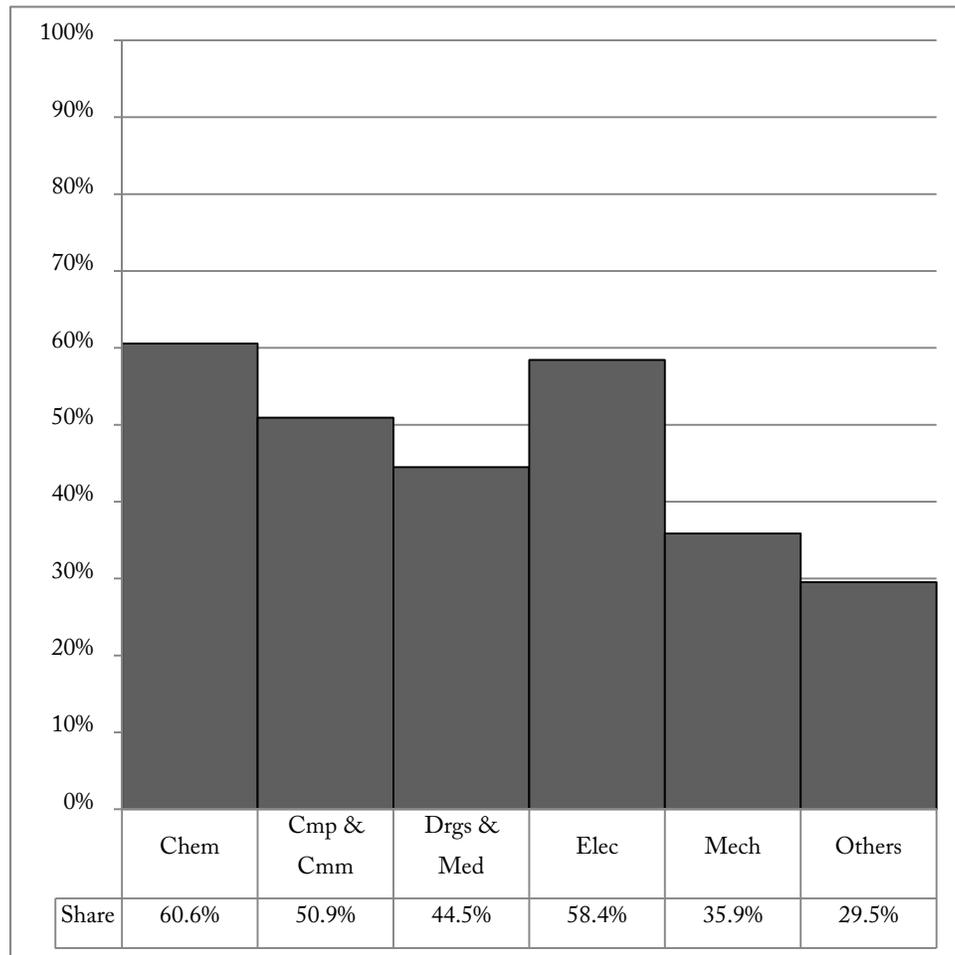


Figure 8: Share of IPR-Challenged Patents in Each Technology Area That Were the Subject of Multiple Petitions

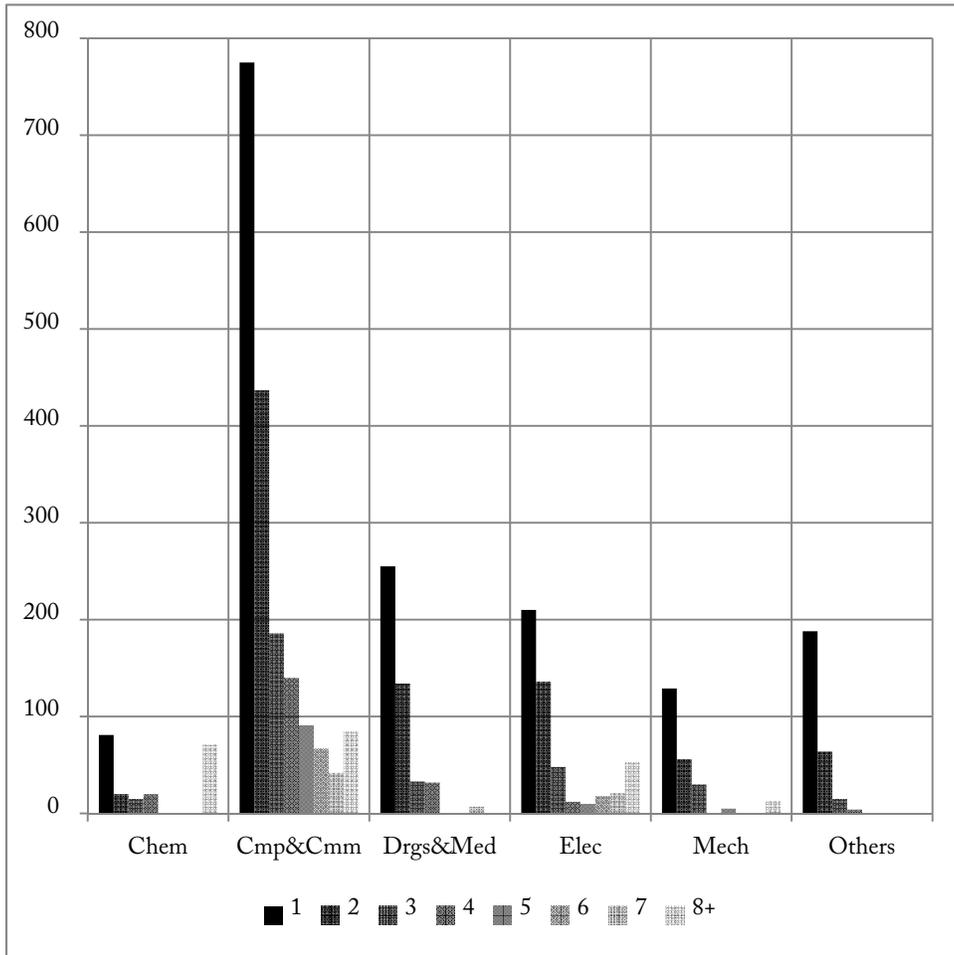


Figure 9: Number of IPR-Challenged Patents Across Technology Area, by Number of IPR Petitions

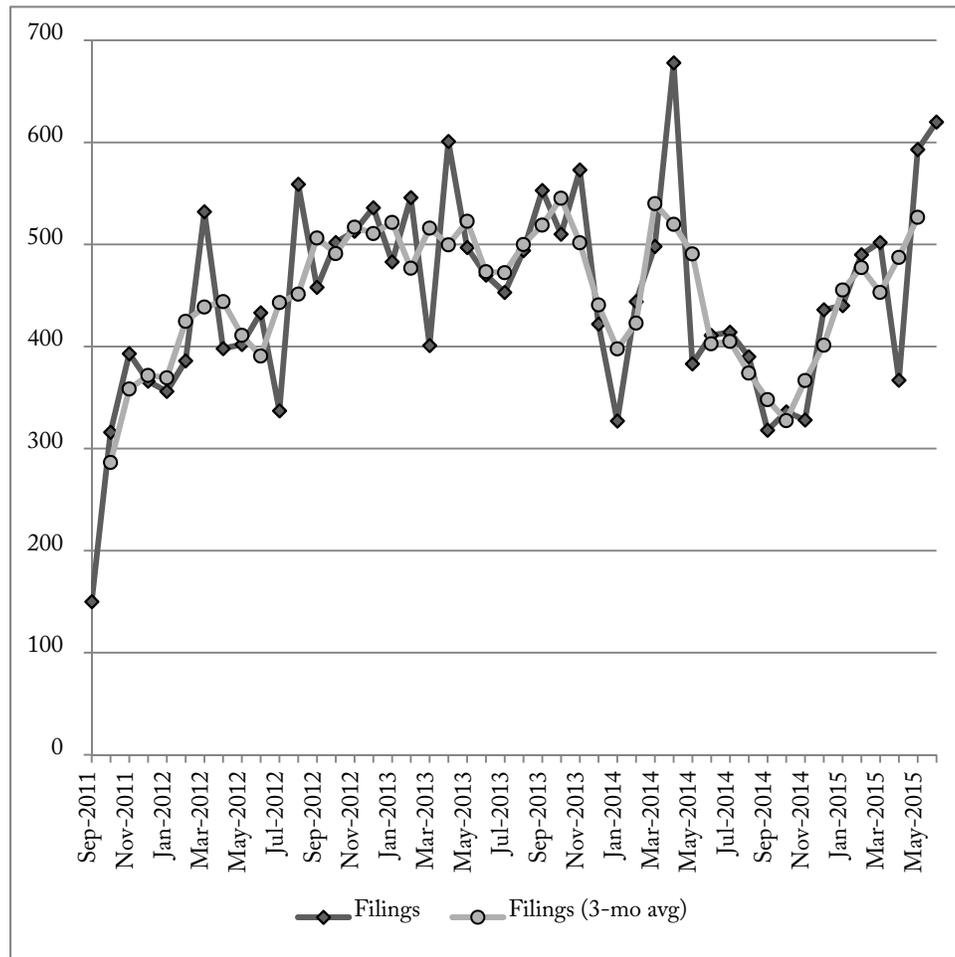


Figure 10: Patent Case Filings by Month

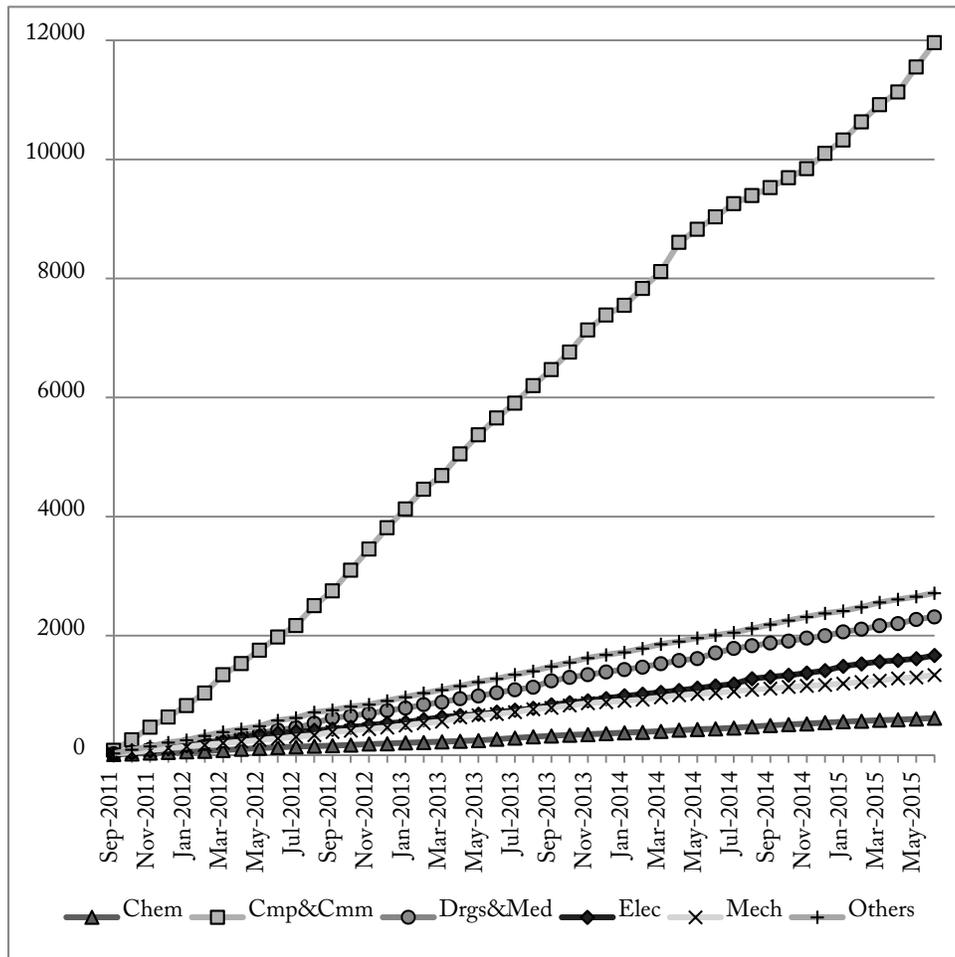


Figure 11: Cumulative Patent Case Filings Across Technology

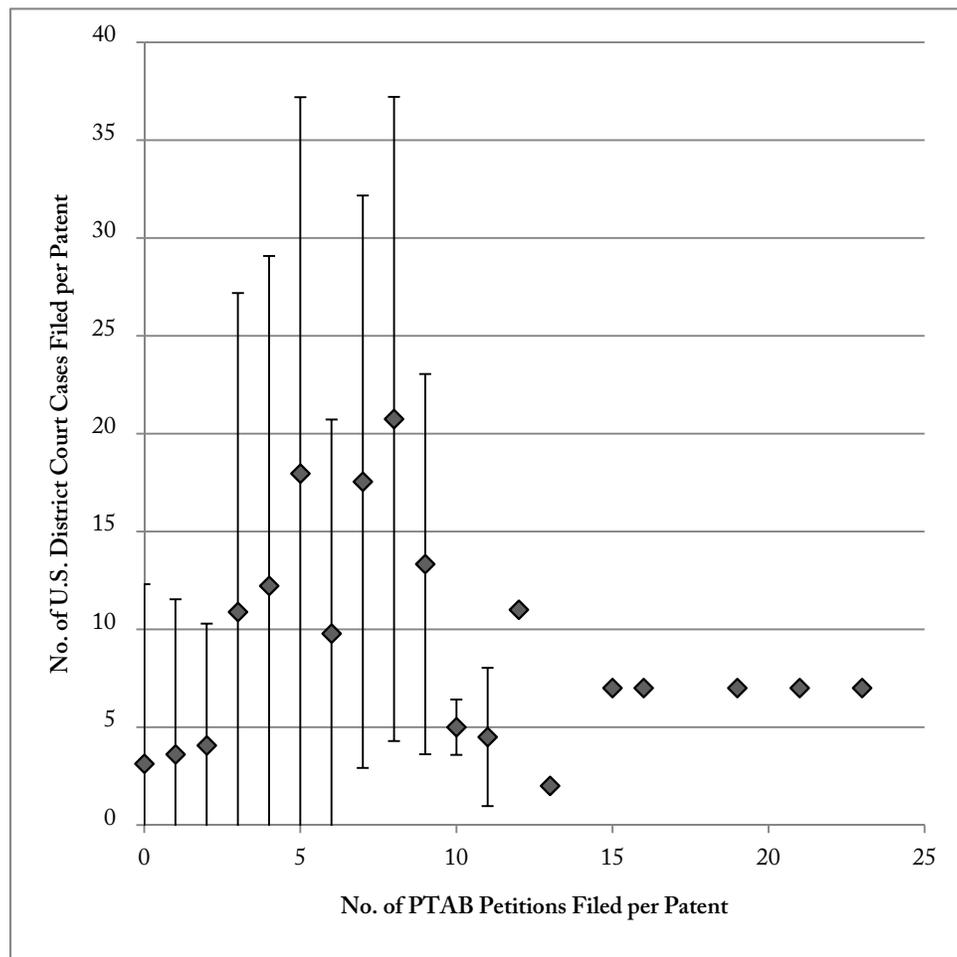


Figure 12: Average Number of Federal-Court Assertion of Patents Challenged in IPR or CBM Petitions

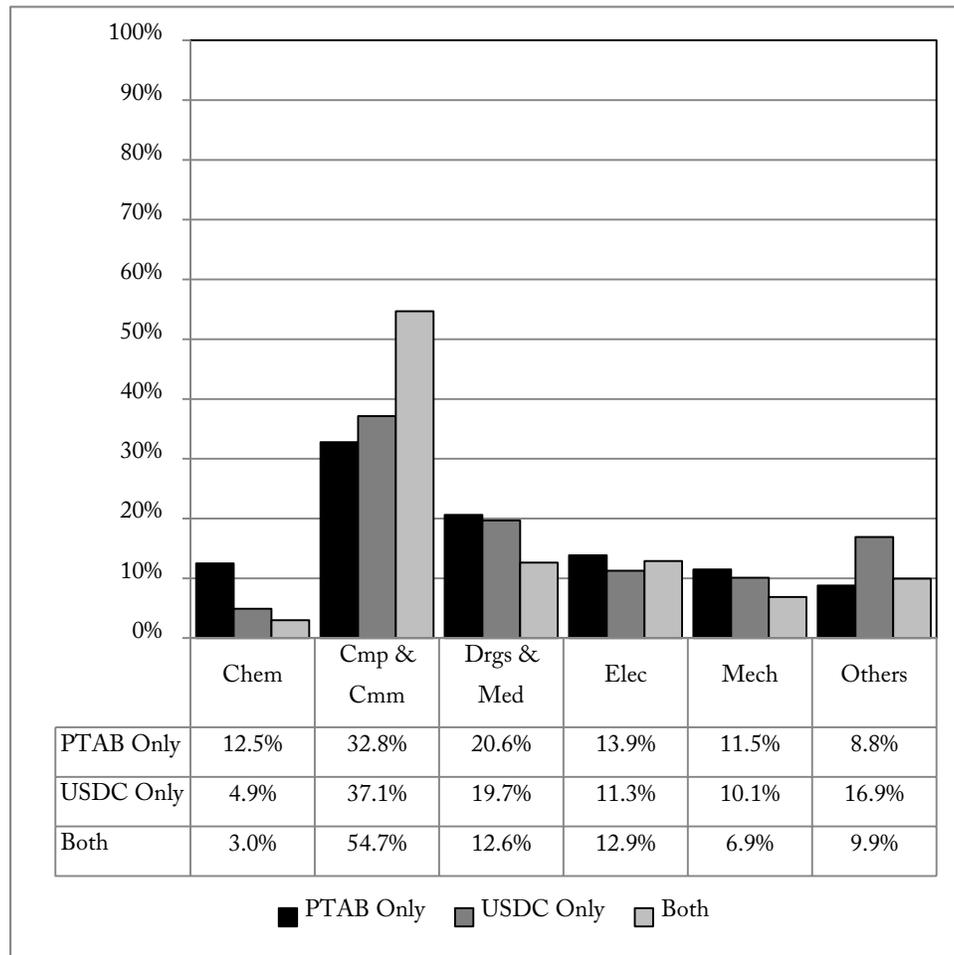


Figure 13: Technology Proportions Among PTAB-Only, Federal Court-Only, and PTAB-and-Federal Court Patents

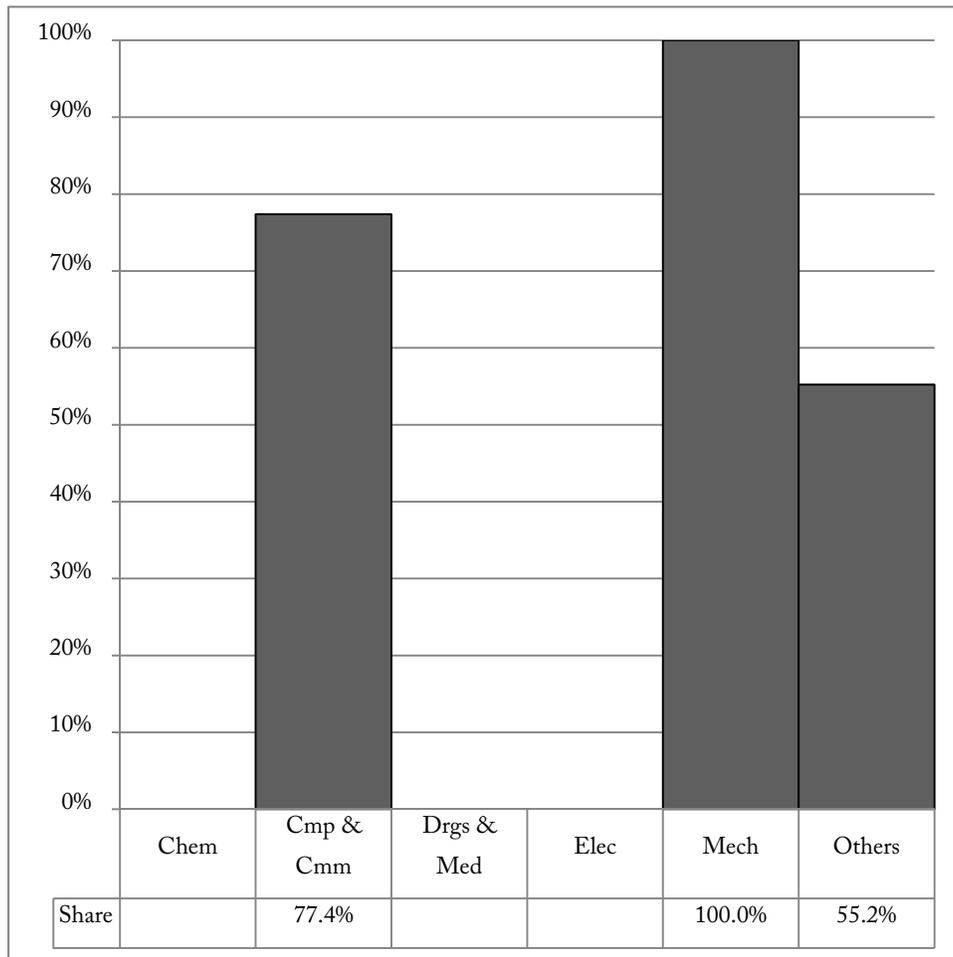


Figure 14a: Share of CBM Petitioners That Were Defendants in a Prior Suit on the Same Patent, by Technology

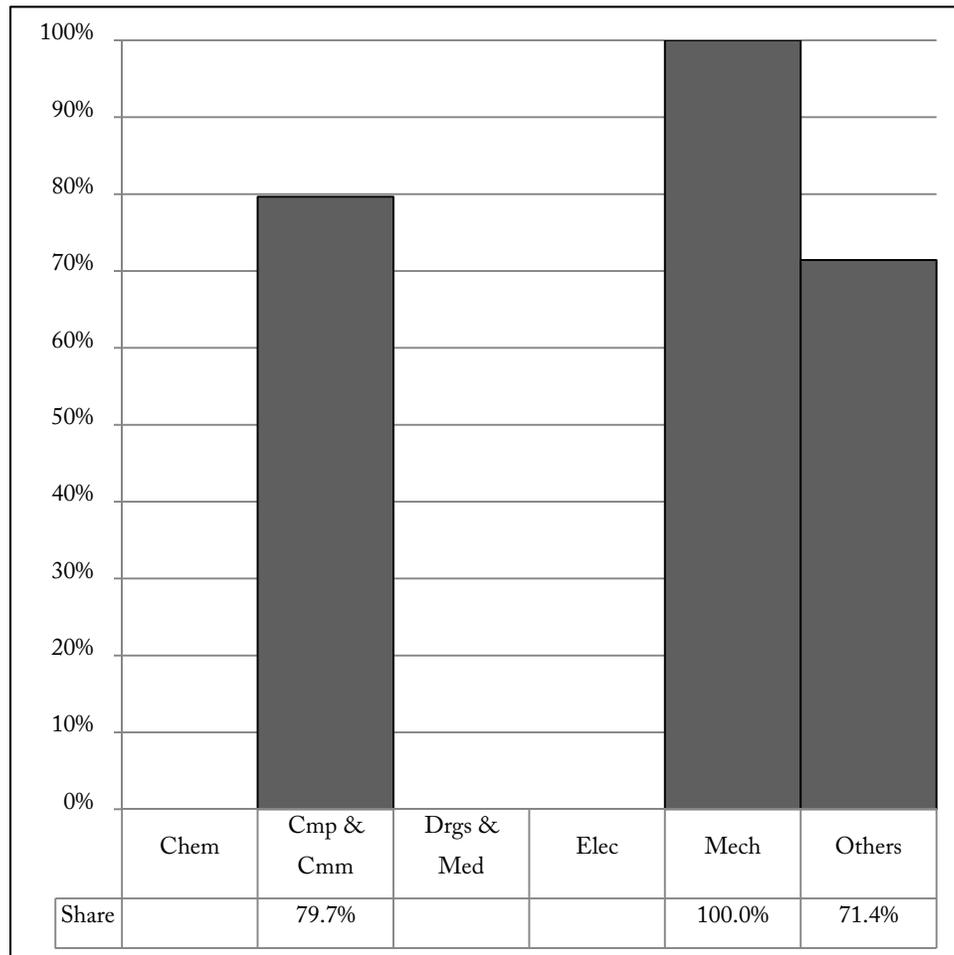


Figure 14b: Share of CBM Petitions in Which At Least One Petitioner Was a Defendant in a Prior Suit on the Same Patent, by Technology

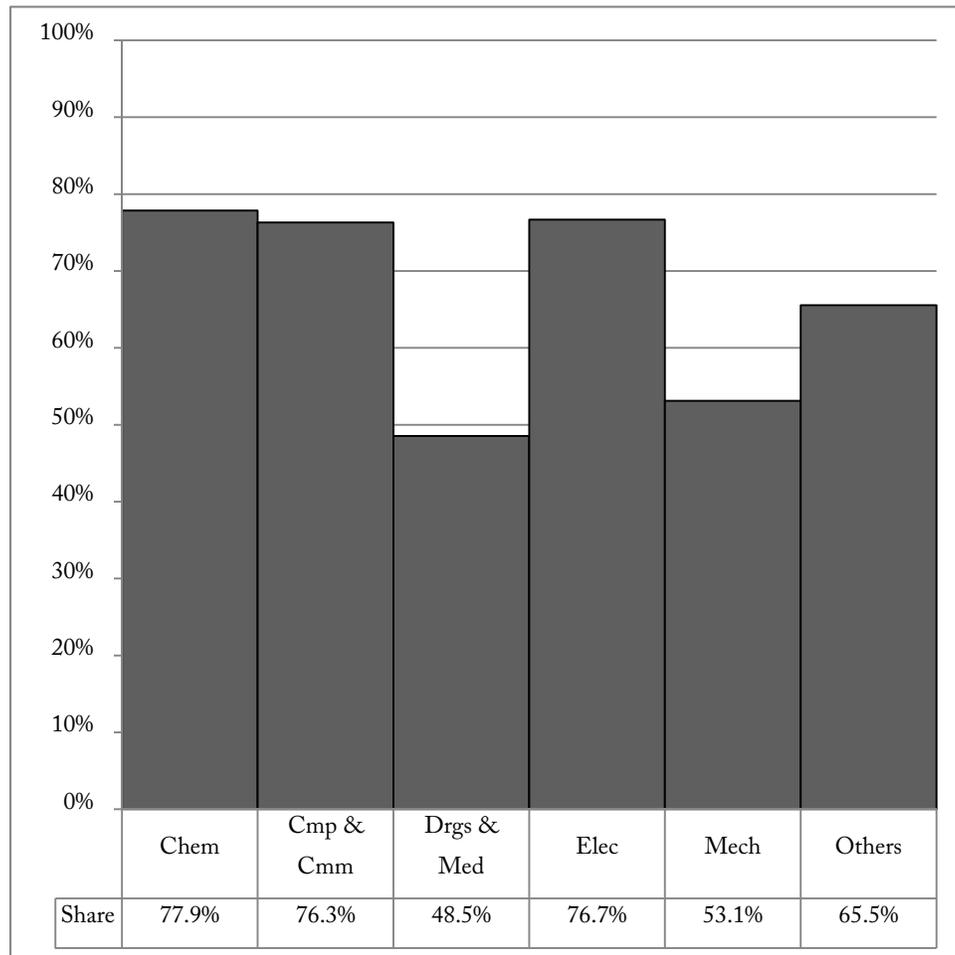


Figure 15a: Share of IPR Petitioners That Were Defendants in a Prior Suit on the Same Patent, by Technology

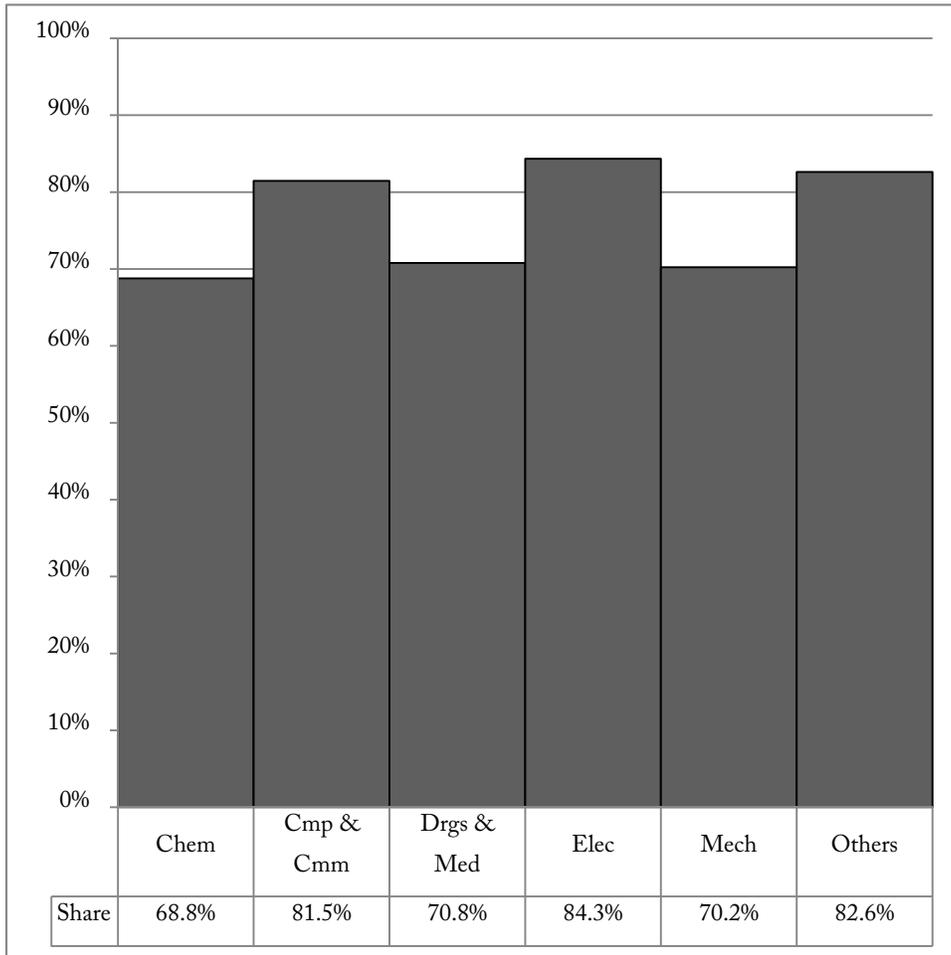


Figure 15b: Share of IPR Petitions in Which At Least One Petitioner Was a Defendant in a Prior Suit on the Same Patent, by Technology

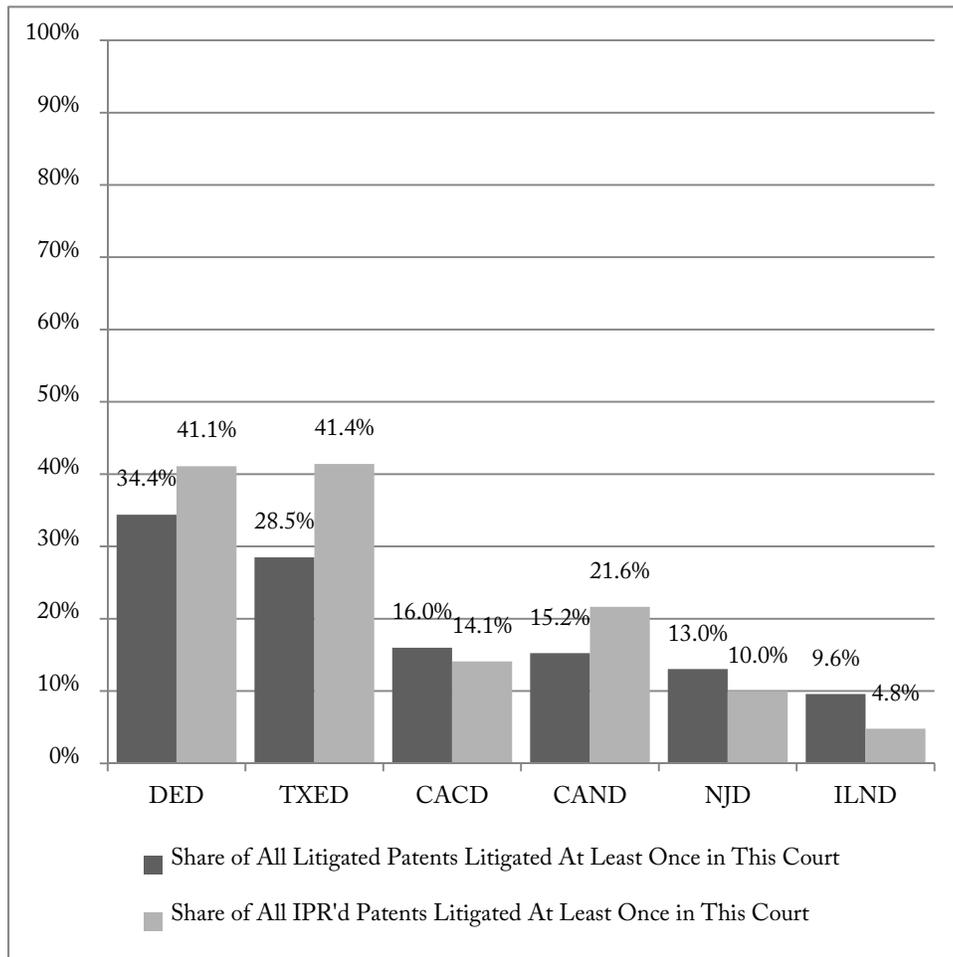


Figure 16: District-Specific Effects

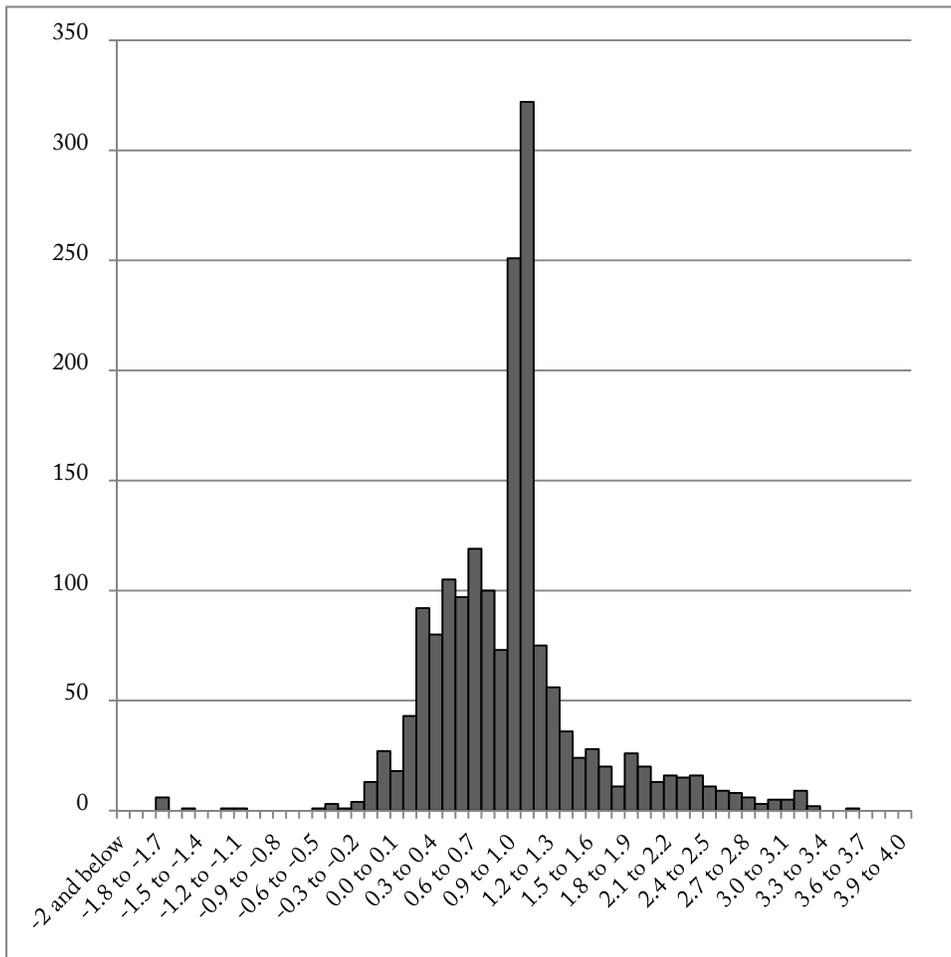


Figure 17: Distribution of Lag Between First IPR Petition and First Federal Court Litigation, in Years

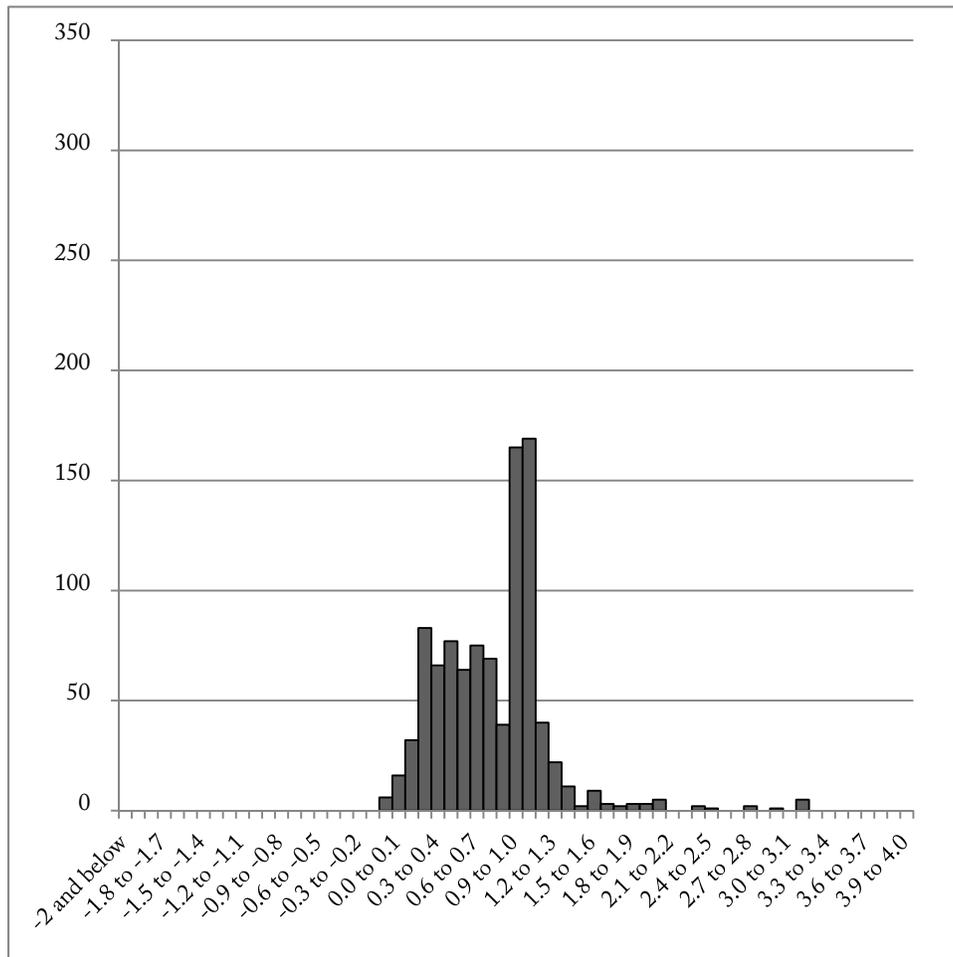


Figure 18: Distribution of Lag Between First IPR Petition and Last Federal Court Filing Prior to Petition, in Years

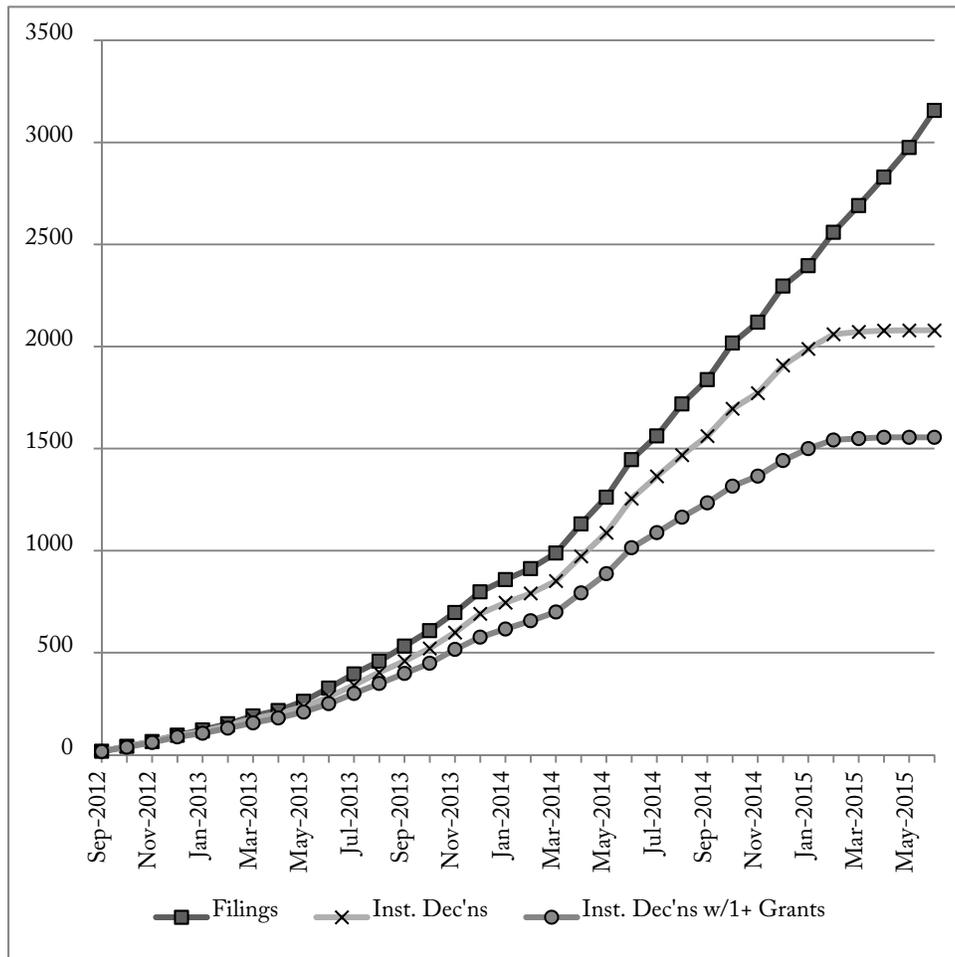


Figure 19: IPR Filings, Institution Decisions, and Institution Decisions Granting At Least One Challenged Claim

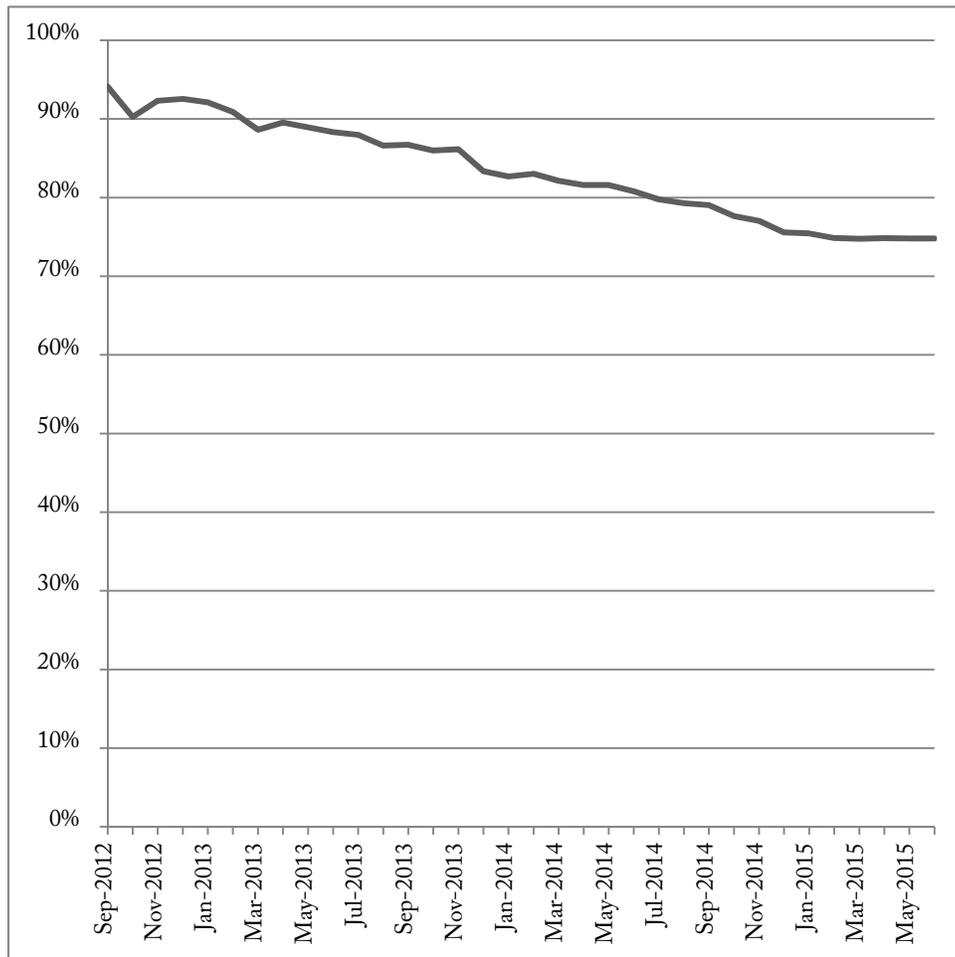


Figure 20: Rate of IPR Institution over Time, by Month

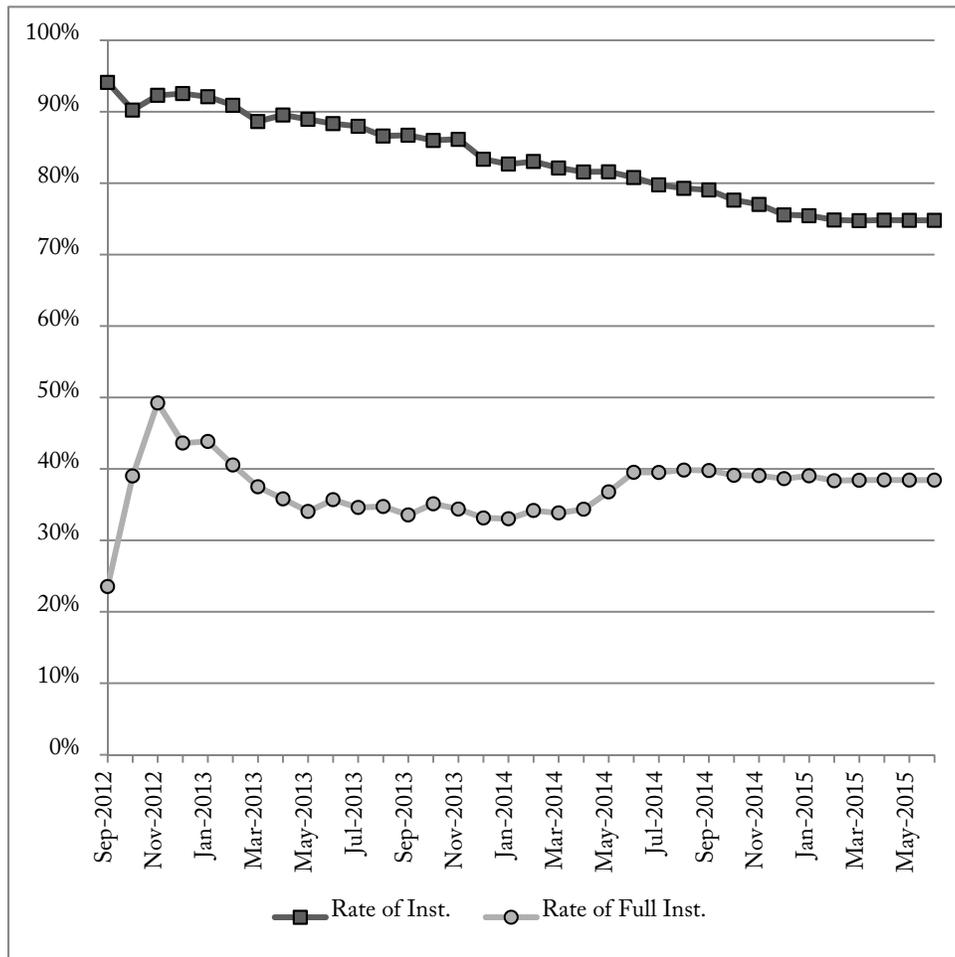


Figure 21: Rate of At-Least-Partial Institution and Full Institution of IPR Petitions over Time, By Month

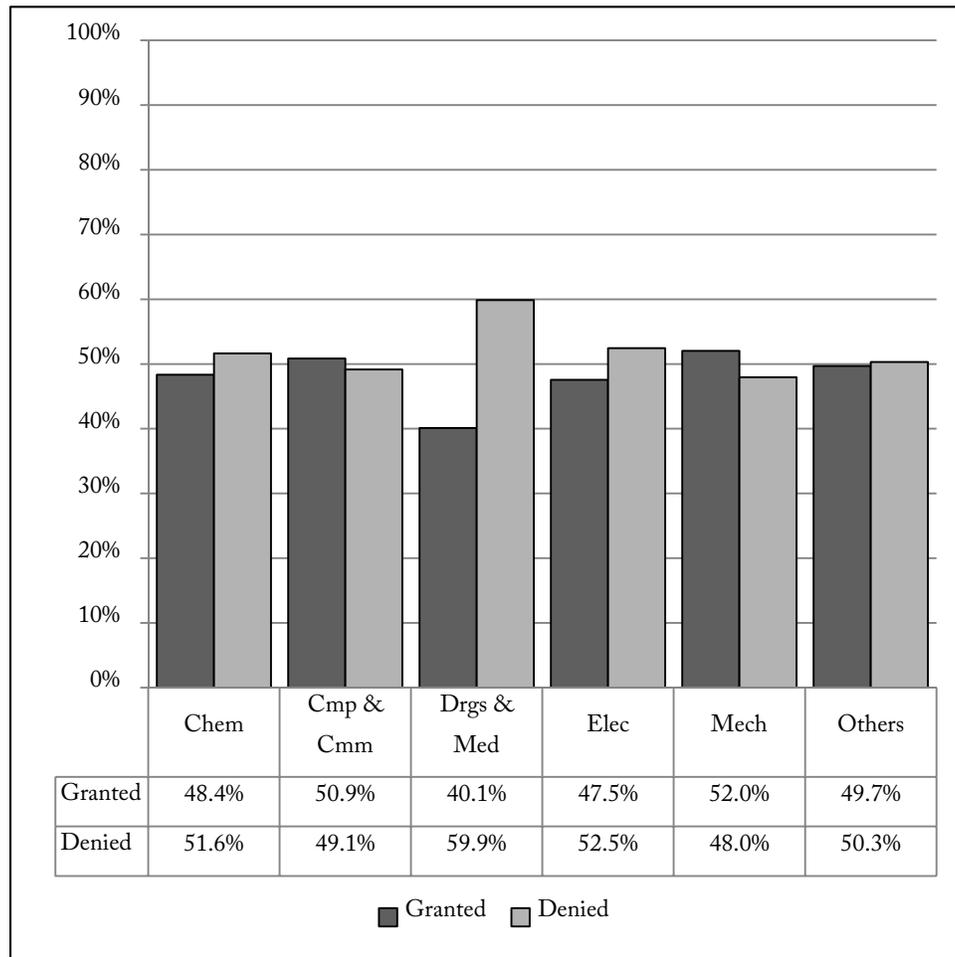


Figure 22: Institutions of IPR Petitions Based on Novelty, by Technology

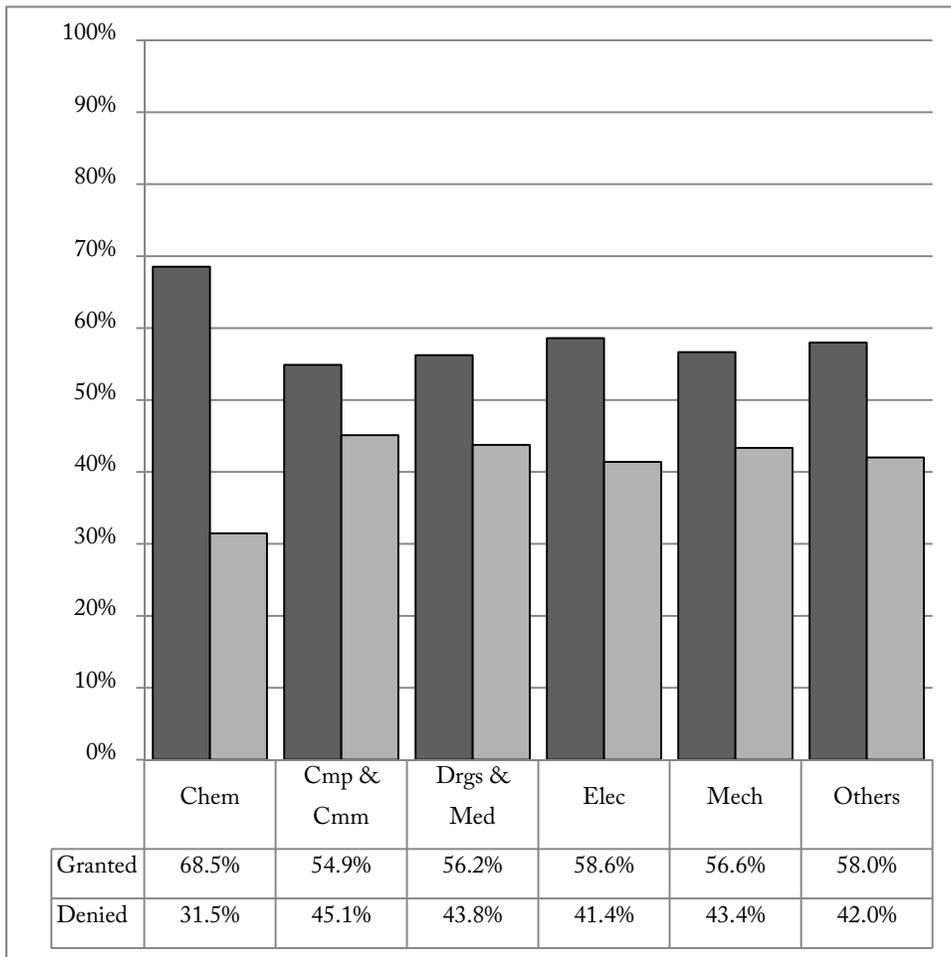


Figure 23: IPR Institutions of Petitions Based on Nonobviousness, by Technology

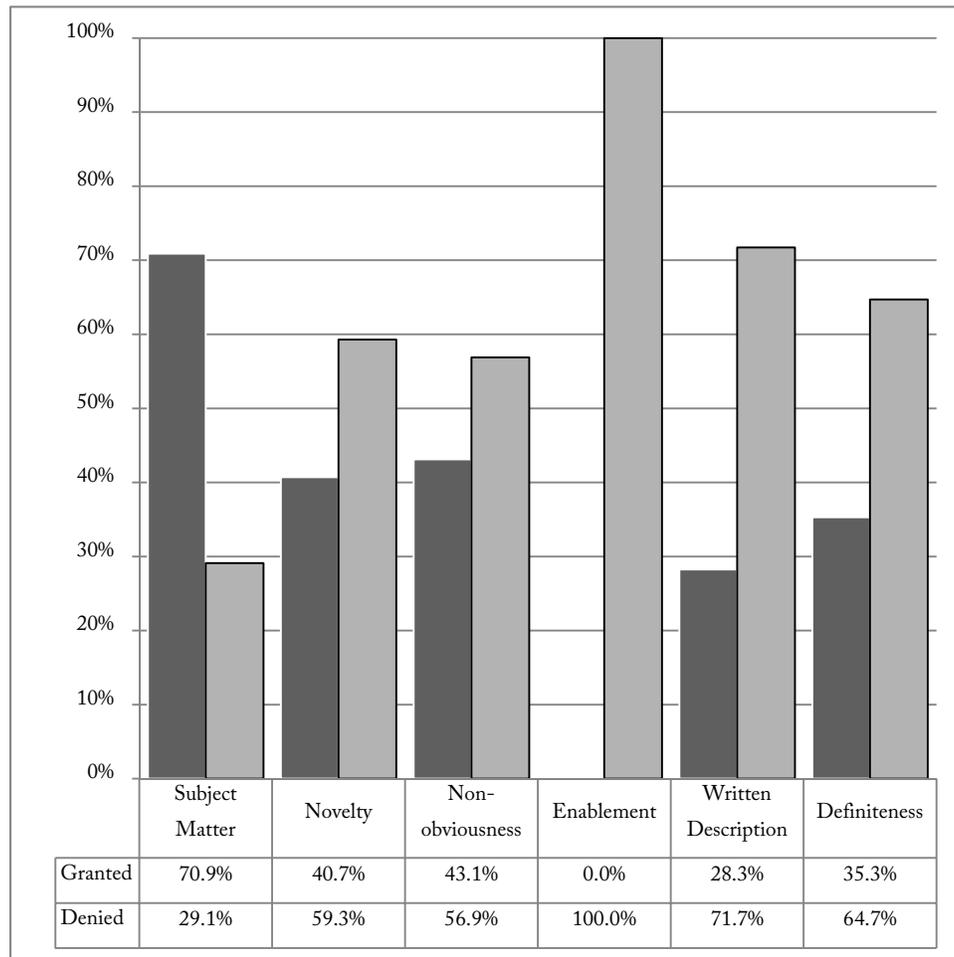


Figure 24: CBM Institutions of Petitions, by Grounds

APPENDIX C: TABLES

Table 1: *Inter Partes* petitions filed by Intel in the *Zond* cases

Petition For	IPR Number	Filing Date
Intel Corporation	IPR2014-00443	Feb. 20, 2014
Intel Corporation	IPR2014-00444	Feb. 20, 2014
Intel Corporation	IPR2014-00445	Feb. 20, 2014
Intel Corporation	IPR2014-00446	Feb. 20, 2014
Intel Corporation	IPR2014-00447	Feb. 20, 2014
Intel Corporation	IPR2014-00455	Feb. 27, 2014
Intel Corporation	IPR2014-00456	Feb. 27, 2014
Intel Corporation	IPR2014-00468	Feb. 28, 2014
Intel Corporation	IPR2014-00470	Mar. 7, 2014
Intel Corporation	IPR2014-00473	Mar. 7, 2014
Intel Corporation	IPR2014-00494	Mar. 13, 2014
Intel Corporation	IPR2014-00495	Mar. 13, 2014
Intel Corporation	IPR2014-00496	Mar. 13, 2014
Intel Corporation	IPR2014-00497	Mar. 13, 2014
Intel Corporation	IPR2014-00498	Mar. 13, 2014
Intel Corporation	IPR2014-00520	Mar. 27, 2014
Intel Corporation	IPR2014-00521	Mar. 27, 2014
Intel Corporation	IPR2014-00522	Mar. 27, 2014
Intel Corporation	IPR2014-00523	Mar. 27, 2014
Intel Corporation	IPR2014-00598	Apr. 9, 2014
Intel Corporation	IPR2014-00686	Apr. 24, 2014
Intel Corporation	IPR2014-00765	May 16, 2014
Intel Corporation	IPR2014-00820	May 27, 2014
Intel Corporation	IPR2014-00843	May 29, 2014
Intel Corporation	IPR2014-00913	Jun 6, 2014
Intel Corporation	IPR2014-00923	Jun 10, 2014
Intel Corporation	IPR2014-00945	Jun 12, 2014

Table 2: *Inter Partes* Review Petitions filed by Defendants in *Zond* cases

Petition For	IPR Number	Filing Date
Advanced Micro Devices, Inc.	IPR2014-01037	June 30, 2014
Advanced Micro Devices, Inc.	IPR2014-01075	June 30, 2014
Advanced Micro Devices, Inc.	IPR2014-01071	June 30, 2014
Advanced Micro Devices, Inc.	IPR2014-01069	June 30, 2014
Fujitsu Semiconductor Limited	IPR2014-00848	May 29, 2014
Fujitsu Semiconductor Ltd.	IPR2014-00850	May 29, 2014
Fujitsu Semiconductor Ltd.	IPR2014-00844	May 29, 2014
Fujitsu Semiconductor Ltd.	IPR2014-00846	May 29, 2014
Fujitsu Semiconductor Ltd.	IPR2014-00845	May 29, 2014
Fujitsu Semiconductor Ltd.	IPR2014-00849	May 29, 2014
Fujitsu Semiconductor Ltd.	IPR2014-00855	May 30, 2014
Fujitsu Semiconductor Ltd.	IPR2014-00866	May 30, 2014
Fujitsu Semiconductor Ltd.	IPR2014-00851	May 30, 2014
Fujitsu Semiconductor Ltd.	IPR2014-00865	May 30, 2014
Fujitsu Semiconductor Ltd.	IPR2014-00856	May 30, 2014
Fujitsu Semiconductor Ltd.	IPR2014-00859	May 30, 2014
Fujitsu Semiconductor Ltd.	IPR2014-00858	May 30, 2014
Fujitsu Semiconductor Ltd.	IPR2014-00863	May 30, 2014
Fujitsu Semiconductor Ltd.	IPR2014-00864	May 30, 2014
Fujitsu Semiconductor Ltd.	IPR2014-00867	May 30, 2014
Fujitsu Semiconductor Ltd.	IPR2014-00918	June 09, 2014
GlobalFoundries U.S., Inc.	IPR2014-01042	June 27, 2014
GlobalFoundries U.S., Inc.	IPR2014-01059	June 27, 2014
GlobalFoundries U.S., Inc.	IPR2014-01047	June 27, 2014
GlobalFoundries U.S., Inc.	IPR2014-01083	June 30, 2014
GlobalFoundries U.S., Inc.	IPR2014-01086	June 30, 2014
GlobalFoundries U.S., Inc.	IPR2014-01076	June 30, 2014
GlobalFoundries U.S., Inc.	IPR2014-01061	June 30, 2014
GlobalFoundries U.S., Inc.	IPR2014-01087	June 30, 2014
GlobalFoundries U.S., Inc.	IPR2014-01073	June 30, 2014
GlobalFoundries U.S., Inc.	IPR2014-01088	July 01, 2014
GlobalFoundries U.S., Inc.	IPR2014-01098	July 01, 2014
GlobalFoundries U.S., Inc.	IPR2014-01099	July 01, 2014
GlobalFoundries U.S., Inc.	IPR2014-01089	July 01, 2014
GlobalFoundries U.S., Inc.	IPR2014-01100	July 01, 2014

Petition For	IPR Number	Filing Date
Renesas Electronics Corp.	IPR2014-01057	June 27, 2014
Renesas Electronics Corp.	IPR2014-01046	June 27, 2014
Renesas Electronics Corp.	IPR2014-01066	June 30, 2014
Renesas Electronics Corp.	IPR2014-01063	June 30, 2014
Taiwan Semiconductor Mfg. Co.	IPR2014-00805	May 23, 2014
Taiwan Semiconductor Mfg. Co.	IPR2014-00782	May 19, 2014
Taiwan Semiconductor Mfg. Co.	IPR2014-00781	May 19, 2014
Taiwan Semiconductor Mfg. Co.	IPR2014-00800	May 22, 2014
Taiwan Semiconductor Mfg. Co.	IPR2014-00799	May 22, 2014
Taiwan Semiconductor Mfg. Co.	IPR2014-00803	May 22, 2014
Taiwan Semiconductor Mfg. Co.	IPR2014-00802	May 22, 2014
Taiwan Semiconductor Mfg. Co.	IPR2014-00807	May 23, 2014
Taiwan Semiconductor Mfg. Co.	IPR2014-00808	May 23, 2014
Taiwan Semiconductor Mfg. Co.	IPR2014-00819	May 27, 2014
Taiwan Semiconductor Mfg. Co.	IPR2014-00821	May 27, 2014
Taiwan Semiconductor Mfg. Co.	IPR2014-00818	May 27, 2014
Taiwan Semiconductor Mfg. Co.	IPR2014-00828	May 28, 2014
Taiwan Semiconductor Mfg. Co.	IPR2014-00829	May 28, 2014
Taiwan Semiconductor Mfg. Co.	IPR2014-00827	May 28, 2014
Taiwan Semiconductor Mfg. Co.	IPR2014-00861	May 30, 2014
Taiwan Semiconductor Mfg. Co.	IPR2014-00917	June 09, 2014
The Gillette Co.	IPR2014-00479	Mar. 4, 2014
The Gillette Co.	IPR2014-00477	Mar. 4, 2014
The Gillette Co.	IPR2014-00974	June 18, 2014
The Gillette Co.	IPR2014-00975	June 18, 2014
The Gillette Co.	IPR2014-00972	June 18, 2014
The Gillette Co.	IPR2014-00973	June 18, 2014
The Gillette Co.	IPR2014-00992	June 19, 2014
The Gillette Co.	IPR2014-00986	June 19, 2014
The Gillette Co.	IPR2014-00981	June 19, 2014
The Gillette Co.	IPR2014-00991	June 19, 2014
The Gillette Co.	IPR2014-00984	June 19, 2014
The Gillette Co.	IPR2014-00990	June 19, 2014
The Gillette Co.	IPR2014-00988	June 19, 2014
The Gillette Co.	IPR2014-00985	June 19, 2014
The Gillette Co.	IPR2014-01003	June 20, 2014

Petition For	IPR Number	Filing Date
The Gillette Co.	IPR2014-00996	June 20, 2014
The Gillette Co.	IPR2014-01000	June 20, 2014
The Gillette Co.	IPR2014-00995	June 20, 2014
The Gillette Co.	IPR2014-01004	June 20, 2014
The Gillette Co.	IPR2014-01012	June 23, 2014
The Gillette Co.	IPR2014-01017	June 23, 2014
The Gillette Co.	IPR2014-01016	June 23, 2014
The Gillette Co.	IPR2014-01015	June 23, 2014
The Gillette Co.	IPR2014-01019	June 23, 2014
The Gillette Co.	IPR2014-01014	June 23, 2014
The Gillette Co.	IPR2014-01013	June 23, 2014
The Gillette Co.	IPR2014-01020	June 23, 2014
The Gillette Co.	IPR2014-01022	June 23, 2014
The Gillette Co.	IPR2014-01025	June 23, 2014
Toshiba Corp.	IPR2014-01072	June 30, 2014
Toshiba Corp.	IPR2014-01070	June 23, 2014
Toshiba Corp.	IPR2014-01067	June 23, 2014
Toshiba Corp.	IPR2014-01074	June 23, 2014
Toshiba Corp.	IPR2014-01065	June 23, 2014

Table 3: *Inter Partes* Review Petitions filed by Defendants in *E-Watch v. LG Electronics*

Petition For	IPR Number	Filing Date
HTC Corp.	IPR2014-00987	June 19, 2014
Sony Mobile Comm. (USA) Inc.	IPR2015-00402	Dec. 10, 2014
LG Electronics, Inc.	IPR2015-00404	Dec. 10, 2014
Kyocera Communications, Inc.	IPR2015-00406	Dec. 10, 2014
Apple Inc.	IPR2015-00411	Dec. 11, 2014
Apple Inc.	IPR2015-00412	Dec. 11, 2014
Apple Inc.	IPR2015-00413	Dec. 11, 2014
Samsung Electronics Co.	IPR2015-00541	Jan. 7, 2015
Samsung Electronics Co.	IPR2015-00610	Jan. 23, 2015
Samsung Electronics Co.	IPR2015-00612	Jan. 23, 2015
ZTE (USA) Inc.	IPR2015-01366	June 09, 2015

CHILLING EFFECTS: ONLINE SURVEILLANCE AND WIKIPEDIA USE

Jonathon W. Penney[†]

ABSTRACT

This Article discusses the results of the first empirical study providing evidence of regulatory “chilling effects” of Wikipedia users associated with online government surveillance. The study explores how traffic to Wikipedia articles on topics that raise privacy concerns for Wikipedia users decreased after the widespread publicity about NSA/PRISM surveillance revelations in June 2013. Using an interdisciplinary research design, the study tests the hypothesis, based on chilling effects theory, that traffic to privacy-sensitive Wikipedia articles reduced after the mass surveillance revelations. The Article finds not only a statistically significant immediate decline in traffic for these Wikipedia articles after June 2013, but also a change in the overall secular trend in the view count traffic, suggesting not only immediate but also long-term chilling effects resulting from the NSA/PRISM online surveillance revelations. These, and other results from the case study, not only offer evidence for chilling effects associated with online surveillance, but also offer important insights about how we should understand such chilling effects and their scope, including how they interact with other dramatic or significant events (like war and conflict) and their broader implications for privacy, U.S. constitutional litigation, and the health of democratic society. This study is among the first to evidence—using either Wikipedia data or web traffic data more generally—how government surveillance and similar actions may impact online activities, including access to information and knowledge online.

DOI: <http://dx.doi.org/10.15779/Z38SS13>

© 2016 Jonathon W. Penney.

† The author would like to thank Victoria Nash, Urs Gasser, Joss Wright, Ron Deibert, J. Nate Mathias, Paul Kelly, Ryan Budish, Jamie Baxter, Kendra Albert, Andy Sellars, Molly Sauter, Masashi Crete-Nishihata, Yana Welinder, Michelle Paulson, Geoff Brigham, David Abrams, Bernie Black, Mordechai Kremnizer, Adam Chilton, Adi Leibovitch, Kate Litvak, James W. Coleman, Aurelie Ouss, Ben Perryman, Dahlia Othman, Elissa Berwick, Ari Waldman, Meredith Whittaker, Vern Paxon, Daniel Villar-Onrubia, Wenming Xu, Yehonatan Givati, Jeff Knockel, Jed Crandall, Phillipa Gill, and Derek Bambauer for comments, advice, and/or suggestions. The author would also like to thank participants for his talks at the 2015 Internet Law Work in Progress Conference, Santa Clara Law School, Santa Clara University, March 2015; the Berkman Center for Internet & Society, Harvard University, May 2015; the 2015 Free and Open Communications on the Internet (FOCI) workshop, USENIX Security Symposium, Washington, D.C., August 2015; and the Inaugural Society for Empirical-Legal Studies (SELS) Global Workshop for Junior Empirical Legal Scholars, Hebrew University, Jerusalem, Israel, December 2015. The author, in particular, would like to thank Josh Furman and the BTLJ editorial team for their invaluable assistance and suggestions.

TABLE OF CONTENTS

I.	INTRODUCTION	119
II.	CHILLING EFFECTS THEORY AND RESEARCH AFTER SNOWDEN	125
A.	CHILLING EFFECTS THEORY	125
B.	POST-SNOWDEN: NEW URGENCY, NEW OPPORTUNITIES FOR RESEARCH	129
C.	PRISM/NSA REVELATIONS: A REFERENCE POINT FOR STUDY	130
D.	HYPOTHESIS.....	134
III.	METHOD AND DESIGN	134
A.	WHY WIKIPEDIA TRAFFIC?	134
B.	RESEARCH DESIGN AND DATA SELECTION	137
C.	METHOD OF ANALYSIS	143
IV.	RESULTS.....	145
A.	NON-MODEL EMPIRICAL FINDINGS	145
B.	MODEL-BASED EMPIRICAL FINDINGS.....	146
1.	<i>First Set of Results</i>	147
2.	<i>Outliers: The “Exogenous Shock” of War</i>	149
3.	<i>Second Set of Results—A Lasting Chilling Effect?</i>	151
4.	<i>Final Results with Comparator Groups</i>	153
V.	IMPLICATIONS.....	161
A.	EMPIRICAL EVIDENCE FOR REGULATORY CHILLING EFFECTS.....	161
B.	THE IMPACT OF WAR AND OTHER EXOGENOUS EVENTS	164
C.	CONSTITUTIONAL LITIGATION	165
D.	SURVEILLANCE, WIKIPEDIA, AND DEMOCRATIC SOCIETY.....	169
VI.	LIMITATIONS.....	171
VII.	CONCLUSION AND FUTURE DIRECTIONS	172
	APPENDIX.....	174

I. INTRODUCTION

On March 10, 2015, the American Civil Liberties Union, on behalf of the Wikimedia Foundation and eight other organizations, filed a lawsuit against the United States Department of Justice and the National Security Agency (NSA) challenging the constitutionality of NSA online surveillance:

This lawsuit challenges the suspicionless seizure and searching of internet traffic by the National Security Agency (“NSA”) on U.S. soil. The NSA conducts this surveillance, called “Upstream” surveillance, by tapping directly into the internet backbone inside the United States—the network of high-capacity cables, switches, and routers that today carry vast numbers of Americans’ communications with each other and with the rest of the world. In the course of this surveillance, the NSA is seizing Americans’ communications en masse while they are in transit, and it is searching the contents of substantially all international text-based communications—and many domestic communications as well—for tens of thousands of search terms. The surveillance exceeds the scope of the authority that Congress provided in the FISA Amendments Act of 2008 (“FAA”) and violates the First and Fourth Amendments.¹

An Op-Ed published the same day in *The New York Times*, co-authored by Wikipedia Founder Jimmy Wales and Wikimedia Foundation’s Executive Director Lila Tretikov, explained the lawsuit was necessary because “pervasive surveillance” caused “a chilling effect” that stifled the “freedom of expression” and “free exchange” of ideas on Wikipedia,² the collaborative online encyclopedia that is global in both content and scope—it contains over 30 million articles available in over 200 languages and is among the ten most visited websites globally.³

1. Complaint for Declaratory and Injunctive Relief at 1, Wikimedia Found. v. NSA, No. 1:15-cv-00662-RDB, 2015 WL 1033734, (D. Md. Mar. 10, 2015), https://www.aclu.org/files/assets/wikimedia_v2c_nsa_-_complaint.pdf [<https://perma.cc/3YNL-3BQR>] [hereinafter Complaint]. For a definition of mass surveillance, see Ben Beaumont, *Easy Guide to Mass Surveillance*, AMNESTY INTERNATIONAL (Mar. 18, 2015), <https://www.amnesty.org/en/latest/campaigns/2015/03/easy-guide-to-mass-surveillance> [<https://perma.cc/6YC2-23CN>].

2. Jimmy Wales & Lila Tretikov, Opinion, *Stop Spying on Wikipedia Users*, N.Y. TIMES (Mar. 10, 2015), <http://www.nytimes.com/2015/03/10/opinion/stop-spying-on-wikipedia-users.html> [<https://perma.cc/CV36-P4XG>].

3. See Judit Bar-Ilan & Noa Aharoni, *Twelve Years of Wikipedia Research*, PROCEEDINGS OF THE 5TH ACM CONF. ON WEB SCI. 243, 243 (2014); Stefanie Hilles, *To Use or Not to Use? The Credibility of Wikipedia*, 10:3 PUB. SERV. Q. 245 (2014); David J. McIver & John S. Brownstein, *Wikipedia Usage Estimates Prevalence of Influenza-Like Illness in the United States in Near Real-Time*, 10:4 PLOS

However, like previous constitutional challenges to NSA surveillance,⁴ the lawsuit was not heard on the merits but dismissed in October for lack of standing.⁵ Wikimedia Foundation, the lead complainant, intends to appeal.⁶

The idea that government surveillance is harmful to free expression and association is not new, nor is skepticism about its empirical and legal basis. In the 1972 Supreme Court case *Laird v. Tatum*, for example, the complainants argued that broad government surveillance and data gathering unconstitutionally chilled their rights.⁷ The Court rejected the

COMPUTATIONAL BIOLOGY 1, 1 (2014). Wikipedia provides a wealth of information about its number of articles, editors, page views, etc. See *Wikimedia Statistics*, WIKIMEDIA, <http://stats.wikimedia.org/#fragment-14> [<https://perma.cc/8DYN-CRP8>].

4. See, e.g., *Clapper v. Amnesty Int'l*, 133 S. Ct. 1138, 1152 (2013) (dismissing a constitutional challenge to NSA surveillance practices for lack of standing).

5. *Wikimedia Foundation v. NSA*, Case 1:15-cv-00662-RDB, 2015 WL 6460364 (D. Md. Oct. 23, 2015).

6. Michelle Paulson & Geoff Brigham, *District Court Grants Government's Motion to Dismiss Wikimedia v. N.S.A., Appeal Expected*, WIKIMEDIA BLOG (Oct. 23, 2015), <http://blog.wikimedia.org/2015/10/23/wikimedia-v-nsa-lawsuit-dismissal> [<https://perma.cc/88XF-3V9T>], stating:

Unfortunately, the court did not actually rule on whether the NSA's upstream surveillance is legal or illegal. Judge T.S. Ellis III, the presiding judge, dismissed the case on standing grounds. We respectfully disagree with the Court's decision to dismiss. There is no question that Upstream surveillance captures the communications of both the user community and the Wikimedia Foundation itself. We believe that our claims have merit. In consultation with our lawyers at the ACLU, we will review the decision and expect to appeal to the Fourth Circuit Court of Appeals.

7. *Laird v. Tatum*, 408 U.S. 1, 13–14 (1972). *Laird* stated:

[The respondents'] claim, simply stated, is that they disagree with the judgments made by the Executive Branch with respect to the type and amount of information the Army needs, and that the very existence of the Army's data-gathering system produces a constitutionally impermissible chilling effect upon the exercise of their First Amendment rights. That alleged 'chilling' effect may perhaps be seen as arising from respondents' very perception of the system as inappropriate to the Army's role under our form of government, or as arising from respondents' beliefs that it is inherently dangerous for the military to be concerned with activities in the civilian sector, or as arising from respondents' less generalized yet speculative apprehensiveness that the Army may at some future date misuse the information in some way that would cause direct harm to respondents. Allegations of a subjective 'chill' are not an adequate substitute for a claim of specific present objective harm or a threat of specific future harm.

Chilling effects theory is discussed in more detail in Section II.A of this Article.

claim due to lack of standing, finding that the surveillance did not constitute an “objective harm or a threat of specific future harm.”⁸ The decision reflected a deep skepticism about both the potential chilling effects and attendant harms of surveillance.⁹ Such “judicial skepticism” has persisted over the decades.¹⁰ In a 2013 case, *Clapper v. Amnesty International*, the Court cited to *Laird* to dismiss a challenge to the legality of NSA surveillance authorized by the Foreign Intelligence Surveillance Act (FISA), and noted that chilling effects fears were “too speculative.”¹¹

Skepticism about chilling effects is not confined to courts. Legal commentators have long questioned the existence or scope of surveillance related chilling effects, and they have also expressed skepticism as to whether the premises of chilling effects theory can be empirically substantiated. Even Frederick Schauer, who offered an early classic statement of chilling effects theory and doctrine, admitted in 1978 that its empirical assumptions about human behavior were “most likely unprovable.”¹² Nearly a decade after Schauer, Vincent Blasi observed that the notion of “chilling effects” on supposed “fearful and overly risk-averse” speakers was “oft-criticized” and based on “crude behavioral speculation.”¹³ More recently, Leslie Kendrick, after surveying both literature and case law, emphasized the theory’s “weak” and “flimsy” empirical basis and concluded additional research was required for the “unsubstantiated empirical judgments” of chilling effects claims.¹⁴ Also recently, Margot Kaminski and Shane Witnov have acknowledged certain social science studies that corroborate forms of chilling effects, but nevertheless call for more empirical work on surveillance and its impact in a “number of critical areas,” including the existence, magnitude, and persistence of surveillance related chilling effects.¹⁵

8. *Id.* at 15.

9. See Margot E. Kaminski & Shane Witnov, *The Conforming Effect: First Amendment Implications of Surveillance, Beyond Chilling Speech*, 49 U. RICH. L. REV. 465, 480 (2015).

10. *Id.* at 482.

11. *Clapper v. Amnesty Int’l*, 133 S. Ct. 1138, 1152 (2013).

12. Frederick Schauer, *Fear, Risk, and the First Amendment: Unraveling the “Chilling Effect,”* 58 B.U. L. REV. 685, 730 (1978).

13. Vincent Blasi, *The Pathological Perspective and the First Amendment*, 85 COLUM. L. REV. 449, 482 (1985).

14. Leslie Kendrick, *Speech, Intent, and the Chilling Effect*, 54 WM. MARY L. REV. 1633, 1657 (2013).

15. Kaminski & Witnov, *supra* note 9, at 517 (calling for further research on the “types of surveillance and surveillance cues that cause chilling effects,” as well as the strength and persistence of such chilling effects).

Privacy theorists, security researchers, and social scientists have also expressed skepticism about the possibility of large scale chilling effects caused by online surveillance.¹⁶ One reason for such skepticism is increasing public acceptance of, or desensitization to, privacy and surveillance concerns, particularly in new technological contexts.¹⁷ Indeed, some research in the field suggests that any chilling effects would, at the very most, be temporary or ephemeral, as online users have changed their behavior in response to shifting norms.¹⁸

16. See, e.g., Daniel Solove, *The First Amendment as Criminal Procedure*, 82 N.Y.U. L. REV. 112, 155 (2007) (“Determining the existence of a chilling effect is complicated by the difficulty of defining and identifying deterrence. It is hard to measure the deterrence caused by a chilling effect because it is impossible to determine with certainty what people would have said or done in the absence of the government activity. Often, the primary evidence will be a person’s own assertions that she was chilled, but merely accepting such assertions at face value would allow anyone claiming a chilling effect to establish one. At the same time, demanding empirical evidence of deterrence is impractical because it will often be impossible to produce.”).

17. See Sandro Nickel, *The Double-Edged Effects of Social Media Terror Communication: Interconnection and Independence vs. Surveillance and Human Rights Calamities*, in NEW OPPORTUNITIES AND IMPASSES: THEORIZING AND EXPERIENCING POLITICS 255, 263 (Zeynep Guler ed., 2014) (“The majority of the population will most probably not fall into self-censoring behavior, a reason for this possibly being the desensitization concerning privacy in general, at least co-constituted by the very digital experiences of the past decade(s).”); BRUCE SCHNEIER, *DATA AND GOLIATH: THE HIDDEN BATTLES TO CAPTURE YOUR DATA AND CONTROL YOUR WORLD* 95–99 (2015) (Schneier, a leading information security expert, speaks of how surveillance leads to “conformity” and, in Chapter 6, generally discusses the need to change lax and accepting public attitudes about increasing surveillance and its harms); David Lyon, *Surveillance, Snowden, and Big Data: Capacities, Consequences, Critique*, 1 BIG DATA & SOC’Y 1, 51 (2014) (noting that the constant “ratcheting up” of government surveillance in recent times is not just a product of the growth of new technologies, but also broader cultural trends accommodating increasing amounts of societal surveillance).

18. Bernhard Debatin & Jennette P. Lovejoy, *Facebook and Online Privacy: Attitudes, Behaviors, and Unintended Consequences*, 15 J. COMPUTER-MEDIATED COMM. 83 (2009), <http://dx.doi.org/10.1111/j.1083-6101.2009.01494.x> [<https://perma.cc/DS47-9ABX>] (documenting Facebook users’ “lax” attitudes about privacy concerns). For research or works suggesting online chilling effects would be temporary or ephemeral, see Laura Bernescu, *When is a Hack not a Hack: Addressing the CFAA’s Applicability to the Internet Service Context*, U. CHI. LEGAL F. 633 (2013) (arguing that users will quickly adopt to changes in the regulatory environment in relation to the Computer Fraud and Abuse Act, rendering any “chilling effects” temporary); Chris Rose, *The Security Implications of Ubiquitous Social Media*, 15 INT’L J. MGMT. & INFO. SYS. 35, 37 (2011) (noting that increased comfort with using the Internet has led many consumers to conform to new norms, particularly on privacy); see also Alessandro Acquisti, Leslie K. John & George Loewenstein, *What is Privacy Worth?*, 42 J. LEGAL STUD. 249, 267–70 (2013), <https://www.cmu.edu/dietrich/sds/docs/loewenstein/WhatPrivacyWorth.pdf> [<https://perma.cc/LF5V-QRSJ>] (advancing explanations for the disconnect between privacy attitudes and the lax or loose approach to privacy in practice).

So, the empirical basis for chilling effects theory, and its different dimensions, remain controversial. Part of the challenge, as privacy scholars like Leslie Kendrick,¹⁹ Daniel Solove,²⁰ and Neil Richards²¹ have noted, is the often “intractable empirical difficulties” in designing research to demonstrate or measure chilling effects.²² Showing the impact and harms of surveillance involves dealing with counterfactuals or proving a negative—self-censorship. As such, it is “difficult to establish either the presence or the absence of a chilling effect, let alone to measure the extent of such an effect.”²³ With the absence of empirical research to substantiate chilling effects, compounded by the methodological challenges for designing and carrying out such research, it is unsurprising skepticism about the theory persists.

Furthermore, with the revelation of widespread Internet surveillance by the United States and other Western governments (thanks to the leaks and disclosures of Edward Snowden), the need for empirical and theoretical study has taken on even greater urgency. This is particularly true because of the range of lawsuits filed by companies, citizen groups, and organizations to challenge government surveillance and related laws,²⁴ but beyond the legal arena, it is crucial for understanding the potential harms of such surveillance to activities online.

The empirical case study discussed in this Article attempts to help address this research void. Building on a recent study of Google search

19. Kendrick, *supra* note 14, at 1638 (“But there are reasons to doubt the chilling effect account. A claim of a chilling effect necessarily rests upon suppositions about the deterrent effects of law. These suppositions rest in turn upon predictions about the behavior of speakers under counterfactual conditions. Meanwhile, the selection of a remedy for chilling—such as an intent requirement—rests on similar predictions about the remedy’s speech-protective effects. In short, both the detection of a problem and the imposition of a remedy involve intractable empirical difficulties.”).

20. Solove, *The First Amendment as Criminal Procedure*, *supra* note 16, at 155 (“Often, the primary evidence will be a person’s own assertions that she was chilled, but merely accepting such assertions at face value would allow anyone claiming a chilling effect to establish one. At the same time, demanding empirical evidence of deterrence is impractical because it will often be impossible to produce.”).

21. Neil Richards, *The Dangers of Surveillance*, 126 HARV. L. REV. 1934, 1964 (2013) (“This is not to say that individual determinations of the chilling of intellectual activities will always be easy. Determining whether a chill to intellectual privacy is substantial would certainly present difficult cases at the margins.”).

22. Kendrick, *supra* note 14, at 1675.

23. *Id.* at 1638.

24. For a review of the broad range of constitutional litigation arising since the Snowden leaks, see generally Edward C. Liu, Andrew Nolan & Richard M. Thompson, CONG. RESEARCH SERV., R43459, OVERVIEW OF CONSTITUTIONAL CHALLENGES TO NSA COLLECTION ACTIVITIES AND RECENT DEVELOPMENTS 7-5700, 12-18 (2014).

traffic and Internet surveillance,²⁵ this Article discusses the first original empirical study of the impact such surveillance has had on Wikipedia use. Consistent with the recent growth of empirical approaches in legal research,²⁶ the study uses an interrupted time series (ITS) design²⁷ to determine whether traffic for articles that may raise privacy concerns for Wikipedia users decreased after the widespread publicity about NSA online surveillance activities. In short, this case study asks: Did Wikipedia traffic for articles on privacy-sensitive topics decrease after the “exogenous shock” of widespread publicity surrounding the surveillance programs in June 2013? A hypothesis based on chilling effects theory would hold that Internet users will be less likely to view or access such privacy-sensitive Wikipedia articles after the revelations. Ultimately, this case study provides results consistent with surveillance related chilling effects, among other findings. The context of the study is also important. Wikipedia was chosen as the focus of this case study for a number reasons, but most importantly because any chilling effect on Wikipedia users has far-

25. Alex Marthews & Catherine Tucker, *Government Surveillance and Internet Search Behavior* (MIT Sloane Working Paper No. 14380, 2015).

26. For discussion of the empirical and experimental turn in legal research, see Daniel E. Ho & Larry Kramer, *Introduction: The Empirical Revolution in Law*, 65 STAN. L. REV. 1195 (2013); see also Adam Chilton & Dustin Tingley, *Why the Study of International Law Needs Experiments*, 52 COLUM. J. TRANS. L. 173, 187–90 (2013) (discussing the “growth” of experimental and quasi-experimental methods in legal research). See generally Gregory Shaffer & Tom Ginsburg, *The Empirical Turn in International Legal Scholarship*, 106 AM. J. INT’L L. 1 (2012) (also discussing the empirical “turn” in international law research).

27. For discussion of interrupted time series research design, see DONALD T. CAMPBELL, JULIAN C. STANLEY & NATHANIEL L. GAGE, EXPERIMENTAL AND QUASI-EXPERIMENTAL DESIGNS FOR RESEARCH 37–43 (1966) (discussing the components of time series designs and their methodological advantages and limitations); Melvin M. Mark, Charles S. Reichardt & Lawrence J. Sanna, *Time-Series Designs and Analyses*, in HANDBOOK OF APPLIED MULTIVARIATE STATISTICS AND MATHEMATICAL MODELING 354–55 (2000) (discussing the use of time series designs to assess the impact of interventions); see also Carlotta Ching Ting Fok, David Henry & James Allen, *Research Designs for Intervention Research with Small Samples II: Stepped Wedge and Interrupted Time-Series Designs*, PREVENTATIVE SCI. 1, 4 (2015) <http://link.springer.com/article/10.1007/s11121-015-0569-4> [<https://perma.cc/FFX5-6CPB>] (offering some suggestions to strengthen the methodological dimensions of ITS designs to study the impact of health interventions); Robert B. Penfold & Fang Zhang, *Use of Interrupted Time Series Analysis in Evaluating Health Care Quality Improvements*, 13:6 ACAD. PEDIATRICS S38 (2013) (discussing the advantages and limitations of employing time series analysis to understand and explore the impact of health policy changes); A.K. Wagner et al., *Segmented Regression Analysis of an Interrupted Time Series in Medication Use Research*, 27 J. CLINICAL PHARMACY & THERAPEUTICS 299 (2002) (discussing advantages of using of segmented regression analysis, along with ITS design, in the context of health research).

reaching implications. The site, which is growing both in popularity and scope, serves as an essential source of information and knowledge online, and functions as an important public tool to complement the democratic process in promoting collective understanding, decision-making, and deliberation.²⁸

Part II of the Article provides additional context related to chilling effects research and the impact of the Edward Snowden disclosures. Part III sets out and justifies the case study's methodology and research design, including its focus on Wikipedia. Part IV discusses the results of the study: consistent with chilling effects theory, (1) Wikipedia traffic to privacy-sensitive articles showed a statistically significant reduction after June 2013, and (2) there was a long lasting change in the overall secular trend in traffic to such articles. The study's implications and limitations are discussed in Part V and Part VI, respectively. Part VII concludes and considers possible directions for future research.

II. CHILLING EFFECTS THEORY AND RESEARCH AFTER SNOWDEN

This Part begins with a more in-depth discussion of chilling effects theory, including leading accounts of its dimensions and assumptions, along with an overview of related studies. From there, the Snowden NSA/PRISM revelations widely covered in June 2013 are discussed and re-framed as presenting a research opportunity to study chilling effects theory. Finally, the latter sections develop the research question and hypothesis—centered on the Snowden leaks—that form the basis of the study discussed in this Article.

A. CHILLING EFFECTS THEORY

The idea that government laws or actions might chill people's free activities gained its most prominent early expression in the United States during the Cold War. The "chilling effects doctrine," a legal doctrine in First Amendment jurisprudence, took shape in a series of cases decided in the 1950s and 60s that dealt with anti-communist state measures. Essentially, the doctrine encouraged courts to treat rules or government actions that "might deter" the free exercise of First Amendment rights "with suspicion."²⁹

28. See *infra* Section III.A.

29. Richards, *supra* note 21, at 1949–50. For early cases recognizing the chilling effects doctrine, see, for example, *Dombrowski v. Pfister*, 380 U.S. 479 (1965); *Wieman v. Updegraff*, 344 U.S. 183 (1952).

But underlying this legal doctrine was a deeper theory with empirical assumptions about behavior in relation to government acts—that certain state acts may chill or deter people from exercising their freedoms or engaging in legal activities. This theory of “chilling effects” received its first comprehensive exploration in Schauer’s *Fear, Risk, and the First Amendment: Unraveling the “Chilling Effects Doctrine*,³⁰ described as the “definitive treatment” of the theory.³¹ Schauer conceived of chilling effects as primarily resulting from people’s fear of prosecution or legal sanction and the uncertainties of the legal process.³² Here, government surveillance may chill or deter people from engaging in certain legal (or even desirable) online activities because they fear legal punishment or criminal sanction, and do not trust the legal system to protect their innocence. Daniel Solove’s work has broadened chilling effects theory by theorizing and exploring modern surveillance and data gathering, explaining how such practices can create a kind of regulatory “environmental pollution” that encourages chilling effects and self-censorship.³³ While Solove’s approach does not discount the sorts of chilling effects Schauer targets, he focuses primarily on how government surveillance of online activities creates a broader atmosphere of conformity and self-censorship; he is concerned with the way regulatory actions—particularly information gathering and surveillance—enhance the risk that a person may suffer harms in the future (e.g., gathering information about a person’s activities may increase the risk they are later “victimized” by identity theft or fraud).³⁴ On this account, people are chilled not because they fear actual punishment for engaging in certain online activities (as Schauer theorizes), but to avoid risks of other kinds, such as the stigma of being labeled or tracked by state

30. Schauer, *supra* note 12.

31. Julie Cohen, *A Right to Read Anonymously: A Closer Look at ‘Copyright Management’ in Cyberspace*, 28 CONN. L. REV. 981, 1011 n.117 (1996) (suggesting Schauer’s work was the “definitive treatment”).

32. Schauer, *supra* note 12, at 687–89. For applications of chilling effects theory to online contexts, scholars at the Berkman Center for Internet & Society at Harvard University have been particularly prolific. *See, e.g.*, JONATHAN ZITTRAIN, *THE FUTURE OF THE INTERNET—AND HOW TO STOP IT* 116, 216 (2008) (exploring the potential chilling effects of perfect enforcement of legal norms by technology measures as well as those caused by citizen surveillance due to the proliferation of devices like smartphones); Yochai Benkler, *Through the Looking Glass: Alice and the Constitutional Foundations of the Public Domain*, 66 LAW & CONTEMP. PROBS. 173, 216–18 (2003) (arguing that the NET Act and Digital Millennium Copyright Act expand protections for certain legal rights online in such a way that will chill expression); Wendy Seltzer, *Free Speech Unmoored in Copyright’s Safe Harbor: Chilling Effects of the DMCA on the First Amendment*, 24 HARV. J.L. & TECH. 171 (2010) (analyzing chilling effects and the DMCA).

33. Daniel Solove, *A Taxonomy of Privacy*, 154 U. PENN. L. REV. 477, 488 (2006).

34. *Id.* at 487.

actors as non-conformists, deviants, or criminals, or the broader concern that information gathered about such activities may be leaked or disclosed publicly, leading to embarrassment or used for nefarious purposes by third-parties.³⁵ Such risks and considerations create a societal context that encourages self-censorship.³⁶ Both of these accounts will be useful to understand and theorize any surveillance related chilling effects observed.³⁷

Part of the broader picture is the impact of *covert* surveillance, in which people are either unaware of surveillance or are only aware of the general possibility of it.³⁸ The Snowden leaks and disclosures, which rendered previously covert surveillance public, have placed concerns like those Solove explores in urgent and concrete terms,³⁹ and a range of public opinion polls and survey-based studies have been conducted to study the effects of the disclosures. Studies by PEN America⁴⁰ and Pew Research Center⁴¹ provide some empirical foundation for the claim that surveillance

35. *Id.* at 496 (discussing the example of how information obtained by surveillance was used to discredit and blackmail Martin Luther King, Jr.).

36. *Id.* at 495.

37. The findings in the empirical legal case study discussed in this article may also provide insights, in turn, for these theories as well. For example, Schauer theorizes chilling effects primarily as a product of individual concerns for actual legal punishment or prosecution in an uncertain legal system, while Solove's account captures broader risks and concerns that may also chill—where online users may not actually fear prosecution, but prefer not to have governments looking over their shoulder or tracking and compiling data about their online activities (even if legal). Both such approaches will likely explain or account for instances of surveillance-related chilling effects, but one explanation may prove more common than the other.

38. Solove, *A Taxonomy of Privacy*, *supra* note 33, at 494–96 (relating such surveillance to Jeremy Bentham's idea for 19th century "Panopticon" prison design, also known as the "Panopticon effect").

39. Christopher Slobogin, *Standing and Covert Surveillance*, 41 PEPP. L. REV. 517, 520 (2014) (noting that as a result of Snowden's disclosures, the U.S. federal government has been forced to acknowledge previously covert surveillance practices).

40. FDR GROUP & PEN AMERICAN CENTER, CHILLING EFFECTS: NSA SURVEILLANCE DRIVES U.S. WRITERS TO SELF-CENSOR 3–4 (2013), http://www.pen.org/sites/default/files/Chilling%20Effects_PEN%20American.pdf [https://perma.cc/5TFK-Q8MF] (noting that 28% of the writers surveyed had "curtailed or avoided" certain online activities due to "fear of surveillance"); FDR GROUP & PEN AMERICAN CENTER, GLOBAL CHILLING: THE IMPACT OF MASS SURVEILLANCE ON INTERNATIONAL WRITERS (2015), http://www.pen.org/sites/default/files/globalchilling_2015.pdf [https://perma.cc/GJ88-TMY2] (noting that the international community is similarly engaging in forms of self-censorship).

41. KEITH N. HAMPTON ET AL., PEW RES. CTR., SOCIAL MEDIA AND THE 'SPIRAL OF SILENCE' 4 (2014), http://www.pewinternet.org/files/2014/08/PI_Social-networks-and-debate_082614.pdf [https://perma.cc/QWVP2-5QJS] (finding, for example, 86% of respondents less willing to discuss NSA surveillance revelations online, than off); LEE RAINIE ET AL., PEW RES. INTERNET PROJECT, AMERICANS' PRIVACY STRATEGIES POST-SNOWDEN 4 (Mar. 16, 2015), <http://www.pewinternet.org/files/>

has a chilling effect on people's activities online, but these survey-based studies have important limitations.

Social science research has long illustrated that self-reported or expressed concerns about privacy do not necessarily reflect people's actual behavior online,⁴² a phenomenon sometimes referred to as the "privacy paradox."⁴³ The reasons for this paradox remain contested. Some attribute the disconnect between privacy concerns and actual behavior to uninformed decisions, while others point to faulty research and survey design. But few disagree that compared to how they actually act, people tend to exaggerate privacy concerns, leading to biased or inaccurate results in research that relies primarily on self-reported privacy behaviors.⁴⁴ In

2015/03/PI_AmericansPrivacyStrategies_0316151.pdf [https://perma.cc/2RQR-MKU4] (noting that 25% of those aware of surveillance have "changed the patterns" of their use of "technological platforms"); Lee Rainie et al., *Anonymity, Privacy, and Security Online*, PEW RES. CTR. (2013), <http://www.pewinternet.org/2013/09/05/anonymity-privacy-and-security-online> [https://perma.cc/JA3N-Q22E].

42. See Spyros Kokolakis, *Privacy Attitudes and Privacy Behaviour: A Review of Current Research on the Privacy Paradox Phenomenon*, COMPUTERS & SOC'Y 1 (2015), <http://www.sciencedirect.com/science/article/pii/S0167404815001017#bib0215> [https://perma.cc/BQ6K-HRKS] (providing a comprehensive explanation and review of "information privacy paradox" literature); see also Alessandro Acquisti & Ralph Gross, *Imagined Communities: Awareness, Information Sharing, and Privacy on the Facebook*, PROC. 6TH WORKSHOP ON PRIVACY ENHANCING TECHNOLOGIES (2006) (finding that Facebook user attitudes concerning privacy differed from their actual behavior and privacy practices on the platform); J. Alessandro Acquisti, *Privacy in Electronic Commerce and the Economics of Immediate Gratification*, PROC. 5TH ACM CONF. ON ELECTRONIC COMMUNICATION (2004), <https://www.heinz.cmu.edu/~acquisti/papers/privacy-gratification.pdf> [https://perma.cc/W5SG-AC93]; Acquisti et al., *supra* note 18 (advancing, among other things, explanations for the disconnect between privacy attitudes and the lax or loose approach to privacy in practice); Bettina Berendt, Oliver Günther & Sarah Spiekermann, *Privacy in E-commerce: Stated References vs. Actual Behavior*, 48 COMM. ACM 101, 104 (2005), <http://www.wiwi.hu-berlin.de/professuren/quantitativ/wi/personen/hl/downloads/BGS.pdf> [https://perma.cc/V578-88DX] (finding Web users disclosure practices online were inconsistent with "stated privacy preferences"); Danah Boyd & Nicole Ellison, *Social Network Sites: Definition, History, and Scholarship*, 13:1 J. COMPUTER-MEDIATED COMM. 210, 222 (2007), <http://onlinelibrary.wiley.com/doi/10.1111/j.1083-6101.2007.00393.x/epdf> [https://perma.cc/ER64-G6KV] (noting literature on the privacy paradox); Jim Harper & Solveig Singleton, *With a Grain of Salt: What Consumer Privacy Surveys Don't Tell Us*, COMPETITIVE ENTERPRISE INST. (2001), http://www.cei.org/PDFs/with_a_grain_of_salt.pdf [https://perma.cc/T28Z-2K7E].

43. Susan Barnes, *A Privacy Paradox: Social Networking in the United States*, FIRST MONDAY (Sept. 4, 2006), <http://firstmonday.org/article/view/1394/1312> [https://perma.cc/PQF2-KNM9].

44. See Kokolakis, *supra* note 42 (on "privacy paradox" more generally). Alessandro Acquisti, for example, has argued that the difference can be explained by the fact that people's privacy decisions are irrational and based on flawed or incomplete information. See generally Acquisti et al., *supra* note 18; see also Harper & Singleton, *supra* note 42

short, though these survey-based studies provide some helpful empirical foundation for chilling effects, more work needs to be done to uncover chilling effects in practice.

B. POST-SNOWDEN: NEW URGENCY, NEW OPPORTUNITIES FOR RESEARCH

The Snowden disclosures about NSA surveillance provide new opportunities for chilling effects research. On June 6, 2013, stories in *The Guardian* and *The Washington Post* detailed previously undisclosed information and leaked classified documents about the surveillance practices of the United States and other Western governments.⁴⁵ The leaked documents also suggested a range of major technology companies were involved with the PRISM program.⁴⁶ The revelations about PRISM were followed by stories in June and subsequent months covering a vast array of government surveillance practices and operations, including the

(arguing that survey designs have been flawed, leading to exaggerated self-reported concerns).

45. The June 2013 Snowden leaks centered in large part on “PRISM,” a secret mass electronic surveillance program operated by the NSA, but they also revealed equivalent programs operated by the United Kingdom and other countries. The original June 6, 2013 stories detailed NSA collection of phone records and the PRISM surveillance program. Barton Gellman & Laura Poitras, *U.S., British Intelligence Mining Data from Nine U.S. Internet Companies in Broad Secret Program*, WASH. POST (June 6, 2013), http://www.washingtonpost.com/investigations/us-intelligence-mining-data-from-nine-us-internet-companies-in-broad-secret-program/2013/06/06/3a0c0da8-cebf-11e2-8845-d970ccb04497_story.html [<https://perma.cc/Z5YF-8FB8>]; Glenn Greenwald, *NSA Collecting Phone Records of Millions of Verizon Customers Daily*, GUARDIAN (June 6, 2013), <http://www.theguardian.com/world/2013/jun/06/nsa-phone-records-verizon-court-order> [<https://perma.cc/RC4B-NTFF>]. For discussion and analysis of subsequent news stories and revelations about other surveillance practices by the U.S. and other governments, see David Lyon, *Surveillance, Snowden, and Big Data: Capacities, Consequences, Critique*, 1 BIG DATA & SOC’Y 1, 2 (2014); Marthews & Tucker, *supra* note 25, at 5; see also Amy Wu et al., “Whistleblower or Leaker?” *Examining the Portrayal and Characterization of Edward Snowden in USA, UK, and HK Posts, in NEW MEDIA, KNOWLEDGE PRACTICES & MULTILITERACIES* 53 (Will W.K. Ma et al. eds., 2014); Vian Bakir, *Agenda Building, and Intelligence Agencies: A Systematic Review of the Field from the Discipline of Journalism, Media, and Communications*, 20 INT’L J. PRESS/POL. 131 (2015), <http://hij.sagepub.com/content/20/2/131.abstract> [<https://perma.cc/6EDX-3URG>]; Keir Giles & Kim Hartmann, *Socio-political Effects of Active Cyber Defence Measures*, 6TH INT’L CONF. ON CYBER CONFLICT (CYCON 2014) (2014), http://ieeexplore.ieee.org/xpl/articleDetails.jsp?reload=true&arnumber=6916393&sortType%3Dasc_p_Sequence%26filter%3DAND%28p_IS_Number%3A6916383%29 [<https://perma.cc/4R8M-WTRF>]; Jie Qin, *Hero on Twitter, Traitor on News: How Social Media and Legacy News Frame Snowden*, 20 INT’L J. PRESS/POL. 166 (2015), <http://hij.sagepub.com/content/20/2/166.abstract> [<https://perma.cc/6TGC-KGB7>].

46. Bakir, *supra* note 45, at 132; Lyon, *supra* note 45, at 2–3; Marthews & Tucker, *supra* note 25, at 5–6; Qin, *supra* note 45, at 171.

monitoring of phone records, e-mails, online chats, and browser histories.⁴⁷ The revelations caused a “media and political storm,” receiving widespread coverage both in traditional and new media outlets, and sparking a “heated international debate” in the United States, Europe, Russia, and beyond.⁴⁸

Governments cited the “War on Terror” to defend the surveillance programs, and this justification was reflected in media coverage of the Snowden revelations, particularly by “legacy” news media.⁴⁹ The Snowden leaks and coverage, as media scholar Vian Bakir notes, highlighted the previously limited public awareness about government surveillance activities while also augmenting that awareness.⁵⁰ Indeed, at least in the United States, the widespread media coverage has led to greater awareness and concern among the general public about government surveillance activities and anti-terrorism efforts more generally. A Pew study in 2014 found that 87% of U.S. adults had heard something about “the government collecting information about telephone calls, e-mails, and other online communications” as part of “efforts to monitor terrorist activity” (with 43% hearing “a lot” and 44% hearing “a little”); another 80% agreed or strongly agreed that “Americans should be concerned” about government surveillance.⁵¹ This increased awareness of online government surveillance—and the focal point provided by the June 2013 revelations—presents a unique opportunity for research.

C. PRISM/NSA REVELATIONS: A REFERENCE POINT FOR STUDY

The NSA/PRISM surveillance revelations in June 2013 (“June 2013 revelations”) and widespread surrounding publicity constituted a kind of “exogenous shock”—an intervening “focusing event”—that provides a

47. Lyon, *supra* note 45, at 2–3; Marthews & Tucker, *supra* note 25, at 5–6.

48. Bakir, *supra* note 45, at 132; Giles & Hartmann, *supra* note 45, at 24; Marthews & Tucker, *supra* note 25; Qin, *supra* note 45, at 166.

49. Bakir, *supra* note 45, at 133 (“Governments insist that their methods are legal, if secret, and necessary to fight the War on Terror and organized crime.”); Marthews & Tucker, *supra* note 25, at 2–6; Nickel, *supra* note 17, at 255 (“The mentioned surveillance programs have always—if ‘revealed’ or publicly debated from start with—been justified by their assumed worth in preventing terror attacks, e.g., by Obama after the NSA’s PRISM program was revealed.”); Qin, *supra* note 45, at 178 (finding that a predominant “framing” in traditional news media coverage of the Snowden surveillance disclosures focused on national security terrorism, along with international relations).

50. Bakir, *supra* note 45, at 133.

51. MARY MADDEN, PUBLIC PERCEPTIONS OF PRIVACY AND SECURITY IN THE POST-SNOWDEN ERA, PEW RES. INTERNET PROJECT 2–3 (Nov. 12, 2014), http://www.pewinternet.org/files/2014/11/PI_PublicPerceptionsofPrivacy_111214.pdf [<https://perma.cc/UA7R-QTDZ>]. See also Bakir, *supra* note 45, at 133–34.

helpful reference point for study.⁵² In policy research, the most prominent empirical studies of policy change often focus on the impact of such “triggering” events, typically involving significant “unplanned jolts” or “shocks” like natural disasters or major economic changes.⁵³ But studies have even shown that changes in the tone of media coverage can create exogenous pressures that lead to important policy and behavioral changes.⁵⁴ To understand the impact of these focusing events, the event is taken as a reference point for study, and observable data before and after the event took place are compared.⁵⁵

A recent MIT study on Google search traffic by Alex Marthews and Catherine Tucker used such a framework to provide an important contribution to chilling effect research. Their innovative research design treated the June 2013 revelations as an exogenous focusing event, and tracked the relative number of searches for certain privacy-sensitive search terms before and after June 2013.⁵⁶ Marthews and Tucker found a statistically significant 5% reduction in Google searches for certain privacy-sensitive search terms after June 2013.⁵⁷ Their study not only provides evidence of chilling effects, but also offers a research design that may be employed to study chilling effects in other online contexts.

However, the study had its limitations. First, the dataset in the Google study only included search term data through mid-December 2013. Without more recent data, it is unclear whether the effects tracked in the

52. Graeme Boushey, *Punctuated Equilibrium Theory and the Diffusion of Innovations*, 40 POL’Y STUD. J. 127, 130 (2012); Marthews & Tucker, *supra* note 25, at 3.

53. See William Lowry, *Potential Focusing Projects and Policy Change*, 34 POL’Y STUD. J. 313, 313–15 (2006) (discussing research analyzing the impact on policy (and other social and political factors) caused by focusing or intervening events). See also Boushey, *supra* note 52, at 130 (discussing the use of focusing events as reference points for policy change studies). See generally FRANK R. BAUMGARTNER & BRYAN D. JONES, *AGENDAS AND INSTABILITY IN AMERICAN POLITICS* (2010) (arguing that dramatic policy shifts can, in part, be attributed to important triggering/focusing events); THOMAS A. BIRKLAND, *AFTER DISASTER: AGENDA SETTING, PUBLIC POLICY, AND FOCUSING EVENTS* 30–35 (1997) (setting out a framework for studying the impact of “focusing events” on policy changes, including the important role, and impact, of news coverage of such focusing events); PAUL A. SABATIER & HANK C. JENKINS-SMITH, *THE ADVOCACY COALITION FRAMEWORK: ASSESSMENT, REVISIONS, AND IMPLICATIONS FOR SCHOLARS AND PRACTITIONERS* (1993) (examining, more generally, the role and impact of external events on policy shifts).

54. See generally BAUMGARTNER & JONES, *supra* note 53 (asserting that dramatic policy shifts can, at least in part, be attributed to important triggering events).

55. Boushey, *supra* note 52, at 130 (discussing how focusing events, and their impact, can help understand policy shifts and other changes over time).

56. Marthews & Tucker, *supra* note 25, at 5–9.

57. *Id.* at 3.

study had a permanent, or at least longer term, impact. Second, the authors obtained their data from Google Trends, which provides Google search data in “normalized” or adjusted format.⁵⁸ The search data is normalized in two ways. First, the data represents only a percentage of total Google searches for any given term.⁵⁹ Second, Google “adjusts” the search data to render comparisons across regions more easily; these results are further “scaled to a range of 0 to 100.”⁶⁰ This, the authors admitted, meant it was “harder to make projections” based on the findings of the study (such as resulting “economic outcomes” due to the reduction in specific search-related ads) because without raw and unadjusted search data, it is difficult to measure on a granular level how people’s Google search activities were impacted.⁶¹ A third limitation of the study was the lack of a genuine control group. The study examined trends before and after June 2013, but there was no opportunity to control the PRISM/NSA revelations like a true experimental intervention.⁶² A fourth limitation was the possibility that users were still searching for the same search terms but simply using an alternative search engine to Google (presumably one not expressly linked to the NSA’s PRISM program).⁶³ Despite these limits, however, Marthews and Tucker *did* provide evidence of chilling effects in a concrete online context—search.

Sören Preibusch’s more recent study employed Marthews and Tucker’s design centered on the June 2013 revelations, examining Bing search term trends and Tor usage data as a proxy for users engaging in “privacy-enhancing” activities (Tor is a browser designed to protect privacy and anonymity online).⁶⁴ Preibusch found that while users’ behavior did change immediately after the June 2013 revelations, those privacy

58. *Id.* at 8.

59. *Where Trends Data Comes From*, GOOGLE (2016) https://support.google.com/trends/answer/4355213?hl=en&ref_topic=4365599 [<https://perma.cc/7TVA-SWMK>] (“Google Trends analyzes a percentage of Google web searches to figure out how many searches were done over a certain period of time. For example, if you search for tea in Scotland in March of 2007, Trends analyzes a percentage of all searches for tea within the same time and location parameters.”).

60. *How Trends Data is Adjusted*, GOOGLE (2016) https://support.google.com/trends/answer/4365533?hl=en&ref_topic=4365599 [<https://perma.cc/V4Z9-QF8C>] (“Google Trends adjusts search data to make comparisons between terms easier. Otherwise, places with the most search volume would always be ranked highest. To do this, each data point is divided by the total searches of the geography and time range it represents, to compare relative popularity. The resulting numbers are then scaled to a range of 0 to 100.”).

61. Marthews & Tucker, *supra* note 25, at 8.

62. The authors note that they only have “quasi” controls. *Id.* at 6.

63. *Id.* at 6–8.

64. Sören Preibusch, *Privacy Behaviors After Snowden*, 58 COMM. ACM 48, 48–52 (2015).

behaviors “faded quickly.”⁶⁵ Like Marthews and Tucker, Preibusch acknowledged important limitations. For example, the use of Bing data likely “biased” his results, and his selection of data sources was “partly pragmatic” in this sense.⁶⁶ And also like Marthews and Tucker, his dataset was temporally limited—only extending from May 2013 to January 2014.⁶⁷ In a different but related 2013 study, Yoan Hermstrüwer and Stephan Dickert found little evidence of significant chilling effects associated with privacy and reputational risks of embarrassing online disclosures, leading them to conclude that “dystopian” concerns often expressed by privacy scholars about chilling effects and the conforming impact surveillance were overstated.⁶⁸ Again, the researchers acknowledged a number of important “caveats” to their findings, most notably that they were likely biased due to self-selection by participants who had already bound themselves to conforming behavior through their choices and involvement in the study itself.⁶⁹

The study in this Article builds on the Marthews and Tucker design. To document how government surveillance has affected user behavior online, the case study’s interrupted time series research design approaches the June 2013 revelations as the interrupting “exogenous shock” or “focusing event,” and examines whether Wikipedia article traffic for certain topics that reasonably raise privacy concerns for Internet users decreased following those revelations. But the study also aims to address some of the aforementioned limitations of studies of this nature. For example, the dataset employed will include data that starts earlier (January 2012) and extends later (August 2014). Furthermore, the Wikipedia article traffic data employed is raw and unadjusted, providing a more

65. *Id.* at 48, 55.

66. *Id.* at 55 (“My analysis of Web search behavior through Microsoft’s Bing search engine may have introduced a bias impossible to quantify, should it exist.”).

67. *Id.* at 48.

68. Yoan Hermstrüwer & Stephan Dickert, *Tearing the Veil of Privacy Law: An Experiment on Chilling Effects and the Right to be Forgotten* 22–23 (Preprints of the Max Planck Institute for Research on Collective Goods, Working Paper No. 2013/15, 2013), <http://www.econstor.eu/bitstream/10419/84983/1/757205445.pdf> [<https://perma.cc/U7L3-TRQ9>] (detailing an experimental study on chilling effects finding that risks of “networked publicity” (exposure online of users’ embarrassing activities) did not affect users’ “privacy valuations,” nor “dampen” either “behavioral idiosyncrasies” nor the “panoply of different behaviors” involved in the study).

69. *Id.* at 25 (“A second critique may be that, in our setting, networked publicity is a function of an endogenous choice, making causal inferences about the factors driving social norm compliance more difficult. Individuals may have self-selected into networked publicity because of their stronger inclination to comply with social norms.”).

accurate and granular understanding of any observed changes in data trends.

D. HYPOTHESIS

This case study asks: does the Wikipedia article traffic for the privacy concerning topics tracked decrease after the “exogenous shock” of widespread publicity surrounding the June 2013 revelations? A hypothesis based on chilling effects theory may be stated this way: due to chilling effects caused by increased awareness of government surveillance online, Internet users will be less likely to view Wikipedia articles on topics that raise privacy-related concerns. In providing noteworthy evidence suggesting a NSA/PRISM surveillance related chilling effect, this study is among the first to do so using web traffic data (instead of survey responses or search) and the first to evidence the impact of surveillance chill not only on Wikipedia users but on how people seek, and access, information and knowledge online more generally. The next Part sets out this case study’s research design and methodology, including its focus on Wikipedia.

III. METHOD AND DESIGN

A. WHY WIKIPEDIA TRAFFIC?

This case study focuses on English Wikipedia (i.e., articles with content in the English language) and traffic to specific Wikipedia articles as a means of exploring chilling effects online. Why Wikipedia? First, despite some skepticism as to its accuracy, Wikipedia is an influential resource for information and knowledge online. Over 50% of Internet users use Wikipedia as a source of information,⁷⁰ and over a third of Americans visit Wikipedia annually, making it one of the top ten most popular sites on the Internet. In a study of college students, researchers Alison Head and Michael Eisenberg found 52% used Wikipedia “frequently.”⁷¹ Therefore, if government surveillance is chilling users from accessing Wikipedia, then there are implications beyond Wikipedia’s function as an online encyclopedia. Researchers have used Wikipedia for a broad range of research, relating to both online and offline concerns,

70. Lee Rainie et al., *Wikipedia, Past and Present*, PEW INTERNET & AM. LIFE PROJECT SURV. (Jan. 13, 2013), http://www.pewinternet.org/files/old-media/Files/Reports/2011/PIP_Wikipedia.pdf [<https://perma.cc/M8J9-UYDG>].

71. Alison J. Head & Michael B. Eisenberg, *How Today’s College Students Use Wikipedia for Course-Related Research*, FIRST MONDAY, Mar. 1, 2010, <http://firstmonday.org/article/view/2830/2476> [<https://perma.cc/7TZZ-9FNK>]; Hilles, *supra* note 3, at 247.

including theorizing and understanding peer-production,⁷² mapping online knowledge and patterns of local knowledge production,⁷³ and investigating the subtle ways that popular information platforms like Wikipedia influence far more than just students or researchers seeking knowledge online.⁷⁴ These works all illustrate Wikipedia's importance beyond being a basic source of information, so a chilling effect on Wikipedia users would also threaten or negatively impact these other important uses and contributions of the site—if people were chilled en masse from using Wikipedia over time, it could no longer be used as an important focal point for such research.

Second, there is existing research suggesting media coverage can impact Wikipedia use. Research has shown how media coverage and “breaking news events” impact Wikipedia editors and other collaborations on article content.⁷⁵ If Wikipedia editors and contributors respond to

72. See, e.g., YOCHAI BENKLER, *THE WEALTH OF NETWORKS: HOW SOCIAL PRODUCTION TRANSFORMS* 70–74, 101–23, 287–94 (2006) (discussing Wikipedia in relation to a range of topics, including “networked information economy,” the “economics of social production,” and the nature of Internet culture); Yann Algan, Yochai Benkler, Mayo Fuster Morell & Jérôme Hergueux, *Cooperation in a Peer Production Economy: Experimental Evidence from Wikipedia*, WORKSHOP ON INFO. SYS. & ECON. (2013), https://www.parisschoolofeconomics.eu/IMG/pdf/hergueux_paper-2.pdf [<https://perma.cc/MGR2-TE4J>] (using Wikipedia to study the social foundations of peer contributions and production).

73. See, e.g., Mark Graham, Bernie Hogan, & Ralph K. Straumann, *Uneven Geographies of User-Generated Information: Patterns of Increasing Informational Poverty*, 104 ANNALS ASS'N AM. GEOGRAPHERS 746 (2014) (using Wikipedia to map patterns of global knowledge and information production).

74. See, e.g., Mark Graham, *Internet Geographies: Data Shadows and Digital Divisions of Labour*, in SOCIETY AND THE INTERNET: HOW NETWORKS OF INFORMATION AND COMMUNICATION ARE CHANGING OUR LIVES 99 (Mark Graham & William H. Dutton eds., 2014) (using Wikipedia to understand a “digital division” of labor in global information production); Shun-Ling Chen, *The Wikimedia Foundation and the Self-Governing Wikipedia Community: A Dynamic Relationship Under Constant Negotiation*, in CRITICAL POINT OF VIEW: A WIKIPEDIA READER 351, 362 (Geert Lovink & Nathaniel Tkacz eds., 2011) (discussing elements of Wikipedia's governance structure, including how it is vulnerable to “chilling effects” and other regulatory problems).

75. Brian C. Keegan, *Emergent Social Roles in Wikipedia's Breaking News Collaborations*, in ROLES, TRUST, AND REPUTATION IN SOCIAL MEDIA KNOWLEDGE MARKETS 57, 57–79 (Elisa Bertino & Sorin Adam Matei eds., 2015) (reviewing literature exploring how Wikipedia covers news events, and the impact those events have on Wikipedia collaborative infrastructure and networks, and also providing a brief overview of research examining the nature and structure of Wikipedia editor networks more generally); Brian C. Keegan, *A History of Newswork on Wikipedia*, PROC. 9TH INT'L SYMP. ON OPEN COLLABORATION, ACM: NEW YORK (2013) (noting, among other things, how Wikipedia becomes a focal point for information seekers during breaking news events, both as a source of information and to understand the event and share information about it); Brian C. Keegan, Darren Gergle & Noshir Contractor, *Hot Off the*

media coverage and significant news events, it is reasonable to predict that the widely covered surveillance revelations may likewise affect Wikipedia users. In short, there is an existing empirical foundation that media coverage of an important story like government surveillance could impact Wikipedia and its users.

There are also methodological reasons for this case study's focus on Wikipedia. First, unlike Google Trends, the Wikimedia Foundation provides a wealth of data on key elements of its site, including article traffic data, which can provide a more accurate picture as to any impact or chilling effects identified.⁷⁶ Second, Wikipedia, a "unique, online, collaborative encyclopedia,"⁷⁷ has over 500 million visitors per month, and its collaborative and peer-produced content is growing at a rate of 17,800 articles per day (as of May 2014, English Wikipedia content includes over 4.6 million articles).⁷⁸ In other words, Wikipedia is a massively popular medium and one that is also growing in content and scope. As such, any observed chilling effect would implicate a large number of Internet users doing something wholly legal—accessing information and knowledge in an encyclopedia—and chilled or reduced use would run counter to these Wikipedia use and content trends.

Finally, the public policy impact of any observed Wikipedia chilling effects is also a consideration. Investigating "chilling effect" claims related to Wikipedia use has recently become a matter of important public interest, in light of the Wikimedia Foundation lawsuit alleging NSA surveillance has had a chilling effect on Wikipedia and its users.⁷⁹ This case study will test and explore these claims.

Wiki: Structures and Dynamics of Wikipedia's Coverage of Breaking News Events, 57 AM. BEHAVIORAL SCIENTIST 595 (2013), <http://abs.sagepub.com/content/57/5/595> [<https://perma.cc/TFQ7-UZKT>] (analyzing Wikipedia revision history data to explore the nature and structure of Wikipedia collaborative efforts in relation to breaking news events).

76. Wikipedia provides a wealth of information about its number of articles, editors, page views, etc. *Growth per Wikipedia Wiki*, WIKIMEDIA, <https://stats.wikimedia.org/wikimedia/animations/growth/AnimationProjectsGrowthWp.html> [<https://perma.cc/9PGP-UVCY>].

77. See Bar-Ilan & Aharony, *supra* note 3, at 243.

78. Hilles, *supra* note 3, at 245; McIver & Brownstein, *supra* note 3, at 1.

79. See Wales & Tretikov, *supra* note 3. See generally Complaint, *supra* note 1 (Wikimedia Foundation is the lead complainant and the only complainant in the lawsuit that provides accessible data that can be analyzed for the purposes of this study.).

B. RESEARCH DESIGN AND DATA SELECTION

This case study uses an interrupted time series (ITS) design.⁸⁰ An ITS design uses a time series, which is a series of measurements or observations over time that is “interrupted” by some intervention or exogenous event. Such intervention divides the time series into two segments, resulting in measurements of time series before and after the intervening event. By “comparing” patterns in the time series data before and after the interruption, the study can assess the impact of an interrupting intervention or an event.⁸¹ This study will compare patterns in the data before and after the June 2013 revelations.

Furthermore, this study combines segmented regression analysis with its ITS design. Such combination offers a powerful means for exploring the effects of interventions, events, or policy changes as long as there is a clearly identified time point of intervention.⁸² Segmented regression is useful because it allows a comparison in data levels and trends (like a reduction in views of Wikipedia articles over time) before and after an intervening event, while helping to isolate the impact of that event by controlling for other factors and variables. Because of its capacity to visualize observed and analyzed data in a compelling way, ITS design has been applied to a range of fields⁸³ and is particularly popular among policy researchers.⁸⁴ It has also been used to explore the effects of laws, policing (including surveillance by law enforcement), and other regulatory actions.⁸⁵

80. See sources cited *supra* note 27.

81. CAMPBELL, STANLEY & GAGE, *supra* note 27, at 37; Fok, Henry & Allen, *supra* note 27, at 4; Mark & Reichardt, *supra* note 27, at 354–55; Penfold & Zhang, *supra* note 27, at S39–S40; Wagner et al., *supra* note 27, at 299.

82. Fok, Henry & Allen, *supra* note 27, at 976 (“The [ITS design] is especially useful when there is a clearly identified time point of intervention or policy change.”); Mark & Reichardt, *supra* note 27, at 354–55, 383 (“[I]nterrupted time-series designs can be among the most credible quasi-experimental designs.”); Penfold & Zhang, *supra* note 27, at S38 (stating that ITS design is among the “strongest” where randomized and controlled experiments are not possible); Wagner et al., *supra* note 27, at 299 (“Interrupted time series [with segmented regression analysis] is the strongest, quasi-experimental design to evaluate longitudinal effects of such time-delimited interventions.”).

83. Mylene Lagarde, *How to Do (or Not to Do) . . . Assessing the Impact of a Policy Change with Routine Longitudinal Data*, 27 HEALTH POL’Y & PLAN. 76, 76 (2011) <http://heapol.oxfordjournals.org/content/27/1/76.full.pdf> [<https://perma.cc/HJG7-ZS65>] (describing how “quasi-experimental” ITS designs employing segmented regression have been used in various fields, including environmental studies, economics, and health policy).

84. See, e.g., Benjamin French & Patrick J. Heagerty, *Analysis of Longitudinal Data to Evaluate a Policy Change*, 27:24 STAT. MEDICINE 5005 (2008) (surveying different research designs and methods in policy change research); Lagarde, *supra* note 83, at 76

This case study uses data on English language Wikipedia article view counts from the online service stats.grok.se, a portal maintained by a Wikimedia Foundation member. This portal provides access to a range of Wikipedia analytics, stats, and data.⁸⁶ In particular, the portal aggregates Wikipedia article view data on a daily and monthly basis.⁸⁷ This data at stats.grok.se has been used in a range of research, including studies involving market trends, health information access, and social-political change.⁸⁸

(describing in detail an ITS design that employs segmented regression as a simple but robust method to study policy impact and change); Wagner et al., *supra* note 27.

85. See, e.g., Samuel Cameron, *The Economics of Crime Deterrence: A Survey of Theory and Evidence*, 41:2 KYKLOS 301, 314 (1988) (noting economists had begun using ITS designs, which then had long been used by criminologists); Daniel S. Nagin, *Criminal Deterrence Research at the Outset of the Twenty-First Century*, 23 CRIME & JUST. 1, 8–12 (1998) (discussing a range of laws and police operations that have been using ITS designs); Lynn W. Phillips & Bobby J. Calder, *Evaluating Consumer Protection Laws: II. Promising Methods*, 14:1 J. CONSUMER AFFAIRS 9 (1980) (surveying literature on methods/research designs used to study consumer protection laws, including ITS). For some more recent examples, see Carl Bonander, Finn Nilson & Ragnar Andersson, *The Effect of the Swedish Bicycle Helmet Law for Children: An Interrupted Time Series Study*, 51 J. SAFETY RES. 15 (2014) (used ITS design to explore the impact of a bicycle helmet law by examining inpatient data on injured cyclists before and after the law was enacted); Becky Briesacher et al., *A Critical Review of Methods to Evaluate the Impact of FDA Regulatory Actions*, 22:9 PHARMACOEPIDEMIOLOGY & DRUG SAFETY 986 (2013) (reviewing a range of ITS design studies examining the impact of FDA regulatory actions often by exploring health data before and after the FDA action); Benjamin David Décaré Héту, *Police Operations 3.0: On the Impact and Policy Implications of Police Operations on the Warez Scene*, 6:3 POL'Y & INTERNET 315 (2014) (exploring the impact of police operations and crackdown on the “warez” (online piracy) scene with an ITS design that examined data on the output of different warez communities before and after five different police operations); Jeffrey T. Ward, Matt R. Nobles, Lonn Lanza-Kaduce, Lora M. Levett & Rob Tillyer, *Caught in Their Own Speed Trap: The Intersection of Speed Enforcement Policy, Police Legitimacy, and Decision Acceptance*, 14:3 POLICE Q. 251 (2011) (using an ITS design to study the impact certain policy changes have on public opinion concerning the legitimacy police action, specifically, comparing speeding citation contestation rates before and after the introduction of an advertising campaign labeling the intervention city a “speed trap”).

86. Stats.grok.se is maintained by Domas Mituzas, a Wikipedia developer, past Board of Trustee on the Wikimedia Foundation, and present member of its Advisory Board. See *Frequent Questions*, GROK, <http://stats.grok.se/about> [<https://perma.cc/KQZ3-XGPZ>].

87. *Id.*

88. See, e.g., Michela Ferron & Paolo Massa, *WikiRevolutions: Wikipedia as a Lens for Studying the Real-Time Formation of Collective Memories of Revolutions*, 5 INT'L J. COMM. 1313 (2011) (examining Wikipedia as a “lens” through which to understand real-time social and political upheaval and change); Michaël R. Laurent & Tim J. Vickers, *Seeking Health Information Online: Does Wikipedia Matter?* 16:4 J. AM. MED. INFORMATICS ASS'N 471 (2009) (using Wikipedia traffic data from stats.grok.se to study the relevance

Like the Marthews and Tucker study, this case study uses a list of keywords the U.S. Department of Homeland Security uses to track and monitor social media.⁸⁹ This list categorizes certain search terms in relation to a range of different issues such as “Health Concern,” “Infrastructure Security,” and “Terrorism.” According to the DHS documents themselves, the list is meant to assist analysts to monitor social media to provide “situational awareness and establish a common operating picture.”⁹⁰ Though the methodology for formulating the list is not well known, presumably the terms represent ideas or content people associate with “terrorism” and other national security matters, which is why government officials are interested in tracking the terms online.⁹¹

Using government keyword lists to study government surveillance or censorship is not new.⁹² Here, the DHS keywords provide a helpful basis to select Wikipedia articles for the study. To be clear, this keyword list is non-random, and it is not chosen based on any assumption that the general public is aware of the list or the topics attached. In other words, this study does not assume that people are avoiding topics relating to these keywords due to the DHS’s media monitoring program. Rather, the list is used for pragmatic methodological reasons. Similar to how the list was

of Wikipedia to how people access to health information online); Helen Susannah Moat et al., *Quantifying Wikipedia Usage Patterns Before Stock Market Moves*, 3 SCI. REP. 1 (2013) (investigating Wikipedia article traffic and usage in relation to stock market changes).

89. The keyword list has been publicly available online since 2012, and was updated and re-posted by the DHS in 2013: U.S. DEP’T OF HOMELAND SEC., NATIONAL OPERATIONS CENTER MEDIA MONITORING CAPABILITY ANALYST’S DESKTOP BINDER (2011), <https://epic.org/foia/epic-v-dhs-media-monitoring/Analyst-Desktop-Binder-REDACTED.pdf> [<https://perma.cc/2Z39-XMW9>] [hereinafter ANALYST’S DESKTOP BINDER]. This was later updated and posted online by the DHS. See U.S. DEP’T OF HOMELAND SEC., PRIVACY IMPACT ASSESSMENT FOR THE OFFICE OF OPERATIONS COORDINATION AND PLANNING (2013), https://www.dhs.gov/sites/default/files/publications/privacy/PIAs/privacy_pia_ops_NOC%20MMC%20Update_April2013.pdf [<https://perma.cc/9VJN-YKRL>] [hereinafter PRIVACY IMPACT ASSESSMENT].

90. PRIVACY IMPACT ASSESSMENT, *supra* note 89, at 23 app. B; Marthews & Tucker, *supra* note 25, at 3–4.

91. Marthews & Tucker, *supra* note 25, at 6.

92. Jedidiah R. Crandall & Masashi Crete-Nishihata et al., *Chat Program Censorship and Surveillance in China: Tracking TOM-Skype and Sina UC*, FIRST MONDAY, July 1, 2013, <http://firstmonday.org/ojs/index.php/fm/article/view/4628/3727> [<https://perma.cc/M5FJ-T4D5>]; Jeffrey Knockel, Jedidah Crandall & Jared Saia, *Three Researchers, Five Conjectures: An Empirical Analysis of Tom-Skype Censorship and Surveillance*, 16:4 FOCI ’11: USENIX WORKSHOP ON FREE & OPEN COMM. ON INTERNET (2011), <https://www.cs.unm.edu/~crandall/foci11knockel.pdf> [<https://perma.cc/FH8H-JUBA>].

used in the Marthews and Tucker search trends study,⁹³ the DHS documents, and the keywords therein, are used to select Wikipedia articles that represent the sort of articles that users may be chilled from accessing in light of government surveillance.⁹⁴

This case study selected forty-eight Wikipedia articles that corresponded with the DHS keywords listed as relating to “terrorism.”⁹⁵ The full list of the keywords used (including terms such as “dirty bomb,” “suicide attack,” “nuclear enrichment,” and “eco-terrorism”) and the corresponding English language Wikipedia articles for which “page view” counts were collected via stats.grok.se can be found in Table 8 of the Appendix.⁹⁶ The keywords relating to “terrorism” were used to select the Wikipedia articles because the U.S. government cited terrorism as a key justification for its online surveillance practices. Moreover, much of the media and news coverage framed the revelations around terrorism and national security.⁹⁷ Wikipedia articles coinciding with these terrorism-related topic keywords may include the kind of information or content users may avoid accessing in light of potential government surveillance. This study aggregated Wikipedia article view counts on a monthly basis for the forty-eight Wikipedia articles over a thirty-two month period, from the beginning of January 2012 to the end of August 2014. Those

93. Marthews & Tucker, *supra* note 25, at 3–4.

94. For example, if the government surveillance is focusing on terrorism online, people may have privacy concerns about accessing terrorism-related information online, and are thus “chilled” or deterred from accessing.

95. PRIVACY IMPACT ASSESSMENT, *supra* note 89, at 27. Locating Wikipedia articles coinciding with each keyword was done manually; this was a rather simple exercise as there was a Wikipedia article that corresponded perfectly with the vast majority of keywords in the “terrorism” DHS keyword category. There were a few discrepancies, however: the Wikipedia article “environmental terrorism” was used for the keyword “environmental terrorist”; the keyword “target” was excluded as they were too many potentially corresponding Wikipedia articles; the Wikipedia article “political radicalism” was used for the DHS keyword “radicalism” because there were too many potentially corresponding articles; the keyword “enriched” was excluded as it was redundant with the included Wikipedia article “nuclear enrichment”; and there were also no Wikipedia articles corresponding with DHS keywords “weapons cache,” “suspicious substance,” “plot,” and “homegrown.” Wikipedia articles corresponding with the remaining 48 DHS “terrorism” related keywords were all included in the study.

96. For clarity, the raw Wikipedia article “Page View” statistics track total views or loads of the Wikipedia articles or pages in question, not unique visitors. See *Pageview Statistics*, WIKIPEDIA.ORG, https://en.wikipedia.org/wiki/Wikipedia:Pageview_statistics [https://perma.cc/JSU2-E6PU].

97. Qin, *supra* note 45, at 178 (finding that a predominant “framing” in traditional news media coverage of the Snowden surveillance disclosures focused on national security terrorism, along with international relations).

forty-eight Wikipedia articles corresponded with all DHS keywords listed in the “terrorism” category.⁹⁸

Although forty-eight is not an extraordinarily large sample size, the Wikipedia traffic attracted by these articles represents over 81 million total article page views over the course of the study. This means that the potential number of Internet users tracked in the study could be several millions.⁹⁹ Moreover, to ensure the sample of forty-eight articles could be generalized to a wider sample of content (both terrorism-related and other topics that may raise privacy concerns), the study used “crowdsourcing” to measure the privacy value of the topics in question, following the approach of Marthews and Tucker.¹⁰⁰

Crowdsourcing involves completing certain tasks with the assistance of larger pools of online users or “crowds”—recruited through online services like Amazon’s Mechanical Turk (MTurk)—and has become a common technique for researchers to evaluate research instruments and other measures for privacy or privacy-related concerns.¹⁰¹ MTurk is an “open” online crowdsourcing platform founded in 2005 that provides a means for

98. PRIVACY IMPACT ASSESSMENT, *supra* note 89, at 27. The findings in this study primarily concern the English speaking world, as only English Wikipedia article view counts are tracked in the Wikipedia data.

99. Thus, though the selection of the forty-eight English Wikipedia articles was not random (there is no sampling frame for all terrorism-related Wikipedia articles), the data clearly indicates these articles represent a substantial number of Wikipedia users. Though a precise number cannot be estimated (the Wikipedia data tracks “Page View” statistics, that is, total views or loads of the Wikipedia articles or pages, not unique visitors), the data arguably still involves a large number of Internet users—many millions.

100. Marthews and Tucker similarly recruited independent “raters” to evaluate the privacy value of Google search terms in their study. Marthews & Tucker, *supra* note 25, at 3–5.

101. See, e.g., Berker Agir, Jean-Paul Calbimonte & Karl Aberer, *Semantic and Sensitivity Aware Location Privacy Protection for the Internet of Things*, PRIVACY ONLINE: WORKSHOP ON SOC’Y, PRIVACY & SEMANTIC WEB (PRIVON) (2014), http://ceur-ws.org/Vol-1316/privon2014_paper5.pdf [<https://perma.cc/MB63-55PL>]; Margherita Bonetto et al., *Privacy in Mini-drone Based Video Surveillance*, WORKSHOP ON DE-IDENTIFICATION FOR PRIVACY PROTECTION MULTIMEDIA (2015), <http://infoscience.epfl.ch/record/206109> [<https://perma.cc/Z4ZT-DVMX>]; Pavel Korshunov et al., *Crowdsourcing-based Evaluation of Privacy in HDR images*, SPIE PHOTONICS EUR. (2014), <http://proceedings.spiedigitallibrary.org/proceeding.aspx?articleid=1873752> [<https://perma.cc/YCJ5-YY48>]; Pavel Korshunov et al., *Framework For Objective Evaluation of Privacy Filters*, 8856 PROCEEDINGS SPIE APPLICATIONS OF DIGITAL IMAGE PROCESSING XXXVI (2013), <http://proceedings.spiedigitallibrary.org/proceeding.aspx?articleid=1744325> [<https://perma.cc/XMJ8-D7LB>]; Jialiu Lin et al., *Expectation and Purpose: Understanding Users’ Mental Models of Mobile App Privacy Through Crowdsourcing*, PROC. 2012 ACM CONF. ON UBIQUITOUS COMPUTING (2012), <http://dl.acm.org/citation.cfm?id=2370290> [<https://perma.cc/Y7ZD-2JFF>].

“task creation,” “recruitment,” “compensation,” and “data collection.”¹⁰² Several studies summarized and documented MTurk’s advantages for survey, experimental, and other empirical research. Paolacci and Chandler recently concluded, after extensively canvassing existing evidence, that researchers may use MTurk for “virtually any study that is feasible to conduct online.”¹⁰³ Indeed, the MTurk service has been validated as a tool for a range of research, including research on behavioral economics, and decision-making, collective behavior experiments, linguistic and cognitive psychological experiments, and, importantly for our purposes, conducting survey research.¹⁰⁴ Samples recruited with MTurk have been found to be “at least as diverse as traditional subject pools” in terms of the general U.S. population and “relatively representative” of the U.S. Internet using population.¹⁰⁵

A total of 415 independent Internet users participated in the crowdsourcing project through MTurk, and they rated each of the forty-eight topics, with which the Wikipedia articles in the data set corresponded. The questions were designed to explore the likelihood that the topics would raise privacy-related concerns for Internet users. To minimize self-selection and response bias (a limitation difficult to avoid in non-random sampling), the brief questionnaires were described as merely an “Online Information Study” to potential MTurk participants.

The respondents recruited for the evaluations were similar to other MTurk participant pools that are “relatively representative of the population of U.S. Internet users.”¹⁰⁶ However, the respondents for this study were younger, more educated, had slightly lower incomes than the broader U.S. Internet population, and were slightly more male than female

102. Michael Buhrmester, Tracy Kwang & Samuel Gosling, *Amazon’s Mechanical Turk: A New Source of Inexpensive, Yet High Quality, Data?*, 6:1 PERSP. ON PSYCHOL. SCI. 3, 3 (2011), <http://pps.sagepub.com/content/6/1/3.abstract> [<https://perma.cc/YQ95-ASS7>].

103. Gabriele Paolacci & Jesse Chandler, *Inside the Turk: Understanding Mechanical Turk as a Participant Pool*, 23:3 CURRENT DIRECTIONS PSYCHOL. SCI. 184, 186 (2014), <http://cdp.sagepub.com/content/23/3/184.abstract> [<https://perma.cc/XN2G-LGD8>].

104. Matthew J.C. Crump, John V. McDonnell & Todd M. Gureckis, *Evaluating Amazon’s Mechanical Turk as a Tool for Experimental Behavioral Research*, 8:3 PLOS ONE e57410, e57410 (2013), <http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0057410> [<https://perma.cc/N5YP-PUYZ>].

105. Panagiotis G. Ipeirotis, *Turker Demographics vs. Internet Demographics*, COMPUTER SCIENTIST BUS. SCH. (2009), <http://www.behind-the-enemy-lines.com/2009/03/turker-demographics-vs-internet.html> [<https://perma.cc/NET8-XU47>]; Gabriele Paolacci, Jesse Chandler & Panagiotis G. Ipeirotis, *Running Experiments on Amazon Mechanical Turk*, 5:5 JUDGMENT & DECISION MAKING 411, 411–12 (2010).

106. Paolacci, Chandler & Ipeirotis, *supra* note 105, at 412.

(56% of respondents were male and 44% female). The respondents were also highly likely to use “websites and other online resources” for information more generally (83.8% were “very likely” and another 15% “somewhat likely”) and, in particular, to “stay informed about current events” (73.5% were “very likely” and another 22.9% “somewhat likely”). Respondents were asked to indicate on a scale of 1 to 5 (1 being very unlikely and 5 being very likely): how likely they thought they would be in trouble if the U.S. government found out that they accessed information about the topic in question (Government Trouble Rating); how “privacy-sensitive” they viewed each topic as (in this case, 5 being highly sensitive and 1 not at all) (Privacy-Sensitive Rating); how likely they would be to delete the browser history on their computer after accessing information about the topic; and how likely they would avoid viewing or accessing information on the topic if they knew the Government was monitoring people’s activities online (Avoidance Rating).¹⁰⁷

The results from the crowdsourcing survey are set out in Table 7 of the Appendix. On balance, the results from the first three categories suggest that the topics raised notable privacy concerns for respondents. The results do not suggest, however, that any one category raised overly strong concerns. This is not entirely surprising, since there is no particular reason why someone would expect to be in “trouble” with the government for simply accessing information online when such access is a legal activity. However, the rating scores in response to the fourth category were noteworthy—things appear to change when people are aware of government surveillance. This is apparent from the higher “Avoidance” rating of 2.62, which suggests that respondents were overall more likely to avoid the topics in question if they *knew* the government was monitoring online activities. In short, the ratings suggest that the topics of the forty-eight Wikipedia articles raise privacy concerns for Internet users, particularly when people suspect the government is monitoring them, which may lead them to avoid or be chilled from accessing that information in particular.

C. METHOD OF ANALYSIS

A strength of an ITS design is that there are multiple assessments or measures before and after the event or intervention in the time series; such multiplicity controls for changes in level and secular trends in the data and

107. This is not a standard scale, but one developed for this case study. For methodological consistency, these questions were designed to track ratings categories used by Marthews and Tucker. Marthews & Tucker, *supra* note 25, at 12, 38.

increases the robustness of results.¹⁰⁸ This case study uses Wikipedia article view counts to create a time series over a thirty-two month period from January 2012 to August 2014 (n=32), with the “interruption” or intervening event dividing the time series into two segments: before and after the June 2013 revelations. Two empirical approaches are used to analyze the interrupted time series. The first is a simple comparison of the mean number of views for all the Wikipedia articles in the dataset before and after June 2013. If there is a chilling effect due to surveillance revelations in June 2013, the average or mean number of views for the forty-eight Wikipedia articles should be lower for months following June 2013 than that of the months before. The second is a model-based empirical analysis. That is, segmented regression of an interrupted time series, which is the recommended method of analysis for ITS designs.¹⁰⁹ The health economist Mylene Lagarde, who has analyzed this method comprehensively, has provided an equation that expresses the specification for the regression analysis:¹¹⁰

$$Y_t = \beta_0 + \beta_1 * \text{time} + \beta_2 * \text{intervention} + \beta_3 * \text{postslope} + \varepsilon_t^{111}$$

The ITS design controls for “secular trends”—the long-term and non-periodic trends in the data.¹¹² To strengthen the robustness and validity of

108. CAMPBELL, STANLEY & GAGE, *supra* note 27, at 37; Fok, Henry, & Allen, *supra* note 27, at 7; Lagarde, *supra* note 83; Penfold & Zhang, *supra* note 27, at S39; Wagner et al., *supra* note 27, at 308.

109. Penfold & Zhang, *supra* note 27, at S41–42; Wagner et al., *supra* note 27, at 299.

110. Lagarde, *supra* note 83, at 79–80.

111. In this case study, Y_t , the “outcome” or dependent variable is the raw aggregate total of Wikipedia article views (or “view count”) for the forty-eight articles in the study. The time variable includes thirty-two time points in the time series, representing each of the thirty-two months in the time series data set from January 2012 to August 2014, which is the period of study. So, in this time series data set, the “outcome” or dependent variable is the aggregate views of all forty-eight Wikipedia articles totaled on a monthly basis, for each of the thirty-two months. For greater clarity, in this model, β_0 captures the baseline level of the outcome variable at time 0—here, that would be the expected total views for all forty-eight Wikipedia articles in the data set at the beginning of the study; β_1 estimates the secular trend or growth rate in the total number of views for the forty-eight Wikipedia articles, independently from the “intervention” or intervening event (the June 2013 surveillance revelations); β_2 estimates the immediate impact of the “intervention” or the exogenous shock of PRISM/NSA surveillance publicity in June 2013, by reflecting the change in the “level” or the total number of views for the Wikipedia articles immediately after the June 2013 events; and, finally, β_3 reflects any change in the trend of the data; that is, any growth or decline in total views for the forty-eight Wikipedia articles on a month-to-month basis, after the intervention. Lagarde, *supra* note 83, at 79–80.

112. Lagarde, *supra* note 83, at 79.

results, however, additional controls can be included in the model, and, where appropriate, auto-correlation can be corrected.¹¹³ Here, although a true control group was not possible,¹¹⁴ comparator groups were also included in the analysis to increase its robustness, including both a comparator including security-related Wikipedia articles and another including the most popular (most viewed) Wikipedia in 2012, 2013, and 2014.¹¹⁵ This, as will be explained in Part IV, is done to compare the impact of the June 2013 surveillance revelations on both the terrorism-related Wikipedia articles and other content unlikely to raise privacy concerns. Overall Wikipedia use trends are also considered in the analysis to help isolate the impact of the June 2013 revelations beyond mere shifts in overall English Wikipedia article traffic in the same time period.¹¹⁶ Again, a prediction based on chilling effects theory is that there will be a decrease in the total views of terrorism-related Wikipedia articles after June 2013. If, in addition to an immediate drop, any chilling effects are more substantial and long-term then the overall long-term article view trends in the data may also be affected.

IV. RESULTS

A. NON-MODEL EMPIRICAL FINDINGS

The results discussed in this Section are “non-model” empirical findings, that is, these findings do not rely upon a statistical (regression) model. Instead, a more basic method of analysis is used whereby the average number of Wikipedia article views before and after the “focusing event” of June 2013 are compared (see Figure 1).

113. Lagarde, *supra* note 83, at 79, 81; Mark & Reichardt, *supra* note 27, at 385; Penfold & Zhang, *supra* note 27, at S42; Wagner et al., *supra* note 27, at 305.

114. Online surveillance potentially affects everyone, and there was no opportunity before the June 2013 revelations to isolate a control group.

115. *See infra* Section IV.B.4.

116. As previously noted, Wikimedia provides a wealth of information about page views. *Page Views for Wikipedia, Both Sits, Raw Data*, WIKIMEDIA (Nov. 11, 2015), <http://stats.wikimedia.org/EN/TablesPageViewsMonthlyOriginalCombined.htm> [<https://perma.cc/2JQS-PRG4>]. The segmented regression analysis was performed using the statistical software package Stata and auto-correlation is controlled for using the Prais-Winsten method where necessary. Lagarde, *supra* note 83, at 79 (recommending controlling for auto-correlation when employing this statistical analysis); *see also* GEORGE G. JUDGE ET AL., INTRODUCTION TO THE THEORY AND PRACTICE OF ECONOMETRICS (1985).

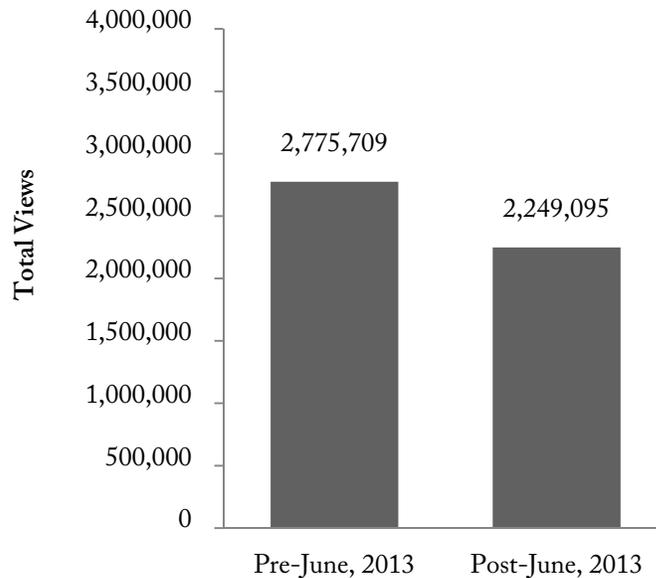


Figure 1. Average Monthly View Counts, Pre and Post June 2013. The reduction after the June 2013 surveillance revelations may suggest a chilling effect.

A lesser average number of views after June 2013 would be consistent with a chilling effect. The difference in mean values before and after June 2013 is notable—a reduction of 526,614 in the average monthly views for the articles, which represents approximately a 19.5% drop in article view counts. This is more than the mean differences found in the Google search terms study before and after June 2013.¹¹⁷ This reduction is also highly statistically significant.¹¹⁸ This itself may constitute evidence of a chilling effect. Of course, there are alternative explanations for these results. One possible explanation is that overall Wikipedia traffic (and thus, all Wikipedia article view counts) decreased after June 2013 for other reasons. Perhaps people are using Wikipedia less and less and this data is simply reflecting this overall declining trend unrelated to any surveillance revelations. So, while these findings are intriguing, a model-based empirical analysis is required to control for such variables as the overall trends in the data and to arrive at more robust empirical results. Therefore, statistical regression models (discussed in the following Section) were used to control for things like overall trends in Wikipedia article view traffic.

117. *But cf.* Marthews & Tucker, *supra* note 25, at 13–14.

118. The Cohen's *d* value was 1.3286.

B. MODEL-BASED EMPIRICAL FINDINGS

As noted earlier, the “outcome” or dependent variable in this analysis represents the raw aggregate total view counts per month for all Wikipedia articles in the data set. Segmented regression of the time series data set—“interrupted” by the June 2013 surveillance revelations—analyzes the impact of these revelations. Several sets of results are reported here to better illustrate findings. Analysis was conducted using Stata statistical software.¹¹⁹

1. *First Set of Results*

The first set of results is represented in Table 1 of the Appendix. Interestingly, the results indicate there was a reduction of 995,085 views immediately following the June 2013 revelations, which is a large, sudden, and statistically significant drop in the total view counts for the forty-eight Wikipedia articles.¹²⁰ The total article views as of May 2013 was 2,960,778, meaning this decline represents an immediate drop-off of over 30% of overall views. The results also indicate that there was no statistically significant change in the secular (or overall long-term) trend in the data. In short, because of the large drop in total view counts for the forty-eight Wikipedia articles, the data supports the existence of an immediate and substantial chilling effect following the June 2013 revelations. Figure 2 is a graphic visualization of the decrease, and includes a scatter plot of the data points in the set and a trend line based on the fitted results produced by the regression analysis:

119. Version 11.1.

120. There is little consensus for the appropriate method to measure effect size for single group ITS designs like the one used for this case study. See Larry Hedges, James Pustejovsky & William Shadish, *A Standardized Mean Difference Effect Size for Single Case Designs*, 3:3 RES. SYNTHESIS METHODS 224, 225 (2012). The most common method in “treatment” studies is the percentage of non-overlapping data (PND), which here is clearly above the 80% threshold for a “large” effect size. See generally Thomas E. Scuggs & Margo A. Mastropieri, *How to Summarize Single-Participant Research: Ideas and Applications*, 9:4 EXCEPTIONALITY 227 (2001) (proposing the use of non-overlapping data metric for summarizing single participant research). However, model diagnostics identified two influential outlier data points. The first outlier concerned view counts for the Wikipedia articles in the data set in November 2012 (Cooks *D* value=0.1644942), and the other was for view counts in July 2014 (Cooks *D* value=0.4121233). Both of these are extreme values and are visible in Figure 2.

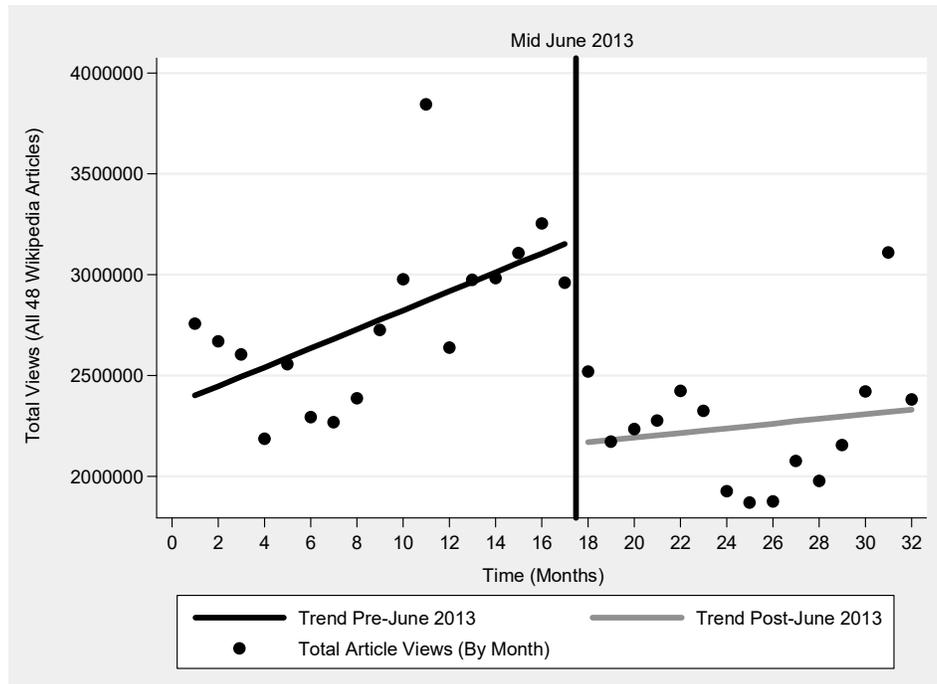


Figure 2. Pre and Post June 2013 Article View Trends (Outliers Included). The sudden drop, and flatter trend or slope in the data, after June 2013 surveillance revelations are consistent with a chilling effect.

In this graph, the large and statistically significant immediate drop (-995,085 page views) over the course of June 2013 can be seen, with the trend line (regression/line of best fit) substantially lower immediately after June 2013, compared to the trend in the data before June. Also, there was no statistically significant change in the secular trend in the data, as the slopes in the data before and after June 2013 are comparable: in each, the trend in view counts for the sample is modestly increasing overall.

Though there was not a statistically significant change in the trend of the data, the graph still suggests something more than an ephemeral chilling effect that dissipates quickly. Rather the data suggests a lasting impact on total article views. For example, the total article views as of August 2014 (month 32) is still lower than the views in April and May 2013 (months 15 and 16), the months prior to the revelations. The results also suggest that the chilling effect did not influence the long-term trend in the data that increased monthly. Though the number of views dropped off after June 2013, the trend in the data still increased until November 2013 (month 23) at a modest but apparent rate on a month-to-month basis.

In sum, the results are consistent with a sharp immediate chilling effect, possibly with a lasting impact on total views. However, the rising

secular trend in the Wikipedia article traffic is inconsistent with a significant long term chilling effect.

2. *Outliers: The “Exogenous Shock” of War*

Two outliers are clearly apparent in the data. The first is at month 11 (November 2012) and the second is at month 31 (July 2014). The view counts for the articles in the dataset skyrocketed in these months, far beyond any other months, either before or after the June 2013 revelations. In November 2012, the total views dramatically and temporarily increased, approaching almost 4,000,000 total views. In July 2014, views exponentially increased far beyond previous or subsequent months. Model diagnostics confirm that these are highly influential outliers with extreme values.¹²¹

What happened during November 2012 and July 2014 that caused the total view counts in the sample to suddenly skyrocket? War and conflict in the Middle East are likely the cause. In November 2012, Israel launched “Operation Pillar of Defense,” an eight day Israeli Defense Force (IDF) operation in Gaza in response to rocket attacks launched by the Palestinian militia group Hamas into southern Israel.¹²² The operation officially began on November 14, 2012, and ended on November 21, 2012 with an Egypt-brokered ceasefire. Then, in July 2014, Israel launched “Operation Protective Edge,” which is an IDF operation in Gaza against

121. Best practices for dealing with outliers in cases like this were observed—the technique used to identify the “influential” outlier should be indicated (here, Cooks *D*), deletion was used as a method to address the outlier, and results are reported with and without the outlier data. See Herman Aguinis, Ryan K. Gottfredson & Harry Joo, *Best-Practice Recommendations for Defining, Identifying, and Handling Outliers*, ORGANIZATIONAL RES. METHODS 8, 20–23 (2014), <http://orm.sagepub.com/content/early/2013/01/11/1094428112470848.abstract> [<https://perma.cc/6S3B-MPQ5>] (Techniques for identifying outliers should be indicated. Cooks *D* is noted (at 8 and 21) as an appropriate technique to identify an outlier’s influence globally in a regression (as here). Also the authors state (at 22) that influential outliers can be dealt with through deletion but “emphasize the importance of reporting the results with and without the chosen handling technique, which includes providing an explanation for any differences in the results, because the mere presence of influential outliers causes a dilemma in determining proper inference about a population based on a sample.”). The Cooks *D* value of 0.1286922 for the November 2012 view count and the Cooks *D* value of 0.3244882 for July 2014 are both extreme values.

122. For a “timeline” of the conflict and the IDF operation against Hamas, see *TIMELINE: Israel Launches Operation Pillar of Defense Amid Gaza Escalation*, HAARETZ (Nov. 20, 2012), <http://www.haaretz.com/news/diplomacy-defense/timeline-israel-launches-operation-pillar-of-defense-amid-gaza-escalation.premium-1.479284> [<https://perma.cc/5FXT-34YB>]; see also *Q&A: Israel-Gaza Violence*, BBC NEWS (Nov. 22, 2012), <http://www.bbc.com/news/world-middle-east-28439404> [<https://perma.cc/4H5L-N79W>].

Hamas.¹²³ This operation ended after fifty days, with widespread media coverage of the thousands of rockets fired from Gaza into Israel and the several thousand strikes by IDF on Hamas targets in Gaza.¹²⁴

These two high profile conflicts coincide with a dramatic and anomalous increase in the view counts during those months for the Wikipedia article on “Hamas” in the dataset. Examining more closely the view counts for the Hamas article over the thirty-two months in the data set, the “Hamas” Wikipedia article view count was 928,533 for November 2012, and then 1,220,490 for July 2014, which are far beyond the mean number of view counts for the article across all months in the study (134,574 monthly views). If we exclude these two outlier months, the contrast between the view counts for the Hamas article during those two months and other months in the dataset is even starker, with the mean being 71,912.¹²⁵ It can be inferred that the media coverage of these two conflicts involving Israel and Hamas led to a dramatic increase in Internet users seeking information about Hamas on English Wikipedia. This conclusion is supported by the fact that view counts for the Wikipedia article “Palestinian Liberation Organization” also increased in those same months.¹²⁶ It is also consistent with the findings of Zeitzoff, Kelly, and

123. Information about the operation is available in a Jerusalem Post article. Ben Hartman, *Fifty Days of Israel's Gaza Operation, Protective Edge—By the Numbers*, JERUSALEM POST (Aug. 28, 2014), <http://www.jpost.com/Operation-Protective-Edge/50-days-of-Israel's-Gaza-operation-Protective-Edge-by-the-numbers-372574> [<https://perma.cc/G64V-LFVW>]; *Gaza Crisis: Toll of Operations in Gaza*, BBC NEWS (Sept. 1, 2014), <http://www.bbc.com/news/world-middle-east-20388298> [<https://perma.cc/QGN6-D536>].

124. Amos Harel, *At the Crossroads of a Gaza Ground Operation*, HAARETZ (Jul. 12, 2014), <http://www.haaretz.com/news/diplomacy-defense/.premium-1.604601> [<https://perma.cc/2M6N-MAPY>] (“Hamas and Israel are waging an image battle. Their moves are the subject of constant media coverage, and, more than in the past, they are using information and photos from civilians, through smart phones and social media.”). In fact, the widespread media coverage of the Gaza conflict in 2012 led to a dramatic increase in social media activity during the 2012 conflict. Thomas Zeitzoff, *Does Social Media Influence Conflict? Evidence from the 2012 Gaza Conflict*, supp. at 2 (Feb. 17, 2015) (unpublished manuscript), http://www.zeitzoff.com/uploads/2/2/4/1/22413724/zeitzoff_socialmedia_2ndgaza_v2.pdf [<https://perma.cc/C8YA-DEJA>] (noting that “international interest” in the Israeli-Palestinian conflict led “to multiple, competing news organizations covering the 2012 Gaza Conflict”).

125. This is confirmed by the *z*-scores for those two data points (3.01 and 4.11, respectively). Both are outlier values. See generally PETER H. WESTFALL & KEVIN S. S. HENNING, UNDERSTANDING ADVANCED STATISTICAL METHODS 247 (2013) (noting the “rule of thumb” that an observation with a *z*-score greater than +3.0 or less than -3.0 is typically considered an outlier).

126. However, these increases, although noticeable in the data, were not so extreme as to constitute outlier observations. The PLO Wikipedia article view count for

Lotan, who have found that major conflicts, including the 2012 Gaza conflict, draw “significantly higher levels” of activity on the social media platform Twitter.¹²⁷ These two influential outliers caused by increased traffic to the Hamas article are excluded in the remaining sets of results.

3. *Second Set of Results—A Lasting Chilling Effect?*

Consistent with best practices for dealing with outliers,¹²⁸ results including the outlier “Hamas” Wikipedia article data were reported above. A second set of results from the analysis, which excludes the outlier data concerning the Hamas article, is presented here and set out in Table 2 of the Appendix. Removing the outliers led to new findings.¹²⁹ Similar to the first reported results, there was an immediate and statistically significant decrease in view counts following the June 2013 revelations: an immediate drop of 693,617 total views. Using the 2,893,553 total article views as of May 2013, this decrease represents an immediate drop-off of just under 25%. This suggests that the revelations in June 2013 are associated with a sharp and sudden decrease in traffic consistent with a chilling effect.

Also importantly, is that after June 2013, there is not only a large and immediate drop in views but also a statistically significant change in the overall trend in the month-to-month views of the Wikipedia articles. Rather than increasing on a monthly basis, the trend after June 2013 has completely changed. Due to the statistically significant decrease of 67,513 monthly views, the overall data trend has shifted from an increase of 41,421 views per month to a decrease of 26,092 per month. This is important because it means that the NSA/PRISM surveillance revelations are associated with a longer term, possibly even permanent, decrease in web traffic to the Wikipedia pages studied, consistent with a longer term (and possibly permanent) chilling effect. Figure 3 illustrates this trend.

The shifting trend of the data, which in this case is a sudden and immediate drop, is particularly consistent with a chilling effect arising

November 2012 had a z-score of 3.0 while for July 2014, it was 2.61. Neither are outliers. For explanation of the usual “rule of thumb” for z-scores and outliers, see *id.* at 247.

127. Thomas Zeitzoff, John Kelly & Gilad Lotan, *Using Social Media to Measure Foreign Policy Dynamics: An Empirical Analysis of the Iranian–Israel Confrontation (2012–13)*, 52 J. PEACE RES. 368, 372 (2015) (among other things, focusing on social media data obtained from Twitter to track foreign policy discussions across languages online).

128. Aguinis et al., *supra* note 121.

129. The second set of results were tested for autocorrelation as recommended, see Lagarde, *supra* note 83, at 79. The Durbin-Watson test statistic showed some possibility of autocorrelation at lag 1, both the Cumby-Huizinga and Breusch-Godfrey tests showed no evidence of autocorrelation across through lag 1-10 (and thus did not reject the null hypothesis of no autocorrelation).

from the June 2013 revelations. If the outlier data relating to Hamas view counts is excluded, the decline in page views is less sudden (e.g., 25% immediate drop-off if the Hamas data are excluded compared to the 30% drop-off if the Hamas data remains in the study). However, regardless of whether the Hamas data is included, there is still a substantial and statistically significant decrease. Moreover, there is a change in the overall trend in the data. Before June 2013, total views of the Wikipedia articles in the dataset slowly increase each month. After June 2013, however, with the widespread “exogenous shock” of publicity surrounding the NSA/PRISM revelations, there is a change in the “slope,” or data trend. Without the outlier “Hamas” view counts in July 2014, the total views are on a downward path.

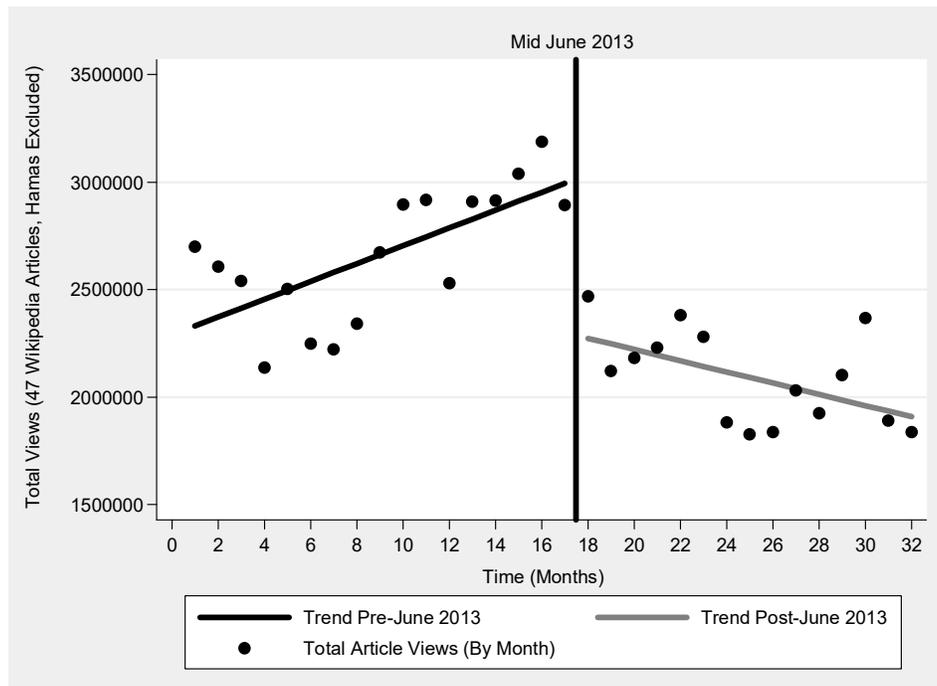


Figure 3. Pre and Post June 2013 Article View Trends (Outliers Excluded). The sudden drop in views and trend shift—from increasing monthly views over time to decreasing after June 2013—is consistent with a significant and long-term chilling effect.

It may be suggested that the reduction in views in June 2013 and then a monthly decrease in traffic to these terrorism-related Wikipedia articles may simply reflect overall Wikipedia article view traffic trends. This is incorrect. Results for an identical ITS for all English Wikipedia article views (across all platforms) for the same thirty-two month period can be found at Table 3 of the Appendix, and show three distinct differences. First, there is no statistically significant shift in the overall article traffic

trend after June 2013. By contrast, there was such a shift in this set of results. Second, the overall article view traffic for all English Wikipedia articles continues to increase month to month after June 2013 (by approximately 114 million views per month). Here, there is a statistically significant monthly decline in article views until the end of the thirty-two month period. Third, while there is a statistically significant reduction in article views over June 2013, the drop off is significantly less (only 15% if you go by actual views in May 2013). So, even assuming that a full 15% of the drop off for the forty-seven terrorism-related Wikipedia articles is simply reflective of background Wikipedia trends, 10% of the reduction in article views over June 2013 remains. This, to be clear, is twice the noteworthy and statistically significant 5% reduction in Google searches for privacy-sensitive terms Marthews and Tucker found after June 2013.¹³⁰ All of these observations suggest these findings reflect far more than mere background Wikipedia trends.

In sum, data visualization shows empirical evidence that is consistent with a long-term chilling effect due to the surveillance revelations, which is not only associated with an immediate drop in views but also a long-term chill on accessing these Wikipedia articles, as users accessed information on these topics less and less frequently.

4. *Final Results with Comparator Groups*

This Section presents a final set of results to strengthen the inference that the reduction in article views after the June 2013 revelations are a result of Wikipedia users' surveillance-related privacy concerns. Two main steps were taken to strengthen this inference in these final results. First, to focus on Wikipedia article content most likely to raise privacy concerns for users, only the thirty-one articles with the highest combined privacy ratings (from the MTurk privacy evaluation) were included.¹³¹ Second, a comparator/quasi-control¹³² group of Wikipedia articles was added to the analysis. In a classic controlled experiment, a control group is randomly

130. Marthews & Tucker, *supra* note 25, at 3.

131. The articles included can be viewed in Table 7 of the Appendix. Of the original 47 articles (48 minus Hamas), the median of the combined privacy score was 2. In this final set of findings any article with a privacy score less than 2 was removed to focus on those articles that should raise the most serious privacy concerns. This left 31 total articles, though the Wikipedia article on "Ammonium nitrate" was ultimately excluded, as it included views in a month with an extreme z-score (4.91) that was skewing regression results. This left a total of 30 articles.

132. The comparator/control group is considered "quasi" as there was no opportunity to isolate a group of Wikipedia users before the June 2013 revelations (because online surveillance potentially affects everyone).

selected from the same population or sampling frame as the experimental group.¹³³ The design's logic is that if you draw from an identical or very similar sampling pool, then the only significant difference between the testing and the control group is that the latter is not exposed to the intervention or treatment; therefore, if the "treatment group" is impacted while the control group is not, the inference that the treatment or intervention caused any observed impact is strengthened.¹³⁴ Though true experiments are rarely found outside laboratories because they require highly controlled settings,¹³⁵ employing quasi-experimental features in research designs like ITS helps strengthen findings and results.¹³⁶ As with experimental designs, a control group employed in an ITS design is ideally identical or very similar to the "experimental" group—if possible drawn from the same population—but would not "experience" the intervention.¹³⁷ Here, results for both the "testing" group—the terrorism-related Wikipedia articles—and comparator groups are compared to better understand the impact associated with the June 2013 revelations.¹³⁸

To create comparator groups as similar as possible to the terrorism-related Wikipedia articles in this study, two groupings of security-related

133. See NOREEN L. CHANNELS, *SOCIAL SCIENCE METHODS IN THE LEGAL PROCESS* 58–60 (1985) (introducing experimental design).

134. *Id.* at 58–60 (describing the classic experimental design and procedure); see also CAMPBELL, STANLEY & GAGE, *supra* note 27, at 13–34 (providing an extensive discussion of different forms of experimental design, and how such designs guard against threats to internal and external validity); DAVID DE VALUS, *RESEARCH DESIGN IN SOCIAL RESEARCH* 53–55, 58 (2003) (introducing classic experimental design and procedures, and also explaining how use of control groups in experiment design help control for unknown factors).

135. See CHANNELS, *supra* note 133, at 61 (noting the difficulty of doing experiments "outside the laboratory"); MATTHEW DAVID & CAROLE D. SUTTON, *SOCIAL RESEARCH: AN INTRODUCTION* 206 (2d ed., 2011) (discussing some of the challenges with using experimental designs in the "social world").

136. See Penfold & Zhang, *supra* note 27, at S43 (noting that while single group ITS designs (with segmented regression) can still be carried out in the absence of a proper control group, the "strength of inference is weaker in the absence of the counterfactual outcome"); Wagner et al., *supra* note 27, at 306–07 (defending single group ITS designs as robust, but also noting and discussing many benefits of employing a control group in an ITS design and analysis).

137. Wagner et al., *supra* note 27, at 306 ("Ideally, a control group that is identical to the study group but does not experience the intervention is followed over the same time period as the intervention group. Comparing the effect in the intervention group with that in the control group then allows separating the intervention effect from others that may have occurred at the same time.").

138. Penfold & Zhang, *supra* note 27, at S40–S41 (noting the importance of comparison between the experimental and control groups); Wagner et al., *supra* note 27, at 306 (noting that the experimental and control groups are compared).

Wikipedia articles was created: one, using the “DHS & Other Agencies” (or domestic security) keyword category from the Privacy Impact Assessment, the same 2013 DHS document used to identify the terrorism related articles,¹³⁹ and a second using the “Infrastructure Security” keyword category from the same document.¹⁴⁰ The logic of this design choice is

139. PRIVACY IMPACT ASSESSMENT, *supra* note 89, at 24. Locating Wikipedia articles coinciding with each keyword was again done manually and similarly was very simple as there was a Wikipedia article that corresponded perfectly with the vast majority of keywords in the “DHS & Other Agencies” keyword category. These were the few discrepancies: the Wikipedia article “Bureau of Land Management” was used for the keyword “National Operations Center (NOC)” (there is no article for the “National Operations Center (NOC)”, but the NOC is located at the Bureau of Land Management, <http://www.blm.gov/noc/st/en.html> [<https://perma.cc/E2VY-3FXB>]); also the National Security Operations Center (NSOC) is located at the NSA and may bias the group post-June 2013; the Wikipedia article “Espionage” was used for the DHS keyword “agent” (being the Wikipedia article for “spy agent”); the Wikipedia article “Task Force 88” was used for the keyword “Task Force” as it is an anti-terrorism task force more consonant with the other security-related keywords in this category; there were a vast range of different articles on emergency/disaster relief organizations the “Red Cross” keyword could refer to, so the Wikipedia article for disaster relief/emergency management was used); two keywords in the DHS document refers to agencies incorrectly; it refers to the “Drug Enforcement Agency (DEA)”, which is the name of an agency in Liberia; this article was excluded but views for the Wikipedia acronym page for “DEA” were included (which *does* refer to a U.S. agency); another keyword refers to “Alcohol Tobacco and Firearms (ATF)”, an agency that was previously split into two separate agencies, the Alcohol and Tobacco Tax and Trade Bureau and the Bureau of Alcohol, Tobacco, Firearms, and Explosives; the Wikipedia article for the former agency was included but not the latter; this is because the reference to “Explosives” in the name of the agency may raise privacy concerns for Internet users therefore biasing the comparator group. Wikipedia articles corresponding with the remaining related keywords were all included in the first set of results involving this domestic security comparator group available in Table 4 of the Appendix. The 25 Wikipedia articles included can be viewed in Table 10 of the Appendix. In the second set of “refined” results, also available in Table 4 and visualized in Figure 4, view counts for the articles “United Nations” and “Federal Bureau of Investigation” were excluded as they included views in certain months constituting significant outliers. Once these extreme outlier articles were excluded, the model achieved good predictive value ($Prob > F = 0.00$, $adj. R^2 = 0.5437$). A similar analysis was undertaken for the terrorism-related articles group with the Wikipedia article on “Ammonium nitrate” excluded as an outlier as well. See *supra* note 143.

140. PRIVACY IMPACT ASSESSMENT, *supra* note 89, at 24. Again, locating articles was simple as there was a Wikipedia article that corresponded naturally with the vast majority of keywords in the “Infrastructure Security” keyword category. These were the few discrepancies: the Wikipedia article “Chemical burn” was used for the keyword “Chemical fire” (there is no “Chemical fire” article); the article “Information infrastructure” was used for the keyword “Computer infrastructure” (there is no “Computer infrastructure” article); the article “Telecommunications network” was used for the keyword “Communications infrastructure” (there is no “Communications infrastructure” article); the article “National Information Infrastructure” was used for the keyword “National infrastructure” (there is no “National infrastructure” article); the

primarily a form of normative matching: while using an identical set of Wikipedia articles for a comparator/control group is impossible, terrorism and domestic/infrastructure security-related articles comprise closely related content and likely attract similar readers and Wikipedia users. The three groups of Wikipedia articles could also be created by matching articles with keywords from the same DHS document, so the articles are being drawn from the same “source,” which renders the groups more similar and closely related for comparative purposes. However, in theory, viewing terrorism-related Wikipedia content is far more likely to raise privacy concerns for Wikipedia users concerned about government surveillance than merely viewing information about domestic security agencies like “Department of Homeland Security” or “Fusion Centers” or articles about infrastructure like “Bridge,” “Port authority,” or “Information infrastructure.” As such, article traffic concerning the domestic security or infrastructure-related Wikipedia articles in the comparator groups should not lead to a June 2013 related chilling effect. The hypothesis, based on chilling effects theory, is that users viewing terrorism-related Wikipedia articles should, by contrast, be more chilled by the surveillance revelations.¹⁴¹ An additional comparator group—Popular Wikipedia pages¹⁴²—is also analyzed and visualized separately for

article “Electrical grid” was used for the keyword “Grid”; the article “Electric power” was used for the keyword “Electric” (there is no article for “Electric”); the article “Power outage” was used for the keyword “outage”; the article “Flight cancellation and delays” was used for the keyword “Cancelled”; there were no existing Wikipedia articles for “CIKR” (Critical Infrastructure & Key Resources) or “NBIC” (National Biosurveillance Integration Center), but view counts for articles for each acronym (NBIC/CIKR) were nevertheless included; “Electric power transmission” was used for the keyword “Power lines” (there is no “Power lines” article); and there was no article corresponding “Transportation security” other than the “Transportation Security Administration,” which was already included in the domestic security-related comparator, so it was not included. There was also no article corresponding with “Service disruption.” All thirty-four Wikipedia articles included can be viewed in Table 11 of the Appendix.

141. This is another reason why the keywords under the “DHS & Other Agencies” and “Infrastructure” categories in the DHS document were used to select comparator groups, as many keywords in the other categories beyond “Terrorism” (e.g., “Cocaine,” “Meth Lab,” “HAZMAT and Nuclear,” “nerve agent,” etc.) concern topics or content that may very well raise privacy concerns for users aware of government surveillance online. As such, these terms would not be appropriate as quasi-controls.

142. Following Marthews and Tucker’s use of “popular” Google Search terms as quasi-controls in their study, this study, for comparative purposes, examined the Wikipedia article traffic for the top ten most popular English Wikipedia articles (in terms of article views) for each of years 2012, 2013, and 2014 (the years included in the thirty-two month study period), according to the Wikimedia Tool Lab’s “Wikitreng” reports. *Trends on Wikipedia*, WIKITRENDS, <https://tools.wmflabs.org/wikitrends> [<https://perma.cc/VXF8-4GQC>]; see also Marthews & Tucker, *supra* note 25, at 7–8. This led to a set

illustrative purposes, with a similar chilling effect hypothesis (unlike privacy concerning terrorism-related content, the June 2013 revelations should have no effect on these pages).

For the 30 terrorism-related article group, results available in Table 4 of the Appendix, the reduction in views and reversal in trend are, once again, consistent with a significant and potentially long-term chilling effect. The immediate drop-off (-225,867) was large, and the trend change from increasing views monthly (26,129) to fewer (-38,160) was statistically significant at the 99% confidence level. Using the predicted views as of May 2013 (854,755), the 225,867 reduction in June represents a sudden drop off of approximately 26%.

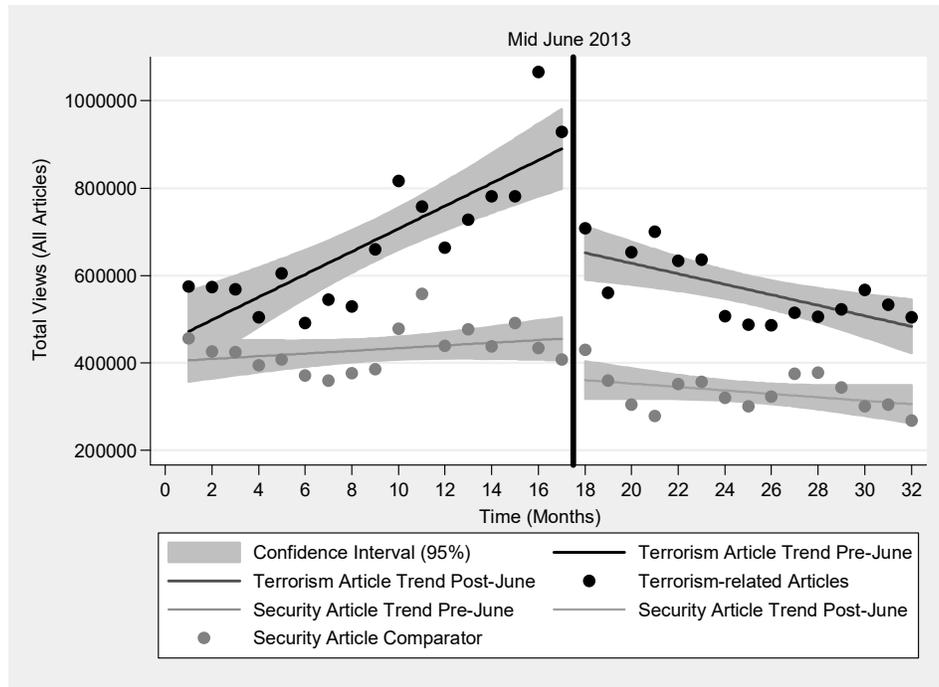
By contrast, article views for the full 25 domestic security-related Wikipedia articles comparator group (results also in Table 4) show little impact, or evidence of chill, associated with the June 2013 revelations. The results show no substantial nor statistically significant reduction in views in June or change in trends, with very high p values (0.531 and 0.551). The small and statistically insignificant reduction of 24,638 in June is dwarfed by the 225,867 drop for the terrorism related views. In fact, the regression model for these results was not statistically significant ($Prob > F = 0.4470$) and thus no predictive value ($Adj. R^2 = -0.0084$) indicating that an analysis based around the June 2013 revelations simply does not fit the actual data on article views for these domestic security-related articles, nor tell us much about it. This is unlike the terrorism-related article group, where the model was highly statistically significant ($Prob > F = 0.0000$) with strong predictive value ($Adj. R^2 = 0.6789$). This is consistent with a chilling effects hypothesis where terrorism-related articles are impacted but security-related articles, that do not raise privacy concerns, are not.

This inference is strengthened through second refined set of results for the domestic-security related articles that addresses outlier¹⁴³ (available in Table 4 and visualized in Figure 4), still suggests no significant immediate or long term impact due to the June 2013 revelations. There is no statistically significant increasing or decreasing monthly view trends for the domestic-security related articles. And while there is a drop in views in June (90,921), it is not statistically significant and far smaller than the 26%

of twenty-six Wikipedia articles comparator group, including articles like “Google,” “Facebook,” “Breaking Bad,” “Game of Thrones,” and “World War II.” Certain Wikipedia articles like “Facebook” and “Google” were in the top ten most popular articles for more than one year, hence twenty-six articles instead of thirty. See Table 12 in the Appendix for all Wikipedia articles in this comparator group.

143. See *supra* note 139.

A. Terrorism Articles Study Group vs. Domestic Security Comparator Group



B. Infrastructure-related Comparator Group

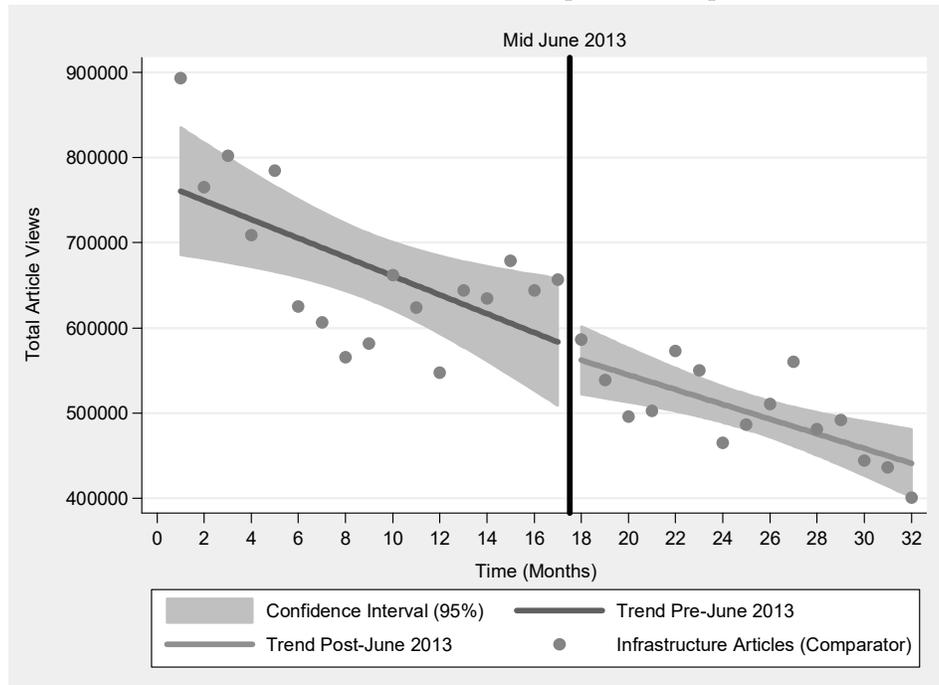


Figure 4: (A) The sudden drop and trend reversal for the terrorism-related articles is consistent with chilling effects, while domestic security articles show little impact. (B) The infrastructure-related article comparator also shows no June 2013 impact.

drop-off for the terrorism-related Wikipedia articles (-225,867). Using predicted article views as of May 2013, this drop represents a reduction of a little over 19%. The common drop in June (even if the magnitude is far different) likely reflects, in part, a smaller overall reduction in English Wikipedia traffic in June 2013 (discussed earlier and apparent in results set out in Table 3 of the Appendix).

In fact, the 15% overall Wikipedia traffic dip in June can almost entirely explain the drop for the security articles (with 4% remaining). The terrorism-related articles, consistent with a chilling effect hypothesis, dropped off an additional 11% in June 2013. Finally, also unlike the terrorism article group—which experienced a statistically significant negative trend change (38,160 fewer views a month)—the comparator group experienced no significant trend change after June 2013.

Figure 4, which compares results for the terrorism-related articles and the refined domestic security-related Wikipedia articles comparator group, suggests no substantial impact for the security articles after the June 2013 revelations—the article view trends suggest a slight dip and then no substantial change in monthly views, and certainly no significant shift in overall trend comparable to the terrorism-related article group. This, as noted, suggests the surveillance revelations had no immediate or significant lasting impact on view trends. This stands in contrast to the terrorism-related articles in the study that noticeably trend downward post-June 2013.

All of these inferences are even further strengthened when autocorrelation is corrected using the Prais–Winsten method (results also in Table 4), with the drop in June 2013 for the terrorism-related articles still remaining large and statistically significant (-219,625) (representing a 25% drop in views) and the overall shift in monthly view trends remains substantial and statistically significant (-37,282). By contrast, the reduction in June for the security-related articles in these results shrink to 28,516—a 7% drop that can be entirely explained by overall Wikipedia view trends in June—and is now no longer statistically significant. The change in monthly view trends also remains insignificant ($p = 0.200$). These results suggest the different impact that June 2013 had on the terrorism and security-related articles is even more apparent with a more robust regression model, further supporting a chilling effects hypothesis.

Similarly, the infrastructure related articles results (available in Table 5 and also visualized in Figure 4 at (B)) indicate no impact from the June 2013 revelations. Unlike the terrorism-related articles, there is no statistically significant drop off in June, only a very small drop of 12,721 views over the month. Using predicted article views as of May 2013

(583,415) that is merely a 2% reduction. Also unlike the terrorism-related articles group, there is no statistically significant change in the overall view trends after June 2013. Before June 2013, there is a decline that continues through that month onwards until the end of the study period.

The popular Wikipedia articles comparator, results in Table 6 and visualized in Figure 5 below, show little negative impact associated with June 2013. Indeed, the model for these results was also not significant, meaning that a model centered on June 2013 does not “fit” data and tells us little about actual changes or trends therein. There was drop in views in June, but it was not statistically significant. And while views increased after June, this was not statistically significant either. Ignoring, for the moment, the June 2013 interrupting line and trend lines in Figure 5, the data as plotted actually suggests views for these articles are fairly constant across the 32 month period and June 2013 plays little to no role in changes or trends. These findings are all consistent with a chilling effects hypothesis. Just as security or infrastructure-related articles would be unaffected by the June 2013 revelations, views for “popular” Wikipedia articles like “Google,” Breaking Bad,” or “2014 FIFA World Cup” would likewise not be negatively impacted as such content is unlikely to raise privacy concerns.

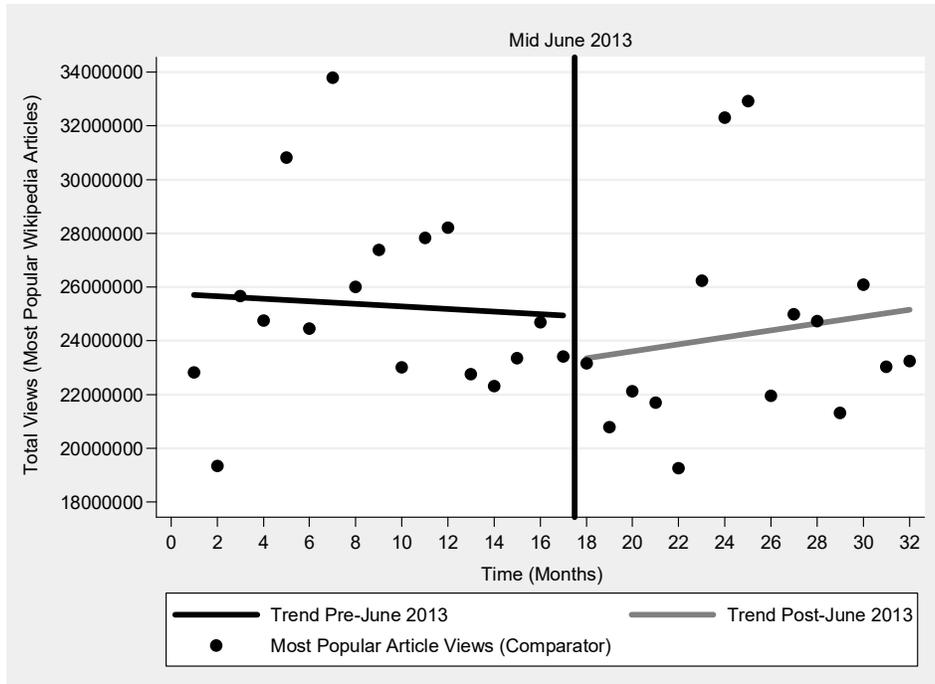


Figure 5: Much like the security and infrastructure comparators, the popular Wikipedia article comparator group also shows little impact from the June 2013 revelations.

All of these findings support a chilling effects theory: that surveillance-related chill caused the sudden drop during and after June 2013, as well as the general trend reversal, for the terrorism-related Wikipedia articles. This hypothesis is supported by the fact that there is no indication in the findings that the security, infrastructure, or popular Wikipedia articles comparator groups were likewise impacted by the June 2013 revelations. The explanation, it may be surmised, is the domestic security and infrastructure related Wikipedia articles, though similar, are simply unlikely to raise privacy concerns for Wikipedia users worried about online surveillance. The same can be said for the “popular” articles. Thus, article views before and after June 2013 show few noteworthy effects.

V. IMPLICATIONS

A. EMPIRICAL EVIDENCE FOR REGULATORY CHILLING EFFECTS

Skepticism among courts, legal scholars, and empirical researchers has persisted about the nature, extent, and even existence of chilling effects due, in large part, to a lack of empirical substantiation.¹⁴⁴ The results in this case study, however, provide empirical evidence consistent with chilling effects on the activities of Internet users due to government surveillance. And, to be clear, the activity here is not only legal—accessing information on Wikipedia—but arguably desirable for a healthy democratic society. It involves Internet users informing themselves about important topics subject to today’s widespread social, political, moral, and public policy debates.¹⁴⁵ The large, statistically significant, and immediate drop in total views for the Wikipedia articles after June 2013, implies a clear and immediate chilling effect. Moreover, the broad and statistically significant shift in the overall trend in the data (e.g., the shift from the second results excluding outliers) suggests any chilling effects observed may be substantial and long-term. This study is among the first to provide evidence of such a chilling effect using web traffic data (instead of survey responses or search), and it is the first to do so in relation to the potential chilling effects on Wikipedia use, thereby demonstrating how government

144. See Kaminski & Witnov, *supra* note 9, at 517 (calling for further research on the “types of surveillance and surveillance cues that cause chilling effects,” as well further research on both the magnitude and persistence of such surveillance related chilling effects); see also Kendrick, *supra* note 14, at 1657; Nickel, *supra* note 18, at 263; Richards, *supra* note 23, at 1964.

145. Clark McCauley, *Terrorism, Research and Public Policy: An Overview*, 3 *TERRORISM & POL. VIOLENCE (SPECIAL ISSUE: TERRORISM RES. & PUB. POL’Y)* 126, 134 (1991) (“Taken together, the financial, social, political, and moral costs of response to terrorism constitute a challenge to the democratic capacity to govern.”).

surveillance potentially affects the way people access and distribute information online.

These results are consistent with chilling effects theory, but arguably contradict other research concerning online privacy behaviors. First, the substantial body of “privacy paradox” research, involving a diverse range of online platforms and contexts, has demonstrated that Internet users’ stated concerns about privacy are often not reflected in their online behavior.¹⁴⁶ There have been a range of explanations for this disconnect,¹⁴⁷ but one common factor offered by behavioral economists is that online users suffer from “incomplete information” and “bounded rationality” in making decisions about privacy, that is, such decisions are often complex and people are limited both by cognitive ability and knowledge.¹⁴⁸ However, the results here, consistent with the chilling effects hypothesis that users are avoiding certain online content due to privacy concerns about surveillance, suggest that users are acting both rationally and logically even with incomplete information about the true nature and scope of covert NSA surveillance practices. In other words, contrary to the “privacy paradox,” privacy concerns are being reflected in online behavior.

Second, as noted earlier in this Article, privacy researchers and legal scholars have expressed skepticism about the possibility of large scale or long-term chilling effects caused by online surveillance due either to a general “desensitization” of privacy concerns in online contexts¹⁴⁹ or due to the fact that online users adapt quickly to shifting norms, rendering chilling effects “temporary.”¹⁵⁰ On this count, research has found, a “lax attitude” among users toward the benefits of online privacy (compared, for example, to the benefits of information disclosure);¹⁵¹ muted user

146. See Annika Bergström, *Online Privacy Concerns: A Broad Approach to Understanding the Concerns of Different Groups for Different Uses*, 53 COMPUTERS IN HUMAN BEHAVIOR 419 (2015), <http://www.sciencedirect.com/science/article/pii/S0747563215300364> [<https://perma.cc/7U2Z-R8BC>]. See generally Kokolakis, *supra* note 42 (providing a comprehensive explanation and review of “information privacy paradox” literature).

147. See, e.g., Kokolakis, *supra* note 42, at 7–9 (reviewing the various interpretations and explanations for paradox).

148. *Id.* at 9.

149. Nickel, *supra* note 18, at 263.

150. See Bernescu, *supra* note 18, at 671 (“However, because consumers in the Internet context quickly adapt to changing norms, any such chilling effect will likely be temporary.”).

151. Debatin, *supra* note 18, at 83, 100–02 (finding that a majority of Facebook users in their study disclosed a great deal of personal information despite being aware of privacy risks; they attribute this to a “lax attitude”); see also Kokolakis, *supra* note 42, at 7

responses to reputational or privacy risks associated with embarrassing behavior being exposed online;¹⁵² and that while negative privacy experiences online prompted users to adjust their sharing practices, their “social or psychological privacy behaviors” online were unaffected.¹⁵³ If there even *were* any privacy related chilling effects, such research suggests they may muted, ephemeral, or short-term. The findings here also contradict these studies, suggesting not only an immediate chilling effect associated with the June 2013 surveillance revelations, but a possible longer term chill as well. A determination of whether this trend will continue further into the future is inherently limited by the data set in this study. Nonetheless, the overall downward trend in the data for the months studied does provide evidence of a more permanent impact.

This case study also provides important insights on how to understand chilling effects, particularly how they operate online. Though Schauer and Solove’s accounts of chilling effects are closely related, there are important distinctions between the two. Schauer approached chilling effects as mainly resulting from uncertainty in the legal system (e.g., vagueness of legislative enactments) and people’s fear of prosecution and legal sanction;¹⁵⁴ Solove, on the other hand, broadened the theoretical outlook by focusing his efforts on surveillance and “executive information gathering.”¹⁵⁵

(discussing research on the “privacy calculus” where people weigh the benefits of privacy over disclosure).

152. Hermstrüwer & Dickert, *supra* note 68, at 22–23 (an experimental study on chilling effects finding that risks of “networked publicity” (exposure online of users’ embarrassing activities) did not affect users’ “privacy valuations,” nor did they “dampen” either “behavioral idiosyncrasies” or the “panoply of different behaviors” involved in the study).

153. Sabine Trepte, Tobias Dienlin & Leonard Reinecke, *Risky Behaviors: How Online Experiences Influence Privacy Behavior*, in VON DER GUTENBERG-GALAXIS ZUR GOOGLE-GALAXIS [FROM THE GUTENBERG GALAXY TO THE GOOGLE GALAXY] 225, 240 (B. Stark, O. Quiring & N. Jakob eds., 2014) (“After encountering harassing or humiliating status posts or messages, users adjusted the information they posted online. However, negative experiences did not affect social or psychological privacy behaviors. It was shown that the ways that users managed their audiences and friends (social privacy) and the kinds of information they shared (psychological privacy) remained unaffected by negative experiences.”).

154. *See* Schauer, *supra* note 12, at 693–95 (discussing, among other things, how fear, risk, and uncertainty in the legal process as contributing to potential chilling effects).

155. *See* Solove, *The First Amendment as Criminal Procedure*, *supra* note 16 (exploring and analyzing cases wherein government information gathering implicates the First Amendment and related “chilling effects”); *see also* Solove, *A Taxonomy of Privacy*, *supra* note 33, at 487–89 (identifying surveillance related chilling effects as a “more modern privacy problem” that “does not fit” with more traditional conceptions of privacy harms).

The findings here are consistent with Solove's approach to chilling effects. On his account, people censor themselves and avoid certain activities not necessarily out of fear of prosecution but out of concern for potential future harms due to privacy violations, embarrassing public disclosures, risks of fraud or identity theft, or being labeled a criminal, deviant, or non-conformist by state authorities.¹⁵⁶ Given the lack of evidence of people being prosecuted or punished for accessing information on Wikipedia or similar sites before, during, or after the June 2013 revelations, it is unlikely that actual fear of prosecution can fully explain the chilling effects suggested by the findings of this study. Rather, Solove's notion of surveillance-related "pollution"—the broader societal context of self-censorship and conformity arising from ubiquitous and large-scale surveillance—may be the better explanation.¹⁵⁷ This inference is supported by the independent privacy evaluation completed by the 415 respondents recruited through MTurk. Respondents' assessments of the forty-eight keyword topics indicated that if they knew the government was monitoring online activities, they would be more likely to avoid the topics in question. In other words, their responses suggested a *potential* for chilling effects relating not to fear of prosecution but the risks, harms, and threats associated with government surveillance. This point should not be taken too far, however, as the independent evaluators were recruited among MTurk users and not Wikipedia users tracked by the article traffic data in this study. Therefore, the findings may not hold for Wikipedia users or Internet users more generally.

B. THE IMPACT OF WAR AND OTHER EXOGENOUS EVENTS

Another important insight from the study is how intervening dramatic external pressures or "exogenous shocks" impact chilling effects. Notwithstanding the evidence of immediate, substantial, and potentially long-term chilling effects due to awareness of government surveillance, those chilling effects can be affected or impacted by other dramatic intervening events or "exogenous shocks" like war. For instance, the impact of the November 2012 and July 2014 Israeli-Hamas conflicts on Wikipedia use can be interpreted in two ways within the chilling effects

156. See Solove, *A Taxonomy of Privacy*, *supra* note 33, at 493–99 (reviewing a range of instances where surveillance and related information gathering activities can lead to chill, self-censorship, inhibition, and other forms of privacy harms).

157. *Id.* at 487–88 (discussing how the broader sets of risks caused by government practices like surveillance, which can be likened to "environmental harm" or "pollution," also should be understood as having broader societal impact beyond any "mental pain and distress" caused to individuals).

framework. First, it could be argued that these instances of high-profile armed conflict “ameliorated” the chilling effects caused by publicity and public awareness of the NSA/PRISM surveillance, since for at least for one of the Wikipedia articles among the forty-eight retrospectively analyzed in the case study, view counts dramatically increased as those conflicts unfolded. People’s desire to learn about the Israeli-Hamas conflicts simply overrode any concerns they may have had about the government monitoring information they were accessing online, thus reducing chilling effects.

A second way of interpreting the results, is that the Israeli-Hamas conflicts essentially masked a broader chilling effect (apparent once the “Hamas” article view data was excluded), by bringing new populations of atypical users to Wikipedia who were less aware of, or perhaps concerned by, government surveillance. This would seem to support the findings of Zeitzoff, Kelly, and Lotan, who have explored how major conflicts resulted in “significantly higher levels” of social media activity. They noted the emergence of what they called “ephemeral” users who seemed to only “tweet” about the 2012 Gaza conflict and nothing else.¹⁵⁸ Perhaps those visiting the Hamas and Palestinian Liberation Organization articles in November 2012 and July 2014 were atypical Wikipedia users and less influenced by the specter of government surveillance, but there is nothing in this study to directly support this assertion. There is probably some truth to both of these explanations. Additional research focusing on this question—how exogenous shocks such as war potentially affect chilling effects—could offer important insights on this dimension of the results.

C. CONSTITUTIONAL LITIGATION

The results in this case study should, first of all, provide empirical support for the chilling effects doctrine in First Amendment law. Skepticism about the chilling effects doctrine dates back decades, and, more recently, scholars have concluded more research is required to support the “unsubstantiated empirical judgments” of chilling effects claims under the First Amendment (and chilling effects more generally).¹⁵⁹ The results presented here meet this call. Second, evidence of chilling effects in this case study may have important implications for a wide array of ongoing constitutional litigation brought in relation to government surveillance practices. Indeed, a significant challenge for recent lawsuits filed against the NSA and the U.S. government, especially those based on

158. Zeitzoff et al., *supra* note 127, at 5; Zeitzoff, *supra* note 124, at 13.

159. Kendrick, *supra* note 14, at 1657.

a chilling effects theory like the Wikimedia Foundation's complaint, is the issue of standing. A key part of this challenge is the nature of government surveillance online—it is covert and secretive, so victims are most often unaware if they have been personally targeted.¹⁶⁰ This is compounded by how the Supreme Court has applied standing in cases involving covert surveillance. The findings of this study may help plaintiffs overcome challenges to standing by providing empirical evidence to ground constitutional claims based on chilling effects and related harms in objective evidence not subjective claims and fears.

The Court's recent decision in *Clapper* reflected existing legal and judicial skepticism concerning chilling effects. In that decision, a five Justice majority dismissed as "too speculative" the plaintiffs' assertion of standing based on a likelihood that their activities would be subject to surveillance in the future.¹⁶¹ The Court similarly dismissed the additional arguments for standing based on "chilling effects," observing that while prior cases found constitutional violations may arise from chilling effects, such violations could not arise "merely" from a person's "knowledge" or "concomitant fear" about government activities.¹⁶² Relying on its 1973 decision in *Laird*, the Court noted that "[a]llegations of a subjective 'chill' are not an adequate substitute for a claim of specific present objective harm or a threat of specific future harm."¹⁶³ Such surveillance related chilling effect claims based on subjective fears were "self-inflicted" injuries, the Court concluded, and thus could not provide standing for the constitutional claims.¹⁶⁴

Clapper is unlikely the final word on standing based on widespread government surveillance. To begin with, the case was decided in February of 2013, several months before the Snowden disclosures and the widespread publicity concerning the PRISM and other government

160. See Richards, *supra* note 23, at 1934 ("Although we have laws that protect us against government surveillance, secret government programs cannot be challenged until they are discovered."). See generally Slobogin, *supra* note 39 (analyzing legal standing issues in relation to constitutional challenges to NSA and other surveillance related practices).

161. *Clapper v. Amnesty Int'l*, 133 S. Ct. 1138, 1143 (2013) ("Respondents assert that they can establish injury in fact because there is an objectively reasonable likelihood that their communications will be acquired under § 1881a at some point in the future. But respondents' theory of future injury is too speculative to satisfy the well-established requirement that threatened injury must be 'certainly impending.'").

162. *Clapper*, 133 S. Ct. at 1152.

163. *Id.* (quoting *Laird v. Tatum*, 401 U.S. 1, 13–14 (1972)).

164. *Id.* at 1152–53 (quoting *Laird v. Tatum*, 401 U.S. 1, 13–14 (1972)).

surveillance programs.¹⁶⁵ Moreover, commentators like Neil Richards, Luke Milligan, and Christopher Slobogin, among others, have offered persuasive criticisms of the Supreme Court's approach to standing in *Clapper*.¹⁶⁶ Richards argues that the *Clapper* approach to standing affirms a "brutal paradox" whereby litigants must prove harms (like chilling effects) arising from secretive covert surveillance but the only party that knows—the government—is not telling.¹⁶⁷ Milligan, on the other hand, offers a compelling argument that the *Clapper* approach to standing and chilling effects claims is inconsistent with the text and history of the Fourth Amendment, which was originally understood to guarantee freedom not just from individual unreasonable searches but also freedom from "fear" of such searches.¹⁶⁸ Lastly, Slobogin offers a strong criticism of *Clapper* based on political process theory and the separation of powers, arguing that chilling effects caused by covert surveillance undermine the political process, and as a result citizens should have standing to challenge such surveillance in court.¹⁶⁹

165. See Slobogin, *supra* note 39, at 520 (noting that "[t]hanks to Edward Snowden, the U.S. federal government has been forced to acknowledge certain surveillance practices, while journalists have shed important additional light).

166. See Luke M. Milligan, *The Forgotten Right to be Secure*, 65 HASTINGS L.J. 713, 732–50 (2014) (arguing for a broader approach to standing than recognized in *Clapper*—that would allow earlier Fourth Amendment challenges to concealed government investigative techniques—based on the Fourth Amendment's original understanding as defined by its text, history, and structure); Richards, *supra* note 23, at 1963–64 (arguing, based on the notion of "intellectual privacy" and its importance to the democratic principle of "self-government," that surveillance privacy harms should be recognized under legal standing doctrines and that *Clapper* fails to do so); Slobogin, *supra* note 39, at 535–41; see also Lexi Rubow, *Standing in the Way of Privacy Protections: The Argument for a Relaxed Article III Standing Requirement for Constitutional and Statutory Cause of Action*, 29 BERKELEY TECH. L.J. 1007 (2014) (analyzing and critiquing current standing doctrine in light of the difficulties of proving privacy harms).

167. See Richards, *supra* note 23, at 1944–45.

168. See Milligan, *supra* note 160, at 750 ("On the basis of both text and history, the Fourth Amendment right "to be secure" can be fairly read to encompass the right to be 'protected' from unreasonable searches and seizures, and quite possibly the right to be 'free from fear' of such government actions. This broader interpretation of 'to be secure' has important implications for prevailing Fourth Amendment rules and procedure (arguing that the text and history of the Fourth Amendment supports standing for technological chilling effects claims).").

169. See Slobogin, *supra* note 39, at 535–41 (drawing on both Richards and Milligan to argue, among other things, that *Clapper* and its standing requirements undermine the "political process," which is what the standing process was meant to protect).

Still, *Clapper* remains the law and thus presents a difficult standard to show injury and standing for “chilling effects” constitutional claims.¹⁷⁰ On this count, this case study’s empirical findings will have implications for present and future litigation. As noted, *Clapper* emphasized the need for evidence beyond “self-inflicted” injuries based on “subjective fears” about chilling effects to support standing.¹⁷¹ This case study provides empirical support for surveillance-related chilling effects on Wikipedia’s users—not on Wikimedia Foundation itself, though it is impacted by virtue of its users being chilled—meaning any constitutional claims are neither subjective, self-inflicted, nor speculative about future harms. The findings also suggest those chilling effects are not trivial or temporary, but may be significant, sudden, and with a long-term impact. The plaintiffs in *Clapper*, whose claims about harms due to chilling effects were based on costs incurred to avoid government surveillance, argued such harms mostly in an empirical vacuum, which left their claims vulnerable to the *Laird* rule that subjective allegations could not create standing. By contrast, this case study provides empirical support for Wikimedia Foundation to assert harm in its lawsuit against the NSA and Justice Department based on chilling effects claims:

The notion that the N.S.A. is monitoring Wikipedia’s users is not, unfortunately, a stretch of the imagination. The harm to Wikimedia and the hundreds of millions of people who visit our websites is clear: Pervasive surveillance has a chilling effect. It stifles freedom of expression and the free exchange of knowledge that Wikimedia was designed to enable.¹⁷²

The results of this case study suggest that the harm produced by chilling effects is not a “stretch of the imagination” at all. These findings imply the June 2013 surveillance revelations, extensively covered by media, had a salient and observable chilling effect on Wikipedia users accessing certain Wikipedia articles. Additionally, this case study provides a more general empirical foundation for companies, organizations, and other institutions whose users may have been “chilled” by government surveillance to assert constitutional harms.

170. Slobogin, *supra* note 39, at 522 (“As the outcome in *Clapper* illustrates, because NSA surveillance is, by design, covert, the standing requirement that plaintiffs allege a ‘concrete’ injury can pose a serious obstacle to parties trying to challenge it.”).

171. *Clapper v. Amnesty Int’l*, 133 S. Ct. 1138, 1152–53 (2013) (quoting *Laird v. Tatum*, 401 U.S. 1, 13–14 (1972)).

172. *Wales & Tretikov*, *supra* note 3.

D. SURVEILLANCE, WIKIPEDIA, AND DEMOCRATIC SOCIETY

This case study also has implications for the health of democratic deliberation among citizens. Surveillance related chilling effects, in deterring people from exercising their rights and freedoms, have clear implications for individual citizens.¹⁷³ However, these same chilling effects also have implications for the broader health of society, threatening what Richards calls “intellectual privacy”—the freedom to read, think, and communicate privately—an essential predicate to democracy and “self government.”¹⁷⁴ Chilling effects are indeed a force for conformity and therefore corrosive to “political discourse.”¹⁷⁵

This, in particular, is a problem for Wikipedia. Democratic theorists have long pointed to public deliberation as an essential tool to enhance collective understanding and decision-making, and Wikipedia has been found to be an important complement to this democratic process.¹⁷⁶ On this count, Wikipedia provides a collaborative model of knowledge production that strengthens democracy. A study by Nathaniel Klemp and Andrew Forcehimes found that Wikipedia offers enhanced democratic deliberation and collective decision-making through its “model” of citizen engagement and information exchange.¹⁷⁷ And, beyond these important contributions, Wikipedia not only remains incredibly popular online—every month, Wikipedia is visited by nearly half a billion people from almost every country on earth¹⁷⁸—but is an increasingly important resource for Internet users to quickly and efficiently inform themselves about government policies, laws, and actions, thus better equipping them to

173. See Richards, *supra* note 23, at 1950 (noting that protection against chilling effects is necessary to preserve freedom of speech and thought, two important First Amendment values).

174. Richards, *supra* note 23, at 1959, 1963.

175. BRUCE SCHNEIER, DATA AND GOLIATH: THE HIDDEN BATTLES TO CAPTURE YOUR DATA AND CONTROL YOUR WORLD 95–99 (2015). See also RON DEIBERT, BLACK CODE: INSIDE THE BATTLE FOR CYBERSPACE 130–32 (2013); Solove, *A Taxonomy of Privacy*, *supra* note 33, at 494–99.

176. Nathaniel Klemp & Andrew Forcehimes, *From Town-Halls to Wikis: Exploring Wikipedia's Implications for Deliberative Democracy*, 6:2 J. PUB. DELIBERATION 1, 27 (2007) (finding that the Wikipedia model of online interaction offers a “powerful” supplement to traditional face-to-face forms of public deliberation).

177. *Id.* at 31–32.

178. See Michelle Paulson & Geoff Brigham, *Wikimedia v. NSA: Wikimedia Foundation Files Suit Against NSA to Challenge Upstream Mass Surveillance*, WIKIMEDIA BLOG (Mar. 10, 2015), <http://blog.wikimedia.org/2015/03/10/wikimedia-v-nsa/> [https://perma.cc/J44Q-VZJN].

“enter into deliberations over political decisions.”¹⁷⁹ Moreover, Wikipedia provides a collaborative model of knowledge production that strengthens democracy. Klemp and Forcehimes have also found that Wikipedia offers enhanced democratic deliberation and collective decision-making through its “model” of citizen engagement and information exchange.¹⁸⁰

The importance of Wikipedia as a source of online knowledge and information is highlighted by the data in this case that showed people taking to Wikipedia’s “ Hamas” article in dramatic numbers in November 2012 and July 2014 to inform themselves about the Israel-Hamas conflicts. Whether these users were typical or atypical Wikipedia users, the findings imply that Wikipedia was a key source of information gathering about a contentious and globally covered armed conflict. This conclusion is consistent with prior research on how people seek information about breaking news stories on Wikipedia and how such events impact its contributor communities and content.¹⁸¹

But these findings have potentially troubling implications too. In contrast to the war-related outliers represented by the Hamas and Palestinian Liberation Organization articles, the case study suggests government surveillance may have a long-term chilling effect on this type of important Wikipedia use. If people are chilled from informing themselves about breaking news stories and other important news events, or from researching matters of law, security, and public policy related to “terrorism” online, then surveillance-related chilling effects will have serious implications for public deliberation about important topics. With people potentially chilled or deterred from such basic acts of information gathering, people will be less informed and our broader processes of democratic deliberation will be weakened. If “intellectual privacy,” as Richards argues, is essential to democracy and “self government” in guaranteeing the space and freedom to read, think, and communicate privately,¹⁸² so too is the freedom (free from insidious surveillance-related chilling effects) to gather basic information from important platforms and resources like Wikipedia that make engaging in acts of thinking, communicating, and decision-making, meaningful.

179. See Klemp & Forcehimes, *supra* note 139, at 31–32 (“Ideally, such potential applications of the Wikipedia model would enhance existing forms of face-to-face deliberation. The information gathered through such political wikis would help to inform citizens and better equip them to enter into deliberations over political decisions.”).

180. *Id.*

181. See sources cited *supra* note 70–72.

182. Richards, *supra* note 23, at 1950, 1963.

Finally, Wikipedia has proven valuable beyond merely being an online source of knowledge. As noted earlier, researchers have employed Wikipedia and information the Wikimedia Foundation makes available online about articles and editing activities for a broad range of online and offline research interests.¹⁸³ As a popular and highly successful collaborative peer-production online platform, Wikipedia is invaluable as a focal point for research exploring collaborative networks and knowledge production, research that is put in doubt if people are chilled from using the site due to government surveillance.

VI. LIMITATIONS

Notwithstanding the significance of this case study's findings and their attendant implications, they have important limitations. First, the period of the study only extends until August 2014. This means that the persistence of any chilling effects beyond that point remains an open question. Though the findings here suggested a long-term, even permanent, chilling effect, this possibility cannot be confirmed or denied using existing data. Additional research using more recent data could shed some light on this aspect of the study.

Second, a true experimental design, one with a true control group—randomly drawn from the identical subject pool or population—was not possible. Given the secrecy surrounding government surveillance practices and their potential wide scope, the research design could not be strengthened by comparing Wikipedia users affected by surveillance with a true control group that had *not* been exposed to online surveillance; the covert nature of the surveillance rendered it impossible to isolate or identify such a group of individuals. This is one of the challenges of studying chilling effects and the impact of surveillance more generally—much of the practices at issue are secret and thus difficult to study systematically. Still, a comparator groups were included in the final analysis, rendering the final results more robust. Relatedly, it is impossible to know (because the data is simply not available) whether Wikipedia users chilled or deterred from viewing the articles included in the study were in fact viewing the very same content elsewhere. Ideally, this, and other confounding factors and variables that may offer alternative explanations, would be controlled for in an experimental setting, but like all naturalistic or observational studies involving data derived from the field (as here), there are variables and factors that cannot be known or

183. See generally sources cited *supra* note 73.

controlled. The ITS design, however, still provides a robust means to analyze data where, as here, true experimental designs are not possible.

Finally, the nature of the Wikipedia data also limited the ways in which the research design could have been strengthened in this case study. For example, though the study focused on English Wikipedia articles, article view counts used to construct the time series data set did not distinguish the geographic origins of article views. That distinction would have provided some insight as to whether evidence of chilling effects varied across geographical regions. Further research attempting to address or overcome these limitations would be valuable.

VII. CONCLUSION AND FUTURE DIRECTIONS

The evidence of chilling effects illustrated in this case study has important implications on multiple fronts. And just as importantly, the study's ITS design, combined with a segmented regression analysis and a comparator group, provides a powerful research design and analytical method that can be employed by researchers in other contexts to explore chilling effects and related regulatory impacts online.

There are also clear future directions for research. This case study has focused primarily on whether the June 2013 surveillance revelations had a chilling effect on Wikipedia users and whether there were any immediate implications of the findings on that question. But the economic impact of those chilling effects, which could shed important light on monetary harms stemming from surveillance and other regulatory actions, was not explored. As Tucker and Marthews observe in relation to chilling effects on Google search,¹⁸⁴ the findings here suggest that NSA/PRISM programs and other forms of government surveillance may have a substantial impact on the bottom line of online service providers and other businesses. Since Wikipedia is a non-profit collaborative effort, this angle was not explored in depth here; but Wikipedia, like many online service providers, survives through user traffic and contributions, both of which may be chilled temporarily or permanently by government surveillance. In fact, Wikimedia Foundation, the ACLU, and other organizations (like *The Nation* magazine and Human Rights Watch) that are party to the

184. Marthews & Tucker, *supra* note 25, at 23 ("From a US competitive standpoint, the longer-run effect observed on international Google users' search behavior indicates that knowledge of US government surveillance of Google could indeed affect their behavior. At the most limited end of the spectrum, it could steer them away from conducting certain searches on US search engines; at the most severe end of the spectrum, they might choose to use non-US search engines.").

lawsuit against the NSA and U.S. Justice Department claim they have incurred significant costs to preserve privacy and confidentiality in response to the NSA surveillance activities.¹⁸⁵ Moreover, it is also unknown whether the findings in this case study hold for comparable forms of online information resources. Was Wikipedia, given its prominence as a popular online information provider, unique in being impacted? And were Wikipedia editors and contributors affected differently from general Wikipedia users (e.g., users who merely read Wikipedia articles, but do not produce or edit them)? Despite the evidence established in this case study concerning the existence and scope of chilling effects, significant gaps remain in the literature. Further work can be done both on Wikipedia and in other online contexts to extend our understanding of chilling effects in both North America and abroad. Though the true scope of chilling effects still remains to be fully explored and analyzed, this case study has aimed to offer a contribution to its deeper understanding.

185. See Complaint, *supra* note 1, at 15, 22, 24, 26.

APPENDIX

Table 1: First Results, 48 Terrorism-related Articles Study Group

Independent Variable	Coefficients	Standard Error	P-value
Coefficient (β_0) Expected Total Views at Beginning of Study	23522364**	171743.1	0.000
Secular trend in data (β_1) Change in Views (Monthly) Before 6/2013	47038.28**	16760.41	0.009
Change in level (β_2) Change in Views Immediately After 6/2013	-995085.2*	241987.6	0.000
Change in slope (β_3) Change in Views (Monthly) After 6/2013	-35517.69	26272.41	0.187

* $p < 0.05$, ** $p < 0.01$

Table 2: Second Results, 47 Terrorism-related Articles (Hamas Excluded)

Independent Variable	Coefficients	Standard Error	P-value
Coefficient (β_0) Expected Total Views at Beginning of Study	2289153**	109751.5	0.000
Secular trend in data (β_1) Change in Views (Monthly) Before 6/2013	41420.51**	10710.65	0.001
Change in level (β_2) Change in Views Immediately After 6/2013	-693616.9**	154640.9	0.000
Change in slope (β_3) Change in Views (Monthly) After 6/2013	-67513.1**	16789.25	0.000

* $p < 0.05$, ** $p < 0.01$

Table 3: Global English Wikipedia Article Views, All Platforms (Millions)

Independent Variable	Coefficients	Standard Error	P-value
Coefficient (β_0) Expected Total Views at Beginning of Study	8313.5**	238.34	0.000
Secular trend in data (β_1) Change in Views (Monthly) Before 6/2013	114.38**	23.26	0.000
Change in level (β_2) Change in Views Immediately After 6/2013	-1535.82**	335.83	0.000
Change in slope (β_3) Change in Views (Monthly) After 6/2013	-46.97	36.46	0.208

* $p < 0.05$, ** $p < 0.01$

Table 4: Final Sets of Results—With Comparators

30 Terrorism-related Wikipedia Articles Study Group
Results correcting auto-correlation (Prais–Winsten method) in parenthesis

Independent Variable	Coefficients	Standard Error	P-value
Coefficient (β_0)	445534.1**	39759.4	0.000
Expected Total Views at Beginning of Study	(455316.5**)	(49923.17)	(0.000)
Secular trend in data (β_1)	26129.9**	3880.12	0.000
Change in Views (Monthly) Before 6/2013	(25243.94**)	(4800.76)	(0.000)
Change in level (β_2)	-225867.4**	56021.35	0.000
Change in Views Immediately After 6/2013	(-219625.8**)	(65833.74)	(0.002)
Change in slope (β_3)	-38160.16**	6082.19	0.000
Change in Views (Monthly) After 6/2013	(-37282.82**)	(7752.107)	(0.000)

* $p < 0.05$, ** $p < 0.01$ **Full 25 Domestic Security-related Wikipedia Articles Comparator Group**Note: This model's fit was not significant ($Prob > F = 0.447$)

Independent Variable	Coefficients	Standard Error	P-value
Coefficient (β_0)	708187.3**	84366.66	0.00
Expected Total Views at Beginning of Study			
Secular trend in data (β_1)			
Change in Views (Monthly) Before 6/2013	11135.07	8233.34	0.187
Change in level (β_2)			
Change in Views Immediately After 6/2013	-24638.34	118873.4	0.837
Change in slope (β_3)			
Change in Views (Monthly) After 6/2013	-20465.87	12905.99	0.124

* $p < 0.05$, ** $p < 0.01$ **Refined 23 Domestic Security-related Articles Group (Outliers Excluded)**

Results correcting auto-correlation (Prais–Winsten method) in parenthesis

Independent Variable	Coefficients	Standard Error	P-value
Coefficient (β_0)	402512.6**	23293.73	0.000
Expected Total Views at Beginning of Study	(424445.7**)	(36816.57)	(0.000)
Secular trend in data (β_1)			
Change in Views (Monthly) Before 6/2013	3090.47	2273.23	0.185
	(145.70)	(3448.036)	(0.967)
Change in level (β_2)			
Change in Views Immediately After 6/2013	-90921.01*	32821.08	0.010
	(28516.32)	(42574.71)	(0.508)
Change in slope (β_3)			
Change in Views (Monthly) After 6/2013	-7022.62	3563.36	0.059
	(7615.21)	(5801.4)	(0.200)

* $p < 0.05$, ** $p < 0.01$

Table 5: 34 Infrastructure Security-related Articles Comparator Group

Independent Variable	Coefficients	Standard Error	P-value
Coefficient (β_0) Expected Total Views at Beginning of Study	771772.3**	30948.71	0.000
Secular trend in data (β_1) Change in Views (Monthly) Before 6/2013	-11079.82**	3020.28	0.001
Change in level (β_2) Change in Views Immediately After 6/2013	-12721.07	43607.01	0.773
Change in slope (β_3) Change in Views (Monthly) After 6/2013	2431.84	4734.38	0.612

* $p < 0.05$, ** $p < 0.01$

Table 6: 26 Most Popular Wikipedia Articles (2012/2013/2014) Comparator Group

Note: This model's fit was not significant ($Prob > F = 0.7938$)

Independent Variable	Coefficients	Standard Error	P-value
Coefficient (β_0) Expected Total Views at Beginning of Study	2.58x10 ⁷ **	1920624	0.000
Secular trend in data (β_1) Change in Views (Monthly) Before 6/2013	-48458.14	187433.7	0.798
Change in level (β_2) Change in Views Immediately After 6/2013	-1716643	2706177	0.531
Change in slope (β_3) Change in Views (Monthly) After 6/2013	177324.7	293807.6	0.551

* $p < 0.05$, ** $p < 0.01$

Table 7: Independent Rating Results

Rating Type	Mean Rating
Government Trouble Rating	1.95
Privacy-Sensitive Rating	2.01
Browser History Delete Rating	2.00
Avoidance Rating	2.62

Table 8: Topic Keyword—48 Article Group

Topic Keyword	Wikipedia Articles	Govern-ment Trouble	Browser Delete	Privacy Sensi-tive	Avoid-ance
Al Qaeda	http://en.wikipedia.org/wiki/Al-Qaeda	2.20	2.11	2.21	2.84
Terrorism	http://en.wikipedia.org/wiki/terrorism	2.19	2.05	2.16	2.79
Terror	http://en.wikipedia.org/wiki/terror	1.98	1.96	2.01	2.64
Attack	http://en.wikipedia.org/wiki/attack	1.92	1.91	1.92	2.56

Topic Keyword	Wikipedia Articles	Govern-ment Trouble	Browser Delete	Privacy Sensi-tive	Avoid-ance
Iraq	http://en.wikipedia.org/wiki/iraq	1.60	1.74	1.76	2.25
Afghanistan	http://en.wikipedia.org/wiki/afghanistan	1.61	1.71	1.75	2.23
Iran	http://en.wikipedia.org/wiki/iran	1.62	1.73	1.78	2.25
Pakistan	http://en.wikipedia.org/wiki/pakistan	1.59	1.71	1.75	2.22
Agro	http://en.wikipedia.org/wiki/agro	1.51	1.80	1.76	2.29
Environmental Terrorism	http://en.wikipedia.org/wiki/Environmental_terrorism	2.20	2.20	2.24	2.92
Eco-Terrorism	http://en.wikipedia.org/wiki/Eco-terrorism	2.22	2.20	2.22	2.92
Conventional Weapon	http://en.wikipedia.org/wiki/Conventional_weapon	2.03	2.16	2.07	2.81
Weapons Grade	http://en.wikipedia.org/wiki/Weapons-grade	2.18	2.22	2.17	2.99
Dirty Bomb	http://en.wikipedia.org/wiki/Dirty_bomb	2.72	2.55	2.50	3.45
Nuclear Enrichment	http://en.wikipedia.org/wiki/Nuclear_enrichment	2.22	2.21	2.21	2.92
Nuclear	http://en.wikipedia.org/wiki/nuclear	1.84	1.97	1.91	2.55
Chemical Weapon	http://en.wikipedia.org/wiki/Chemical_weapon	2.43	2.36	2.39	3.16
Biological Weapon	http://en.wikipedia.org/wiki/Biological_weapon	2.44	2.39	2.39	3.18
Ammonium nitrate	http://en.wikipedia.org/wiki/Ammonium_nitrate	2.49	2.44	2.26	3.24
Improvised Explosive Device	http://en.wikipedia.org/wiki/Improvised_explosive_device	2.82	2.64	2.53	3.46
Abu Sayyaf	http://en.wikipedia.org/wiki/Abu_Sayyaf	2.02	1.96	1.99	2.57
Hamas	http://en.wikipedia.org/wiki/hamas	1.90	1.93	1.97	2.49
FARC	http://en.wikipedia.org/wiki/FARC	1.83	1.88	1.90	2.46
Irish Republican Army	http://en.wikipedia.org/wiki/Irish_Republican_Army	1.62	1.77	1.83	2.24
Euskadi ta Askatasuna	http://en.wikipedia.org/w/Euskadi_ta_Askatasuna	1.86	1.88	1.88	2.43
Hezbollah	http://en.wikipedia.org/wiki/hezbollah	1.86	1.90	1.96	2.46
Tamil Tigers	http://en.wikipedia.org/wiki/Tamil_Tigers	1.76	1.86	1.87	2.39
PLO	http://en.wikipedia.org/wiki/Palestine_Liberation_Organization	1.77	1.87	1.91	2.42
Palestine	http://en.wikipedia.org/wiki/	1.81	1.89	1.95	2.47

Topic Keyword	Wikipedia Articles	Govern- ment Trouble	Browser Delete	Privacy Sensi- tive	Avoid- ance
Liberation Front	Palestine_Liberation_Front				
Car bomb	http://en.wikipedia.org/wiki/Car_bomb	2.72	2.61	2.50	3.40
Jihad	http://en.wikipedia.org/wiki/jihad	2.15	2.19	2.17	2.89
Taliban	http://en.wikipedia.org/wiki/taliban	2.06	2.03	2.10	2.70
Suicide bomber	http://en.wikipedia.org/wiki/Suicide_bomber	2.25	2.31	2.24	2.97
Suicide attack	http://en.wikipedia.org/wiki/Suicide_attack	2.30	2.36	2.29	3.04
AL Qaeda in the Arabian Peninsula	http://en.wikipedia.org/wiki/Al-Qaeda_in_the_Arabian_Peninsula	2.01	1.98	2.06	2.63
Al Qaeda in the Islamic Maghreb	http://en.wikipedia.org/wiki/Al-Qaeda_in_the_Islamic_Maghreb	2.05	1.98	2.06	2.60
Tehrik-i-Taliban Pakistan	http://en.wikipedia.org/wiki/Tehrik-i-Taliban_Pakistan	1.96	1.96	1.97	2.59
Yemen	http://en.wikipedia.org/wiki/yemen	1.60	1.72	1.74	2.18
Pirates	http://en.wikipedia.org/wiki/pirates	1.44	1.67	1.67	2.10
Extremism	http://en.wikipedia.org/wiki/extremism	1.64	1.90	1.86	2.40
Somalia	http://en.wikipedia.org/wiki/somalia	1.50	1.68	1.67	2.12
Nigeria	http://en.wikipedia.org/wiki/nigeria	1.48	1.66	1.64	2.07
Political radicalism	http://en.wikipedia.org/wiki/Political_radicalism	1.75	1.91	1.97	2.48
Al-Shabaab	http://en.wikipedia.org/wiki/Al-Shabaab	1.84	1.89	1.89	2.48
Nationalism	http://en.wikipedia.org/wiki/nationalism	1.48	1.71	1.73	2.20
Recruitment	http://en.wikipedia.org/wiki/recruitment	1.74	1.90	1.87	2.54
Fundamentalism	http://en.wikipedia.org/wiki/fundamentalism	1.60	1.79	1.80	2.32
Islamist	http://en.wikipedia.org/wiki/islamist	1.79	1.89	1.93	2.45
MEAN		1.95	2.00	2.01	2.62

Table 9: Topic Keyword—30 Terrorism-related Article Study Group

Topic Keyword	Wikipedia Articles	Combined Privacy Rating
Al Qaeda	http://en.wikipedia.org/wiki/Al-Qaeda	2.34
Terrorism	http://en.wikipedia.org/wiki/terrorism	2.30
Terror	http://en.wikipedia.org/wiki/terror	2.15

Topic Keyword	Wikipedia Articles	Combined Privacy Rating
Environmental Terrorism	http://en.wikipedia.org/wiki/Environmental_terrorism	2.39
Eco-terrorism	http://en.wikipedia.org/wiki/Eco-terrorism	2.39
Conventional weapon	http://en.wikipedia.org/wiki/Conventional_weapon	2.27
Weapons Grade	http://en.wikipedia.org/wiki/Weapons-grade	2.39
Dirty Bomb	http://en.wikipedia.org/wiki/Dirty_bomb	2.81
Nuclear Enrichment	http://en.wikipedia.org/wiki/Nuclear_enrichment	2.39
Nuclear	http://en.wikipedia.org/wiki/nuclear	2.07
Chemical Weapon	http://en.wikipedia.org/wiki/Chemical_weapon	2.59
Biological Weapon	http://en.wikipedia.org/wiki/Biological_weapon	2.60
Improvised Explosive Device	http://en.wikipedia.org/wiki/Improvised_explosive_device	2.86
Abu Sayyaf	http://en.wikipedia.org/wiki/Abu_Sayyaf	2.14
FARC	http://en.wikipedia.org/wiki/FARC	2.02
Euskadi ta Askatasuna	http://en.wikipedia.org/w/Euskadi_ta_Askatasuna	2.01
Hezbollah	http://en.wikipedia.org/wiki/hezbollah	2.05
Palestine Liberation Front	http://en.wikipedia.org/wiki/Palestine_Liberation_Front	2.03
Car Bomb	http://en.wikipedia.org/wiki/Car_bomb	2.81
Jihad	http://en.wikipedia.org/wiki/jihad	2.35
Taliban	http://en.wikipedia.org/wiki/taliban	2.22
Suicide Bomber	http://en.wikipedia.org/wiki/Suicide_bomber	2.44
Suicide Attack	http://en.wikipedia.org/wiki/Suicide_attack	2.50
AL Qaeda in the Arabian Peninsula	http://en.wikipedia.org/wiki/Al-Qaeda_in_the_Arabian_Peninsula	2.17
Al Qaeda in the Islamic Maghreb	http://en.wikipedia.org/wiki/Al-Qaeda_in_the_Islamic_Maghreb	2.17
Tehrik-i-Taliban Pakistan	http://en.wikipedia.org/wiki/Tehrik-i-Taliban_Pakistan	2.12
Political Radicalism	http://en.wikipedia.org/wiki/Political_radicalism	2.03
Al-Shabaab	http://en.wikipedia.org/wiki/Al-Shabaab	2.03
Recruitment	http://en.wikipedia.org/wiki/recruitment	2.01
Islamist	http://en.wikipedia.org/wiki/islamist	2.02
MEAN		1.95

Table 10: Topic Keyword—25 Domestic Security-related Article Comparator Group

Topic Keyword	Wikipedia Articles
Department of Homeland Security	https://en.wikipedia.org/wiki/United_States_Department_of_Homeland_Security
Federal Emergency Management Agency	https://en.wikipedia.org/wiki/Federal_Emergency_Management_Agency
Coast Guard	https://en.wikipedia.org/wiki/Coast_guard
Customs and Border Protection	https://en.wikipedia.org/wiki/Customs_and_Border_Protection
Border patrol	https://en.wikipedia.org/wiki/Border_Patrol
Secret Service	https://en.wikipedia.org/wiki/Secret_Service
Bureau of Land Management	https://en.wikipedia.org/wiki/Bureau_of_Land_Management
Homeland defense	https://en.wikipedia.org/wiki/Homeland_defense
Agent/Espionage	https://en.wikipedia.org/wiki/Espionage
Task Force 88	https://en.wikipedia.org/wiki/Task_Force_88_(anti-terrorist_unit)
Central Intelligence Agency	https://en.wikipedia.org/wiki/Central_Intelligence_Agency
Fusion center	https://en.wikipedia.org/wiki/Fusion_center
DEA	https://en.wikipedia.org/wiki/DEA
Secure Border Initiative	https://en.wikipedia.org/wiki/Secure_Border_Initiative
Federal Bureau of Investigation	https://en.wikipedia.org/wiki/Federal_Bureau_of_Investigation
Alcohol and Tobacco Tax and Trade Bureau	https://en.wikipedia.org/wiki/Alcohol_and_Tobacco_Tax_and_Trade_Bureau
U.S. Citizenship and Immigration Services	https://en.wikipedia.org/wiki/United_States_Citizenship_and_Immigration_Services
Federal Air Marshal Service	https://en.wikipedia.org/wiki/Federal_Air_Marshal_Service
Transportation Security Administration	https://en.wikipedia.org/wiki/Transportation_Security_Administration
Air Marshal	https://en.wikipedia.org/wiki/Air_marshal
Federal Aviation Administration	https://en.wikipedia.org/wiki/Federal_Aviation_Administration
National Guard	https://en.wikipedia.org/wiki/National_Guard
Disaster Relief / Emergency Management	https://en.wikipedia.org/wiki/Emergency_management
U.S. Immigration and Customs Enforcement	https://en.wikipedia.org/wiki/U.S._Immigration_and_Customs_Enforcement
United Nations	https://en.wikipedia.org/wiki/United_Nations

Table 11: Topic Keyword—34 Infrastructure Security-Related Article Comparator

Topic Keyword	Wikipedia Articles
Information security	https://en.wikipedia.org/wiki/Infrastructure_security
Airport	https://en.wikipedia.org/wiki/Airport
Airplane	https://en.wikipedia.org/wiki/Airplane
Chemical burn	https://en.wikipedia.org/wiki/Chemical_burn
CIKR	https://en.wikipedia.org/wiki/CIKR
AMTRAK	https://en.wikipedia.org/wiki/Amtrak
Collapse	https://en.wikipedia.org/wiki/Collapse
Information infrastructure	https://en.wikipedia.org/wiki/Information_infrastructure
Telecommunications Network	https://en.wikipedia.org/wiki/Telecommunications_network
Telecommunication	https://en.wikipedia.org/wiki/Telecommunication
Critical infrastructure	https://en.wikipedia.org/wiki/Critical_Infrastructure
National Information Infrastructure	https://en.wikipedia.org/wiki/National_Information_Infrastructure
Metro	https://en.wikipedia.org/wiki/Metro_station
WMATA	https://en.wikipedia.org/wiki/Washington_Metropolitan_Area_Transit_Authority
Subway	https://en.wikipedia.org/wiki/Subway
BART	https://en.wikipedia.org/wiki/Bay_Area_Rapid_Transit
MARTA	https://en.wikipedia.org/wiki/Metropolitan_Atlanta_Rapid_Transit_Authority
Port Authority	https://en.wikipedia.org/wiki/Port_authority
NBIC	https://en.wikipedia.org/wiki/NBIC
Power grid	https://en.wikipedia.org/wiki/Electrical_grid
Power	https://en.wikipedia.org/wiki/Power
Smart	https://en.wikipedia.org/wiki/Smart
Full body scanner	https://en.wikipedia.org/wiki/Full_body_scanner
Electric power	https://en.wikipedia.org/wiki/Electric_power
Failure	https://en.wikipedia.org/wiki/Failure
Power outage	https://en.wikipedia.org/wiki/Power_outage
Blackout	https://en.wikipedia.org/wiki/Blackout
Brownout	https://en.wikipedia.org/wiki/Brownout
Port	https://en.wikipedia.org/wiki/Port
Dock (maritime)	https://en.wikipedia.org/wiki/Dock_(maritime)
Bridge	https://en.wikipedia.org/wiki/Bridge
Flight Cancellation and	https://en.wikipedia.org/wiki/

Delay	Flight_cancellation_and_delay
Delay	https://en.wikipedia.org/wiki/Delay
Electric power transmission	https://en.wikipedia.org/wiki/Electric_power_transmission

Table 12: 26 Most Popular Articles in 2012, 2013, & 2014 Comparator Group

Topic Keyword	Wikipedia Articles
Facebook	https://en.wikipedia.org/wiki/Facebook
Wiki	http://en.wikipedia.org/wiki/Wiki
Deaths in 2012	https://en.wikipedia.org/wiki/Lists_of_deaths_by_year#2012
One Direction	https://en.wikipedia.org/wiki/One_Direction
The Avengers (2012 film)	https://en.wikipedia.org/wiki/The_Avengers_(2012_film)
Fifty Shades of Grey	https://en.wikipedia.org/wiki/Fifty_Shades_of_Grey
2012 phenomena	https://en.wikipedia.org/wiki/2012_phenomenon
Google	https://en.wikipedia.org/wiki/Google
The Dark Knight Rises	https://en.wikipedia.org/wiki/The_Dark_Knight_Rises
The Hunger Games	https://en.wikipedia.org/wiki/The_Hunger_Games
Deaths in 2013	https://en.wikipedia.org/wiki/Lists_of_deaths_by_year#2013
Breaking Bad	https://en.wikipedia.org/wiki/Breaking_Bad
G-force	https://en.wikipedia.org/wiki/G-force
World War II	https://en.wikipedia.org/wiki/World_War_II
Youtube	https://en.wikipedia.org/wiki/YouTube
List of Bollywood Films 2013	https://en.wikipedia.org/wiki/List_of_Bollywood_films_of_2013
United States	https://en.wikipedia.org/wiki/United_States
Online shopping	https://en.wikipedia.org/wiki/Online_shopping
Java	https://en.wikipedia.org/wiki/Java
Alive	https://en.wikipedia.org/wiki/Alive
Deaths in 2014	https://en.wikipedia.org/wiki/Lists_of_deaths_by_year#2014
Climatic Research Unit email controversy	https://en.wikipedia.org/wiki/Climatic_Research_Unit_email_controversy
Amazon.com	https://en.wikipedia.org/wiki/Amazon.com
2014 FIFA World Cup	https://en.wikipedia.org/wiki/2014_FIFA_World_Cup
Ebola virus disease	https://en.wikipedia.org/wiki/Ebola_virus_disease
Game of Thrones	https://en.wikipedia.org/wiki/Game_of_Thrones

FACILITATING COMPETITION BY REMEDIAL REGULATION

Kristelia A. García[†]

ABSTRACT

In music licensing, powerful music publishers have begun—for the first time ever—to withdraw their digital copyrights from the collectives that license those rights, in order to negotiate considerably higher rates in private deals. At the beginning of the year, two of these publishers commanded a private royalty rate nearly twice that of the going collective rate. This result could be seen as a coup for the free market: Constrained by consent decrees and conflicting interests, collectives are simply not able to establish and enforce a true market rate in the new, digital age. This could also be seen as a pathological form of private ordering: Powerful licensors using their considerable market power to impose a supracompetitive rate on a hapless licensee. While there is no way to know what the market rate looks like in a highly regulated industry like music publishing, the anticompetitive effects of these withdrawals may have detrimental consequences for artists, licensees and consumers. In industries such as music licensing, network effects, parallel pricing and tacit collusion can work to eliminate meaningful competition from the marketplace. The resulting lack of competition threatens to stifle innovation in both the affected, and related, industries.

Normally, where a market operates in a workably competitive manner, the remedy for anticompetitive behavior can be found in antitrust law. In music licensing, however, some concerning behaviors, including both parallel pricing and tacit collusion, do not rise to the level of antitrust violations; as such, they cannot be addressed by antitrust law. This

DOI: <http://dx.doi.org/10.15779/Z38NZ8H>

© 2016 Kristelia A. García.

[†] Associate Professor, University of Colorado Law School; Director-Content Initiative, Silicon Flatirons Center for Law, Technology, and Entrepreneurship. The author would like to thank the following people for their especially generous and helpful contributions: Michael Abramowicz, Robert Brauneis, Ben Depoorter, Peter DiCola, Mark Lemley, Doug Melamed, Paul Ohm, Gideon Parchomovsky, Matthew Sag, Pamela Samuelson, Philip Weiser, and Christopher Yoo, as well as participants in the University of Pennsylvania Copyright Scholarship Roundtable, BYU Law's Law & Entrepreneurship Colloquium, the Spangenberg Center for Law, Technology & the Arts at Case Western Reserve University School of Law, the University of Colorado Faculty Colloquium, the Silicon Flatirons' Content Conference Workshop, the Intellectual Property Scholar's 2015 Conference, the Silicon Flatirons/Searle Center 2015 Cable Workshop, and Lewis & Clark Law School's IP in the Trees Workshop for their helpful insights, and Jane Thompson in the Colorado Law Library for stellar research assistance. This work was supported by a Mark Twain Fellowship. Accolades and compliments, and criticism couched as such, are welcome at kristelia@colorado.edu.

is no small irony. At one point, antitrust served as a check on the licensing collectives by establishing consent decrees to govern behavior. Due to a series of acquisitions that have reduced the music publishing industry to a mere three entities, the collectives that are being circumvented by these withdrawals (and whose conduct is governed by consent decrees) now pose less of a competitive concern than do individual publishing companies acting privately, or in concert through tacit collusion. The case of intellectual property rights, which defer competition for creators and inventors for a limited period of time, is particularly challenging for antitrust.

Running contrary to conventional wisdom, this Article posits that regulation—not antitrust—is the optimal means of enabling entry and innovation in the music licensing market. While regulation is conventionally understood to restrict new entry and to interfere with competition, this Article demonstrates that where a market becomes highly concentrated, regulation can actually encourage competition by ensuring access to key inputs at competitive rates. While not without its drawbacks, including an increase in the cost of private action, remedial regulation in music licensing corrects anticompetitive behavior and ensures ongoing access to content and fair payment to artists, while supporting continued innovation in content distribution.

TABLE OF CONTENTS

I.	INTRODUCTION	185
II.	PRIVATE ORDERING AS PATHOLOGY	191
A.	(NEARLY) EVERYTHING YOU NEVER WANTED TO KNOW ABOUT MUSIC LICENSING	192
1.	<i>Music Licensing 101</i>	192
2.	<i>Collectives, Consent Decrees & Rate Courts</i>	193
3.	<i>Music Publishing Consolidation</i>	196
B.	PRIVATE LICENSING IN THE SHADOW OF A COLLECTIVE: PANDORA V. ASCAP	197
1.	<i>Music Publishers Attempt to Withdraw Digital Public Performance Rights</i>	198
2.	<i>Amendment of ASCAP's Governing Articles</i>	200
3.	<i>"Negotiation"</i>	203
4.	<i>The Rate Court Determination</i>	206
5.	<i>The Consent Decree</i>	208
C.	COLLECTIVE LICENSING V. PRIVATE ORDERING	209
1.	<i>Diminished Advantages of Collective Action</i>	209
2.	<i>Advantages of Private Ordering</i>	211
3.	<i>Potential Drawbacks to Private Ordering</i>	212
III.	AN ANTITRUST PARADOX	218
A.	ANTICOMPETITIVE BEHAVIOR	218
B.	CONSEQUENCES OF ANTICOMPETITIVE BEHAVIOR	222
IV.	CONVENTIONAL SOLUTIONS & THEIR DISCONTENTS	227

A.	THE FREE MARKET APPROACH.....	228
B.	ANTITRUST LAW.....	230
	1. <i>Inapplicability</i>	230
	2. <i>Challenges</i>	236
V.	FACILITATING COMPETITION THROUGH REGULATION.....	242
A.	PENALTY DEFAULTS, ALTERING RULES & COMPETITION.....	244
	1. <i>Default Theory</i>	244
	2. <i>Application to Regulation</i>	246
B.	REMEDIAL REGULATION.....	246
	1. <i>Logistics</i>	247
	2. <i>Applicability</i>	250
	3. <i>Challenges</i>	251
	4. <i>Advantages</i>	253
VI.	CONCLUSION.....	255
	APPENDIX.....	257
I.	INTRODUCTION	

On March 30, 2015, rapper Jay-Z and a dozen of his closest artist-friends came together at a press release event in New York City to announce a new music streaming service that was widely billed as “revolutionary.”¹ The service called Tidal was touted as artist-owned and promised to pay artists more than any other music streaming service. While the logistics around Tidal’s business model were foregone in favor of star power and flashy showmanship, this much is clear: The only way Tidal is going to command \$19.99/month (twice the monthly cost of competing services) from customers is if the participating artists—which include Jack White, Beyoncé, Daft Punk, and Kanye West—withdraw their content from all other music streaming services and thereby allow Tidal to create scarcity and command a higher subscription rate.

While this reduction in the number of distribution outlets may seem undesirable from a consumer perspective, this result is precisely what unregulated intellectual property (IP) rights do when they protect the

1. See, e.g., Ben Sisario, *Jay Z Reveals Plans for Tidal, a Streaming Music Service*, N.Y. TIMES (March 30, 2015), <http://www.nytimes.com/2015/03/31/business/media/jay-z-reveals-plans-for-tidal-a-streaming-music-service.html> [<http://perma.cc/M4ZX-CJ9W>]

rights holder from competition for a limited period of time.² The real concern, and the subject of this Article, is the potential to stifle innovation in the content distribution space (here, the burgeoning music streaming industry) and to lock listeners into to the current distribution technology. Arguably, this is *not* what IP seeks to do—or at least not what it should do.

From a regulatory (if not a contractual) perspective,³ the withdrawal scenario posited above is a real possibility. Take the recent American Society of Composers, Authors and Publishers (ASCAP) withdrawals, for example. ASCAP is a “performance rights organization” (PRO) that handles the collection and administration of public performance royalties, which are royalties incurred from the playing of a song on terrestrial or digital radio. At the beginning of last year, two of the nation’s three major music publishers⁴ withdrew their digital copyrights from ASCAP in order to license them through private deals.

The idea that a music publisher might forego the ease and convenience of collective licensing of its IP rights in order to do all of the work itself is counterintuitive. Conventional wisdom justifies collectives because they

2. Concededly, not a limited enough time for some. *See, e.g.*, Jane C. Ginsburg et al., *The Constitutionality of Copyright Term Extension: How Long is Too Long?*, 18 CARDOZO ARTS & ENT. L.J. 651 (2000) (discussing the constitutionality, and overall desirability, of copyright term extension from the perspectives of incentives, natural rights, international law, and legislative history). To be clear, it is not suggested that IP prevents Taylor Swift from having to “compete” with Katy Perry for listeners, but rather that IP prevents Taylor Swift from having to compete with others for the reproduction and distribution of her own content.

3. In order to withdraw his content from Spotify, for example, Kanye West would have to rely upon the terms of his contracts with each of his publishing company and record label, and, in turn, their respective contractual obligations to Spotify, making the whole enterprise rather unlikely, though Taylor Swift—who controls her own publishing—has successfully done just that. *See, e.g.*, Steve Knopper, *Taylor Swift Abruptly Pulls Entire Catalog From Spotify*, ROLLING STONE, (Nov. 3, 2014), <http://www.rollingstone.com/music/news/taylor-swift-abruptly-pulls-entire-catalog-from-spotify-20141103> [<http://perma.cc/LN3H-VSXA>].

4. A “music publisher” is a company that owns the copyrights on various musical compositions and licenses the use of those compositions to such entities as radio stations, sports stadiums, filmmakers, restaurants, and record labels—traditionally through a collective, but more recently via private negotiation. The three major music publishers in the U.S. today are Sony Music Publishing, LLC (“Sony”), Universal Music Publishing Group (“Universal”) and Warner Music Publishing (“Warner”). *See, e.g.*, Ed Christman, *First-Quarter Music Publishing Rankings: SONGS Surges Again*, BILLBOARD, (May 12, 2014), <http://www.billboard.com/biz/articles/news/publishing/6084783/first-quarter-music-publishing-rankings-songs-surges-again> [<http://perma.cc/SR7Z-WZ4N>] (showing Sony, Universal, and Warner with 31.9%, 18.1%, and 13.3% market shares, respectively).

are able to reduce transaction costs and consolidate bargaining power.⁵ The purported circumvention of ASCAP undermines these justifications and suggests that they no longer hold sway in the digital age. The rights withdrawals marked the first ever attempt at withdrawal and private ordering⁶ in the music publishing industry and resulted in a negotiated royalty rate nearly double ASCAP's going rate.

One could view this result as a coup for the free market: Constrained by an antiquated consent decree and faced with conflicting member interests, ASCAP artificially depresses the "market" rate for music licensing. This result could also be viewed as a pathological form of private ordering in which two powerful companies wielded their considerable market power to coerce a supracompetitive rate from a hapless licensee.

Unfortunately, it is difficult to determine what the market rate looks like in a highly regulated industry like music publishing⁷ because there has never been a true "market" for music licensing in any economically meaningful sense. While it is true that a songwriter can technically license her own songs, this is not logistically possible (at least in an analog world). Nearly every artist uses a collective for all relevant intents and purposes. The establishment of the music publishing industry arose more or less simultaneously with the creation of collectives that consolidated composition copyrights and set a royalty rate, thereby foreclosing the

5. See, e.g., Robert P. Merges, *Contracting into Liability Rules: Intellectual Property Rights and Collective Rights Organizations*, 84 CALIF. L. REV. 1293, 1302–3 (1996) ("The high costs of contracting . . . drive the right holders to pool their property rights in a collective organization. . . . It is the high transaction costs associated with the initial entitlements that lead the parties to establish the organization—an organization that then dramatically lowers the costs of exchanging the rights."). That bargaining power is, of course, checked by consent decree, see *infra* Part II.B.5.

6. Here, "private ordering" refers to negotiation and deal making between private parties, outside of any statutory or regulatory regime.

7. There exists an ongoing debate about copyright's status as a form of regulation versus as a property right regime. Compare Adam Mossoff, *Is Copyright Property?*, 42 SAN DIEGO L. REV. 29, 41 (2005) (proposing that "a person's right to control the disposition of his creation, and thereby enjoy the fruits—the profit—of his labors, is central to the legal definition and protection of property entitlements."), with Siva Vaidhyanathan, *COPYRIGHTS AND COPYWRONGS: THE RISE OF INTELLECTUAL PROPERTY AND HOW IT THREATENS CREATIVITY* 12 (2003) (noting that "copyright issues are now more about large corporations limiting access to and use of their products We must seek a balance. . . . Instead of bolstering 'intellectual property,' we should be forging 'intellectual policy.'") For the limited purposes of content licensing, at least, the role of copyright law as regulation is undeniable, and this Article will treat it as such. Not only do the copyright laws contain statutory licenses for cable and satellite television, for example, but they also ultimately entrust royalty rate setting to governmentally authorized entities, including the copyright royalty board, and the rate court.

development of a free market for the licensing of musical compositions. For this reason, this Article takes no issue with the rate obtained, but rather with the potential for unchecked anticompetitive consequences for artists, licensees, and consumers. Notably, the two music publishers at issue—Sony and Universal—together control roughly 50% of all digital music publishing copyrights. Without this content, an Internet radio service would be hard-pressed to compete for listeners. This was the argument Pandora advanced in a rate court⁸ proceeding that challenged the privately obtained rate as coerced.⁹

In an industry like music licensing, network effects, parallel pricing and tacit collusion can exacerbate the effects of consolidation by removing the threat of meaningful competition¹⁰ from the marketplace.¹¹ This allows individual companies—acting alone or in tacit collusion with

8. In this context, a “rate court” is a court with exclusive jurisdiction to oversee licensing arrangements between a collective like ASCAP, and a potential licensee, like Pandora. The rate court’s jurisdictional authority comes from the consent decree governing the collective. In a recent report on “Copyright and the Music Marketplace,” the Copyright Office called the rate court “expensive and time-consuming” and noted an inability to “keep up with the pace set by the new digital marketplace.” U.S. COPYRIGHT OFFICE, COPYRIGHT AND THE MUSIC MARKETPLACE: A REPORT OF THE REGISTER OF COPYRIGHTS 93–94 (2015), <http://copyright.gov/policy/musiclicensingstudy/copyright-and-the-music-marketplace.pdf> [<http://perma.cc/45KF-LLUA>].

9. *In re* Petition of Pandora Media, Inc., 6 F. Supp. 3d 317, 343 (S.D.N.Y. 2014) [hereinafter Order] (noting that “[u]nless Pandora could do without [Sony’s content] and remove [the content] from its repertoire . . . Pandora had to obtain a license from Sony or face crippling copyright infringement claims.”). It is also worth noting that withdrawal wasn’t necessary from a licensing perspective: ASCAP’s authority to license the publishers’ content is non-exclusive, meaning the publishers could have readily competed with ASCAP in the market to license their songs without withdrawing those rights from the collective. But the withdrawals worked to reduce the field of competition—afterward, there was only one option for licensing that content, and that was through that publisher itself, sans consent decree.

10. Much can be made about the definition of “competition,” and about whether it ought to be used in an economic or antitrust sense, in layman’s terms or aspirationally. In the words of prominent antitrust scholar Rebecca Haw Allensworth, “‘Competition’ has no single monolithic meaning; the values traded off in [antitrust] trade in different currencies, and balancing them thus requires judgment beyond the realm of economics.” Rebecca Haw Allensworth, *The Commensurability Myth in Antitrust*, 69 VAND. L. REV. 1, 4 (2016). This Article takes a broad view of competition, and considers both the procompetitive and anticompetitive effects of antitrust law.

11. “Network effects” refer to the increased value to a consumer deriving from an increase in the number of other consumers enjoying the same product or service. “Parallel pricing” is the difficult-to-detect situation wherein competitors recognize that it is in their collective best interest to informally cooperate as to price. “Tacit collusion” refers to low-level, coordinated effort between similarly situated competitors that does not rise to the level of a Sherman Act violation. *See infra* Part III.A.

similarly situated competitors—to engage in parallel pricing of performance rights, or to withhold rights altogether, thereby potentially barring entry to prospective licensees, all while avoiding antitrust scrutiny.

The case of intellectual property, where regulation defers competition or creates monopolies for creators and inventors for a limited period of time, is particularly challenging. Professor Suzanne Scotchmer attributed this dilemma to the fact that “under the Sherman Act, attempts to acquire or maintain a monopoly are illegal, even though merely being a monopolist is not. This distinction is especially important in markets with intellectual property, since intellectual property already grants a ‘legal monopoly.’”¹² It is this unique potential for anticompetitive behavior that does not meet the definition of antitrust “harm” (i.e., does not explicitly violate §§ 1 or 2 of the Sherman Act), and therefore is not addressed by that body of law, that is the focus of this Article.

Normally, where a market operates in a workably competitive manner, the remedy for anticompetitive behavior can be found in antitrust law. In music licensing, however, a breakdown in structural competition reduces the ability of antitrust to function, in large part because antitrust law does not provide a remedy to address parallel pricing and tacit collusion.

The failure of antitrust in this marketplace is no small irony. At one point, antitrust law served as a check on the power of the collective licensing entities by establishing a consent decree to govern their behavior. No such consent decrees apply to individual publishing companies. This discrepancy has resulted in a situation where the collective that is being circumvented now poses less of a competitive concern than the individual, powerful copyright holders acting privately.

Previous scholarship offers qualified encouragement of private ordering as a means of increasing efficiency in the shadow of an inefficient statutory license.¹³ This Article reaches a different conclusion in the collective context, where collectives are governed by consent decree while powerful individual rights holders are not. Currently, the two largest collectives—ASCAP and Broadcast Music, Inc. (BMI)¹⁴—function under

12. Suzanne Scotchmer, *Innovation and Incentives* 290 (2004).

13. See Kristelia A. García, *Private Copyright Reform*, 20 MICH. TELECOM. & TECH. L. REV. 1 (2013).

14. Like ASCAP, BMI is a non-profit music licensing collective that competes with ASCAP for music publishing rights. In 2014, both ASCAP and BMI reported administration of approximately \$1 billion each in public performance royalties. See ASCAP, ASCAP 2014 Annual Report at 4 (2014), http://www.ascap.com/~media/files/pdf/about/annual-reports/ascap_annual_report_2014.pdf [<http://perma.cc/YD24-UMAJ>];

consent decrees that govern their operations. These consent decrees originated from concerns about anticompetitive behavior, including supracompetitive pricing and barriers to entry, that can result from the consolidation of a large number of copyrights in only a handful of entities. Notably, these same concerns apply equally to powerful, independent music publishing companies like Sony and Universal (which also consolidate large quantities of copyrights). The terms of the consent decree, however, do not govern those companies' behavior, nor do they constrain their ability to harm consumers and innovators who depend on access to the content.

Prior work also examines the phenomenon of private ordering around an unpalatable statutory license, or "penalty default license," as the inevitable result of inherent inefficiencies and uncertainty.¹⁵ Similarly, the proliferation of private ordering in public performance rights licensing—where rights are traditionally licensed by collectives—challenges the conventional view that collectives are the optimally efficient licensing mechanism. Instead, this phenomenon demonstrates that collectives suffer many of the same inefficiencies as statutory licenses. While regulation is often criticized as restricting new entry and interfering with competition in the marketplace, this Article argues that where a market is structurally noncompetitive, remedial regulation can open the market to entry and innovation and can maintain competition therein. This approach is consistent with prior IP regulatory reform efforts, such as existing compulsory licensing regimes that ensure ongoing access to content. Indeed, from a historical perspective, IP rights have often been subject to regulatory limits—such as compulsory licensing—designed to facilitate and protect competition in the content distribution marketplace.¹⁶

The analysis proceeds in four Parts. Part II begins with a music licensing primer, then analyzes the recent rate court proceeding between Pandora and ASCAP to illustrate the emergence of a pathological form of private ordering in an industry historically dominated by collective licensing. In addition, Part II describes the rise and fall of collectives as the preferred means of licensing copyrighted content. Part III discusses the paradox revealed by this narrative; namely, the fact that the licensing

and *BMI Reports Record-Breaking Revenues of Over \$1 Billion*, BMI (Sept. 10, 2015), <http://www.bmi.com/press/entry/572180> [<http://perma.cc/H36T-YCPS>].

15. See Kristelia A. García, *Penalty Default Licenses: A Case for Uncertainty*, 89 N.Y.U. L. REV. 1117 (2014).

16. See generally Timothy Wu, *Copyright's Communications Policy*, 103 MICH. L. REV. 278, 279 (2004) (discussing "copyright's poorly understood role in regulating competition among rival disseminators")

collectives that are being circumvented—once a cause for antitrust concern due to their consolidation of content, but whose conduct is now governed by consent decree—now pose less of a competitive concern than the conduct of powerful copyright holders acting privately. Part IV makes the case for music as unique among other forms of content in order to explain the failings of antitrust in the music licensing space and the need for a different approach. Running contrary to conventional wisdom, Part V suggests that remedial regulation can encourage and maintain competition in the music industry in a way that antitrust cannot. In so doing, the Article provides a powerful case study of how the absence of structural competition leaves antitrust law unable to police competition and instead calls for the introduction of remedial regulatory oversight.

II. PRIVATE ORDERING AS PATHOLOGY

At the beginning of 2014, two of the country's three major music publishers attempted for the first time ever to withdraw digital copyrights that they had previously assigned to ASCAP, in order to license those rights through private negotiation. This preference for private over collective licensing¹⁷—which resulted in a negotiated royalty rate nearly double ASCAP's collective rate—challenges the traditional justifications for collective rights organizations, such as consolidation of bargaining power and minimization of transaction costs, and questions the conventional view of collectives as the optimal means of licensing copyrighted content.

Those who disfavor the consent decree under which ASCAP operates would conclude that private ordering predictably leads to rates and terms that more accurately represent the parties' respective valuations. This Part challenges that conclusion and instead posits that despite its perceived advantages, private ordering in the shadow of a regulated collective can have negative consequences, including various forms of anticompetitive behavior and other adverse market impacts. Beginning with a brief overview of music copyrights and licensing, this Part presents the recent

17. Not just “over,” but in exclusion of, collective licensing (since the grant of rights to a collective like ASCAP is a non-exclusive one, such that the publishing companies could have easily left the digital rights with ASCAP while simultaneously licensing them privately). See *United States v. ASCAP*, No. 41-1395, 2001 WL 1589999, § IV(B) (S.D.N.Y. 2001) *cert. denied* 132 S. Ct. 366 (2011). In theory, at least, this would allow for competition between the music publishers themselves, and even between the publishers and ASCAP (though this would require that the content not be withdrawn).

Pandora–ASCAP rate proceeding as an example of what can go right and what can go wrong in private ordering in the shadow of a collective.

A. (NEARLY) EVERYTHING YOU NEVER WANTED TO KNOW
ABOUT MUSIC LICENSING

1. *Music Licensing 101*

The world of music licensing is incredibly (and, many would argue, unduly) complicated.¹⁸ This is as much the result of piecemeal legislation and powerful lobbyists as it is a reflection of well-intended, yet largely unsuccessful, attempts to anticipate and accommodate ever-evolving technologies and consumer preferences.¹⁹ For the purposes of this Article, it suffices to understand that music is unique among copyrighted content (e.g., films, television programs, novels) in that a single song is actually comprised of two distinct copyrights: One on the underlying musical composition²⁰ and another on the sound recording of that musical composition.²¹ For each of these copyrights, there exists a companion right—that of “public performance”—that gives a licensee the right to play (i.e., perform) a song publicly, such as at a restaurant, sporting event, or over the terrestrial radio airwaves.

These performance rights may be further broken down into digital performance rights (for online plays via an Internet radio service like Pandora), and terrestrial performance rights (for analog plays via an FM radio station). Copyrights on a musical composition and copyrights on a sound recording (and their respective public performance rights) are typically held by different parties and pay royalties differently depending on the platform and the use within that platform. This dichotomy (quadchotomy?) results in a number of interesting and counterintuitive phenomena in music licensing. For example, while the sale of a digital download of The Beach Boys’ hit song “Surfin’ USA” on Apple’s iTunes platform triggers a statutory royalty paid on both the underlying musical composition and the sound recording, playing the same song on an FM radio station results in a public performance royalty payment to Chuck Berry, the song’s composer, but no such royalty payment to the Wilson

18. For the especially motivated (or confused) reader, Appendix A offers a more in-depth, graphical representation of the music licensing concepts presented in this section.

19. *See generally* JESSICA LITMAN, DIGITAL COPYRIGHT (2006) (detailing a legislative process that relies on lobbyists for the construction and proposal of copyright laws).

20. *See* 17 U.S.C. § 102(a)(2) (2012). Formerly known simply as “sheet music,” a musical composition can range from a formally notated work to an informal transcription or approximation of a musical work.

21. *See id.*

brothers, performers of the sound recording. This Article is concerned only with public performance rights on musical compositions (i.e., Chuck Berry in the above example) and with the collectives that administer those rights—specifically, ASCAP, the nation’s largest music licensing collective.

The collective licensing of performance rights for musical compositions evolved in response to a perceived gap in copyright law. § 115 of the Copyright Act of 1976 (hereinafter, the “Copyright Act”)²² provides for a compulsory license that allows recording artists to record cover versions of the musical compositions of other artists. There is no comparable compulsory license for the public performance of musical compositions, however. This means that the licensing of public performance rights for musical compositions must be negotiated between a content owner and prospective licensee. In other words, § 115 allows for the creation of a new sound recording of a composition, but not for the public performance of that composition.

While individual songwriters can technically negotiate and administer their own public performance rights,²³ this is largely impractical—at least in an analog world. For example, there is no practical way for The Beatles’ estate to monitor—much less invoice and enforce against—every restaurant, coffee shop, and retail clothing store nationwide that plays any piece of the band’s extensive repertoire. For this reason, public performance rights for musical compositions have traditionally been assigned to and managed by a collective like ASCAP.²⁴

2. *Collectives, Consent Decrees & Rate Courts*

ASCAP is a non-profit organization that licenses public performance rights for, and collects and distributes royalties on behalf of, its members.

22. 17 U.S.C. § 115(a)(1).

23. *Cf.* 17 U.S.C. § 114(e)(1) (public performance rights for digital sound recordings, which are subject to a compulsory license).

24. In the United States, the collective management of performance rights for musical compositions vests primarily in three organizations: ASCAP, BMI and SESAC. Most of these rights are held by ASCAP and BMI, the nation’s two largest PROs. The Society of European Stage Authors and Composers (SESAC) is a much smaller, privately-held company that handles the remainder or, in rare instances, the individual songwriters manage the rights themselves and operate under consent decree. This comports with Professor Wu’s characterization of PROs as more like governmentally regulated compulsory licensors than as private entities. *See* Wu, *supra* note 16, at 280.

Those members range from individual songwriters with a mere handful of copyrights to large music publishing companies with thousands of copyrights. Wielding expertise acquired through repeat negotiation and utilizing consolidated bargaining power deriving from its large quantity of content, collectives like ASCAP offer prospective licensees various licensing options of different configurations, durations, and price points. The most popular offering, known as a “blanket license,” offers a licensee the rights to all of the content in a collective’s catalog.²⁵ In exchange for this service, members pay ASCAP a share of the royalties collected on their behalf.²⁶ The price at which ASCAP offers its various license configurations to licensees is determined in the first instance by its membership, and failing that, by a rate court as detailed below.

ASCAP formed in 1914 in response to a desire for musicians to be paid when their music was played in public spaces, like restaurants.²⁷ The popular account recalls Italian composer Giacomo Puccini and American composer Victor Herbert lunching in New York City when a band in the restaurant began playing one of Herbert’s hit songs. Puccini allegedly grew enraged upon learning that Herbert was not being paid for the use of his song as he would have been in Europe. And so ASCAP was born. Its founding members included such musical giants as John Philip Sousa, Irving Berlin, George and Ira Gershwin, and Cole Porter.²⁸ Today, ASCAP represents over half a million members and has a catalog of over 10 million songs.²⁹

ASCAP operated as the sole public performance rights collective until 1940, when its ongoing battle with the radio industry reached its peak, and the radio stations pulled all ASCAP-licensed content from the

25. *Licensing Help*, ASCAP, <http://www.ascap.com/licensing/termsdefined.aspx> [<http://perma.cc/VTU2-E6TG>].

26. As of the publication of this Article, the formula for calculating this share of royalties is: (use weight)×(licensee weight)×(“follow the dollar” factor)×(time of day weight)+(audio feature premium credits)+(tv premium credits)=credits. For further, yet still unsatisfactory, explanation of this calculation, see *ASCAP Payment System*, ASCAP, <http://www.ascap.com/members/payment/royalties.aspx> [<http://perma.cc/9ZF8-2Z48>]. Needless to say, the formula is often critiqued for its opacity.

27. To be sure, the Copyright Act of 1909 had a public performance right for sound recordings, but it had yet to be interpreted doctrinally and so there was no practical means of enforcing the payment of royalties earned. See Copyright Act of 1909, Pub. L. No. 60-349, 35 Stat. 1075 (1909).

28. See Jeff Lunden, *Collecting Money for Songwriters, A 100-Year Tug of War*, NPR MUSIC (Feb. 13, 2014), <http://www.npr.org/2014/02/13/275920416/collecting-money-for-songwriters-a-100-year-tug-of-war> [<http://perma.cc/VY8L-79VA>].

29. *Happy Birthday, ASCAP!*, ASCAP, <http://www.ascap.com/100.aspx#ascaptoday> [<http://perma.cc/X5WH-25FV>].

airwaves.³⁰ During the eight months that it took for ASCAP and the broadcasters' association to renegotiate a rate, the broadcasters formed their own collective, BMI, as an ostensible competitor to ASCAP. The potential for anticompetitive behavior resulting from the consolidation of performance rights had been exposed, however, and the Department of Justice (DOJ) took notice.

The long-recognized propensity of collectives to raise concerns about anticompetitive behavior has been recognized by the academic literature³¹ and the DOJ alike. The DOJ first alleged ASCAP was an "unlawful combination" in the early 1930s.³² In 1940, a follow-on lawsuit brought by the DOJ against both ASCAP and BMI alleging unreasonable restraint of trade resulted in the imposition of a consent decree to govern the collectives' operation.³³ ASCAP's consent decree, originally entered into in

30. Another, less well-documented complaint against ASCAP was its exclusion of certain genres; specifically, jazz, blues, R&B and "hillbilly" (or country) music. *See, e.g.*, CHOSEN CAPITAL: THE JEWISH ENCOUNTER WITH AMERICAN CAPITALISM 144–45 (Rebecca Kobrin ed., 2012).

31. For more on the history of confrontation between ASCAP and radio, see Peter DiCola & Matthew Sag, *An Information-Gathering Approach to Copyright Policy*, 34 CARDOZO L. REV. 173, 203–08 (2012). In another example, Glynn Lunney, *Copyright Collectives and Collecting Societies: The United States Experience*, in COLLECTIVE MANAGEMENT OF COPYRIGHT AND RELATED RIGHTS 340 (Daniel J. Gervais ed., 2010), notes:

In the United States, [collectives] are viewed as something of a necessary evil. By reducing the transaction costs entailed in enforcing and licensing the public performance of musical works, they create a market in which otherwise there would be only infringement. But they do not merely reduce the transaction costs associated with the public performance right, they also eliminate competition between the individual copyright owners over public performance licensing terms and pricing. Because of this anti-competitive potential, copyright collectives in the United States have faced recurring litigation over whether their licensing practices violate the anti-trust laws.

32. *Id.*

33. For ASCAP's original consent decree, see *United States v. ASCAP*, 41 Civ. 1395 (S.D.N.Y. 1941). For BMI's, see *United States v. BMI*, 64 Civ. 3787 (S.D.N.Y. 1941). Due to its relatively small size and market power, SESAC is not subject to a consent decree. It is worth noting that SoundExchange, the designated collective for sound recordings, differs from the music publishing collectives in that SoundExchange was designated by the Librarian of Congress as the sole collective for the collection and administration of sound recording royalties under its regulatory authority. *See* Digital Performance Right in Sound Recordings and Ephemeral Recordings, 80 Fed. Reg. 59588 (Oct. 2, 2015) (to be codified at 37 C.F.R. § 380). As such, it does not pose the same threat of anticompetitive behavior, and therefore does not operate under a consent decree.

1941, arose from a lawsuit brought by the DOJ under § 1 of the Sherman Act of 1890 (hereinafter, the “Sherman Act”).³⁴

As amended in 2001, ASCAP’S consent decree (hereinafter, the “Second Amended Final Judgment” or “AFJ2”) names the U.S. District Court for the Southern District of New York (the “SDNY”) as rate court.³⁵ In that capacity, the SDNY provides a forum for the resolution of rate disputes between content licensors like ASCAP and content licensees like Pandora. Where a licensee and ASCAP are not able to reach a fee agreement, the licensee can petition the SDNY to set a “reasonable fee.”³⁶ At the rate court proceeding, ASCAP has the burden of proving that the rate it seeks is reasonable.³⁷

Among other things, the AFJ2 requires that ASCAP grant the right to perform *all* of the compositions in its repertoire to any prospective licensee who requests to do so.³⁸ It also prohibits ASCAP from engaging in price discrimination between similarly situated licensees.³⁹ Notably, the consent decree governs only the behavior of ASCAP and not that of its individual members.

3. *Music Publishing Consolidation*

Songwriters assign their copyrights to music publishing companies, who in turn license those compositions to recording artists, record labels, and various analog and digital distribution outlets. When the music publishing industry first began in the 1890s, royalties earned from the licensing of musical compositions comprised the majority of music industry revenues and that continued until the sale of sheet music slowed during the Great Depression.⁴⁰ The growing popularity of radio saw public performance revenues take the lead during World War II.⁴¹ The advent of rock and roll led to a massive increase in album sales, which pushed royalties earned on sound recordings to the forefront during the 1960s and 1970s.⁴² Before long, the recording industry dwarfed music

34. 15 U.S.C. § 1 (2004). The Sherman Act is one of three primary acts that together constitute U.S. antitrust law. *See infra* Part IV.

35. *See* United States v. ASCAP, No. 41-1395, 2001 WL 1589999, § IX (S.D.N.Y. June 11, 2001).

36. *Id.*

37. *Id.*

38. *Id.* §§ VI, IX(E).

39. *Id.* § IX(C).

40. *See* GEOFFREY P. HULL, THOMAS HUTCHISON & RICHARD STRASSER, *THE MUSIC BUSINESS AND RECORDING INDUSTRY* 112 (3d ed. 2011).

41. *Id.*

42. *Id.* at 113.

publishing, which in turn morphed from an industry that focused on creating musical works, into primarily a “copyright industry” that focused on licensing musical works.

Given the miniscule bargaining power afforded the owner of a copyright on a single musical composition (however popular the song), it is not surprising that music publishing is an industry built on the acquisition and consolidation of lots of musical compositions. Indeed, the three major music publishers today are the result of extensive consolidation among many smaller publishing companies. Universal resulted from a merger between PolyGram, MCA, Rondor, BMG, and Zomba. EMI Music Publishing (EMI, acquired by Sony in 2012) began in 1974 as a combination of Ardmoor, Beechwood, Keith Prowse, and Central Songs, followed by the addition of Screen Gems and Colgms, SBK, CBS and Jobette. Warner began in 1929 with its acquisition of Chappell-Harms. Sony—currently the nation’s largest publisher—got its start in 1989 when the company purchased CBS (then known as Tree), and then scored a coup with the purchase of Michael Jackson’s ATV catalog, which included the Beatles’ repertoire, for a purchase price of \$100 million.⁴³

The most plausible explanation for the dramatic consolidation observed in the music publishing industry is money—or, more accurately, the lack thereof. New technological developments including the Internet, the iPod, and the smartphone have changed the way that people consume music. This has resulted in a shift away from traditional revenue streams like CD sales to the sale of digital tracks, and, more recently still, to the sale of streaming subscriptions. These developments have not only dramatically reduced absolute revenue in dollars,⁴⁴ but have also forced ongoing changes to traditional business models and to the role of collectives like ASCAP.

B. PRIVATE LICENSING IN THE SHADOW OF A COLLECTIVE:
PANDORA V. ASCAP

At the beginning 2014 two of the nation’s three major music publishers attempted to withdraw their digital performance rights from ASCAP, the collective that governs those rights. Presumably, publishers had not attempted withdrawal of these rights before for a couple of reasons: First, because ASCAP’s governing documents didn’t allow for it;

43. *Id.* at 115.

44. *See, e.g.*, David Pakman, *The Price of Music*, RECODE (March 18, 2014), <http://recode.net/2014/03/18/the-price-of-music> [http://perma.cc/28ZJ-5FK7].

and second, because publishers didn't want to. This section follows the temporal progression of the publishers' withdrawal and the ensuing response.

1. *Music Publishers Attempt to Withdraw Digital Public Performance Rights*

On November 5, 2012 Pandora asked the SDNY to settle a rate dispute with ASCAP over digital performance royalties.⁴⁵ ASCAP administers the digital performance rights for two of the country's three major publishers, Sony and Universal. These companies currently hold 31.9% and 18.1% market share, respectively.⁴⁶ Pandora is a non-interactive⁴⁷ digital radio service that licenses digital public performance rights for the songs offered through its various stations. Pandora users create personalized digital radio stations in which they request to hear songs of a certain genre, or that sound like a particular song or artist. These users can then further refine the listening experience by assigning the songs played either a "thumbs up" or "thumbs down."

In July 2005, Pandora and ASCAP negotiated for a blanket license known as a "form license" that ran from 2005 until Pandora canceled it in 2010.⁴⁸ The form license, which the parties named the "5.0 License,"

45. See Order, *supra* note 9, at 6.

46. See Ed Christman, *First Quarter Music Publishing Rankings: SONGS Surges Again*, BILLBOARD (May 12, 2014), <http://www.billboard.com/biz/articles/news/publishing/6084783/first-quarter-music-publishing-rankings-songs-surges-again> [<http://perma.cc/3UYW-DZAT>]. Already a dominant presence in the industry, Sony became even more powerful when it acquired EMI's publishing assets in 2012. See, e.g., Alex Pham, *Sony completes EMI acquisition, creating largest music publisher*, L.A. TIMES (June 29, 2012), <http://articles.latimes.com/2012/jun/29/entertainment/la-et-ct-sony-completes-emi-acquisition-20120629> [<http://perma.cc/C6KG-S9WV>] (citing the close of the merger, and noting a catalog of over 2 million makes the combined entity the largest publisher in the world). See also Ed Christman, *Sam Smith's 'Stay With Me' Is Music Publishing's Big Winner in Q3*, BILLBOARD (Nov. 8, 2014), <http://www.billboard.com/biz/articles/6304471/sam-smiths-stay-with-me-is-music-publishings-big-winner-in-q3> [<http://perma.cc/7JP3-SP96>] (citing the Sony-EMI combined market share at 31.22% in the third quarter of 2014).

47. "Non-interactive" refers to a music service in which the user cannot choose a specific song to listen to, but instead can set parameters such as genre, and artists or albums or songs that the selections should "sound like." The significance of this categorization is that non-interactive services can take advantage of the statutory license. Interactive services – in which users can hear a specific track or artist on demand—cannot use the statutory license and instead must engage in direct negotiations with rights holders. In this way, the Internet radio experience offered by services like Pandora is distinct from digital music platforms like iTunes, on which consumers purchase specific songs and albums, and also from interactive streaming services like Spotify, on which users can populate playlists with specific tracks.

48. See Order, *supra* note 9, at 30.

required Pandora to pay the greater of either 1.85% of revenue, or a per-session rate.⁴⁹ Citing the difficulty of calculating a per-session rate, Pandora notified ASCAP of its intent to terminate the form license on October 28, 2010, at which point the parties negotiated for a new, “through-to-the-audience” blanket license to run from January 1, 2011 through December 31, 2015.⁵⁰

In accordance with the consent decree under which it operates, ASCAP granted the license on a provisional basis, with rates to be determined after the fact by negotiation.⁵¹ Pandora filed its rate court petition after nearly two years of negotiation with ASCAP without an agreement as to rate.⁵² In addition to setting a rate in accordance with Section 9 of the AFJ2, the rate court’s opinion details a disturbing level of cooperation among the major music publishing companies. The opinion also highlights a new and significant development in the licensing of public performance rights: partial withdrawal of rights from a collective in favor of private ordering.⁵³

EMI⁵⁴ was the first of the major music publishers to propose partial withdrawal of its rights from ASCAP. Citing a desire to unify its publishing and recording rights in order to offer licensees a single point of access, as well as a general dissatisfaction with ASCAP’s high administrative costs, EMI informed ASCAP in September 2010 that it was considering withdrawing either its digital rights only, or, failing that, its public performance rights entirely, from ASCAP.⁵⁵ Shortly thereafter,

49. *Id.* at 30–31.

50. *Id.* at 33.

51. *See* United States v. ASCAP, No. 41-1395, 2001 WL 1589999, § 5 (S.D.N.Y. 2001).

52. *See* Order, *supra* note 9, at 34.

53. One might wonder why, if the music publishers were unhappy with ASCAP and BMI’s rates, they didn’t simply take their business to SESAC, the remaining collective. As a small, privately-held collective, SESAC is invitation only, foreclosing this option to all but a select few rights holders.

54. EMI’s musical composition catalog was acquired by Sony in 2012, and EMI no longer exists. *See supra* II.A.3.

55. *See* Order, *supra* note 9 at 39. As discussed in Part II.A.1 *supra*, two different copyrights protect musical compositions and sound recordings, and these copyrights are often held by different entities, and then licensed by different collectives. This was the case with EMI, whose publishing rights were administered by ASCAP, while its sound recording rights were administered by SoundExchange. If EMI were to withdraw its rights from ASCAP and SoundExchange, and instead license both in-house, prospective licensees wouldn’t have to strike two different deals (or so the argument went). In reality, those licensees would likely already have blanket deals in place with each of ASCAP and

Sony and Universal expressed interest in withdrawal as well. There was only one problem at the time: ASCAP's governing articles didn't allow for partial rights withdrawal. As things stood, the publishers either had to withdraw their rights entirely (i.e., digital and analog), or leave them all with ASCAP.

2. *Amendment of ASCAP's Governing Articles*

This state of affairs didn't bode well for ASCAP. Fearing the loss of its largest members (and the associated royalty shares), ASCAP responded with an unprecedented proposal to amend its governing articles so as to allow members to withdraw their digital (also known as "new media") rights, while allowing ASCAP to retain administrative authority over their analog rights.⁵⁶ The proposal was met with skepticism from ASCAP's independent writer-members, and rightfully so. The loss of digital revenue from ASCAP's largest publisher-members would leave the much smaller (and less powerful) independent songwriters footing a much larger portion of the collectives' overhead. This overhead was largely attributed to analog collections—physical monitoring of bars, invoicing of sports parks, etc.—such that those costs would not be diminished much by the digital withdrawals.

In an effort to convince the smaller member hold-outs to nonetheless pass the amendment allowing for withdrawal, the major music publishers took a two-pronged approach: First, they agreed to let ASCAP handle the *distribution* (but not negotiation or collection) of digital royalties, albeit at a much lower administrative fee.⁵⁷ This would keep at least a little of the digital money coming in. Second, the withdrawing publishers argued that by allowing them to go out into the market and obtain higher royalty rates through private negotiation with licensees, those rates could then be

SoundExchange, and this withdrawal would merely necessitate yet another deal, this time directly with EMI. But I digress.

56. The digital rights withdrawals required multiple amendments to ASCAP's governing documents, effectively removing the withdrawing publishers from the auspices of the consent decree that governs ASCAP. See Order, *supra* note 9, at 39. This withdrawal is referred to herein as "partial withdrawal." As it happens, only digital rights might potentially benefit from EMI's purported "rights unification," since only digital rights on the sound recording side are entitled to a public performance royalty. There is no public performance right for sound recordings on the analog side. That is to say, terrestrial radio broadcasters do not pay recording artists when they play their music (they do pay the songwriters), but digital radio broadcasters do. For a full explanation of this conundrum, see García *supra* note 15 at 1134–36.

57. See Order, *supra* note 9, at 53.

presented to the rate court as evidence of a new “market rate,” allowing ASCAP in turn to demand a higher rate for all of its members.⁵⁸

The arguments hit the mark and on April 27, 2011, the ASCAP Board voted to amend its governing articles to allow a member to withdraw its new media rights only.⁵⁹ This amendment is especially significant given that technological developments are likely to result in the conversion of most analog ASCAP licensees to digital in the next few years, rendering ASCAP’s remaining functions obsolete. In addition, the amendment allows a withdrawing publisher to rejoin at any point, making withdrawal a very low risk option.⁶⁰

Within days of the passage of the amendment allowing it to do so, EMI publicly announced the withdrawal of its digital rights, followed by Sony’s announcement in July 2012,⁶¹ and Universal’s in February 2013.⁶² By early 2013, two of the three major music publishers had done the unthinkable—they had withdrawn their digital rights from ASCAP in an effort to increase profits.

The major music publishers correctly understood the AFJ2’s requirement that ASCAP license its repertoire to any and all prospective licensees, preventing ASCAP from using its inherent market power to demand a higher rate.⁶³ This hand-tying frustrated the music publishing companies because it kept their going rate below that of sound recordings. A digital music service like Spotify currently pays \$0.0052 per play to SoundExchange for the public performance rights to a sound recording, while paying only \$0.00052 per play—or one tenth of the rate paid to SoundExchange—to ASCAP for the same public performance rights to the underlying musical compositions.⁶⁴ This disparity reflects a

58. See Order, *supra* note 9, at 45. As explored further at Part II.C.3.b *infra*, this would result in the adoption and enforcement of a potentially misrepresentative “market” rate.

59. See ASCAP, COMPENDIUM OF ASCAP RULES AND REGULATIONS, Rule 1.12 at 10–14 (2014), <http://www.ascap.com/~/.media/files/pdf/members/governing-documents/compendium-of-ascap-rules-regulations.pdf> (hereinafter COMPENDIUM) [<http://perma.cc/9ATQ-QUWM>].

60. See Order, *supra* note 9, at 48.

61. Effective January 1, 2013. See *id.* at 60–61.

62. Effective July 1, 2013. See *id.* at 72.

63. See *id.* at 41.

64. There is little transparency as to rate—neither for licensees nor artists—but the good folks at Trichordist have extrapolated these commonly accepted figures. See *The Streaming Price Bible – Spotify, YouTube & What 1 Million Plays Means to You*, TRICHORDIST (Nov. 12, 2014), <http://thetrichordist.com/2014/11/12/the-streaming-price>

determination by Congress that the nature of the industries for sound recordings and musical compositions vary in material and substantial ways. For example, the overhead required to commission and record an album has traditionally outpaced the cost of signing a songwriter.⁶⁵

In a further effort to differentiate between the two rights, Congress included in the Digital Performance Right in Sound Recordings Act (DPSRA) a clause prohibiting the rate court from taking sound recording license rates into account when setting the rates for musical compositions.⁶⁶ This congressional prohibition on the rate court, however, did not keep the major publishers from taking note. At the Pandora–ASCAP proceeding, Sony’s EVP of Business & Legal Affairs, Peter Brodsky, cited the “massive unfair disparity” between what Pandora pays for sound recordings and what it pays for musical compositions as the principal reason for the company’s withdrawal from ASCAP: “It was the ‘differential’ between the rates paid to the labels and the publishers that was the problem[.]”⁶⁷

By December 2012, the withdrawing publishers had succeeded in wringing yet another amendment from ASCAP—this time, one that would allow them to withdraw their new media rights selectively (on a licensee-by-licensee basis). Publishers could then focus their negotiating efforts on larger, more commercially successful digital partners such as Pandora while leaving the smaller, bootstrap entities—largely believed to be fly-by-night anyway⁶⁸—or ASCAP to deal with.⁶⁹

The second amendment met with even more resistance from ASCAP’s embattled writer-members than the first. Nonetheless, it eventually passed

-bible-spotify-youtube-and-what-1-million-plays-means-to-you [http://perma.cc/9U9R-EKGT].

65. For example, a typical record costs between \$500K–\$2M to produce. *See, e.g., Typical investment by a major record company in a newly signed artist*, IFPI, <http://www.ifpi.org/how-record-labels-invest.php> [http://perma.cc/M6ZC-7WUC]. A typical songwriter’s advance ranges from \$18,000–\$100,000. *See, e.g., DONALD S. PASSMAN, ALL YOU NEED TO KNOW ABOUT THE MUSIC BUSINESS* 269 (7th ed. 2009). *See infra* note 88 and accompanying text.

66. Specifically, the DPSRA provides that “[l]icense fees payable for the public performance of sound recordings . . . shall not be taken into account in any . . . proceeding to set or adjust the royalties payable to copyright owners of musical works for the public performance of their works.” 17 U.S.C. § 114(i) (2012).

67. Order, *supra* note 9, at 63.

68. ASCAP’s VP for New Media and Technology Matthew DeFilippis offered this explanation: “Given the rapidly changing marketplace and the low barriers to entry, new digital music services launch quite frequently. Many will never gain traction with listeners or generate substantial revenue.” *Id.* at 52.

69. *See id.* at 51.

permitting a withdrawing member to indicate that it wished to leave with ASCAP its digital rights only for services paying less than \$5,000 per year in licensing fees.⁷⁰ These smaller services now faced a choice: Accept ASCAP's 5.0 form license without negotiation or go without the content.⁷¹ As a larger digital service, Pandora faced an even more dire fate: Private "negotiation" with the major publishers.⁷²

3. "Negotiation"

Pandora's filing of a rate court petition against ASCAP angered the heads of major music publishing companies, like Universal's CEO, Zach Horowitz, who—through a series of emails and other communications—encouraged ASCAP's then-CEO John LoFrumento to "be strong . . . You can really push Pandora . . . They will pay more . . ."⁷³ Horowitz also called Pandora's counsel on multiple occasions to encourage them to drop the case by suggesting that in representing Pandora they were losing "huge goodwill with writers and artists."⁷⁴ He sent similar communications to ASCAP board members, to BMG Publishing CEO Laurent Hubert, to National Music Publishers Association (NMPA)⁷⁵ President David Israelite, and to Sony's CEO Martin Bandier, whose company had also withdrawn its digital rights and was at that time in the middle of private rate negotiations with Pandora.⁷⁶

Meanwhile, by late November 2012, Pandora believed that it had finally reached an agreement with ASCAP, and that it was just waiting for final approval from the ASCAP Board to terminate the rate court

70. See *id.* at 52, and COMPENDIUM, *supra* note 59, § 1.12.2(b) stating:

If in part, the Member's New Media Transmission Licensing Rights will continue to be licensed and administered by ASCAP for only those 'New Media Services' . . . that: (i) qualify for and are licensed by way of any of ASCAP's form new media license agreements, or, (ii) qualify for and are licensed by an ASCAP 'affiliate' or 'multi-site' license agreement.

71. See Order, *supra* note 9, at 52.

72. Pandora's negotiation with EMI post-withdrawal was reportedly amicable, and the parties settled on the going rate of 1.85% in May 2011 (leaving Pandora paying the same as it would under a license with ASCAP, but saving EMI the administrative fees and royalty shares it would otherwise have had to pay to ASCAP). See Order, *supra* note 9, at 53–54. Because EMI was acquired by Sony shortly thereafter, however—thereby severely limiting the duration of the agreement—the remainder of this Article focuses on the Sony and Universal withdrawals for all intents and purposes.

73. Order, *supra* note 9, at 57.

74. *Id.*

75. The NMPA is the U.S. trade association for music publishing.

76. See Order, *supra* note 9, at 57.

proceeding. But that approval never came because Sony had threatened to sue ASCAP if it reached an agreement with Pandora that encompassed the Sony repertoire.⁷⁷ Sony also hinted that it would withdraw the rest of its rights (i.e., its analog rights) if ASCAP otherwise licensed its repertoire to Pandora.⁷⁸ As a result of pressure from Sony and from Universal, ASCAP's CEO LoFrumento told the ASCAP Board that he planned to reject the terms they had negotiated with Pandora. None of the Board members asked for an explanation, nor requested to discuss the rejection further.⁷⁹

By this time, Sony had acquired EMI's publishing catalog and now controlled roughly 30% of the market for musical compositions.⁸⁰ On October 25, 2012, Sony set the tone for its negotiation with Pandora with this opening remark: "[I]t's not our intention to shut down Pandora."⁸¹ In other words, unless Pandora agreed to Sony's rate or removed all of Sony's content from its service by January 1, 2013, it would face crippling copyright infringement litigation.⁸² Interestingly, Pandora anticipated, and had in fact specifically raised, this concern to the Federal Trade Commission (FTC) during its review of Sony's acquisition of EMI, noting that the combination would leave Pandora "no choice" but to agree to whatever terms Sony demanded as it "could not survive without access to the combined Sony and EMI catalogues."⁸³

In an effort to gain a foothold in its negotiation with Sony, Pandora requested a list of Sony's tracks so that it would know which songs would need to be removed from the service should a rate not be agreed to in time. No such list was forthcoming. According to Judge Cote, "Sony had a list readily at hand," but "understood that it would lose an advantage in its negotiations with Pandora if it provided the list of works and deliberately

77. *Id.* at 59.

78. *Id.*

79. *Id.*

80. *See id.* at 61.

81. *Id.* at 62.

82. *Id.* at 62–63.

83. *Id.* at 62. Nonetheless, in the fall of 2012, the FTC announced that it would not challenge the merger. As discussed in the sections below, this counterintuitive determination is most likely the result of the FTC's mandate under the antitrust laws, which call for intervention in any merger that would "substantially lessen competition." *See* U.S. DOJ & FTC, HORIZONTAL MERGER GUIDELINES ¶¶4–6 (2010) (hereinafter HORIZONTAL MERGER GUIDELINES). In this case, there was already a breakdown in structural competition, so that the proposed merger—while not creating or protecting competition—also did not worsen it.

chose not to do so.”⁸⁴ Pandora turned next to ASCAP, making the same request for a list of Sony’s tracks, but Sony (unsurprisingly) would not give ASCAP permission to release the list.⁸⁵

On January 17, 2013, Sony and Pandora agreed to an undisclosed advance and a rate 25% above ASCAP’s going rate.⁸⁶ In violation of extant confidentiality agreements, Sony shouted this rate from the rooftops, crediting its size for the dramatic rate increase it achieved, and calling the rate “quite reasonable. When you compare it to the rate record companies are getting, it was really miniscule.”⁸⁷

Instead of proffering an efficiency rationale—namely, that digital plays of a song are easily tracked via readily available technologies, and that moving the administration of those rights in-house saves the company royalty shares that it would otherwise have to pay to ASCAP—Sony cited dissatisfaction with the disparity between the publishing royalty rate and the sound recording royalty rate: “We were struck by the vast disparity between what the record companies received from digital music services for the sound recording rights that they conveyed and what was paid for the [musical composition] performance right.”⁸⁸ Historically, the royalty rate for sound recordings has always been higher. This disparity reflects the significantly higher overhead costs involved in granting artist advances, hiring studio time, and pressing and delivering a physical product. In contrast, music publishing does not require investment in a physical product, and little in the way of advances, thereby reducing the overall risk as well. In addition, publishing enjoys a royalty on the terrestrial performance side that does not extend to recordings at all.⁸⁹ Nonetheless, Sony determined that, given its size and market power, it could “really push Pandora and get a much better settlement as a result . . . [Pandora] will pay more, a lot more than they originally intended”⁹⁰

And Sony was right. Sony’s boast prompted Universal to begin its negotiations by demanding even more. In February 2013, Universal

84. *Id.* at 65–66.

85. *Id.* at 68.

86. *Id.* at 69; *see also, e.g., Sony Negotiates 25% Royalty Increase from Pandora: Report*, BILLBOARD (Jan. 17, 2013), <http://www.billboard.com/biz/articles/news/publishing/1510421/sonyatv-negotiates-25-royalty-increase-from-pandora-report> [<http://perma.cc/3SPQ-5ZNS>].

87. *See* Order, *supra* note 9, at 71.

88. *Id.* at 40 (testimony of Martin Bandier).

89. For a full discussion of the legislative history behind terrestrial performance rights, *see* García, *supra* note 15.

90. Order, *supra* note 9, at 57.

announced its withdrawal of new media rights from ASCAP and began its private negotiations with Pandora.⁹¹ Universal's Horowitz knew well the rate his recording counterparts were receiving, as he previously ran that side of the business for Universal. Following Sony's lead, Horowitz opened negotiations with Pandora by saying "we want Pandora to survive," followed by "how did you get Marty [Bandier] at Sony to agree to such a low payment?"⁹²

In response to Universal's prompt demand for an effective rate 50% higher than the prevailing ASCAP rate, Pandora repeated the request it had made of Sony—a request for a list of Universal content for removal from the Pandora service.⁹³ Unlike Sony, Universal provided Pandora with the list, but subject to a non-disclosure agreement (NDA) such that Pandora could not use the list for the purposes of removing any content (because to do so would require sharing it with programmers). As a result, Pandora's options were to either accept Universal's proffered rate, or face a flurry of copyright infringement claims.⁹⁴

On July 1, 2013, the parties provisionally "agreed" to a rate of 7.5% (just under double the ASCAP rate) for a six-month period (i.e., enough time to allow for a rate court decision).⁹⁵ By way of comparison, Universal charged iHeartRadio—a competitor of Pandora, and subsidiary of radio conglomerate Clear Channel—a rate of 1.7%.⁹⁶ This rate is both lower than ASCAP's going rate and nearly a quarter of the rate Universal demanded of Pandora. Like Sony, Universal refused Pandora's request that the provisional agreement not be submitted as evidence of a "market rate" at the rate court proceeding.⁹⁷ Indeed, ASCAP ultimately presented, among other things, both the Sony and the Universal rates as benchmarks for determining a market rate. Meanwhile, Pandora presented 1.70%, or slightly less than the ASCAP rate.⁹⁸

4. *The Rate Court Determination*

After a three-week bench trial that included direct testimony, deposition testimony, and a variety of submitted affidavits, Judge Cote, writing on behalf of the rate court, held that "ASCAP has . . . failed to

91. *Id.* at 72.

92. *Id.* at 74.

93. *See id.* at 74–75.

94. *Id.* at 76–77.

95. *Id.* at 79.

96. *Id.* at 22.

97. *Id.* at 80.

98. *Id.* at 7.

demonstrate that Pandora’s direct licenses with Sony and [Universal] constitute fair market benchmarks.”⁹⁹ Specifically, the rate court determined that “Sony and [Universal] each exercised their considerable market power to extract *supra*-competitive prices,” and that “the evidence at trial revealed troubling coordination between Sony, [Universal], and ASCAP.”¹⁰⁰ Importantly, “ASCAP, Sony, and [Universal] did not act as if they were competitors with each other in their negotiations with Pandora. Because their interests were aligned against Pandora, and they coordinated their activities with respect to Pandora, the very considerable market power that each of them holds individually was magnified.”¹⁰¹

Judge Cote called the Universal-Pandora rate the result of “virtually no meaningful negotiations,” and noted that because Universal “control[s] roughly 20% of the music market, [it] began with and insisted upon a demand that bore no relation to the then-existing market price,” and as such “this license rate cannot be said to represent a bargain arrived at by a willing buyer and seller.”¹⁰² As to Sony’s rate, the S.D.N.Y. found that even if it represented a plausible market valuation, it could not be used as a benchmark because of “the coercive process by which it was negotiated.”¹⁰³

In the end, the rate court set a rate of 1.85% of revenues—the same rate currently available under ASCAP’s 5.0 form license—for every year of the license term (from 2011 through 2015). An interim determination by the rate court additionally granted Pandora’s motion for partial summary judgment by holding that the major publishers’ purported withdrawals of new media rights—and, by extension, ASCAP’s amended Compendium allowing them to do so—violates the consent decree.¹⁰⁴

Specifically, the court noted that Section IV(C) of the AFJ2 prohibits ASCAP from licensing a composition to some services (i.e., smaller digital, and also non-digital) and not others (i.e., larger digital services like

99. *Id.* at 94.

100. *Id.* at 97.

101. *Id.* at 98–99.

102. *Id.* at 105–06.

103. *Id.* at 104.

104. *See In re* Petition of Pandora Media, No. 12 Civ. 8035, 2013 WL 5211927, at 1 (S.D.N.Y. Sept. 17, 2013). The court held:

Because the language of the consent decree unambiguously requires ASCAP to provide Pandora with a license to perform all of the works in its repertory, and because ASCAP retains the works of ‘withdrawing’ publishers in its repertory even if it purports to lack the right to license them to a subclass of New Media entities, Pandora’s motion for summary judgment is granted.

Pandora).¹⁰⁵ The court also interpreted “repertory” as used in the consent decree to include all of the rights pertaining to a work.¹⁰⁶ In other words, the rate court held that publishers wishing to withdraw their digital rights must withdraw all of their rights (i.e., digital and analog). This effectively invalidated Sony and Universal’s attempted withdrawal of digital rights only, putting to the publishers the more difficult determination of whether to withdraw analog rights as well, and how to administer them.

On appeal, the Second Circuit affirmed the rate court’s decision in its entirety.¹⁰⁷ The appeals court confirmed as to rate, finding that “ASCAP failed to carry its burden of proving that its proposed rate was reasonable.”¹⁰⁸ With regard to the preclusion of withdrawals by the consent decree, the Second Circuit confirmed that “[t]he partially withdrawn works at issue remain in the ASCAP repertory pursuant to the plain language of the consent decree.”¹⁰⁹

5. *The Consent Decree*

The Second Circuit’s interpretation of ASCAP’s consent decree was supported by a statement from the DOJ confirming “the consent decree governing the licensing activities of [ASCAP] does not permit ASCAP to accept partial grants of public performance rights.”¹¹⁰ In the face of mounting criticism of the consent decrees as outdated, the Senate Judiciary Committee has undertaken a review of the consent decrees.¹¹¹ As part of this review, the Committee has solicited public comments on, among other things, whether if “partial or limited grants of licensing rights to ASCAP and BMI are allowed, should there be limits on how such grants are structured?”¹¹²

105. *Id.* at 4 (“AFJ2 also contains a provision preventing ASCAP from discriminating in pricing or with respect to other terms or conditions between similarly situated licensees.”)

106. *Id.* at 12–14.

107. *Pandora Media, Inc. v. ASCAP et al.*, 785 F.3d 73 (2d Cir. 2014).

108. *Id.* at 78.

109. *Id.*

110. Brief for U.S. Department of Justice, *Pandora Media, Inc. v. ASCAP et al.*, 785 F.3d 73 (2d Cir. 2014) (No. 14-1158-cv(L)), <http://www.justice.gov/atr/cases/f312700/312748.pdf> [<http://perma.cc/N6FV-TPT9>].

111. See *Antitrust Consent Decree Review*, DOJ <http://www.justice.gov/atr/ascap-bmi-decree-review> [<http://perma.cc/Z3Y2-3KAV>].

112. *Antitrust Consent Decree Review—ASCAP and BMI 2014*, DOJ, <http://www.justice.gov/atr/cases/ascap-bmi-decree-review.html> [<http://perma.cc/4Q8G-HCDS>]. For a review of public comments submitted in response to this, and other, inquiries, see *ASCAP and BMI Consent Decree Review Public Comments*, DOJ, <http://www.justice.gov/atr/public-comments> [<http://perma.cc/GUE3-9HCV>].

On March 10, 2015, the Senate Judiciary Committee held a hearing on the consent decrees. In what is undoubtedly a first for the music industry, comments from such traditionally opposed parties as the National Association of Broadcasters and Public Knowledge (i.e., everyone except ASCAP and the music publishers) reached a common conclusion: The consent decrees are necessary to protect against anticompetitive behavior and to promote competition in online music distribution.¹¹³ The question remains whether this conclusion is true without application to the individual publishers.

C. COLLECTIVE LICENSING V. PRIVATE ORDERING

The mass exodus of the major publishers from ASCAP begs the question: Why would a company who voluntarily joins a collective to handle the difficult and time-consuming work of royalties collection suddenly want to withdraw and do all of the work itself? This question can be answered in two parts: In the first instance, by considering the decline in advantages afforded by collective action in the digital age; and second, by detailing the gains to be had from private over collective action. The final subsection will address potential drawbacks to withdrawing from the collective, including adverse selection, manipulation of market rates, and access concerns.

1. *Diminished Advantages of Collective Action*

The fact that the major music publishers withdrew their digital rights from ASCAP challenges the conventional wisdom that lauds collective rights management as the optimal means of licensing copyrighted content.¹¹⁴ Despite the potential for anticompetitive harm stemming from the aggregation of copyrights,¹¹⁵ collectives like ASCAP have traditionally

113. *Hearing on "How Much for a Song? The Antitrust Decrees that Govern the Market for Music" Before the S. Subcomm. On Antitrust, Competition Policy and Consumer Rights, 114th Cong. (2015) (statement of Mike Dowdle, Bonneville International Corporation) <http://www.judiciary.senate.gov/imo/media/doc/Dowdle%20Testimony.pdf> [http://perma.cc/9PLE-ZLXP].*

114. *See, e.g., Merges, supra note 5.* To some extent, the existence of the consent decrees further cuts against the notion of collective licensing as an optimal solution to the transaction cost problem. *See DiCola & Sag, supra note 31 at 208. See also Ariel Katz, The Potential Demise of Another Natural Monopoly: Rethinking the Collective Administration of Performance Rights, 1 J. COMPETITION L. & ECON. 541 (2005) (challenging the traditional case for PROs).*

115. Some scholars have gone so far as to call collectives' blanket licenses "illegal restraints of trade" insofar as they fix prices for a product that cannot be obtained elsewhere. *See Ivan Reidel, The Taylor Swift Paradox: Superstardom, Excessive Advertising and Blanket Licenses, 7 N.Y.U. J.L. & BUS. 731, 737 (2011) (declaring blanket licenses to*

been justified on the same bases as other collective bargaining models. First, they minimize transaction costs associated with identification of, and negotiation with, multiple contracting partners. In his seminal work on property rules and IP licensing, Professor Robert Merges explains that “[i]t is the high transaction costs associated with the initial [rights] entitlements that lead the parties to establish the [collective]—an organization that then dramatically lowers the costs of exchanging the rights.”¹¹⁶ In other words, collectives “conserve on transaction costs by either making it easier to identify and locate rightholders, or by creating the occasion for repeat-play, reciprocal bargaining.”¹¹⁷

In her Order, Judge Cote explains how those costs are lowered for both licensees and licensors:

Employing ASCAP to perform these [licensing] functions is efficient for music users and copyright holders. A music user can license an enormous portfolio of copyrighted music through the execution of a single license without having to contact each copyright holder. Copyright holders benefit from ASCAP’s expertise and resources in policing the market, negotiating licenses, and distributing the revenue from a vast array of licenses promptly and reliably among the multiple owners of the public performance copyrights in each work. The ability of ASCAP and other performing rights organizations . . . to grant licenses covering a large number of compositions creates significant economies of scale in the market for music licensing.¹¹⁸

Second, collectives help level the playing field in negotiations by concentrating bargaining power and by setting (and enforcing) a collective valuation.¹¹⁹ In an industry like digital music licensing—whose inception came after the establishment of the collectives and for which there is not, and never has been, a real market—collectives have a lot of power in setting the “market” rate. Indeed (with the exception of cases that go to

be “illegal restraints of trade” and calling for urgent reform of DOJ enforcement practices). While this Article recognizes the growing strength of these (and other) arguments, they do not negate the valuable role that collectives serve in ensuring access to content, and in funneling royalties to artists. *See infra* Part V.B.4.

116. *See* Merges, *supra* note 5, at 1302–03.

117. *Id.* at 1294. Unfortunately, this doesn’t mean that collectives fare any better than statutory licenses at differentiating amongst individual content valuations. *See* García, *supra* note 15, at 1128–29. In the collectives’ defense, however, it may be the case that tailoring is not worth the effort in the bundled licensing context.

118. Order, *supra* note 9, at 11.

119. *See* Merges, *supra* note 5 at 1294 (noting that CROs function by “promulgat[ing] rules and procedures for placing a monetary value on members’ property rights.”)

the rate court), ASCAP sets the rate for the content that it administers. But this can hardly be called a “market rate” in the absence of a market. To the extent opt-out speaks to inefficiency, we need look no further than the private “success” of publishers like Sony and Universal for a challenge to the current system.

Technological development and concomitant changes in consumer behavior and business models have challenged the traditional justifications for collectives. For example, searchable online databases have made identifying content owners easier, greatly reducing the transaction costs associated with pairing licensees and licensors. For better or for worse, industry consolidation has decreased the number of smaller, independent content owners that stand to benefit from pooling their content and bargaining power.

With respect to digital content, audio matching technologies have made it both possible and cost-effective for content owners to track their repertoire’s usage themselves. As more traditional analog outlets—for example, brick-and-mortar retail stores and restaurants—transition to digital streaming, the proportion of royalties attributed to analog sources continues to shrink, eventually allowing a licensor to administer most of its rights itself. In short, the need for collectives is waning in the digital world. Additionally, as Professor Daniel Gervais has put it: “Initially, CMOs developed out of necessity; it was not feasible for authors and publishers to maintain a direct relationship with users. With the advent of new technologies, however, authors and publishers are increasingly able to initiate and maintain a direct relationship with users.”¹²⁰

2. *Advantages of Private Ordering*

To be clear, the partial withdrawal of rights from ASCAP by the major music publishers is a form of private ordering. One interpretation of the events described in the Pandora–ASCAP proceeding is consistent with a functioning free market in which a powerful party (or parties) naturally emerges, through ingenuity and hard work, and in which a market rate reflecting supply and demand is eventually established. In other words, proof positive of free market economics.

Under this view, Sony came to a position of market power through decades of diligent content acquisition and strategic consolidation. Also under this view, the rate achieved in private negotiation with Pandora reflects the value of Sony’s content to Pandora’s Internet radio service.

120. Gervais, *supra* note 31, at 27.

Otherwise, Pandora would simply take its business elsewhere or shut down its service.

Another (admittedly oversimplified) reason a company might choose to withdraw from a collective to negotiate privately is that the company may be able to do better in the private market (where “better” means that they can make more money). Previous scholarship examined the potential advantages of private ordering around an inefficient statutory regime, and this is arguably another example of improvement on the status quo. There, the potential gains from private deal making include a higher rate, the ability to tailor terms to a specific piece of content, use, or partner, and the flexibility to more easily and readily adjust deal terms in response to new technologies or business models.¹²¹

In the music industry, these advantages have already been observed on the sound recordings side of the business, where private ordering around a statutory license initially emerged in 2012. When iHeartMedia¹²²—a large media conglomerate and owner of hundreds of radio stations nationwide—and Big Machine—a relatively small record label and home of recording artist Taylor Swift—circumvented § 114 in favor of private ordering, the parties benefited in several ways. First, Big Machine gained the right to be paid on terrestrial plays (a nonexistent right under current laws), while iHeartMedia secured a lower-than-statutory rate on digital plays. This scheme was far better tailored (pun intended!) for a radio-friendly artist like Taylor, and its short, renewable term allowed the parties to make easy adjustments in the future. Finally, and happily for the parties, circumvention and private ordering vis-à-vis one partner did not necessitate private negotiation across the board with all partners. In addition to lowering the risk associated with circumvention, private ordering allowed for greater specificity as to the content and uses contemplated by the parties.

3. *Potential Drawbacks to Private Ordering*

Despite its purported advantages, private ordering in the shadow of a collective, and particularly in the context of a highly regulated market like music publishing, presents a variety of concerns—including adverse

121. See García, *supra* note 15, at 1146–51.

122. Formerly, Clear Channel Communications, Inc. On September 16, 2014, the company changed its name to reflect a new commitment to digital media. See Andrew Hampp, *Clear Channel Changes Name to iHeartMedia*, BILLBOARD (Sept. 16, 2014), <http://www.billboard.com/articles/business/6251599/clear-channel-changes-name-iheartmedia> [<http://perma.cc/A3V2-F3ZM>].

selection, misrepresentation of market rates, and denial of access to content.

a) Adverse Selection

As noted in Part II.B.2 *supra*, not everyone at ASCAP was on board with the amendments to ASCAP's governing documents that permitted selective withdrawal of new media rights. Notably, the individual songwriter-members expressed concern about the loss of transparency in the accounting and distribution of royalties to songwriters if the individual publishers took over those functions. Specifically, "[t]hey were concerned . . . that the publishers would not manage with as much care the difficult task of properly accounting for the distribution of fees to multiple rights holders, and might even retain for themselves certain monies, such as advances, in which writers believed they were entitled to share."¹²³ Moreover, ASCAP's writer-members recognized that a withdrawal of digital rights represented a withdrawal of cash (i.e., lost royalty shares) from the collective, leaving the small, independent songwriters and publishing companies footing a larger share of the cost of daily operations at ASCAP.¹²⁴

Resistance intensified with introduction of the second amendment allowing for selective withdrawal. The independent songwriter members were unhappy with the monies that ASCAP was losing as a result of the new media withdrawals and did not want to allow the majors to continue to weaken the ASCAP organization.¹²⁵ They were also upset by what they viewed as a bait-and-switch by the major publishers who had promised to go out into the market and secure for the industry a higher "market rate." Instead, Sony directly licensed DMX (a company that provides digital music services for retail stores) at a rate considerably lower than the going ASCAP rate, in exchange for a hefty advance.¹²⁶ DMX then took that

123. See Order, *supra* note 9, at 42.

124. See *id.* at 45. These costs include administrative, litigation, and advocacy expenses. As the bulk of administrative costs stem from the analog side of the business (which rights were retained by ASCAP), withdrawal of digital rights did little to reduce overhead costs while dramatically reducing revenue.

125. See *id.* at 51.

126. See Steve Gordon, *Direct Licensing Controversy: Will Publishers Be Able To License Public Performing Rights To Digital Music Services Directly (Instead of through the PROs) and What Are the Consequences for Songwriters?* FUTURE OF THE MUSIC BUSINESS (May 27, 2014) <http://www.futureofthemusicbusiness.biz/2014/05/direct-licensing-controversy-will.html> [<http://perma.cc/4DPR-WPE4>] ("In 2007 Sony negotiated a direct deal with DMX, the digital background music service. In doing so, it received an advance payment

rate—sans mention of the advance—to the rate court to secure a lower rate for itself across the board.¹²⁷ The primary proponent of this second amendment to ASCAP's Compendium was Sony, who pushed for its passage despite receiving legal advice raising “antitrust concerns with the carve-out proposal.”¹²⁸

As a result of the major publisher withdrawals, ASCAP's remaining members—the songwriters themselves, and the smaller, independent publishers who lack the ability to demand private negotiation—were left with a weaker organization. Not only was ASCAP left touting a less robust and less valuable blanket license, but it was also collecting less in royalty shares. This put the full burden of covering the costs of operation—including litigation, lobbying, and advocacy—on the smaller members left behind.

Prior work has examined the phenomenon of private ordering around an unpalatable statutory license, or “penalty default license,” as the inevitable result of inherent inefficiencies and bounded uncertainty.¹²⁹ The proliferation of private ordering in the market for collectively licensed IP rights challenges the conventional view of collectives as the optimally efficient licensing mechanism, and instead demonstrates that they suffer many of the same inefficiencies as statutory licenses.¹³⁰ Specifically, it shows that in some cases, an individual party can privately negotiate a higher rate than a collective can command. As a consequence, dominant market players who opt out of the collectives, leave behind weaker players with a weaker organization and set off a downward spiral that has

of 2.7 million dollars. It is doubtful whether Sony's writers received any portion of this money.”).

127. *See id.*; Order, *supra* note 9, at 51 n.36.

128. *Id.* at 53 n.37.

129. *See* García, *supra* note 15.

130. N.B.: ASCAP has argued that these inefficiencies stem from the consent decrees themselves:

It is clear that the legal and regulatory restrictions imposed on ASCAP by the Consent Decree and the Copyright Act severely limit ASCAP's members from achieving competitive rates for their works. Indeed, as the Copyright Office noted in its recent report on Copyright and the Music Marketplace, “[t]here is substantial evidence to support the view that government-regulated licensing processes imposed on publishers and songwriters have resulted in depressed rates.”

Hearing on Performance Rights Organization Consent Decrees, 114th Cong. 14 (2015) (statement of Elizabeth Matthews, ASCAP (citing U.S. COPYRIGHT OFFICE, COPYRIGHT AND THE MUSIC MARKETPLACE (2015)), <http://www.judiciary.senate.gov/imo/media/doc/Matthews%20Testimony.pdf> [<http://perma.cc/HD4C-BMFG>]).

negative overall welfare consequences despite the apparent benefits to early-departers.

Adverse selection—the term used herein to describe this defection of powerful members from their respective collectives—decreases the efficiency and efficacy of the abandoned collective. These consequences stem, in large part, from the fixed costs associated with collective rights administration and enforcement. In addition to leaving collectives with reduced income from membership fees, adverse selection may leave remaining members to suffer reduced royalties as a result of a diminished catalog, reduced bargaining power, and weaker enforcement capabilities.

These fixed costs are difficult to amortize and may result in a system with higher gross transaction costs—including litigation and lobbying expenses—that must be borne by the remaining members. It is true that a publishing company has no legal or economic obligation to forego increased revenues in order to avoid making things more difficult for individual writer-members. Nonetheless, there are potential trickle-down effects for the entire system. Where the financial burden of sustaining the collective becomes unsustainable for the weaker members, adverse selection can lead to diminished distribution of content to consumers. It also affords defecting members the opportunity for anticompetitive behavior to the detriment of innovation in both the creation and dissemination of content.

b) Manipulation of “Market” Rate

The promised upside of this adverse selection—the notion that the higher rate negotiated by the majors would go on to be adopted as the “market” rate to the benefit of all—not only proved false, but actually went the other way when DMX presented its *lower* rate (sans advance) for adoption instead.¹³¹ The DMX example illustrates how circumvention of a collective in favor of private can lead to misrepresentation of “market” rates. This is especially true in an industry like music publishing, where no

131. See *Broad. Music, Inc. v. DMX, Inc.*, 683 F.3d 32, 49 (2d Cir. 2012) (confirming adoption of DMX’s proposed rate); and Steve Gordon, *DMX v. BMI Demonstrates that Digital Services May Use Direct Licensing to Reduce Their Payments to the PROs but the Decision May Be Reversed on Appeal*, THE FUTURE OF THE MUSIC BUSINESS (July 12, 2011), <http://www.futureofthemusicbusiness.biz/2011/07/dmx-vs-bmi-demonstrates-that-digital.html> [<http://perma.cc/72T3-6HPR>] (discussing DMX’s successful rate reduction campaign based on the negotiated Sony rate and asking “how can 550 direct licenses [the number held by Sony] be a benchmark for the true value of the [collective’s] blanket licenses when those 550 licenses represent, in probability, only a tiny fraction of the songs represented by the [collective]”).

real market exists outside of the collectives that have operated more or less since the industry's inception.

It is unclear whether the rates obtained by Sony and Universal are “right” or whether they are supracompetitive (as Pandora has alleged). The going ASCAP rate may not be “right” either—it may well be, as the music publishing industry has alleged, the result of Pandora’s manipulations. Indeed, in a move labeled “cynical and shameless,” Pandora recently purchased a small, terrestrial radio station in South Dakota in an unabashed effort to take advantage of the lower statutory rate offered Internet radio services owned by broadcast radio stations.¹³²

In the context of the ASCAP withdrawals, the major music publishers also enjoy a first-movers advantage. As first to withdraw their digital rights and first to go out into the “market” to negotiate a private rate, they are able to set a baseline rate that subsequent private deals are most likely to adopt, either by example or by peer pressure. In other words, if Pandora agrees to a rate double that of ASCAP’s, so too might their competitors (if they want to remain competitors). This can lead to the establishment of an industry norm or custom that is based on little more than coercion and desperation as other digital distribution services are pressured to pay the same rates or shut down.¹³³

c) Collectives’ Role in Ensuring Access To Content

With all of their faults and potential drawbacks, collectives like ASCAP serve an important function that private ordering might otherwise foreclose: access to content for all prospective licensees willing to pay an agreed-upon rate and meet agreed-upon terms. This access is enforced by consent decree. The AFJ2 directs ASCAP to “grant to any

132. David Israelite, *Statement of National Music Publishers Association (NMPA) CEO David Israelite*, NMPA <http://www.billboard.com/articles/news/6553964/pandora-radio-station-fcc-approval-kxmz> [<http://perma.cc/5U68-6262>]. For more on the radio station purchase, see Glenn Peoples, *Pandora Buys Terrestrial Radio Station in South Dakota, Aims for Lower ASCAP Royalties*, BILLBOARD (June 11, 2013), <https://www.billboard.com/biz/articles/news/radio/1566479/pandora-buys-terrestrial-radio-station-in-south-dakota-aims-for> [<https://perma.cc/AM4F-FKB8>]. The very existence of a different, lower rate structure for an Internet radio station owned by a broadcast radio station versus a free-standing Internet radio station illustrates the byzantine nature of the heavily lobbied music industry.

133. For a critique of the influence of custom on law in the IP context, see Jennifer E. Rothman, *The Questionable Use of Custom in Intellectual Property*, 93 VA. L. REV. 1899 (2007). For the same critique in other contexts, see, for example, Lisa Bernstein, *The Questionable Empirical Basis of Article 2’s Incorporation Strategy: A Preliminary Study*, 66 U. CHI. L. REV. 710 (1999) (criticizing the incorporation of commercial customs into Article 2 of the UCC).

music user making a written request therefor a non-exclusive license to all of the works in the ASCAP repertory.”¹³⁴ As a result, ASCAP cannot license its repertory to one Internet radio service and refuse to license another. Individual music publishers like Sony and Universal can make such refusals, however. This is particularly concerning where the major publishers effectively control an asset essential to any music streaming service: music.¹³⁵

The concern about access to content is nothing new. Congress has long recognized the importance of access to the maintenance of competition in the content industries. For example, the cable industry operated for many years under so-called “program access rules,” which required owners of cable programming to make that content available to rival distributors.¹³⁶ The program access rules were promulgated as a central part of the Cable Television Consumer Protection and Competition Act of 1992 (the “Cable Act”), and have been credited with driving the growth of the satellite television industry.¹³⁷

When the program access rules were enacted, lawmakers were concerned that content owners would refuse to license content to competitors.¹³⁸ As a result of these rules, satellite companies like Dish and

134. *United States v. ASCAP*, No. 41-1395, 2001 WL 1589999, § VI (2001).

135. It may be reasonably argued that while *some* music is essential to a music streaming service, one publisher’s content is largely interchangeable with another’s. After all, if Honey Nut Cheerios suddenly refused to sell its product to Grocery Store A (and it is allowed to do so), Grocery Store A could simply sub in generic Nutty O’s. Diehard fans of the original may balk, but would hardly view Grocery Store A as being competitively disadvantaged. Part IV.B.2.c *infra* will attempt to show that, for a variety of reasons, it’s a bit more complicated than that in the music context.

136. *See* 47 U.S.C. § 548 (2012).

137. *See, e.g.*, Comments of the Satellite Broadcasting and Communications Association Before the Federal Communications Commission at 17 (July 29, 2002), http://www.sbca.com/PublicAffairdocs/Competition_2002Comments.pdf (“Industry observers credit the program access rules . . . as one of the most important factors in the rise of DBS as a successful consumer service and competitor to cable.”) [<http://perma.cc/V3MY-NNZK>]; Joe Flint, *FCC lets program access rules expire*, L.A. TIMES (Oct. 5, 2012), <http://articles.latimes.com/2012/oct/05/entertainment/la-et-ct-fcc-program-access-20121005> [<http://perma.cc/EN3L-FJFG>].

138. *See* JONATHAN E. NUECHTERLEIN AND PHILIP J. WEISER, DIGITAL CROSSROADS: TELECOMMUNICATIONS LAW AND POLICY IN THE INTERNET AGE 343–47 (2d ed. 2013) [hereinafter DIGITAL CROSSROADS]. Nuechterlein and Weiser note:

The concern underlying these requirements is that a cable incumbent, left to its own devices, might withhold affiliated programming from its MVPD [multichannel video program distributor] rivals in the hope that the programming is so indispensable to the television experience of

DirecTV grew and proved to be viable competitors. With the growth of these companies, moreover, the FCC has retained these rules, but now questions how much longer the rules will stay in place.¹³⁹

Unfortunately, the music publishing industry does not currently enjoy robust competition, and collectives play a valuable role in providing—though not necessarily guaranteeing—access. As Part V *infra* will discuss, only a mandatory, compulsory license can guarantee access. But first, the next Part discusses a final, paradoxical category of concern stemming from the withdrawal, and private negotiation, of public performance rights: anticompetitive behavior.

III. AN ANTITRUST PARADOX

A. ANTICOMPETITIVE BEHAVIOR

Perhaps the most unexpected development highlighted by the Pandora–ASCAP rate proceeding is the growing anticompetitive threat posed not by ASCAP, but by the individual music publishers acting in loose coordination to demand more money, or, failing that, to deny access to their content altogether—something that ASCAP, per the consent decree, cannot do. The coordination of effort among the major music publishers and ASCAP went beyond written “encouragement” during negotiation—Universal’s Horowitz emailed both LoFrumento at ASCAP and fellow publishing heads at Sony and Warner/BMG, urging them to “[b]e strong You can really push Pandora and get a much better settlement as a result. They are reeling. They will pay more, a lot more than they originally intended” to include public announcement of terms afterward, even where such disclosure was in clear contravention of extant

many viewers that they will forgo the rivals’ service in favor of the cable incumbent’s service.

139. In the Matter of Revision of the Commission’s Program Access, 27 FCC Rcd. 12605, No. 12-123, at 4 (2012) (“[A] preemptive prohibition on exclusive contracts is no longer ‘necessary to preserve and protection competition and diversity in the distribution of video programming.’”) (quoting *Cablevision Systems Corp. v. F.C.C.*, 597 F.3d 1306, 1308 (D.C. Cir. 2010)). In *Cablevision*, the court went on to note that “We anticipate that cable’s dominance in the MVPD [multichannel video program distributor] market will have diminished still more by the time the Commission next reviews the prohibition, and expect that at that time the Commission will weigh heavily Congress’s intention that the exclusive contract prohibition will eventually sunset.” *Id.* at 1314. See also DIGITAL CROSSROADS, *supra* note 138, at 344 (“In October 2012, the FCC denied a third five-year extension, allowed the flat ban on exclusive contracts to expire, and announced that it would continue to review such exclusive contracts on a case-by-case basis under section 628(b).”)

NDA^s.¹⁴⁰ “By mid-January 2013, and despite the existence of a confidentiality agreement, Sony leaked the key terms of the Pandora license to the press. . . . [Sony’s] Bandier was quoted as saying that ‘[a]t the end of the day, we got a terrific deal for our songwriters. Our thinking has been vindicated.’”¹⁴¹

This announcement by Sony not only informed the other publishers of the rate Sony had achieved, but also informed the songwriters that those publishers compete for. In other words, if Universal were to obtain less, its songwriters could simply go over to Sony, making Sony’s announced rate the *de facto* floor. Together with a demand that Pandora remove the respective publishers’ content, coupled with a refusal to identify that content, and amid threats of massive copyright litigation,¹⁴² this low-level, coordinated effort is referred to herein as “tacit collusion.”¹⁴³

A highly concentrated industry like music licensing is especially vulnerable to this form of anticompetitive conduct. The more concentrated the market, the greater risk to competition posed by coordination.¹⁴⁴ Some industries—like telecommunications, or utilities—face a set of circumstances in which multi-firm production is more costly than single-firm production, and so are said to tend toward “natural monopoly,” an economic condition in which optimal efficiency is reached with only one firm in the market.¹⁴⁵ The typical natural monopoly industry faces very high start-up costs. A prospective entrant in such an industry would be forced (at prohibitive cost and great inefficiency) to recreate necessary infrastructure already put into place by the dominant firm.¹⁴⁶

An oligopoly, on the other hand, is found where a few large, powerful firms dominate a market, so that the actions of each firm impact the other firms, and in which each firm recognizes a strategic interdependence with

140. Order, *supra* note 9, at 57.

141. *Id.* at 71.

142. *See supra*, Part II.B.3.

143. As discussed in detail in Part III.B *infra*, tacit collusion is distinct from explicit collusion in that the former is not a violation of antitrust law.

144. *See, e.g.*, Carl Shapiro, *Theories of Oligopoly Behavior*, in THE HANDBOOK OF INDUSTRIAL ORGANIZATION 329 (R. Schmalensee & R.D. Willig, eds. 1989); George Stigler, *A Theory of Oligopoly*, J. POL. ECON., 72:1 (1964).

145. *See generally*, William J. Baumol, *On the Proper Cost Tests for Natural Monopoly in a Multiproduct Industry*, AM. ECON. REV. 67 (1977).

146. *See generally* LOUIS KAPLOW & STEVEN SHAVELL, MICROECONOMICS 42 (2004) (noting, for example, that in a natural monopoly, “the competitor’s profit would likely not be sufficient to cover the cost of building another network”)

the other firms.¹⁴⁷ As these firms' combined market share increases, their strategic interdependence increases. When the minimum efficient scale—defined as the lowest level of output at which a firm can operate at minimum cost per unit—is a relatively large percentage of the market, smaller firms cannot compete, and the market becomes a natural oligopoly.¹⁴⁸ This is the case in music publishing: An industry comprised of only three firms, each with double-digit market share, demonstrates a willingness and ability to tacitly collude with respect to price. The oligopoly condition allows for “a group of firms to reach a collusive outcome without overt acts of detectable communication. Such tacit collusion results from a ‘meeting of the minds,’ whereby competitors recognize that it is in their collective best interests to set price or quantity equal to the collusive level.”¹⁴⁹ This behavior is commonly known as parallel pricing. While not technically illegal—indeed, and as discussed further in the next Part, “application of the antitrust laws [in this context] becomes challenging”¹⁵⁰—parallel pricing is nonetheless problematic precisely because it is difficult to monitor and enforce against.¹⁵¹

Network effects have further exacerbated the effects of consolidation in the music industry by removing the threat of meaningful competition from the marketplace. Network effects (also called “network externalities”) describe the phenomenon—especially prevalent in technology companies—of an increased value to a consumer deriving from an increase in the number of other consumers enjoying the same product or service. One is not likely to choose a laundry detergent based on which brand his closest friends use, for example, nor is he likely to begin getting more out of his laundry detergent if his neighbors adopt the same brand. In the case of a video game console, however, people are more likely than not to adopt the brand—for example, the Xbox—that their friends are using, as this allows them to play with together on the same platform. As more and more people also purchase an Xbox, individual enjoyment of the device grows,

147. *See generally* HAL R. VARIAN, INTERMEDIATE MICROECONOMICS: A MODERN APPROACH 468–91 (5th ed. 1999) (discussing oligopoly and oligopolistic pricing behavior).

148. *See id.* at 427 (defining minimum efficient scale as “the level of output that minimizes average cost, relative to the size of demand”)

149. Alan Devlin, *A Proposed Solution to the Problem of Parallel Pricing in Oligopolistic Markets*, 59 STAN. L. REV. 1111, 1112–13 (2007).

150. *Id.* at 1113.

151. *Cf.* RICHARD A. POSNER, ANTITRUST LAW 94 (2d ed. 2001) (arguing the antitrust laws ought to be able to identify and regulate parallel pricing). This view has not been adopted.

both in terms of new game development and a community to play with. This is an example of network effects.

In music, network effects play a role on both the content and the distribution side. In distribution, as in the Xbox example, consumers are more likely to join the music streaming platform that their friends are on so that they can share playlists, for example. When it comes to the songs themselves, network effects have been shown to have a significant influence on consumer tastes and preferences, leading to what this article will term psychological, as opposed to technological, lock-in. For example, a 2007 study of over 14,000 listeners took a unique approach to measuring what types of things influence whether a song will become a “hit.”¹⁵² The authors created an experiment in which the first set of listeners heard a series of songs and—without any information other than the name of the artist and title of the song—were asked to rank them in order of preference. The second set of listeners heard the same series of songs, only this time they were also given the number of times each song had been downloaded by others in their social circle. With the second group, listeners showed a clear preference for songs that others had preferred. The study concluded that “social influence played as large a role in determining the market share of successful songs as differences in quality.”¹⁵³

For this reason, songs are not truly substitutable. Listeners are as locked-in to the songs their social group listens to as they are to the gaming platform their social group plays on. As a result, a fledgling streaming service cannot compete without a catalog of popular songs. This enables the owners of popular songs (i.e., the major music publishers)—acting alone or in tacit collusion with similarly situated entities—to act anticompetitively by threatening to withhold their content from a service or by offering favorable rates to one service over another.¹⁵⁴

152. See Salganik et al., *Experimental Study of Inequality and Unpredictability in an Artificial Cultural Market*, 311 *SCIENCE* 854 (2006).

153. Duncan J. Watts, *Is Justin Timberlake a Product of Cumulative Advantage?*, *N.Y. TIMES* (April 15, 2007), <http://www.nytimes.com/2007/04/15/magazine/15wwlnidealab.t.html> [<http://perma.cc/S8GR-U898>].

154. In a slightly different context, Francesco Parisi and Ben Depoorter have suggested that in the case of complementary inputs to a derivative work, “price coordination and monopolistic pricing do not in all circumstances produce inefficient equilibria.” *The Market for Intellectual Property: The Case of Complementary Oligopoly*, in *THE ECONOMICS OF COPYRIGHT: DEVELOPMENTS IN RESEARCH AND ANALYSIS* 25, (W. Gordon & R. Watt, eds., 2003). This suggests that the welfare effects of price coordination and competition may depend on the type and nature of the goods involved, and that structural competition doesn’t necessarily solve the problem either.

In the case of music publishing, Sony and Universal tacitly colluded to charge Pandora a higher rate than other Internet radio services were paying through ASCAP. Given the publishers' combined market share of roughly 50%, a prospective entrant into the music streaming space could be easily blocked by the major publishers' joint refusal to deal.¹⁵⁵

B. CONSEQUENCES OF ANTICOMPETITIVE BEHAVIOR

There are a variety of reasons to care about anticompetitive behavior on the part of music publishers. First, it represents a very real threat to innovation in both the creation and distribution of content. If publishers go after the most successful companies in a given space—as they did with Pandora in the music streaming space—that risks a situation in which the companies that control an essential input (in this case, music) in one industry, can threaten to stifle innovation in another, related industry (in this case, streaming distribution). To be clear, this is not an argument in favor of Pandora's preservation, but rather for the preservation of multiple distribution options.

These content owners can also block access to content altogether. Without a statutory license, there is no longer guaranteed access to content, and nothing to stop a content owner who wishes to start its own streaming service from withholding its content from potential competitors. Instead of requiring thousands of individual negotiations, a digital radio service today can obtain upward of 80% of all of the music publishing rights that they need from a mere three companies. And if those three companies cooperate on rate, explicitly or tacitly, the remaining rights holders will eventually be forced to take that rate in order to remain competitive. Again, this is not an argument in favor of ASCAP's preservation necessarily, but rather an argument against the foreclosure of access to content that enables new entry into the market.

155. This is so despite the fact that most streaming customers actually consume a mere 5% of the content offered by a typical music service. *See, e.g.,* Paul Resnikoff, *95 Percent of Streaming Music Catalogs Are 'Irrelevant' to Consumers, Study Finds*, DIGITAL MUSIC NEWS (Sept. 10, 2015), <http://www.digitalmusicnews.com/2015/09/10/95-percent-of-streaming-music-catalogs-are-irrelevant-to-consumers-study-finds> [<http://perma.cc/53XV-TZTU>]. Rather, it is consumers' perspective of what they're getting that matters: a user simply feels better paying \$9.99 for "millions of songs" than for what they actually use, which is probably closer to "tens of songs." For an intriguing argument against more content for the sake of more content, see Michael Abramowicz, *An Industrial Organization Approach to Copyright Law*, 46 WM. & MARY L. REV. 33, 37–38 (2004) (suggesting that "changes along the edges of copyright law that lead to slight reductions in the number of works produced but greater dissemination of other works could increase social welfare")

Some may argue that this behavior is not a real cause for concern. After all, free market theory tells us that if a company like Sony or Universal is indeed over-pricing its product or otherwise behaving badly, the market will correct such behavior by taking its business elsewhere, by admitting new entrants, and/or by otherwise reacting in such a way as to drive the price back to a mutually agreeable level. In that case, the argument could be made that concerns about tacit collusion or parallel pricing are unwarranted. For example, in the IP context the very purpose of IP rights is to protect copyright owners from competition for a limited period of time. To the extent competition is suppressed, then, this is nothing more than the system at work.¹⁵⁶

These arguments based on the free market theory assume a competitive market, and the problem is that the protection from competition afforded content owners by copyright law does not extend (or, is not intended to extend) to distributors of that content. For example, as a content owner, Sony could decide to favor a particular distribution company (whether by affiliation, or by contractual agreement, or for kicks) and charge other distribution companies a different amount or decline to license its content to other distribution companies altogether. This protects Sony's favored distributor via an unintended extension of its lawful copyright protections. Notably, under the AFJ2, ASCAP is specifically prohibited from exhibiting such favoritism, while no such restriction applies to Sony as an individual publisher. In other words, the problem with the free market theory as applied to music licensing is that there is no competitive market.

While economic theory is not generally hostile to price discrimination in a competitive marketplace,¹⁵⁷ where structural competition has been compromised such discrimination can cut against innovation by shrinking the field of competitors. In other words, the problem is not the price itself, but rather the lack of competition. If a music publisher like Sony decides to go after Pandora, the most successful distribution company in the Internet radio space, the result is a situation in which a company with control over a necessary input—in this case, music—can drive a leading

156. Indeed, Professor Lydia Pallas Loren has suggested that “[t]he entire structure of the sound recording digital public performance right is statutorily geared to protect incumbents.” *The Dual Narratives in the Landscape of Music Copyright*, 52 HOUS. L. REV. 537, 577 (2014).

157. See Joseph Farrell & Philip J. Weiser, *Modularity, Vertical Integration, and Open Access Policies: Towards a Convergence of Antitrust and Regulation in the Internet Age*, 17 HARV. J.L. & TECH. 85, 108–09 (2003).

innovator out of the market. After all, together Sony and Universal own roughly half of all digital music copyrights. A digital music service would be hard pressed to compete effectively without this content.¹⁵⁸

Indeed, this threat has arisen before. In the wake of Napster's shut down in 2001, the major labels formed MusicNet and PressPlay with the intent of distributing their own content. At that time, there were five major labels: AOL Time Warner, Bertelsmann and EMI formed MusicNet, and Vivendi Universal and Sony formed PressPlay. Together, the five labels—through the newly-formed MusicNet and PressPlay entities—controlled over 80% of all commercial music.¹⁵⁹ This consolidation of market power predictably led to a DOJ investigation of allegations from small online music services that the companies “refused licenses . . . because they did not pony up hundreds of thousands of dollars for negotiations,” and that “MusicNet and Pressplay could potentially exclude them from online distribution deals.”¹⁶⁰ Although the inquiry ended in 2003 without further action on behalf of the DOJ, both companies were promptly broken up and sold off.¹⁶¹

In the short term, Pandora suffers the most obvious injury in the form of higher licensing costs resulting from a higher royalty rate—to be clear, a rate both higher than Pandora was previously paying and higher than that faced by its peers. Importantly, this is not strictly a private harm vis-à-vis Pandora, but rather a harm that extends to the broader market for music licensing. Some music consumers can expect to see higher subscription rates and/or more advertising as Pandora passes along some of its increased costs, thereby reducing consumer welfare. Other consumers,

158. Collectively, the withdrawal of new media rights by Sony (with the acquisition of EMI) and Universal reduced ASCAP's revenues by half. *See* Order, *supra* note 9, at 342.

159. *See* Kelly Donohue, *MusicNet & PressPlay: To Trust or Antitrust?*, 1 DUKE L. & TECH. REV. 1, 2 (2001) (noting, in addition, that since “the labels have either bought or sued remaining competitors, MusicNet and PressPlay virtually control the online music industry through the use of its copyrights”).

160. *Justice Dep't Begins Probe into MusicNet, PressPlay*, BILLBOARD (Aug. 7, 2001) <http://www.billboard.com/articles/news/78850/justice-dept-begins-probe-into-musicnet-pressplay> [<http://perma.cc/X7TF-D5E3>].

161. *See, e.g.*, Jon Healey, *N.Y. Firm Buys Out MusicNet*, L.A. TIMES (Apr. 13, 2005) <http://articles.latimes.com/2005/apr/13/business/fi-musicnet13> (“A New York venture capital firm said Tuesday that it had taken over MusicNet, an online music service formed in 1999 by three major record companies and RealNetworks Inc.”) [<http://perma.cc/4RHY-F6ME>]; and *Roxio Rebranding Pressplay As Napster*, BILLBOARD (May 20, 2003), <http://www.billboard.com/articles/news/70995/roxio-rebranding-pressplay-as-napster> (“Roxio has acquired online music service Pressplay from Universal Music Group (UMG) and Sony Music Entertainment and plans to revive it under the Napster name before March 2004.”) [<http://perma.cc/8A2S-PXXL>].

unable to afford the higher subscription rates, may find themselves forced either into the ad-supported environment or out of the market altogether.

In the long term, tacit price collusion may reduce allocative efficiency by affording the coordinating companies the opportunity to engage in parallel pricing so as to raise prices supracompetitively¹⁶² or discriminatorily. The societal harm resulting from parallel pricing takes several forms. First, as seen previously, some consumers may be priced out of the market. Second, innovation may decrease if prospective licensees cannot afford the new “market” rate, and so are deterred from entering in the first place.

In a functioning market, one might expect that this harm would be short-lived since a truly supracompetitive rate will eventually adjust downward. In a highly regulated market like music licensing, however, this adjustment is far from guaranteed.¹⁶³ Those entrants who do manage to enter the market will face tacitly collusive pricing in the form of a so-called “price umbrella.” A price umbrella allows a dominant firm (or tacitly colluding dominant firms) to set a price which all prospective competitors must either meet or beat (but not exceed) in order to attract buyers.¹⁶⁴ Finally, returning to the example, Pandora was targeted for being the largest, most commercially successful digital licensee in the space. Collusion of this sort can stifle innovation as potential market entrants decline to make themselves the new target of powerful content owners.¹⁶⁵

In addition to the potentially negative impact on consumers and innovation, there are also consequences for ASCAP, ASCAP’s writer-members, and smaller licensees. In the interest of self-preservation, ASCAP was pressured into amending its governing documents in ways that were not clearly in the interest of the organization as a whole. As a result, ASCAP’s writer-members are left bearing the lion’s share of

162. Any argument to the effect that the resulting price increase was merely adjusting for a “below market” price is rebutted first by the lack of a competitive marketplace (required for such an adjustment to take place); and second, by the fact that if this were true, there would have been no reason for the companies to work together to raise the price; the (competitive) market would have self-corrected.

163. Although this isn’t the case in the Pandora–ASCAP example, the publishing companies could alternately tacitly collude to depress prices in order to bolster a favored entrant (or entrants), thereby potentially misleading new entrants with artificial price signals.

164. *See* AM. BAR ASSOC., PROVING ANTITRUST DAMAGES: LEGAL AND ECONOMIC ISSUES 227 (2d ed. 2010).

165. *See generally, e.g.,* Kevin M. Lemley, *The Innovative Medium Defense: A Doctrine to Promote the Multiple Goals of Copyright in the Wake of Advancing Digital Technologies*, 110 PA. ST. L. REV. 111, 112 (2005) (extending innovation from a patent law concern to a copyright one, and considering the ways in which “copyright owners may stifle innovation in ways never before contemplated”)

ASCAP's costs. If and when the withdrawing majors take their privately negotiated rate—nearly twice that of the going ASCAP rate—to the rate court for adoption as the new “market rate,” smaller licensees (mostly start-ups and regional services) will face even higher costs of entry.

This is not just an efficiency argument: Unequal treatment of services also offends freedom of speech principles. Differential treatment of certain services “harms consumers by delaying the rollout of new distribution technologies.”¹⁶⁶ The greater this differential, the greater the threat to “the very existence of newer technologies for music distribution.”¹⁶⁷ It is further a disservice to assume that economic effect is the only goal of antitrust, when in fact antitrust regulation is in significant part an “exercise in judgment.”¹⁶⁸ This judgment includes consideration of both mathematical and social values.

Another consequence of this breakdown in structural competition is the potential for technological lock-in. Technological lock-in “is made possible when a critical mass of interdependent users accepts a standard.”¹⁶⁹ Specifically, “[w]hen switching costs are sufficiently high and technology-specific network externalities strong, the market may be subject to excess inertia, or ‘lock-in’ to a particular technology.”¹⁷⁰

Closely related to the concept of network effects, technological lock-in occurs when the incentive to improve an existing technology or to innovate is diminished by the degree of coordination between users of an existing technology. In music licensing, the major music publishers can effect a lock-in to an inferior technology by offering it a better rate than that offered to other technologies, or by denying access to their content to other companies altogether. In this way, for example, consumers may be stuck with a single clunky, outdated music streaming service favored by the major music publishers, while the developers of superior technology are either discouraged or pushed out.¹⁷¹

A final, oft-overlooked consequence of this brand of anticompetitive behavior is the loss of revenues for artists. Without a statutory license that

166. Peter DiCola, *Copyright Equality: Free Speech, Efficiency, and Regulatory Parity in Distribution*, 93 B.U. L. REV. 1837, 1839 (2013).

167. *Id.*

168. Allensworth, *supra* note 10 at 1, 53–54.

169. Clayton P. Gillette, *Lock-in Effects in Law and Norms*, 78 B.U. L. REV. 813, 817 (1998).

170. Catherine Fazio & Scott Stern, *Innovation Incentives, Compatibility, and Expropriation as an Antitrust Remedy: The Legacy of the Borland/Ashton-Tate Consent Decree*, 68 ANTITRUST L.J. 45, 52 (2000).

171. *See id.*

guarantees artists a share of royalties,¹⁷² creators are potentially denied a portion of revenues stemming from private deals, with predictable consequences for the disincentivization of creation. In addition to reducing transaction costs and consolidating bargaining power, collectives (in the absence of a statutory license) frequently serve in the role of “protector” for artists whose interests diverge from those of the distribution intermediary to whom they have assigned their copyrights.

Previous work has noted, for example, that SoundExchange—as the designated collective for public performance royalties for sound recordings—enforces § 114(g)(2) of the Copyright Act’s mandatory royalties distribution to recording artists (a distribution that does not necessarily occur under privately negotiated deals).¹⁷³ Without this statutory protection, artists are left to whatever distribution they negotiated in their contract. More often than not, for recording artists signing onto deals before making a name for themselves, this amount (where extant) will be significantly lower than the statutory rate.

Similarly, a composer may never receive a share of public performance royalties paid directly to her music publisher (as opposed to paid through a collective) as standard music publishing contracts typically deny songwriters a share in royalties stemming from privately negotiated licenses.¹⁷⁴ ASCAP, on the other hand, offers transparent royalties collection and administration to all of its members, be they music publishing companies or individual songwriters. Together, the potential for diminished economic incentives for both content creators and content distributors, coupled with a reduction in, or even elimination of access to that content, represent a real threat to innovation.

IV. CONVENTIONAL SOLUTIONS & THEIR DISCONTENTS

The Pandora–ASCAP example demonstrates the potential for unchecked anticompetitive behavior to stifle innovation and to lock consumers into existing technology while potentially denying artists their fair share of revenues. There are a few possible options for avoiding these outcomes. This Part discusses two: First, we can do nothing, and allow the

172. For example, 17 U.S.C. § 114(g)(2) mandates a 50/50 split of royalties earned on digital performance of sound recordings between artists and copyright assignees. 17 U.S.C. § 114(g)(2) (2012). See *infra* Part III.B for more detail on this oft-overlooked consequence of statutory circumvention.

173. See García, *supra* note 15, at 1151–52; *infra* Part V.B.4.

174. See PASSMAN, *supra* note 65 at 264 (“[A]ll songwriter contracts say the writer doesn’t share in any performance monies received by the publisher.”).

market to correct itself. Alternately, we can turn to antitrust law. As detailed herein, neither of these options is particularly effective in the music licensing context, nor, indeed, in highly regulated industries generally.

A. THE FREE MARKET APPROACH

Classic free market theory, or economic liberalism theory, suggests one possible response to anticompetitive behavior: Do nothing, and allow the market to correct itself.¹⁷⁵ Starting from basic supply and demand, the free market theory tells us that with little to no government intervention, producers and consumers will come together in the marketplace to set optimal prices. The general rule is that profit-maximizing firms will ultimately settle on the price where marginal cost (the cost of producing one additional unit) equals marginal revenue (the additional revenue that can be earned on the sale of one additional unit).

Under this theory, where a firm has over- or under-shot the equilibrium price, consumers will respond by moving toward the firm, or toward its competition, until eventually the equilibrium price is reached. In a two-entity market for widgets, for example, if Company A prices its widgets too high, free market theory says that consumers will purchase their widgets from Company B, unless and until Company A lowers its price (or, alternately, differentiates its widgets in some way). Company B may respond by further lowering its price, and so on, but neither company will go below the equilibrium price.¹⁷⁶ Similarly, if Company B refuses to sell its widgets to a downstream entrant, Company A can step in and fill the order. In this way, a market can correct itself, and anticompetitive behavior is avoided.

Unfortunately, this self-correction is only possible in a workably competitive market, or one with low to no barriers to entry (such that the monopoly won't be sustainable). What we have instead in the music publishing industry is a highly concentrated market, with high barriers to

175. See generally MILTON FRIEDMAN, CAPITALISM AND FREEDOM (2002) (making the case for competition capitalism).

176. Some companies will engage in so-called "loss leading" behavior, where they may actually sell at a loss for some period of time, or for certain product lines, in an effort to build interest in some other product line, or in the company generally. Amazon's Prime delivery service—grossly underpriced, but catering to the company's biggest spenders—is a good example of this. See, e.g., Greg Bensinger, *Amazon's Spending Leads to Biggest Quarterly Loss in 14 Years*, WALL ST. J. (Oct. 23, 2014), <http://www.wsj.com/articles/amazons-spending-leads-to-another-loss-1414095239> [<http://perma.cc/BF37-8NT4>].

entry. In his testimony to the Copyright Royalty Board (CRB)¹⁷⁷ on behalf of Pandora, Professor Carl Shapiro notes that:

A moderately or highly concentrated market in which the leading suppliers tacitly collude is not workably competitive. For example, if the leading suppliers have settled into some form of coordinated interaction, e.g., by refraining from competing actively to poach each other's customers, the market will fail to be workably competitive. More generally, if the leading suppliers are colluding—either expressly or tacitly—the market is not workably competitive.¹⁷⁸

This is precisely what exists in music licensing—leading suppliers of public performance rights tacitly colluding to demand higher royalty rates. Applying these principles directly to interactive music streaming services, Shapiro continues:

If interactive streaming services indeed 'must carry' the music from each of several major record companies to be competitive, and if these services have a limited ability to control the mix of music played by their customers because customers pick which songs to listen to, the market for recorded music licensed to interactive streaming services is *not* workably competitive.¹⁷⁹

Notwithstanding the efforts of Congress, the DOJ, and the relevant courts, consolidation in the music industries has only intensified. Currently, a mere three record labels constitute 65% of the music found on Pandora,¹⁸⁰ while the three largest music publishers control roughly 65% of the market for musical compositions.¹⁸¹ Notably, the cooperation

177. The CRB consists of three judges appointed by the Librarian of Congress. The Board determines and adjusts statutory rates under the Copyright Act. *See Copyright Royalty Judges*, LIBRARY OF CONGRESS, <http://www.loc.gov/crb/background/> [<http://perma.cc/KA8Q-XBDV>].

178. *In re* Determination of Royalty Rates and Terms for Ephemeral Recording and Digital Performance of Sound Recordings, 14-CRB-001-WR, at 11 (U.S.C.R.B. Oct. 6, 2014) (Written Direct Testimony of Carl Shapiro), http://www.loc.gov/crb/rate/14-CRB-0001-WR/statements/Pandora/14_Written_Direct_Testimony_of_Carl_Shapiro_with_Appendices_PUBLIC.pdf [<http://perma.cc/4RQ6-7Y3L>].

179. *Id.* at 12. Pandora, on the other hand, is not an interactive music streaming service because its consumers aren't able to choose specific songs to listen to. Shapiro goes on to differentiate Pandora's service from the interactive services for the purposes of suggesting that any rate reached vis-à-vis the interactive services should not be extended to Pandora, which he argues *is* able to steer consumers toward or away from certain songs.

180. *See* Shapiro Written Testimony, *supra* note 178, at 13 n.19.

181. *See* Christman, *supra* note 46 (offering first quarter 2015 rankings for music publishing).

among the major music publishers that Judge Cote in her Order calls “collusion” happened in spite of extant antitrust laws and multiple opportunities for antitrust review. Without a workably competitive market, the forces of supply and demand are not able to correct a supracompetitive price, nor check such anticompetitive practices as refusal to deal. The anticompetitive behavior revealed by the Pandora–ASCAP proceeding occurred in a highly regulated industry, in the shadow of a consent decree, and under the auspices of a rate court and two different regulatory agencies: the DOJ and the FTC. This points to an extensive breakdown in structural competition.

B. ANTITRUST LAW

The significance of a loss of structural competition is that antitrust, the usual go-to solution for checking anticompetitive behaviors, doesn’t work well (if at all) without it. In the absence of a workably competitive market, antitrust law is largely ineffective, for the reasons described in this section. This is especially true when, as here in music licensing, there is a highly regulated market with a differentiated good, like songs.

1. *Inapplicability*

The biggest challenge for antitrust in this context is the inapplicability of its usual arsenal: the Sherman Act, merger review, and consent decrees.

a) Limitations of the Sherman Act

The most significant challenge for antitrust in the content licensing context is that neither of the concerning behaviors at issue—i.e., neither tacit collusion nor parallel pricing—rise to the level of a Sherman Act violation, yet both carry significant competitive downsides. Section 1 of the Sherman Act forbids all contracts and business combinations made “in restraint of trade,” while § 2 prohibits monopolization and attempts to monopolize.¹⁸² Unlike explicit collusion, tacit collusion is not a violation of § 1.¹⁸³

182. 15 U.S.C. §§ 2, 3 (2012).

183. See, e.g., William H. Page, *A Neo-Chicago Approach to Concerted Action*, 78 ANTITRUST L.J. 173, at Part I (2012) (concluding that “Section 1 does not reach tacit collusion”). Cf. Richard A. Posner, *Oligopoly and the Antitrust Laws: A Suggested Approach*, 21 STAN. L. REV. 1562, 1598–605 (1968) (arguing that tacit collusion *should* violate § 1). Judge Posner’s position has not been adopted, however, as he himself has acknowledged. See, e.g., *In re High Fructose Corn Syrup Antitrust Litigation*, 295 F. 3d 651, 654 (2002) (acknowledging that “it is generally believed . . . that an express, manifested agreement, and thus an agreement involving actual, verbalized communication, must be proved in order for a price-fixing conspiracy to be actionable under the Sherman Act”). This piece

In their treatise on antitrust law, Professors Areeda, Hovenkamp and Elhauge write that “[t]he courts are nearly unanimous in saying that mere interdependent parallelism does not establish the contract, combination, or conspiracy required by Sherman Act § 1.”¹⁸⁴ The Supreme Court confirmed this interpretation:

Because § 1 of the Sherman Act “does not prohibit [all] unreasonable restraints of trade . . . but only restraints effected by a contract, combination, or conspiracy,” . . . “[t]he crucial question” is whether the challenged anticompetitive conduct “stem[s] from independent decision or from an agreement, tacit or express[.]” . . . While a showing of parallel “business behavior is admissible circumstantial evidence from which the fact finder may infer agreement,” it falls short of “conclusively establish[ing] agreement or . . . itself constitut[ing] a Sherman Act offense.”¹⁸⁵

The Court went on to hold that “[e]ven ‘conscious parallelism,’ a common reaction of ‘firms in a concentrated market [that] recogniz[e] their shared economic interests and their interdependence with respect to price and output decisions’ is ‘not in itself unlawful.’”¹⁸⁶ Unfortunately for antitrust regulators, the coordination between Sony and Universal is precisely this sort of “conscious parallelism” that can be contributed to “shared economic interests” and “interdependence.”

Parallel pricing similarly does not constitute an antitrust violation. While parallel pricing is not technically illegal, it is indicative of a highly concentrated market.¹⁸⁷ In *Matsushita Elec. Indus. Co. v. Zenith Radio Corp.*, the Supreme Court considered a case that American television producers brought against their Japanese counterparts in which they alleged that the Japanese firms:

had illegally conspired to drive American firms from the market by engaging in a scheme to fix and maintain artificially high prices for television sets sold by petitioners in Japan and, at the

takes no position on whether or not these behaviors *should* fall under the auspices of the Sherman Act, but rather takes their exclusion as well-established.

184. AREEDA ET AL., *ANTITRUST LAW: AN ANALYSIS OF ANTITRUST PRINCIPLES AND THEIR APPLICATION* ¶ 1433(a) (2003).

185. *Bell Atlantic Corp. v. Twombly*, 550 U.S. 544, 553–54 (2007) (alterations in original; internal citations omitted).

186. *Id.*

187. *See generally*, Donald F. Turner, *The Definition of Agreement Under the Sherman Act: Conscious Parallelism and Refusals to Deal*, 75 HARV. L. REV. 655 (1962) (discussing parallel behavior among firms); *infra* Part III.A.

same time, to fix and maintain low prices for the sets exported to and sold in the United States.¹⁸⁸

There, the Court concluded that “conduct as consistent with permissible competition as with illegal conspiracy does not, standing alone, support an inference of antitrust conspiracy.”¹⁸⁹

As with tacit collusion, the difficulty with parallel pricing lies in its ambiguity. The question remains whether Sony and Universal could have independently decided to withdraw and demand a higher rate from Pandora. Arguably—and herein lies the rub for antitrust: it is inapplicable to some instances of anticompetitive behavior. While a colorable argument could be made that Sony and Universal’s conduct with regard to the partial withdrawals goes beyond mere parallel pricing and tacit collusion into actionable anticompetitive behavior, the fact remains that the DOJ and the FTC have both declined to act.

b) Limitations of Merger Review

Both the DOJ and the FTC enforce the antitrust laws as laid out in three distinct, but related, acts: the Sherman Act,¹⁹⁰ the Clayton Antitrust Act of 1914 (the “Clayton Act”),¹⁹¹ and the Robinson-Patman Act of 1936 (the Robinson-Patman Act).¹⁹² Broadly speaking, the Sherman Act prohibits all combinations “in restraint of trade” and makes it unlawful to monopolize an industry.¹⁹³ The Clayton and Robinson-Patman Acts focus specifically on price discrimination.¹⁹⁴

The DOJ and the FTC typically allocate merger reviews according to their relative expertise in different subject matter areas. Most recently, the FTC has handled merger review for the music industry. In the case of Sony’s acquisition of EMI—a move that took the already concentrated music publishing industry from four firms to three—the FTC declined to challenge the merger. While approval of 4-to-3 mergers is not unique in the antitrust context, it is unusual to see a merger approved despite explicit concerns voiced about the majors’ ability to influence pricing: In the wake of the FTC’s 2012 review of the proposed Sony-EMI merger, Pandora

188. 475 U.S. 574, 574 (2012).

189. *Id.* at 588.

190. 15 U.S.C. §§ 1–7 (2012).

191. 15 U.S.C. §§ 12–27.

192. 15 U.S.C. § 13.

193. 15 U.S.C. §§ 1–2. For more on the legislative history behind its enactment, see Robert H. Bork, *Legislative Intent and the Policy of the Sherman Act*, 9 J.L. & ECON. 7 (1966).

194. See Bork, *supra* note 193.

filed an opposition cautioning that “the combination of the Sony and EMI catalogs would give Pandora ‘no choice’ but to enter into a direct license for the content. While Pandora ‘could survive without access to Sony’s musical content,’ it ‘could not survive without access to the combined Sony and EMI catalogues.’”¹⁹⁵

At the same time Sony was looking to acquire EMI’s publishing catalog, Universal sought to acquire EMI’s recorded music business. In that case too, the FTC declined to intervene.¹⁹⁶ Because the FTC only publishes its findings when a merger is challenged, the agency’s decision-making process is not fully transparent in either case. The most likely explanation for the agency’s decision lies in its mandate that calls for intervention in any merger that would “substantially lessen competition.”¹⁹⁷ In the music publishing industry, there was already a breakdown in structural competition, such that the proposed merger—while neither creating nor protecting competition—also did not worsen it. The FTC most likely approved Sony’s acquisition of EMI’s catalog because it didn’t make matters worse: The competitive environment in music licensing was dismal before the Sony-EMI merger and remained dismal afterward. In other words, “the FTC was not able to establish that the merger would lessen competition in the market . . . because of a lack of evidence of pre-merger competition.”¹⁹⁸

Another possible explanation for the agency’s decision is that the companies’ catalogs were viewed as complements, rather than substitutes. In that case, a music streaming service would require both companies’ catalogs, such that the FTC did not view the combination as a horizontal merger.¹⁹⁹ With a mandate allowing for intervention only where a

195. Order, *supra* note 9, at 62.

196. Notably, the European Union’s antitrust regulators also approved both mergers, but only with divestitures. In the case of Universal/EMI, this included divestiture of 60% of EMI’s European business. See Vanessa Mock et al., *Universal to Win EMI – After Selling Key Assets*, WALL ST. J. (Sept. 20, 2012), <http://www.wsj.com/articles/SB10000872396390444032404578006343033008694> [<http://perma.cc/W5BA-594F>]. The FTC required no such divestitures.

197. See HORIZONTAL MERGER GUIDELINES, *supra* note 83 at ¶¶ 4–6.

198. Shapiro Written Testimony, *supra* note 178, at 13 n.17 (referring to the Universal/EMI merger approval).

199. See *id.* at n.15 (calling it obvious that “the FTC saw the repertoires of Universal and EMI as complements, not substitutes, for interactive streaming services. Therefore, for this group of buyers, the Universal/EMI merger was not a horizontal merger, and the normal loss of direct competition that occurs in a horizontal merger was not present.”).

combination will worsen competition, antitrust can do little in an industry that has lost its structural competition.²⁰⁰

While it is too late for the music licensing industry—antitrust law cannot institute structural competition once it is lost—an important lesson may be gleaned for antitrust law and industry consolidation going forward: Where antitrust fails to prevent consolidation up front, it may be powerless to “fix” the problem afterward. The music industry offers a powerful cautionary tale as to why structural competition matters, and why the structural presumption does not always apply.

c) Limitations of Consent Decrees

One of antitrust’s most effective tools, the consent decree, has likewise proven ineffective in the music licensing space. As discussed at Part I(A)(2) *supra*, the DOJ first imposed a consent decree on ASCAP in 1941.²⁰¹ That decree, last amended in 2001, required ASCAP “to grant to any music user making a written request therefor a non-exclusive license to perform all of the works in the ASCAP repertory.”²⁰²

Prompted by anticompetitive concerns about ASCAP’s consolidation of content, the consent decree restricts how ASCAP licenses content in a variety of ways. Among other things, it makes a copyright owner’s grant of licensing authority to ASCAP non-exclusive—which means that a rights holder can always license its own content—and requires that ASCAP charge all similarly-situated entities the same rate.²⁰³ These requirements were intended to assuage concerns about anticompetitive behavior in music licensing. But the consent decree acts only on ASCAP, and not on ASCAP’s individual members, thus limiting its ability to thwart anticompetitive behavior on the part of individual entities.

The FTC, through its repeated determination not to intervene in the series of mergers that have reduced the music publishing industry to a mere three entities, has effectively foreclosed a finding of worsened competition stemming from anticompetitive behavior on behalf of the individual publishers: ASCAP’s consent decree requires all grants from rights holders be made on a non-exclusive basis, thereby allowing a party to opt out and negotiate privately. This assumes that the individual

200. As Part IV.B.2.c *infra* explains, music goods are complementary and not substitutable, making this situation different from the standard 4-to-3 merger.

201. See *supra* note 33 and accompanying text.

202. See *United States v. ASCAP*, No. 41-1395, 2001 WL 1589999, § VI (S.D.N.Y. 2001).

203. *Id.* § IV(B)–(C).

publisher, acting alone, does not have the ability to act anticompetitively. Unfortunately, this assumption is wrong. Sony, the nation's largest music publisher, currently enjoys a market share of roughly 30%, and may well be able to exercise unilateral monopoly power in some circumstances. When working in tacit collusion with Universal, the combined market share is raised to nearly 50%.

Antitrust law does not provide a remedy for breaking up monopolist (or oligopolist) firms unless and until they engage in predatory conduct. Even if one firm is not found to be monopolistic on its own, two or more firms may tacitly collude to set prices or to bar entry to a new service by withholding content altogether. It is well established that antitrust law does not address the oligopolist problem of tacit collusion.²⁰⁴ This is why the maintenance of structural competition is so important. Where it is not maintained—as in the case of the music industry—it cannot be rebuilt.

In light of the discontent stemming from the attempted withdrawals by Sony and Universal, the Senate Judiciary Committee has begun a review of ASCAP's consent decree.²⁰⁵ In his request for comments, Senator Leahy called for “renewed attention to the consent decrees” in order to “ensure the decrees’ purposes are still being met.”²⁰⁶ By and large, the publicly available comments received to date recognize the role and import of a consent decree for the purposes of creating and protecting competition in the music licensing space.

In its written comments, ASCAP focuses on three overarching suggestions for changes to its consent decree: First, an expedited rate-setting process to replace the rate court.²⁰⁷ Second, ASCAP seeks a congressional blessing to allow for partial digital withdrawals, claiming that “this approach would result in competitive market transactions that would then provide informative benchmarks for the rate-setting tribunal.”²⁰⁸ Unfortunately for ASCAP, “competitive market transactions” between “truly willing buyers and willing sellers” is not possible without a competitive market.²⁰⁹ Finally, ASCAP asks for permission to license not only public performance rights, but also mechanicals, synchronization, and

204. *See id.* (proposing that tacit collusion is not—and should not—constitute a violation of the antitrust laws).

205. *See* discussion *supra* Part II.B.5.

206. *How Much For A Song?: The Antitrust Decrees That Govern the Market for Music*, 114th Cong. (2015) (statement of Senator Patrick Leahy), <http://www.judiciary.senate.gov/imo/media/doc/03-10-15LeahyStatement.pdf> [<http://perma.cc/HYC2-ASEN>].

207. *See* Matthews, *supra* note 129 at 19.

208. *Id.* at 19–20.

209. *Id.* For further rebuttal of this claim, *see supra* Part II.C.3.b.

print rights. This would arguably create a “one-stop shop’ for musical work rights.”²¹⁰ While this would undoubtedly improve upon the labyrinth that potential music licensees currently navigate, it could also exacerbate the potential for anticompetitive behavior on behalf of the collective.

Pandora and ASCAP agree that the consent decrees aren’t working, but for different reasons. Pandora’s comments, not surprisingly, urge caution and note several judicially recognized instances of “egregious misconduct” on behalf of the music publishers and collectives.²¹¹ In addition to Judge Cote’s finding of “troubling coordination” in the Pandora–ASCAP proceeding, Pandora references a December 2013 proceeding between Pandora and BMI in which the judge held that “BMI cannot combine with [music publishers] by holding in its repertory compositions that come with an invitation to a boycott attached.”²¹² Notably, both of these findings occurred under extant consent decrees. So long as the consent decrees do not apply to individual entities, and those entities are allowed to withdraw, the consent decrees are powerless to correct anticompetitive behavior arising from private ordering of the type seen here. For this reason, the consent decrees no longer function to curb anticompetitive behavior.²¹³

2. Challenges

Even where antitrust is applicable, it is an *ex post* review—that is, the review is conducted only when the alleged harm is already evident.²¹⁴ At that point, both the review and enforcement of a remedy—if one is even

210. See Matthews, *supra* note 129 at 20.

211. See *How Much For A Song?: The Antitrust Decrees That Govern the Market for Music*, 114th Cong. 4 (2015) (statement of Christopher S. Harrison, Pandora Media, Inc.) <http://www.judiciary.senate.gov/imo/media/doc/Harrison%20Testimony.pdf> [<http://perma.cc/47KJ-BU4W>].

212. *Id.*

213. The problem of anticompetitive behavior unchecked by antitrust is not exclusive to music publishing. Sound recordings have also devolved into a natural oligopoly situation, with a mere three entities controlling nearly all sound recordings copyrights. See *supra* Part III.A. In the case of sound recordings, however, piracy serves as a check on most anticompetitive behavior. Unlike music publishers, the record labels are at all times arguably “competing with free,” which keeps them honest in a way that music publishers don’t face.

214. This is not to say that antitrust cannot have *ex ante* effects as well. See, e.g., Parisi & Depoorter, *supra* note 154 at 21 (discussing the effect of antitrust rulings on competition between direct and intermediary licensors).

available—are notoriously cumbersome, time-consuming, and costly.²¹⁵ Antitrust cases are also notoriously difficult to prove. Even if there was conduct rising to the level of an antitrust violation, antitrust is ill-suited to high-tech industries like music licensing where antitrust enforcers' ability to understand and predict industry evolution is especially limited.²¹⁶ In addition to the inapplicability of antitrust law to certain anticompetitive behaviors and its procedural shortcomings, antitrust faces particular challenges in the context of highly-regulated industries, in the IP field, and music in particular.

a) In the Regulatory Context

Regulation frequently works to encourage investment in an industry by protecting the firms within it. This is especially true in industries with high start-up costs, or with high risks of failure, and for companies that rely on a distribution system. Public utilities, for example, face all of these challenges, and new entry into these markets requires substantial investment. In the early years of electricity generation, “companies saw clearly the advantages of a regulated marketplace to protect their investments from unbridled competition.”²¹⁷

Extensive regulatory regimes tend to undermine antitrust by giving firms an implied immunity from antitrust review.²¹⁸ For this reason,

215. See, e.g., Comment of the Staff of the Federal Trade Commission, FERC Docket No. RM11-14-000 at 1 (2011) (noting that “[i]nconsistent approaches may make the antitrust review process longer, more confusing, and more costly than necessary”).

216. See, e.g., Ronald A. Cass, *Antitrust for High-Tech and Low: Regulation, Innovation, and Risk*, 9 J.L. ECON. & POL'Y 169, 169–70 (2013) (claiming that “[t]raditional problems of regulation generally, and of antitrust enforcement specifically, are exaggerated in high-technology sectors, where antitrust enforcers' abilities to understand and predict industry evolution are most limited and where enforcement actions are most likely to rest on debatable predicates about the effects of specific conduct.”). Cf. *United States v. Microsoft*, 253 F. 3d 34 (D.C. Cir. 2001) (offering an example of sophisticated antitrust analysis in a high-tech context).

217. See MARY M. TIMNEY, *POWER FOR THE PEOPLE: PROTECTING STATES' ENERGY POLICY INTERESTS IN AN ERA OF DEREGULATION* 47 (2004).

218. See, e.g., Stacey L. Dogan & Mark A. Lemley, *Antitrust Law and Regulatory Gaming*, 87 TEX. L. REV. 685, 686 (2009) (asserting that “antitrust laws are impliedly repealed by government regulation of a particular industry”); Daniel F. Spulber & Christopher S. Yoo, *Mandating Access to Telecom and the Internet: The Hidden Side of Trinko*, 107 COLUM. L. REV. 1822, 1851 (2007) (“Courts have long recognized that the enactment of a federal regulatory scheme can immunize particular conduct from antitrust scrutiny.”); see also *Credit Suisse Sec. (USA) LLC v. Billing*, 551 U.S. 264, 276 (citing four conditions that the court will consider in determining whether a regulatory regime precludes antitrust, and determining that “the securities law impliedly precludes the application of the antitrust laws”).

regulation is frequently considered to have anticompetitive consequences, and to work at odds with antitrust. Music's regulatory regime—comprising an entire Act²¹⁹ and meriting congressional reconsideration²²⁰—likewise reduces the role of antitrust, not least of all because IP regulation protects creators and inventors from competition (albeit for a limited period of time).²²¹

b) In the Intellectual Property Context

An important challenge for antitrust in the music industry is its juxtaposition to the intellectual property context. IP law necessarily contravenes antitrust law insofar as it grants limited “monopolies” to inventions, content, and brands via property rights afforded by patent, copyright, and trademark law, respectively. In each case, these legally sanctioned monopolies are limited in both scope and duration, and are traditionally justified by the proposition that on balance, they produce enough social utility—by encouraging creation and dissemination of new content—to offset the societal costs of deterring competition.²²²

Of course, the exclusive rights granted by copyright do not a true economic monopoly make. For example, most novels are copyrighted, but no one novel is considered to have a monopoly on the genre as a whole. This is because while imperfect, content is generally considered

219. See The Copyright Act, 17 U.S.C. §§ 101–1332 (1976).

220. See, e.g., Maria A. Pallante, *The Next Great Copyright Act*, 36 COLUM. J.L. & ARTS 315 (2013) (calling for “comprehensive review and revision of U.S. copyright law”).

221. See, e.g., Herbert Hovenkamp, *Antitrust and the Regulatory Enterprise*, 2004 COLUM. BUS. L. REV. 335, 341 (2004) (“One consequence of regulation is a reduced role for the antitrust laws. When the government makes rules about price or output, market forces no longer govern. To that extent antitrust is shoved aside.”).

222. Professors Hovenkamp, Janis and Lemley explain the justification this way:

Because intellectual property rights impose costs on the public, the intellectual property laws can be justified by the public goods argument only to the extent that the laws on balance encourage enough creation and dissemination of new works to offset those costs. One of the reasons that intellectual property rights are limited in scope, in duration, and in effect is precisely in order to balance these costs and benefits The key to economic efficiency lies in balancing the social benefit of providing economic incentives for creation and the costs of limiting the diffusion of knowledge.

HERBERT HOVENKAMP ET AL., *IP AND ANTITRUST: AN ANALYSIS OF ANTITRUST PRINCIPLES APPLIED TO INTELLECTUAL PROPERTY LAW* § 1.3 (2002) (hereinafter TREATISE).

substitutable.²²³ If a bookstore doesn't have a title that a friend recommends, or if it is selling a comparable title for half the price, the customer may choose to purchase the comparable title. This consolation purchase may not prevent the customer from purchasing the recommended book in the future, but the choice demonstrates the limited power of copyright's "monopoly" grant. Most individual copyrights (for example, a copyright to a single song) are not viewed as conferring market power on individual owners, but in the aggregate they may—for example, a music publisher with rights to millions of songs.

The highly technical nature of many IP-related disputes challenges antitrust still further. Some commentators suggest that antitrust's impotence in the IP context owes further to its overly narrow analysis that ignores this "marketplace of ideas."²²⁴ Academics, industry, and legislators have all lamented the difficulties of policing the so-called "marketplace of ideas," defined as "a sphere in which intangible values compete for acceptance," and also as "a form of nonprice competition."²²⁵ The idea underlying this push for expanded merger review is to recognize that technologically inclined firms compete on a number of dimensions other than price such as the pace and breadth of innovation, customer service, and overall quality.²²⁶

Much of the emphasis in case law at the intersection of antitrust and IP has been on compulsory licensing as a theoretical equitable remedy for the anticompetitive exploitation of IP rights. For example, 17 U.S.C. § 115 contains a compulsory cover license, under which licensees willing to pay the statutory rate may license the use of a musical composition for the purpose of creating a sound recording. This represents an exception to the copyright holder's general right to exclude. In this sense, compulsory licenses reflect a balance between exclusive rights for copyright owners and ensure access to creative works. Notably, this balance stems from a regulatory, and not an antitrust, solution.

223. See, e.g., Michael Abramowicz, *An Industrial Organization Approach to Copyright Law*, 46 WM. & MARY L. REV. 33, 37 (2004) ("Copyrighted works can serve as imperfect substitutes for one another . . .").

224. See Maurice E. Stucke & Allen P. Grunes, *Antitrust and the Marketplace of Ideas*, 69 ANTITRUST L.J. 249, 249 (2001) (arguing for antitrust merger review to be expanded "to include its impact on the 'marketplace of ideas'").

225. *Id.* at 251, 297. For another example of antitrust's impotence in the IP context, see Aaron K. Perzanowski, *Rethinking Anticircumvention's Interoperability Policy*, 42 U.C. DAVIS L. REV. 1549 (2009) (suggesting that in lieu of antitrust, the DMCA's interoperability exemption be expanded)

226. *Id.* at 279.

c) In the Music Context

One of the biggest challenges to the application of the free market theory to the music publishing industry is that free market economics assumes a competitive marketplace for rival and excludable goods. Unfortunately, that is not the state of affairs for the music industry, which instead touts a single, public good—songs—in a marketplace that is not workably competitive.²²⁷ As a noncompetitive market, music is uniquely susceptible to anticompetitive concerns.²²⁸

First, music is a differentiated product—that is, each song is unique, and consumers of music value variety. This demonstrates that the model of perfect competition does not apply because that model “assume[s] that many suppliers offer a homogenous product.”²²⁹ This is not the case in music, where no one song is a perfect substitute for another. Sellers of differentiated products, like songs, can typically demand a price above marginal cost and how much above depends on the buyer’s elasticity of demand.

In music, an interactive service—one that allows consumers to select specific songs to play—is said to have nearly inelastic demand, and so typically faces higher prices than noninteractive services in which consumers can merely guide the song selection by dictating genre, or indicating “sounds like” choices. This is because, as Professor Shapiro describes in detail in his written testimony, a noninteractive service like Pandora has the ability to “steer” playlists toward some songs and away from others.²³⁰ This makes non-interactive services’ demand largely elastic, thereby lowering the price they can demand.

In addition, in the music-licensing context, one publisher’s catalog is frequently complementary with—and not substitutable for—another’s.²³¹

227. Unlike standard goods—including those traditionally protected by property rights, like a house—songs, once released, are not excludable, nor is their worth diminished through their use by others. N.B.: Professor Yoo has persuasively argued that copyrighted goods are more accurately viewed as “impure public goods,” or goods that are non-rival but excludable, or non-excludable but rival. Christopher S. Yoo, *Copyright and Public Good Economics: A Misunderstood Relation*, 155 U. PA. L. REV. 635, 635 (2007).

228. See, e.g., Frank H. Easterbrook, *The Limits of Antitrust*, 63 TEX. L. REV. 1, 1 (1984) (stating that “[t]he goal of antitrust is to perfect the operation of *competitive* markets.”) (emphasis added).

229. Shapiro Written Testimony, *supra* note 178, at 4.

230. Shapiro Written Testimony, *supra* note 178, at 5–7 (“[T]he more easily a music service can steer listeners toward or away from specific sound recordings, the lower will be the price that music service will be able to negotiate.”).

231. See *supra* note 179 and accompanying text.

This circumstance cuts to the heart of the conflict between IP and antitrust in the form of unilateral refusals to license.²³² In the absence of a compulsory license, IP owners are generally under no obligation to license the intellectual property, nor, indeed, to use it at all.²³³ This is problematic from an economic perspective because licensing (at least in theory) allows the market to transfer IP to its most productive use.²³⁴ Meanwhile, the DOJ Antitrust Division's Intellectual Property Guidelines take the position that content licensing is essentially pro-competitive:

Licensing, cross-licensing, or otherwise transferring intellectual property . . . can facilitate integration of the licensed property with complementary factors of production. This integration can lead to more efficient exploitation of the intellectual property, benefiting consumers through the reduction of costs and the introduction of new products. Such arrangements increase the value of intellectual property to consumers and to the developers of the technology. By potentially increasing the expected returns from intellectual property, licensing also can increase the incentive for its creation and thus promote greater investment in research and development.²³⁵

In other words, licensing is the most efficient means of extracting value from an IP right, and circumstances that allow for lower levels of licensing, such as withdrawals and refusals to deal, may lead to lower value levels as well.²³⁶

232. See TREATISE, *supra* note 222, § 13.1 (stating that “[u]nilateral refusal to license cases . . . cut to the heart of the intellectual property owner’s right to exclude others from practicing the intellectual property” and noting that “[a]s such, efforts to invoke antitrust law in this context deserve special scrutiny.”).

233. *Id.* § 13.2a (stating that “an intellectual property owner has no obligation to use its right at all.”); *id.* § 13.2b (adding that the “‘right’ to refrain from using intellectual property would be a hollow thing indeed if the intellectual property owner could not prevent others from infringing the right.”). Indeed, the common justification for compulsory licenses is their role as a point of guaranteed access. While outside the scope of the immediate Article, future work will consider the theory and practice of this type of engineered access.

234. *Id.* § 13.2 (“Economic theory encourages licensing because it allows the market to transfer the intellectual property right to the most productive user of that right.”).

235. U.S. DOJ & FTC, *Antitrust Guidelines for the Licensing of Intellectual Property*, JUSTICE.GOV, § 2.3 (Apr. 6, 1995), <http://www.justice.gov/atr/antitrust-guidelines-licensing-intellectual-property> (hereinafter *Antitrust Guidelines for IP*) [<http://perma.cc/ZL8A-GXGV>].

236. The courts have also recognized the importance of licensing, primarily through the doctrine of copyright misuse. Copyright misuse is an affirmative defense akin to “unclean hands.” See, e.g., *In re Napster, Inc. Copyright Litigation*, 191 F.Supp.2d 1087, 1102 (2002) (“Copyright misuse as a defense to an infringement action finds its origins in

Congress has recognized the special situation of the music industry as a peddler of complimentary, but differentiated goods. It is also susceptible to refusals to license. Concern over the latter in particular is reflected in the compulsory license scheme prevalent in copyright law as it applies to music specifically. For example, § 115—the compulsory license for cover songs—recognizes a conflict between the societal value inherent in the creation of different versions of a song and the propensity for copyright owners to deny the realization of this value. The compulsory license resolves this conflict by limiting the copyright holder's ability to refuse to license in exchange for money in the form of the statutory rate. Professor Jane Ginsburg has written:

[T]he real purpose of a compulsory license is to reduce the extent to which copyright ownership of the covered work conveys monopoly power, so that the copyright owner must make the work available to all who wish to access and exploit it.²³⁷

The existence of a compulsory license also supports the claim that in the absence of a workably competitive market, some form of regulation is called for. In copyright, regulation already exists and can be fairly readily augmented to make up for the short reach of antitrust in the music context. In other words, a market that is not workably competitive cannot self-correct, and antitrust cannot curb behaviors outside its purview, but regulation can. The next Part suggests that contrary to the conventional view of regulation as competition-reducing, “remedial regulation” can encourage competition in a way that antitrust cannot.

V. FACILITATING COMPETITION THROUGH REGULATION

There is a third option for checking anticompetitive behavior, maintaining competition, encouraging innovation, preventing technological lock-in, and ensuring payment to artists: regulation. The conventional view of regulation is as a system that works against competition; one that thwarts new entry and protects incumbents.²³⁸ Indeed, the Telecommunications Act of 1996—intended to mark the deregulation of the telecommunications industry—proclaims as its

the equitable defense of unclean hands and is similar to the patent law defense of the same name.”)

237. Jane C. Ginsburg, *Creation and Commercial Value: Copyright Protection of Works of Information*, 90 COLUM. L. REV. 1865, 1926 (1990).

238. See, e.g., *supra* Part IV.B.2.a.

purpose: “To promote competition and *reduce regulation* in order to secure lower prices and higher quality services for American telecommunications consumers and encourage the rapid deployment of new telecommunications technologies.”²³⁹ The goal of this Part is to challenge the conventional view and to present regulation as potentially pro-competitive.

Conventional thinking about how to approach the competition problem, or bargaining breakdown, in content generally falls into two divergent points of view: There are those who would reduce dependence upon (or in some cases do away with altogether) the current statutory licensing regime in favor of private ordering and/or other, preferable mechanisms such as fair use, patent pools, and collectives;²⁴⁰ and those who favor compulsory licensing over private deal making for avoiding bottlenecks and for more robust information exchange.²⁴¹ The former view ignores the important role compulsory licenses play in ensuring access to content; the latter ignores the potential informational value derived from private rate setting. Both of these perspectives ignore the competitive market.

This Article departs from both of these perspectives, proposing instead a new model for maintaining competition in the licensing of intellectual property rights. This proposal calls for adherence to a mandatory, compulsory license by default, but embraces private ordering where (and

239. Title page of 47 U.S.C. (1996) (emphasis added).

240. See, e.g., Wendy J. Gordon, *Fair Use as Market Failure: A Structural and Economic Analysis of the Betamax Case and its Predecessors*, 82 COLUM. L. REV. 1600, 1613 (1982) (“In extreme instances, Congress may correct for market distortions by imposing a regulatory solution such as a compulsory licensing scheme,” but maintaining that “the broad brush of this regulatory solution is too sweeping for most cases.”); Mark A. Lemley, *The Economics of Improvement in Intellectual Property Law*, 75 TEX. L. REV. 989, 1061–67, 1068–83 (1997) (pushing for new rules around derivative works and fair use as preferable to compulsory licensing for avoiding bargaining breakdowns in copyright); Merges, *supra* note 5, at 1295–96 (arguing against compulsory licensing as subject to “legislative lock-in” and instead favoring such devices as patent pools and collectives like ASCAP for overcoming bargaining obstacles); Robert P. Merges, *Intellectual Property Rights and Bargaining Breakdown: The Case of Blocking Patents*, 62 TENN. L. REV. 75, 75–76 (1994) (arguing for expanded use of reverse doctrine of equivalents over compulsory licensing for the avoidance of bargaining breakdowns in patent law).

241. See, e.g., Ian Ayres & Eric Talley, *Solomonic Bargaining: Dividing a Legal Entitlement to Facilitate Coasean Trade*, 104 YALE L.J. 1027, 1092–94 (1995) (arguing that compulsory licensing can induce parties to reveal their “truthful revelation,” thereby overcoming bargaining breakdown); see also Neil Weinstock Netanel, *Copyright and a Democratic Civil Society*, 106 YALE L.J. 283 (1996), 334–35 nn.247–48 (discussing Ayres & Talley’s theory).

only when) real competition can be shown to exist between rival content licensors. This proposal, referred to herein as the “remedial regulation model,” utilizes existing mechanisms—specifically, statutory licenses, a collective administrator, and existing regulatory authorities—to correct anticompetitive behavior at minimal cost.

The current competition policy for the licensing of intellectual property assumes robust competition, and so allows for private ordering in the shadow of the statutory license. For example, § 114 of the Copyright Act allows copyright owners to *either* use the statutory license, *or* to negotiate their own royalty rates and license terms for the public performance of sound recordings.²⁴² As a result, conventional antitrust mechanisms—like ASCAP’s consent decree—are wholly ineffective against anticompetitive behavior perpetrated by individuals, who can merely opt-out.

The remedial regulation model updates copyright’s competition policy by reversing this assumption. Instead, it assumes monopolistic (or oligopolistic) market power, thereby converting the existing, circumventable statutory licenses into mandatory, compulsory licenses under which parties may petition for permission to deal privately. Requiring only minimal statutory amendment and utilizing existing regulatory agencies and collectives, the remedial regulation model offers licensors and licensees a compromise: Continued access to content for all at a predictable rate and the flexibility to negotiate private terms, so long as industry consolidation has not reached a point so as to call into question the arms-length nature of any such transactions. This proposal builds, in part, on the existing literature on penalty defaults and altering rules. After a brief review of default theory, this Part will show its application in the regulatory context and will detail a remedial regulatory solution to copyright’s competition problem.

A. PENALTY DEFAULTS, ALTERING RULES & COMPETITION

1. *Default Theory*

A “penalty default” is an undesirable fall back option designed to penalize those who, through failure to do or to not do some thing (be it negotiate, or share information), do not otherwise negotiate around it. The concept of “penalty default rules” was first introduced by Professors

242. See 17 U.S.C. § 114(e) (2012).

Ian Ayres and Robert Gertner,²⁴³ who described them as unpalatable fallback options in contract law that kick in unless the parties negotiate their own terms. Such rules, they argue, induce more knowledgeable parties to “reveal information by contracting around the default penalty.”²⁴⁴ Prior work has extended this concept to licensing and demonstrates that “penalty default *licenses* encourage[] more efficient deal making among otherwise unequal parties by motivating them to circumvent an inefficient statutory license in favor of private ordering.”²⁴⁵

In other words, penalty defaults are a mechanism by which regulators can encourage or discourage a certain behavior without regulating that behavior directly. This is particularly useful where the behavior sought to be modified is not easily regulated, such as to encourage retirement savings, organ donation, and to curb pollution.²⁴⁶ The next section argues that penalty defaults might also prove especially useful for regulating behavior that is not readily ameliorated by existing legal regimes, such as the anticompetitive behavior of the individual music publishing companies whose tacit collusion and parallel pricing activities are not checked by antitrust.

Altering rules establish the “necessary and sufficient conditions for altering default legal consequences.”²⁴⁷ “Impeding” altering rules aim to “deter opt-out by artificially increasing its difficulty.”²⁴⁸ This is effectively what remedial regulation does: By requiring a showing of sufficient competition before private ordering is permitted, the statutory license is made “quasi-mandatory” or sticky.²⁴⁹

243. Ian Ayres & Robert Gertner, *Filling Gaps in Incomplete Contracts: An Economic Theory of Default Rules*, 99 *YALE L. J.* 87, 91 (1989).

244. *Id.* at 94. In other words, where a penalty default rule results in an undesirable outcome for Party A (possessor of information unknown to Party B), Party A may be incentivized to negotiate around the default term, thereby revealing his information not only to Party B, but also to the legislature, which can use that information to draft better rules, default and otherwise.

245. García, *supra* note 15, at 1122 (emphasis added).

246. *See generally* RICHARD H. THALER & CASS R. SUSSTEIN, *NUDGE* (2008) (identifying areas in which behavior modification via incentive or encouragement might be best received).

247. Ian Ayres, *Regulating Opt-Out: An Economic Theory of Altering Rules*, 121 *YALE L.J.* 2032, 2114 (2012).

248. *Id.* at 2084.

249. *Id.* at 2087.

2. *Application to Regulation*

In the regulatory context, the remedial concept behind impeding altering rules works to penalize an undesirable behavior in hopes of encouraging a different behavior. Here, it does so by mandating compliance with a statutory rate—thereby foreclosing private ordering with all of its potential benefits—unless and until sufficient competition can be shown in the relevant marketplace.

There is precedent for this approach. In wholesale electricity, for example, the Federal Energy Regulatory Commission (FERC) sets the applicable rates for energy transmission. A utility company is allowed to charge a “market-based tariff only if [the company] demonstrates that it lacks or has adequately mitigated market power, lacks the capacity to erect other barriers to entry, and has avoided giving preferences to its affiliates.”²⁵⁰

Varying in procedure, but similar in spirit, are patent pools, or the pooling of patents between two or more companies. Patent pooling is generally acceptable, even favored, unless “(1) excluded firms cannot effectively compete in the relevant market for the good incorporating the licensed technologies and (2) the pool participants collectively possess market power in the relevant market.”²⁵¹ Where these conditions exist, the DOJ or the FTC will review the licensing arrangement for anticompetitive effect before determining whether the parties will be allowed to engage in the pooling activity. In both of these examples, a competitive marketplace is not assumed, but must first be shown.

B. REMEDIAL REGULATION

In lieu of antitrust, this Article advocates utilizing remedial regulation—or, regulation that discourages industry consolidation—in

250. *Morgan Stanley Capital Group Inc. v. Public Utility Dist. No. 1 of Snohomish County, Wash.*, 554 U.S. 527, 537 (2008); *see also* Market-Based Rates For Wholesale Sales of Electric Energy, Capacity And Ancillary Services by Public Utilities, Order No. 697, 119 FERC ¶ 61,295, at ¶ 7 (2007), <http://www.ferc.gov/whats-new/comm-meet/2007/062107/E-1.pdf> [<https://perma.cc/T36A-BAVL>], which cites the development of a 4-prong analysis for determining whether a seller can engage in private ordering:

(1) [W]hether the seller and its affiliates lack, or have adequately mitigated, market power in generation; (2) whether the seller and its affiliates lack, or have adequately mitigated, market power in transmission; (3) whether the seller or its affiliates can erect other barriers to entry; and (4) whether there is evidence involving the seller or its affiliates that relates to affiliate abuse or reciprocal dealing.

251. *Antitrust Guidelines for IP*, *supra* note 235, § 5.5.

order to open the market and maintain competition. This model assumes a baseline that tends toward oligopoly, natural or otherwise, and so allows for private ordering only where sufficient competition can first be shown. Otherwise, regulation operates to ensure ongoing access to the relevant input(s) for all prospective consumers or licensees able and willing to meet the statutory requirements and to pay the statutory rate. Because this regulation does not necessarily represent a market rate—nor, indeed, as high a rate as private ordering might obtain—this Article labels it “remedial.” It punishes the lack of a competitive marketplace.

If a company wants to engage in private ordering to obtain a higher rate or better terms, it must first petition to show the existence of sufficient competition in the relevant market. While such “remedial regulation” cannot create a robust competitive market where none exists, it can prevent a few powerful firms from unilaterally controlling the price for an input, or from barring new entry to the market altogether to the detriment of both consumers and innovators in the space. As is the case with other highly regulated industries, the underlying assumption here is that the government has a greater responsibility for checking anticompetitive behavior in the music licensing space owing to its role in the granting of exclusive property rights via copyright.

As with the wholesale electricity example, remedial regulation places the burden of proving a competitive marketplace on the party seeking to get out from under the statutory regime. This resets the baseline assumption and brings competition policy in line with positive market conditions, while at the same time establishing a “safe harbor” that allows for private ordering (and its concomitant advantages) when, and only when, sufficient competition can be shown. The next section outlines one possible path toward implementation of remedial regulation in the music licensing context.

1. *Logistics*

In the music licensing context, remedial regulation does three things. First, it removes the authority for negotiation found in certain statutory licenses,²⁵² thereby converting them into mandatory (i.e., non-

252. For example, 17 USC § 114(e) permits:

[A]ny copyright owners of sound recordings and any entities performing sound recordings affected by this section [to] negotiate and agree upon the royalty rates and license terms and conditions for the performance of such sound recordings and the proportionate division of

circumventable) statutory licenses. Specifically, transition to a remedial regulatory regime begins with the conversion of existing non-mandatory statutory licenses—like § 114²⁵³—into a mandatory, statutory license. There is precedent for this in the Copyright Act itself: Both § 111 (for broadcast cable) and § 119 (for satellite television)²⁵⁴ are mandatory, compulsory licenses that do not allow for circumvention and private ordering.

Next, remedial regulation also establishes a new compulsory license for the public performance of musical compositions. Currently, these rights are negotiated in the market, allowing for the types of anticompetitive conduct seen in the music licensing space. The establishment of a compulsory license provides a guaranteed point of access by removing the threat of unilateral refusals to license. As with other compulsory licenses—such as § 115's cover license—this statutory infringement on the exclusionary rights of copyright owners is justified by the societal benefit of avoiding anticompetitive behavior that threatens to stifle innovation in the distribution space. These compulsory licenses serve as impeding altering rules, or sticky defaults, intended to make opt-out difficult.

These newly mandatory compulsory licenses can be administered by a governmentally authorized collective. As it did with SoundExchange in the context of public performance rights for sound recordings, the government could likewise designate one of the existing collectives—such as ASCAP or BMI—or elect to form an entirely new collective to be charged with the administration and distribution of royalties for these newly-created licenses. In any case, the CRB retains rate-setting authority for all statutory licenses.

Finally, a remedial regulatory regime establishes a procedure through which a rights holder wishing to forego the statutory license and engage in private ordering may petition to do so via a showing of sufficient competition in the market for music licensing. A party wishing to circumvent the statutory license files a petition with the regulatory body (most likely the FTC, given both its institutional competence and specialization in the music industry).²⁵⁵ This designation would require a

fees paid among copyright owners, and [to] designate common agents on a nonexclusive basis to negotiate, agree to, pay, or receive payments.

17 U.S.C. § 114(e) (2012).

253. 17 U.S.C. § 114.

254. 17 U.S.C. §§ 111, 119.

255. One might also suggest that the additional workload for the FTC is its just desert for allowing the market concentration in music licensing to reach its current levels.

congressional delegation of authority to the FTC for petition review.²⁵⁶ The FTC then either accepts the petition, allowing parties to circumvent and engage in private negotiation, or rejects it, in which case the parties must proceed under the statutory rates set by the CRB. Importantly, this ensures continued access to content regardless of the state of competition, while simultaneously eliminating adverse selection concerns resulting from opt-out of the most powerful parties.

A petition to circumvent mandatory statutory licenses must show real and sufficient competition in the relevant market. In the wholesale electricity context, FERC requires a company seeking authorization to engage in “market-based rates” or “MBRs” file an application under § 205 of the Federal Power Act.²⁵⁷ Borrowing liberally from that application’s guidelines, a petition in the music-licensing context might require any or all of the following:

- A letter explaining the basis for the petition, and containing contact information and a description of the petitioning entity’s business.
- A description of the specific type of license to be negotiated for by petitioner.
- A description of any affiliates of the petitioner and their business activities. Or, if the petitioner has no affiliates, the application should include a representation to that effect. Affiliates can be defined as, but not limited to, upstream owners and wholly or partially owned entities.
- Representations of how the petitioner satisfies the reviewing agency’s concerns with regard to horizontal market power.
- Representations of how the petitioner satisfies the reviewing agency’s concerns with regard to vertical market power.
- Representations of how the petitioner satisfies the reviewing agency’s concerns with regard to barriers to entry and unilateral refusals to license.

Congress might consider an appeals process to allow for dispute resolution in the event of dissatisfaction regarding a determination by the

256. For an excellent analysis of the various government institutions that participate in the making and enforcement of copyright law, and the pros and cons of each, *see* DiCola & Sag, *supra* note 31.

257. *See* 16 U.S.C. § 12 (2014).

reviewing agency, or disagreement about the sufficiency of evidence presented. Should an appeals option present too heavy a burden for the reviewing agency, it should be noted that FERC has successfully defended the lack of an appeals process in the MBR context.

In sum, the remedial regulation model establishes a regulatory regime that punishes a lack of competition by prohibiting private ordering, and all of the benefits that it brings, until sufficient competition can be shown. This discourages incumbents from amassing market power, since increased market share will only limit their ability to engage in private negotiation and deal making. It also mitigates some of the potential drawbacks of private ordering, such as the misrepresentation of “market” rates, and adverse selection consequences borne by smaller licensors when their larger counterparts withdraw.

2. *Applicability*

Certain features make an industry particularly well suited for remedial regulation. The first is lack of a functioning market. In the music licensing context, the concern is anticompetitive behavior such as tacit collusion and parallel pricing, but breakdowns in market function are found in other places as well, for example, in industries specializing in differentiated and/or complementary products; i.e., products which do not face perfect competition. A regulatory scheme that punishes dominance by denying dominant players the freedom to negotiate privately is a scheme that discourages such convergence in the first place.

Second is the inapplicability of another legal regime, such as the inapplicability of antitrust law to the anticompetitive behavior exemplified by the individual music publishing companies. Where an extant set of laws is unable to modify unwanted or undesirable behavior, remedial regulation may be able to achieve the desired result with minimal statutory amendment and lowered costs by utilizing existing regulatory bodies and agencies.

Remedial regulation is particularly well suited to industries already subject to extensive regulation. In those cases, the regulatory regime is likely already working at cross-purposes with a more hands-off approach—such as the inherent conflict between copyright’s limited monopoly grant and antitrust law. The existence of an extensive regulatory regime also makes remedial regulation an easy transition: First, the industry players are accustomed to regulation; second, many of the statutory terms that will come to constitute the remedial regulation are likely to already be in place. The same is true of existing agencies, giving

remedial regulation an advantageous starting point, and lowering the cost of implementation.

Finally, remedial regulation can also be useful in certain contexts—such as highly technical fields, like IP—where lawmakers and administrators are unlikely to fully understand the nuances in a field, and/or where the industry players have a distinct advantage over lawmakers in their ability to anticipate and respond to new developments. In those markets, remedial regulation places the burden on the parties to make a case for themselves and the market in which they operate. This puts the informational demands on the party best equipped to provide the relevant information, and at the least cost.

3. *Challenges*

The suggestion to implement additional regulation, remedial or otherwise, in an already highly regulated environment such as music licensing, is not made lightly. As a general matter, regulation begets regulatory gaming, or “private behavior that harnesses procompetitive or neutral regulations and uses them for exclusionary purposes.”²⁵⁸ In their seminal article on this topic, Professors Dogan and Lemley suggest that while regulatory gaming cannot generally be avoided *ex ante*, it may be checked by continued antitrust oversight of regulated markets.²⁵⁹ While not a wholesale fix, the fact that the DOJ and the FTC would continue to have jurisdiction over the music publishing companies should assuage gaming concerns here.

Another concern is the fact that remedial regulation would shift some of the antitrust oversight from the auspices of the Sherman Act to that of the Copyright Act; or, from Title 15 to Title 17. This opens competition policy up to the potentially negative influences of lobbying. Copyright—with its concentrated market power and disparate interests—is particularly susceptible to the influence of lobbyists.

Further, by mandating statutory licensing, remedial regulation forecloses the opportunity for a true market to develop. Without a real market, it is difficult to know whether the statutory rates set are economically efficient. This concern may be mitigated, however, by the uniform application of the rate across all parties.

258. Dogan & Lemley, *supra* note 218 at 687.

259. *See id.* at 688 (While “[r]egulatory agencies and even Congress cannot normally prevent gaming *ex ante* . . . some level of antitrust enforcement . . . provides a necessary check on behavior like product hopping that has no purpose but to exclude competition.”)

In his classic work on organizational discontent, Albert Hirschman describes a dichotomy of reactions from an unsatisfied player: The entity can either continue as a member and voice its complaints, or it can exit the organization and do business elsewhere.²⁶⁰ In the absence of sufficient competition, a remedial regulatory regime removes the exit option. This leaves disgruntled publishers to use their voices within ASCAP (or other designated collective), and/or under the compulsory license, where they may be able to wield their considerable market power to unfair advantage.

Part II.C.2 *supra* described the advantages of private ordering, including the ability to achieve a higher royalty rate, to negotiate licenses tailored to a particular piece of content and use, and to enjoy increased flexibility to respond to changes in technology and consumer preferences. It follows that another downside to remedial regulation is that it makes private ordering and the benefits it brings more difficult, time-consuming, and costly. In some instances, it may discourage private deal making altogether. This cost is arguably offset, however, by the societal gains from the elimination of anticompetitive behavior. Importantly, private ordering is not eliminated, but rather regulated. Abuses are checked, and access to content for all comers—established firms and prospective entrants alike—is ensured.

In addition, the remedial regulation model utilizes an existing regulatory agency, the CRB, for rate setting. To the extent that the existence of, and exercise of rate-setting authority by, the CRB poses separation of power issues, so too would a remedial regulation scheme that assigns rate setting to that agency.²⁶¹ Whether the CRB's functions should be moved to the executive branch, or vested elsewhere, however, is outside the scope of this Article, and not specific, nor limited, to the remedial regulation model.

260. See ALBERT O. HIRSCHMAN, EXIT, VOICE, AND LOYALTY: RESPONSES TO DECLINE IN FIRMS, ORGANIZATIONS, AND STATES (1970).

261. Writing on administrative agencies generally, Professor Jody Freeman has summarized the concern this way:

The combination of executive, legislative, and adjudicative functions in administrative agencies appears to violate the separation of powers principles embodied in the Constitution. Worse yet, despite their considerable discretionary power to impact individual liberty and property rights, allocate benefits and burdens, and shape virtually every sector of the economy, agencies are not directly accountable to the electorate.

Jody Freeman, *The Private Role in Public Governance*, 75 N.Y.U. L. REV. 543, 545–46 (2000).

Finally, and perhaps most importantly, regulatory price setting of the sort advocated herein “has been plagued by complicated valuation, allocation and second-best pricing problems that have bordered on insurmountable.”²⁶² Further, as noted in prior works, the default nature of compulsory licenses makes the set rates sticky, and parties may come to feel entitled to the regulatory rate regardless of its efficiency or appropriateness.²⁶³ But regulatory price setting is already status quo for the music licensing industry, such that its further encouragement does not worsen the situation therein. To the contrary, to the extent that remedial regulation can work to eliminate anticompetitive behavior, its benefits may well outweigh any costs. If the model works—i.e., if it indeed discourages consolidation of market power and entities’ petitions for private ordering are granted—the market will eventually move away from regulatory price setting altogether.

4. *Advantages*

The most obvious advantage of the remedial regulation model is that it discourages anticompetitive behavior. The possibility of private ordering is also preserved, so long as some do not reap its advantages at the expense of others. Importantly, remedial regulation maintains a point of access for content by leaving existing statutory licenses and a collective in place.

The additional costs borne by a party wishing to petition for private ordering are justified by the potential upside of private licensing over licensing under the statute, and are further excused by the optional nature of petition: No one has to petition, and the statutory license is always available as a point of access. Utilizing the FTC for regulatory review makes further use of an existing regulatory body readily possessed of all necessary authority for so doing. Review of petitions for circumvention can also help to further inform the reviewing agency of current market conditions. This information can lead to better merger reviews in the future.

The remedial regulation model is cost-effective in that it requires only minimal statutory amendment to remove the non-mandatory default from existing statutory licenses. The establishment of a new compulsory license for public performance rights likewise imposes minimal legislative burden as it models itself after the existing compulsory license for digital public

262. Christopher S. Yoo, *The Economics of Network Access*, 28 ADMIN. & REG. L. NEWS 5 (2003). In the music context specifically, the history of webcasting royalty determinations is arguably an unmitigated disaster. See DiCola & Sag, *supra* note 31 (suggesting the problems owe to flawed institutional design).

263. See García, *supra* note 15 at 1157–59.

performance of sound recordings, including the designation of a sole collective—quite possibly one already in existence—for royalties administration.

In the long-term, implementing a penalty default regulatory scheme will encourage innovation by ensuring a competitive environment in which new entrants are not discouraged from starting new services by the threat that their success might be punished with rate coercion. In essence, the remedial regulation model calls for a policy of facilitating the emergence and continuation of rival content providers and distributors. Where a robust number of providers cannot exist—owing to market conditions or otherwise—the remedial regulation model ensures all players face the same market conditions. As detailed in previous work, private ordering offers a number of benefits over the statutory regime, but a statutory rate is still better overall than an artificial “market” rate imposed by a single dominant firm.²⁶⁴

The remedial regulation model’s use of mandatory statutory licenses also works to ensure payment to artists. For example, the statutory license for sound recordings contains a section calling for mandatory distribution of royalties to artists.²⁶⁵ Section 114(g)(2) mandates a 50/50 split of incoming royalties between the creator herself and the intermediary to whom the copyright is assigned:

- 50% of receipts shall be paid to the copyright owner;
- 2.5% of receipts shall be deposited in an escrow account for distribution to non-featured musicians;
- 2.5% of receipts shall be deposited in an escrow account for non-featured vocalists; and
- 45% of receipts shall be paid to the featured recording artist on the sound recording.²⁶⁶

Such assurances for artists aren’t limited to statutory licenses. Article XVII(1)(c) of ASCAP’s Articles of Association likewise specifies the distribution of royalties as “one-half thereof to be distributed among the ‘Music Publisher’ members, and one-half among the ‘Composer and

264. *See generally* García, *supra* note 15 (discussing the potential for manipulation of “market” rates).

265. 17 U.S.C. § 114(g) (2012).

266. 17 U.S.C. § 114(g)(2)

Author' members, respectively."²⁶⁷ Article XX, Section 4 further notes that "[t]he royalties, or the right to participate in the royalties, and the rights of the members in the Society, shall not be sold or otherwise disposed of"²⁶⁸ When either § 114, or ASCAP (or both) is circumvented, so too is this guaranteed payment to the artist, who becomes subject to whatever portion their contract with the intermediary allows for (which, in the case of many songwriters, is nothing).

Remedial regulation's newly established compulsory license for the public performance of musical compositions could offer similar payment guarantees that cannot be circumvented unless and until a successful petition is filed. Even then, a minor statutory amendment could dictate that the artist protection portions are not circumventable.²⁶⁹ These artist protections can be made inalienable, such that where a firm successfully petitions circumvention by a showing of sufficient competition, the artist distribution portion of the circumvented statutory license remains in force.

This comports with legislative intent in setting up these protections in the first place: "The Committee intends the language of section 114(g) to ensure that a fair share of the digital sound recording performance royalties goes to the performers according to the terms of their contracts."²⁷⁰ According to Representative Conyers, this provision was adopted to "ensure[] that musicians, vocalists, and artists receive their royalties from digital music directly from the collection agent instead of through other intermediaries."²⁷¹ The concern is obvious: In the absence of statutory protection for those musicians, vocalists, and artists, the intermediary record labels and music publishers might cut them out of their share of royalties, as indeed has been the case in recent deals.

VI. CONCLUSION

Despite the existence of more content now than ever before, there are far fewer *legal* distribution and licensing channels for that content. The dearth of competition in the licensing space is bad for users, artists, and

267. AM. SOC'Y OF COMPOSERS, AUTHORS, AND PUBLISHERS, ARTICLES OF ASSOCIATION OF THE AM. SOC'Y OF COMPOSERS, AUTHORS, AND PUBLISHERS 19–20 (2002), <http://www.ascap.com/~media/files/pdf/members/governing-documents/articles-of-association.pdf> [<http://perma.cc/S54Q-CE4N>].

268. *Id.* at 22.

269. *See* García, *supra* note 13, at 8 (proposing "a fidelity clause requiring parties who circumvent the compulsory license to adhere to the statutorily mandated distributions in order to obviate circumvention of statutory protections for non-parties").

270. H.R. Res. No. 104-274, at 24 (1995).

271. 148 Cong. Rec. H7047 (2002) (statement of Rep. Conyers).

innovators alike. By recognizing the value of private ordering in content licensing and aiming to facilitate it when competition is robust, remedial regulation works to maintain and encourage competition where antitrust law does not.

The question of how to best manage emerging technologies and the challenges they present has never been more pressing. The debate around net neutrality, for example, is essentially a debate about whether antitrust law or regulation is the better means of ensuring competition in the broadband Internet space. Some lawmakers have argued that “vigorous application of the antitrust laws can prevent dominant Internet service providers (ISPs) from discriminating against competitors’ content or engaging in anticompetitive pricing practices.”²⁷²

Others have pointed out that application of antitrust laws to ISPs requires an expansive interpretation of antitrust’s jurisdiction. These critics have expressed concern that antitrust, unlike regulation, “does not address the non-economic goals of net neutrality, including the protection of free speech and political debate.”²⁷³ Remedial regulation may well prove a sustainable option for the maintenance of competition among Internet service providers.

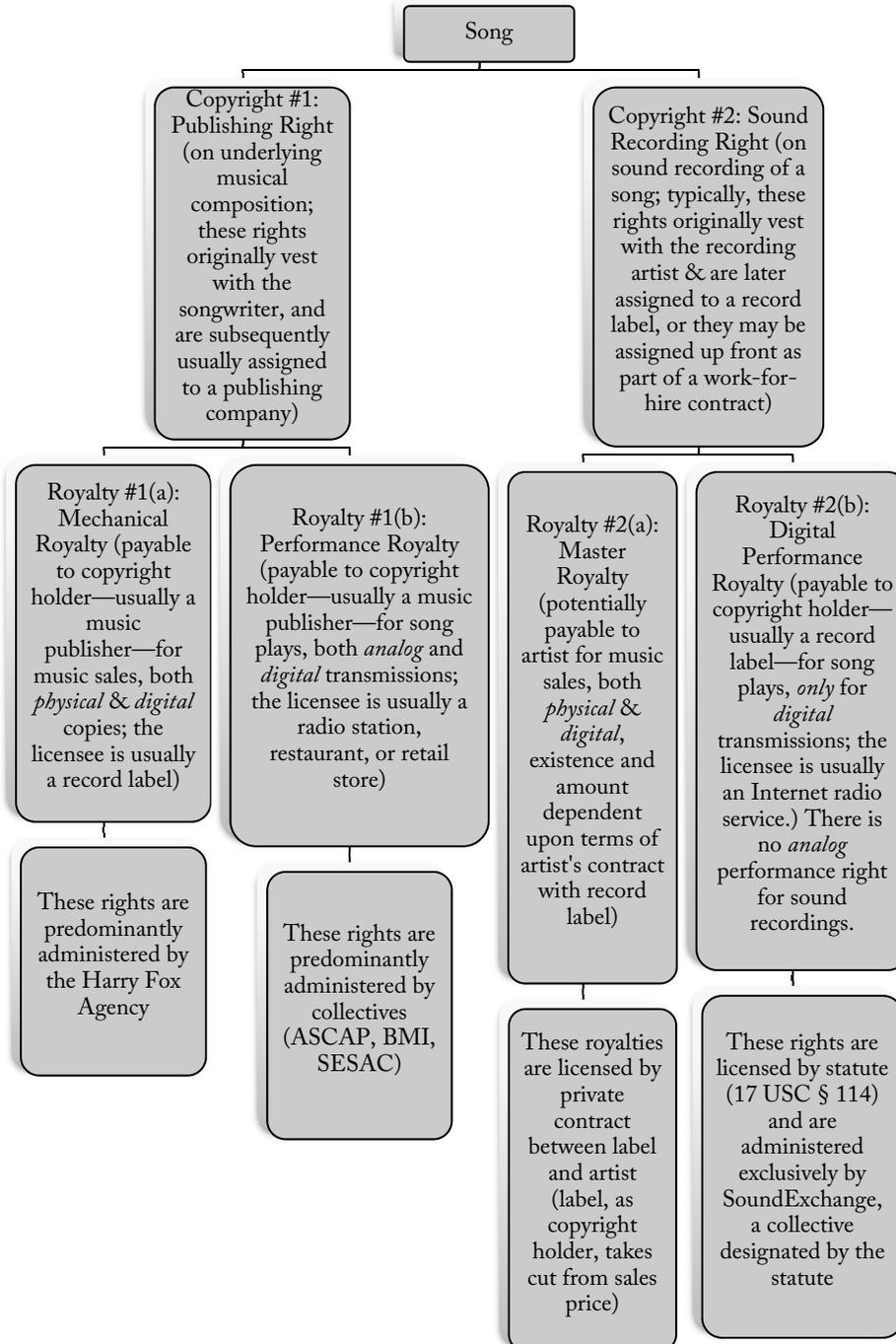
Acknowledging the potential for remedial regulation to maintain and encourage competition in industries, like music licensing, where antitrust has failed is only a starting point. Recognition of regulation as a tool for checking anticompetitive behavior offers lawmakers a means of ensuring continued growth and innovation in high tech sectors that have proven resistant to traditional antitrust enforcement. By shifting the burden of maintaining competition to the firms that wish to get out from under regulation, remedial regulation harnesses the power of private industry in order to serve public goals.

272. *Net Neutrality: Is Antitrust Law More Effective Than Regulation in Protecting Consumers and Innovation?*, 113th Cong. (2014) (statement of Hon. Bob Goodlatte, Chairman House Judiciary Committee), http://judiciary.house.gov/index.cfm/hearings?id=CCF704C2-E445-4D9A-8B84-26F6B9A59CC6&Statement_id=4CB6551B-8469-48C7-A155-8F62D9D172A6 [<http://perma.cc/VT76-QDVL>].

273. *Net Neutrality: Is Antitrust Law More Effective Than Regulation in Protecting Consumers and Innovation?*, 113th Cong. (2014) (statement of Ranking Member John Conyers, Jr.), <http://democrats.judiciary.house.gov/opening-statement/ranking-member-conyers-statement-net-neutrality-hearing> [<http://perma.cc/YB74-DNJG>].

APPENDIX

Music Copyrights Reference Grid



THE INFORMATIONAL VALUE OF PATENTS

Clark D. Asay[†]

ABSTRACT

Traditional patent law theories teach that a patent's rights of exclusion are a patent's key benefit to the patentee and are necessary to make the patent system work. Yet patentees are increasingly giving away such rights, in whole or in part, as part of a growing phenomenon: patent pledges. In these scenarios, patentees voluntarily commit to limit enforcement of their patent rights. This phenomenon seems to contradict traditional patent law theories. After all, if exclusive rights are necessary, why are patentees increasingly sacrificing some or all of those rights?

This Article argues that patentees do so because in patent pledging contexts, patents often entail a different value proposition than what traditional patent law theories posit. That is, patent pledgers use patents as tools to signal information about themselves and their innovation preferences to product, labor, and capital markets. This information may then facilitate a variety of economic purposes behind such pledges. This Article uses concepts from signaling theory in other disciplines to identify several patent law features that help make patents valuable as informational tools. It also reviews several recent Supreme Court cases and their possible implications for the informational value of patents identified in this Article. The Article concludes by arguing that these underappreciated informational roles of patents deserve greater consideration in formulating and tailoring patent law and policy, particularly in industries, such as information technology, where patent pledging is more common.

DOI: <http://dx.doi.org/10.15779/Z38J85N>

© 2016 Clark D. Asay.

[†] Associate Professor of Law, BYU Law School. JD, Stanford Law School. M.Phil, University of Cambridge. Many thanks to Stephanie Bair, Jorge Contreras, Brigham Daniels, Christopher Funk, Matthew Jennejohn, Michael Mattioli, David Moore, Amelia Rinehart, Sharon Sandeen, Andres Sawicki, Brett Scharffs, Katherine Strandburg, Lisa Grow-Sun, Liza Vertinsky, participants at a BYU Law faculty paper workshop, participants at the 2015 Patent Pledges Symposium held at American University College of Law, participants at a CLE presentation for the IP section of the Utah State Bar, and participants at the 2015 IP Scholars Conference for helpful comments on earlier versions of this Article.

TABLE OF CONTENTS

I.	INTRODUCTION	261
II.	TRADITIONAL PATENT LAW THEORIES	268
	A. PATENT LAW'S DISCLOSURE REQUIREMENTS	269
	B. HOW TRADITIONAL PATENT LAW THEORY TREATS THE POSSIBLE INFORMATIONAL VALUE OF PATENTS	270
III.	THE INFORMATIONAL VALUE OF PATENTS.....	275
	A. SIGNALING THEORY AND PATENTS	276
	B. HOW EXCLUSIVE RIGHTS CAN BOLSTER THE INFORMATIONAL FUNCTIONS OF PATENTS	278
	C. THE INFORMATIONAL VALUE OF PATENT LAW'S DISCLOSURE REQUIREMENTS	282
IV.	THE RISE OF PATENT PLEDGES	286
	A. TESLA.....	287
	B. MICROSOFT	294
	C. TWITTER	300
	D. IBM	304
	E. THE PATENT IMPOVERISHED	307
V.	SCOTUS AND THE INFORMATIONAL ROLE OF PATENTS	308
	A. THE INFORMATIONAL IMPACT OF RECENT PATENTABLE SUBJECT MATTER CASES.....	309
	1. <i>Patentable Subject Matter at the Supreme Court</i>	309
	2. <i>The Possible Informational Effects of the Alice Decision</i>	311
	B. DEFINITE CLAIMS	318
VI.	AN INDUSTRY-SPECIFIC PHENOMENON?	320
VII.	CONCLUSION.....	323

I. INTRODUCTION

Tesla made headlines in 2014 when its iconic CEO, Elon Musk, publicly pledged not to enforce the company's patents relating to electronic cars against those practicing the company's patented technologies.¹ Shortly thereafter, Toyota, one of Tesla's main competitors, followed suit with its own patent pledge relating to hydrogen fuel cell car technology.² And while these announcements garnered significant attention at the time, they are only the tip of the iceberg.³ Parties are increasingly engaging in "patent pledging," a phenomenon where parties voluntarily commit to limit enforcement of their patent rights.⁴ Hundreds of parties have made such pledges implicating thousands of patents, including the likes of Google, IBM, Fujitsu, Ford, Apple, Bank of America, Cisco Systems, Comcast, Facebook, Intel, LinkedIn, Twitter, Microsoft, Monsanto, Samsung, and many others.⁵ Furthermore, "FRAND" commitments—where parties voluntarily commit to license certain of their patents on "fair, reasonable and non-discriminatory" terms—have been prevalent in standard-setting contexts for some time, and have continued to grow in importance.⁶

The economic motivations behind patent pledges have been discussed extensively, particularly in the FRAND context.⁷ Often the economic

1. Elon Musk, *All Our Patent Are Belong to You*, TESLA BLOG (June 12, 2014), <http://www.teslamotors.com/blog/all-our-patent-are-belong-to-you> [<http://perma.cc/4KS4-SUXQ>].

2. Charlie Osborne, *Toyota Pushes Hydrogen Fuel Cell Cars with Open Patent Portfolio*, ZDNET (Jan. 6, 2015), <http://www.zdnet.com/article/toyota-pushes-hydrogen-fuel-cell-cars-with-open-patent-portfolio> [<http://perma.cc/XHB9-6N3K>] (discussing Toyota's move).

3. See Program on Info. Justice & Intellectual Prop.—Am. Univ. Wash. Coll. of Law, *Non-SDO Patent Statements and Commitments*, PROGRAM ON INFO. JUSTICE & INTELLECTUAL PROP., <http://www.pijip.org/non-sdo-patent-commitments> [<http://perma.cc/4XVZ-6MF9>] [hereinafter Program on Info. Justice] (cataloguing the known existing non-standards development organization (SDO) patent pledges).

4. See generally Jorge L. Contreras, *Patent Pledges*, 47 ARIZ. ST. L.J. 543 (2016) (discussing this phenomenon).

5. Program on Info. Justice, *supra* note 3; see also Contreras, *supra* note 4.

6. The literature on FRAND commitments is voluminous. For a review of some of this literature, see Jorge L. Contreras, *A Market Reliance Theory for FRAND Commitments and Other Patent Pledges*, 2015 UTAH L. REV. 479 (2015).

7. See, e.g., U.S. DEP'T OF JUSTICE & U.S. PATENT & TRADEMARK OFFICE, POLICY STATEMENT ON REMEDIES FOR STANDARDS-ESSENTIAL PATENTS SUBJECT TO VOLUNTARY F/RAND COMMITMENTS 5 (2013), <http://www.justice.gov/atr/public/guidelines/290994.pdf> [<http://perma.cc/AT4T-5VQJ>] (discussing the primary purposes behind FRAND commitments); Anne Layne-Farrar, A. Jorge Padilla & Richard Schmalensee, *Pricing Patents for Licensing in Standard-Setting Organizations: Making*

rationale behind FRAND commitments is to enable adoption of common technical standards, which grows the economic pie for everyone involved, while also providing patent owners with some economic remuneration for their patent rights.⁸ Outside of the FRAND context, scholars have also begun to identify a variety of economic purposes that may lead parties to give up or otherwise limit their valuable patent rights.⁹ Many of these economic purposes are similar to the purposes underlying FRAND commitments.¹⁰ For instance, one of Tesla's basic purposes in pledging its patents was to encourage others to collaborate with the company in making electronic car technologies a more widely available alternative to the established auto industry.¹¹

Yet while the economic motivations behind patent pledges may be clear in many cases, less clear is the role of patents in promoting such efforts. Do patents help facilitate such purposes, or are they simply an impediment to them? Much of the patent literature, with its extensive focus on patent "hold-up" problems, suggests the former.¹² In other words, patent pledges help address the threat of patent owners using their patents to "hold up" innovation.¹³ Thereafter, innovation can move forward with fewer impediments, and the economic pie can grow for everyone

Sense of FRAND Commitments, 74 ANTITRUST L.J. 671 (2007) (discussing the purposes behind FRAND commitments in general and issues in resolving what FRAND commitments actually mean); Mark A. Lemley & Carl Shapiro, *A Simple Approach to Setting Reasonable Royalties for Standard-Essential Patents*, 28 BERKELEY TECH. L.J. 1135 (2013) (discussing how adopting common technical standards helps facilitate competition and innovation, and indicating that requiring FRAND terms are a primary means of enabling such benefits).

8. See Lemley & Shapiro, *supra* note 7, at 1137.

9. See Contreras, *supra* note 4, at 30–49 (discussing possible economic motivations behind non-SDO patent pledges).

10. *Id.*

11. David Houlihan, *How Tesla Played 'Fox' With Its Patent Pledge*, BLUE HILL RESEARCH (June 23, 2014), <http://bluehillresearch.com/tesla-played-fox-with-its-patent-pledge> [<http://perma.cc/6WQX-GEB6>] (citing to several sources that point to this economic purpose behind Tesla's pledge).

12. See, e.g., Thomas F. Cotter, *Patent Holdup, Patent Remedies, and Antitrust Responses*, 34 J. CORP. L. 1151 (2009) (pointing to a variety of sources that identify patent holdup as a serious problem and rebuffing scholarship questioning the problems of patent holdup); Mark A. Lemley & Carl Shapiro, *Patent Holdup and Royalty Stacking*, 85 TEX. L. REV. 1991 (2007) (discussing the problems of patent hold-up generally); Mark A. Lemley, *Ten Things to Do About Patent Holdup of Standards (And One Not To)*, 48 B.C. L. REV. 149 (2007) (discussing ways to address the hold-up problem).

13. Lemley, *supra* note 12, at 156–59 (discussing this solution in the standard-setting context).

involved.¹⁴ But patents, rather than assisting with such efforts, are often viewed in the literature as impediments to be removed.¹⁵ In other words, they may play little if any positive role, at least at this later, post-invention stage of the innovation cycle.¹⁶ The same conclusions largely hold true in the “knowledge sharing” and “knowledge commons” literature, where patent rights, compared to other factors, are largely viewed as either problematic, or at least less relevant, in fostering innovation.¹⁷

In contrast, this Article argues that the growing patent pledging phenomenon makes manifest that patents have previously underappreciated informational value to patent holders and the relevant public alike. In other words, rather than simply being possible impediments to innovation post-invention, patents may serve a variety of different informational functions that help promote the types of economic

14. See Michael L. Katz & Carl Shapiro, *Network Externalities, Competition, and Compatibility*, 75 AM. ECON. REV. 424 (1985) (discussing the possible benefits of network effects, in which the value of a good or service increases the more other people use the same good or service. And in order for network effects to occur, patent licenses or pledges must often be secured).

15. They may, of course, play a role in incentivizing parties to develop the invention in the first place. See Lemley, *supra* note 12, at 156–59 (suggesting that parties in a standards organization should not be required to license their patents royalty-free, because doing so may discourage such parties from engaging in the inventive behavior in the first place). But once the patents are in place, they become, at least in much of the patent holdup literature, an obstacle to competition and collaborative innovation.

16. See *id.* But not all theories treat these later transactions as simply impediments, and may in fact view the economic opportunities associated with them as a beneficial aspect of the patent system. See, e.g., Robert P. Merges, *A Transactional View of Property Rights*, 20 BERKELEY TECH. L.J. 1477, 1487–90 (2005) (describing how property rights, including patent rights, may induce parties to disclose information before, during, and after contract formation that they otherwise may withhold for fear that the value of their property will be lost).

17. See, e.g., Michael J. Madison, Brett M. Frischmann & Katherine J. Strandburg, *Constructing Commons in the Cultural Environment*, 95 CORNELL L. REV. 657, 681–82 (2010) (“[P]atent and copyright laws construct particular environments with default boundaries governing access to and use of certain forms of knowledge. Commons arrangements grounded in those laws involve contextually specific deviations from the default given by IP law. These constructed cultural commons may lead to creativity, innovation, and improvement that would not be attainable either in the so-called ‘natural’ state of information without intellectual property protection or in the state of information with ‘full’ intellectual property protection.”); James Bessen & Alessandro Nuvolari, *Diffusing New Technology Without Dissipating Rents: Some Historical Case Studies of Knowledge Sharing* 1–3 (Bos. Univ. Sch. of Law, Working Paper No. 14-18, May 6, 2014), http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2433567 [<https://perma.cc/DZ7U-QRUW>] (indicating that knowledge sharing at certain stages of technological innovation increased, rather than dissipated, rents for innovators, while also concluding that patents presented such innovators during the same stages with very little value).

purposes behind patent pledges. This Article discusses these informational functions from a theoretical standpoint and then applies that theory to several case studies of actual patent pledges.

Traditional patent law theories have remained rather agnostic about any possible informational value of patents, at least as pertains to the patent holder.¹⁸ Indeed, those theories that do argue that patents have informational value typically posit that that value belongs to the public, not the inventor.¹⁹ In other words, because obtaining a patent requires the inventor to disclose a significant amount of technical information relating to the invention as part of the patent application process, the public may benefit from those disclosures since they become publicly available once the application is published or the patent issues.²⁰

But according to traditional patent law theories, the real prize for the patent holder is the exclusive rights that a patent bestows upon the patent owner.²¹ These rights may, depending on the theory, incentivize the inventor to do a variety of socially beneficial things—such as invent something in the first place, publicly disclose the invention, or commercialize it later.²² But these rights do not provide the patent holder with any informational value.²³ Indeed, the purported informational value that the public derives from the patent is in direct opposition to the inventor's exclusive rights.²⁴ In other words, the patent holder sacrifices

18. See, e.g., Timothy R. Holbrook, *Possession in Patent Law*, 59 SMU L. REV. 123, 146 (2006) (arguing that “disclosure obligations [are] inconsistent with the theoretical justifications of patent law”); Clark D. Asay, *The Informational Effects of Patent Pledges*, in PATENT PLEDGES—GLOBAL PERSPECTIVES ON PATENT LAW'S PRIVATE ORDERING FRONTIER (Jorge L. Contreras & Meredith Jacob eds., forthcoming 2016) (reviewing this agnosticism).

19. See, e.g., Jeanne C. Fromer, *Patent Disclosure*, 94 IOWA L. REV. 539, 542 (2009); Lisa Larrimore Ouellette, *Do Patents Disclose Useful Information?*, 25 HARV. J.L. & TECH. 545, 557–59 (2012).

20. See generally Fromer, *supra* note 19 (citing information disclosures as a significant possible benefit of the patent system, but arguing that the system, as currently implemented, largely fails to perform this function).

21. Mark A. Lemley, *The Myth of the Sole Inventor*, 110 MICH. L. REV. 709, 736–45 (2012) (describing utilitarian and commercialization theories as the predominant theories under patent law and making clear that in each school the prospect of exclusive rights is the key to encouraging inventors to develop inventions (utilitarian theory) and develop them post-invention (commercialization theory)).

22. *Id.* (outlining each of the predominant patent law theories and their typical rationales).

23. See Clarisa Long, *Patent Signals*, 69 U. CHI. L. REV. 625, 635 (2002) (indicating that traditional theories assume that inventors suffer losses when disclosing information and that exclusive rights are an attempt to balance that loss).

24. *Id.*

informational value in order to obtain her exclusive rights.²⁵ Her informational sacrifice is the “quid pro quo,” in the words of many courts and scholars, for her receiving exclusive patent rights.²⁶

But the patent pledging phenomenon shows that patents can provide patent holders and the public alike with significant informational value. This informational value comes in at least two different, but related, forms. First, patents provide informational value in the patent pledging context by enabling credible “signals” between the patent holder and participants in capital, labor, and product markets.²⁷ These signals may improve a sender’s ability to recruit talented employees, collaborate with competitors, and attract investment.²⁸ This Article draws on “signaling theory” from other disciplines in order to explain how patents can be useful in facilitating informational signals, thereby creating informational value for both patent owners and recipients thereof.²⁹

Second and relatedly, patents have informational value to patent owners and the relevant public based on the information disclosures the patenting process requires.³⁰ That is, these disclosures provide patent holders with a standardized, well-understood means by which to communicate information more generally.³¹ In turn, the public nature of

25. *Id.*

26. *See, e.g.*, *J.E.M. Ag Supply, Inc. v. Pioneer Hi-Bred Int’l, Inc.*, 534 U.S. 124, 142 (2001) (“The disclosure required by the Patent Act is ‘the *quid pro quo* of the right to exclude.’” (quoting *Kewanee Oil Co. v. Bicron Corp.*, 416 U.S. 470, 484 (1974))).

27. Clarisa Long has done some prior work in which she argues that patents signal information to investors about a firm’s quality. *See* Long, *supra* note 23. However, Long limits herself to a patent’s possible signals to capital markets, and she does not consider patents in pledging contexts or what signaling functions patents may play in such circumstances. This Article thus builds on Long’s work while identifying key differences with it.

28. *See infra* Part IV for an extensive discussion of how companies have used patent pledges in order to accomplish such goals.

29. The literature on signaling is rich. For a few examples discussing signaling in a variety of circumstances, *see* James D. Morrow, *The Strategic Setting of Choices: Signaling, Commitment, and Negotiation in International Politics*, in STRATEGIC CHOICE AND INTERNATIONAL RELATIONS 86, 86–91 (David Lake & Robert Powell eds., 1999) (discussing signaling between international actors); ERIC A. POSNER, LAW AND SOCIAL NORMS 19, 22 (2000) (discussing signaling between social actors generally); A. MICHAEL SPENCE, MARKET SIGNALING: INFORMATIONAL TRANSFER IN HIRING AND RELATED SCREENING PROCESSES 107–09 (1974) (discussing signaling in the job market context); David H. Moore, *A Signaling Theory of Human Rights Compliance*, 97 NW. U. L. REV. 879 (articulating a signaling theory to explain human rights compliance by nation-states).

30. For an overview of these disclosure requirements, *see infra* Section II.A.

31. *See, e.g.*, Ouellette, *supra* note 19, at 556 (noting that the patent system is “entrenched” as part of the international patent system).

the disclosures benefits the relevant public because that public has access to and can verify and assess the information.³² Many have argued that patent disclosures have little informational value because patents are often not scrutinized. The lack of scrutiny is for a variety of reasons. For instance, there are simply too many patents to read, the patents may not disclose useful technical information even if they are read, and reading patents may result in enhanced liability under patent law's willful infringement standards.³³ But in the patent pledging context, many of these concerns dissipate because a pledge will often identify specific patents for the public's review and use, thus making the possibility of later willful infringement allegations less of a concern. Consequently, though informational uncertainties may remain,³⁴ the public has greater incentives to scrutinize the disclosures relating to the pledged patents, thereby creating informational value to both patent holders and the affected public.

These informational merits of patents thus suggest that the exclusionary value of patents, which predominant theories typically view as a patent's key value,³⁵ need not and should not dominate debates about how best to reform patent law. Instead, the informational value of patents should become a consideration in assessing patent law theory, cases, and doctrines, too. And this may be particularly so in industries, such as software and information technology (IT) more generally, where patent pledging is more typical. Indeed, some scholars have argued that patent law already is and should become even more industry-specific,³⁶ and the

32. Long, *supra* note 23, at 665–66 (discussing costs associated with verifying patented information, noting that some aspects of verification present low costs while some present high costs).

33. See, e.g., Alan Devlin, *The Misunderstood Function of Disclosure in Patent Law*, 23 HARV. J.L. & TECH. 401, 402–03 (2010) (concluding that patents are ineffective at conveying useful information to the public); Holbrook, *supra* note 18, at 146 (arguing that “disclosure obligations [are] inconsistent with the theoretical justifications of patent law”); Note, *The Disclosure Function of the Patent System (or Lack Thereof)*, 118 HARV. L. REV. 2007 (2005) (concluding that the patent system largely fails to disclose useful technical information to the public).

34. Lemley, *supra* note 21, at 746 (arguing that several factors lead to patent documents that are often opaque and that, consequently, patents may often be of dubious informational value to those reading them).

35. *Id.* at 736–45.

36. See DAN L. BURK & MARK A. LEMLEY, *THE PATENT CRISIS AND HOW THE COURTS CAN SOLVE IT* (2009) (arguing that patent law should be tailored according to industry in order to better foster innovation); Dan L. Burk & Mark A. Lemley, *Is Patent Law Technology Specific?*, 17 BERKELEY TECH. L.J. 1155 (2002) (same); Dan L. Burk & Mark A. Lemley, *Policy Levers in Patent Law*, 89 VA. L. REV. 1575, 1581–95 (2003)

informational roles of patents identified in this Article lend additional support to such arguments.

None of this is to say that the other roles of patents that traditional patent law theories emphasize are irrelevant in light of the patent pledging phenomenon. Indeed, patents owners can and often do use patents in a variety of ways, depending on their economic purposes in any given situation. Sometimes those uses will conform to traditional patent law theories, while in other cases those uses may be more informational in nature. Such multifaceted uses of patents are not contradictory, but instead simply demonstrate that patents have a variety of uses beyond what traditional patent law theories teach. As this Article will argue, greater recognition of such multi-dimensional realities would serve patent law and policy well.

This Article proceeds as follows. Part II provides a brief overview of traditional patent law theories. It shows that these theories treat rights of exclusion as the key value of a patent because of those rights' exclusionary potential. And to the extent that these theories consider a patent's informational value, they posit that the public, not the patent holder, realizes such value. In other words, patents have negative informational impacts on patent owners because patent owners must sacrifice information about their inventions in order to obtain the real prize of exclusive rights. Indeed, such information disclosures may curtail those rights in a variety of circumstances, further making clear—at least according to these theories—the negative informational effects that such disclosures can have on patent holders.

Part III then provides an informational account of patent rights. Contrary to traditional theories, this account shows that patents can provide patent holders and the public alike with significant informational value. This Part thus argues that patent holders increasingly use their patents in patent pledging scenarios, *inter alia*, in order to credibly and efficiently signal information to product, labor, and capital markets about their research and development activities and preferences. These signals may then translate into a variety of economic opportunities for both the signalers and the recipients thereof. In so doing, patent holders and the relevant public both realize significant informational value from patents. This Part also examines what features of the patent system make patents valuable in these regards.

(arguing that innovation is industry specific and that patent law should more readily adapt to that reality).

Part IV then reviews several examples of patent pledges in support of Part III's informational account of patents. In particular, it reviews pledges from Tesla, Microsoft, Twitter, and IBM in order to illustrate the informational value that these companies and the public may realize from the companies' patent pledging activities.

Part V turns to several recent Supreme Court decisions relating to patent law. The Supreme Court has been particularly active in the last few years in tackling key patent law questions³⁷ but has not explicitly taken into account the informational value of patents in rendering its decisions. This Part examines these cases based on this Article's informational account of patents and argues that several of the cases will likely boost the informational value of patents.

Finally, Part VI urges courts and policymakers to take into account the informational value of patents in rendering decisions and changes in patent law, particularly in industries, such as IT, where patent pledging is more typical. Indeed, the informational role of patents in some industries, but not others, provides additional support for the industry-specific patent law tailoring for which others have argued.

II. TRADITIONAL PATENT LAW THEORIES

As this Part will demonstrate, traditional patent law theories largely fail to identify any possible informational value of patent rights. This is particularly so with respect to the patent holder. Indeed, the patent holder is assumed to have suffered an informational loss in pursuit of the real prize of a patent: exclusive rights. Even those theories that focus on the informational value of patents typically suggest that this value belongs to the public, not the patent holder. Hence, these accounts also assume that the patent holder suffers an informational loss as part of the patenting process.

Section II.A first lays out patent law's information disclosure requirements. It does so because these requirements are the basis for how traditional patent law theories assess the informational value of patents (or lack thereof). Section II.B then analyzes traditional patent law theories in light of these disclosure requirements, concluding that traditional theories

37. See Lisa Larrimore Ouellette, *Supreme Court Patent Cases*, WRITTEN DESCRIPTION BLOG, <http://writtendescription.blogspot.com/p/patents-scotus.html> [<http://perma.cc/W522-77JG>] (listing all patent law cases that the Supreme Court has taken since 1952, and indicating that in recent years the Supreme Court has heard "quite a number of patent cases").

largely ignore the informational value of patents that is manifest in patent pledging contexts.

A. PATENT LAW'S DISCLOSURE REQUIREMENTS

Patent law's information disclosure requirements generally come in five parts. First, patent law's "enablement" requirement stipulates that a patent applicant must disclose enough technical details in the application to enable someone of ordinary skill in the particular "art" to practice the invention "without undue experimentation."³⁸ Though this enablement standard does not require patentees to disclose every relevant detail of their invention, it does require that patent applicants disclose sufficient technical information such that some skilled person would be able to replicate the invention without significant obstacles.³⁹

Second, patent law requires that an applicant disclose enough technical details so that a person of ordinary skill in the relevant field would recognize that the applicant had actually invented what she claims to have invented.⁴⁰ While this "written description" requirement often overlaps significantly with the enablement requirement discussed directly above, there are cases where this latter requirement forces the applicant to disclose more technical details than the enablement standard, alone, would require.⁴¹

Third, patent law requires patent applicants to disclose their "best mode" of practicing the invention.⁴² Though recent patent law reforms have significantly weakened this requirement, it technically remains on the books.⁴³ It thus may remain important in ensuring that patent applicants disclose technical details that they otherwise would withhold.⁴⁴

38. 35 U.S.C. § 112(a) (2012); *see also* Sean B. Seymore, *Heightened Enablement in the Unpredictable Arts*, 56 UCLA L. REV. 127, 139–54 (2008).

39. *See, e.g.*, *Invitrogen Corp. v. Clontech Labs., Inc.*, 429 F.3d 1052, 1058, 1070–71 (Fed. Cir. 2005) (finding sufficient enablement based in part on tacit knowledge of one skilled in the art, and clarifying that enablement does not require disclosure of every detail of the claimed invention).

40. 35 U.S.C. § 112(a); *see also* Dennis Crouch, *An Empirical Study of the Role of the Written Description Requirement in Patent Examination*, 104 NW. U. L. REV. 1665, 1669 (2010) (discussing the role of the written description requirement generally).

41. *See, e.g.*, *Gentry Gallery, Inc. v. Berkline Corp.*, 134 F.3d 1473 (Fed. Cir. 1998).

42. § 112(a).

43. *See, e.g.*, Brian J. Love & Christopher B. Seaman, *Best Mode Trade Secrets*, 15 YALE J.L. & TECH. 1 (2012) (discussing the gutting of this requirement, while also noting that it nonetheless remains a requirement that may still be legally enforceable).

44. *Id.*

Fourth, patent law also requires applicants to state their inventions in definite and clear terms in the form of “claims” towards the end of the patent document.⁴⁵ Patent claims are often analogized to the “metes and bounds” of a land deed, whereby the applicant charts out the periphery or boundaries of that to which they purport to have exclusive rights.⁴⁶ The frequently cited rationale for requiring patent claims is that they help provide the public with clearer notice of what the patent holder alleges to own, and what, therefore, remains unencumbered.⁴⁷

Last, generally all of these disclosures eventually become publicly available once a patent application is published and indexed, typically eighteen months from the time the patent application was filed, or once the patent actually issues.⁴⁸ Hence, in most cases the disclosures required as part of the application process become accessible to the public.

B. HOW TRADITIONAL PATENT LAW THEORY TREATS THE POSSIBLE INFORMATIONAL VALUE OF PATENTS

With few exceptions, traditional patent law theories view these information disclosures as a sacrifice that an inventor must make in order to obtain the real prize of a patent: exclusive rights. And if anything, these theories typically view these information disclosures as a potential threat to the rights that inventors do have.

For instance, the dominant theory behind patent law is utilitarian, sometimes called “economic incentives” theory.⁴⁹ This school of thought views patent rights as necessary economic incentives to inventive behavior.⁵⁰ That is, patents are necessary because, without granting these exclusive rights, inventors may be reluctant to engage in inventive activity for fear that they will not be able to internalize the benefits of their

45. § 112(b); *see also* *Nautilus, Inc. v. Biosig Instruments, Inc.*, 134 S. Ct. 2120 (2014) (setting forth a new standard for determining whether patent claims are “definite” enough); *see generally* Jeanne C. Fromer, *Claiming Intellectual Property*, 76 U. CHI. L. REV. 719 (2009) (discussing claiming within patent law generally and contrasting it with the claiming system within copyright law).

46. Richard L. Wynne, Jr., *Patent Law: Warner-Jenkinson Co. v. Hilton Davis Chemical Co.: How Can the Federal Circuit Control the Doctrine of Equivalents Following the Supreme Court’s Refusal to Set the Standard?*, 50 OKLA. L. REV. 425, 436 (1997) (discussing this often-employed analogy and citing cases that use it).

47. *Id.* at 444–45.

48. § 122.

49. Lemley, *supra* note 21, at 736–38 (describing the utilitarian model as “orthodox”).

50. *Id.*

inventive activity.⁵¹ The primary basis for this fear is that inventive ideas are non-rivalrous and non-excludable, that is, inventions have the characteristics of a public good.⁵² Consequently, absent patent rights, competitors could simply practice the inventive idea without incurring the same costs that the inventor did in developing it.⁵³ A patent's rights of exclusion purportedly address this concern and thereby incentivize parties to engage in inventive behavior.⁵⁴

Traditional applications of utilitarian theory thus largely ignore any possible informational value of patents to the patentee. And to the extent that the informational dimensions of patents are taken into account, they are framed negatively. As one scholar notes, inventors "are assumed to suffer losses when information is made public, a loss exclusive rights attempt to compensate."⁵⁵ Other scholars question whether these exclusive rights adequately compensate for the disclosures, given that a patent's information is typically made public before the exclusive rights are actually granted.⁵⁶

Indeed, patents' negative informational impact on inventors may be even more pronounced when disclosures made in compliance with patent law ultimately limit the scope of a patent holder's claims. For instance, while courts are generally not supposed to read limitations from these information disclosures into patent claims,⁵⁷ core patent law doctrine also requires that patent claims be read in light of the patent specification, where these disclosures are found.⁵⁸ As discussed above, patent claims are not supposed to exceed whatever the inventor has enabled others to practice through her disclosures, or what the inventor has shown to have invented through the same disclosures.⁵⁹ So to the extent that an inventor claims more technical territory than is justified on the basis of her information disclosures, the information disclosures may operate to limit

51. Mark A. Lemley, *The Economics of Improvement in Intellectual Property Law*, 75 TEX. L. REV. 989, 993–94 (1997).

52. Holbrook, *supra* note 18, at 132–33 (discussing these aspects of intellectual works as well as their purported consequences); Eric E. Johnson, *The Economics and Sociality of Sharing Intellectual Property Rights*, 94 B.U. L. REV. 1935, 1940–42 (2014) (same).

53. Sources cited *supra* note 52.

54. Lemley, *supra* note 51.

55. Long, *supra* note 23, at 635.

56. Holbrook, *supra* note 18, at 132–35.

57. *Phillips v. AWH Corp.*, 415 F.3d 1303, 1312–23 (Fed. Cir. 2005) (en banc) (reviewing the standards for construing patent claims).

58. *Id.*

59. *See supra* Section II.A.

their rights of exclusion. And since exclusive rights are the key benefit under the utilitarian model, information disclosures and the possible limitations that they impose on those rights come out a net negative under this theoretical framework. In sum, under the utilitarian model patents fail to provide inventors with informational value, instead requiring, if anything, an informational sacrifice.⁶⁰

Other predominant patent law theories are similar to traditional utilitarian theory in treating the informational impact of the patenting process on inventors as either a possible negative (because they may work to limit a patentee's exclusive rights) or an outright sacrifice. For instance, prospect theory treats exclusive rights as the key to incentivizing post-invention research and development.⁶¹ According to this theoretical school, granting broad exclusive rights to inventors early gives them the proper economic incentives to further develop, commercialize, and license their inventions for the benefit of society.⁶² A related school of thought, commercialization theory, posits that additional exclusive rights are often necessary in order to ensure that inventions are commercially developed for society's benefit.⁶³

Hence, in both of these accounts and similar to utilitarian theory, exclusive rights are the key to making the patent system work. And these rights are crucial to inventors because of the direct economic benefits that they promise, not because of any informational value that patent rights may otherwise represent.⁶⁴ For instance, information disclosures, to the extent that they figure into these theoretical accounts, are either a possible

60. Of course, another view may be that utilitarian theory is either agnostic to or even supportive of patent law's information disclosure requirements. After all, if the fit between a patent's claims and information disclosures is properly done, the information disclosures support, rather than possibly diminish, a patentee's exclusive rights. But even under this view, exclusive rights remain the key benefit of the patent. Information disclosures only become a benefit to the extent they bolster such rights; they have no inherent value of their own, at least according to the utilitarian model in patent law.

61. Edmund W. Kitch, *The Nature and Function of the Patent System*, 20 J.L. & ECON. 265, 266–71 (1977).

62. *Id.*

63. See, e.g., Michael Abramowicz & John F. Duffy, *Intellectual Property for Market Experimentation*, 83 N.Y.U. L. REV. 337 (2008) (arguing that current IP law may not adequately incentivize market experimentation and suggesting that it may need to expand in order to provide the proper level of incentives for such experimentation); Ted Sichelman, *Commercializing Patents*, 62 STAN. L. REV. 341 (2010) (proposing a "commercialization" patent, in addition to traditional patents, in order to better incentivize post-invention commercialization of patented technologies).

64. BURK & LEMLEY, *supra* note 36, at 1600 (discussing the economic importance of licensing opportunities under traditional prospect theory).

diminution of those rights or, at best, a support to them.⁶⁵ But even in the supporting role, the disclosures do not appear to have any inherent value of their own. In other words, patents have little informational value to the patent holder under these accounts. And any informational value of patents that may accrue to the public is irrelevant to the functioning of patents that these theories envision.⁶⁶

“Disclosure” and “coordination” theories are even clearer that patent law’s information disclosure requirements are a necessary sacrifice on the part of the inventor. These theories generally posit that inventors will be reluctant to share technical information relating to their inventions, either with the public or other third parties, without having exclusive rights in their inventions.⁶⁷ This hesitancy stems in part from a fear that, without exclusive rights in the inventions, inventors will fail to capture the economic value of their inventions, while others will.⁶⁸ Hence, in order to properly incentivize inventors to share information regarding their inventions, patent law grants inventors exclusive rights in their patented inventions. Those rights purportedly enable them to better capture the economic value of their inventions.⁶⁹ But in return, inventors are required to satisfy patent law’s information disclosure requirements by providing technical information that, according to some strands of these theories, they would otherwise keep secret.⁷⁰ Hence, while the public may obtain

65. Holbrook, *supra* note 18, at 135–36 (arguing that the “teaching function” of patents is irrelevant to prospect theory).

66. *Id.*

67. *See, e.g.,* *Kewanee Oil Co. v. Bicron Corp.*, 416 U.S. 470, 481 (1974) (describing the inventors’ disclosures as the “quid pro quo” for granting the inventor exclusive rights in their invention); WILLIAM M. LANDES & RICHARD A. POSNER, *THE ECONOMIC STRUCTURE OF INTELLECTUAL PROPERTY LAW* 326–30 (2003) (elaborating on the coordination function and value of patents); Paul J. Heald, *A Transaction Costs Theory of Patent Law*, 66 OHIO ST. L.J. 473, 488–89, 497 (2005) (arguing that patents help address concerns about information misappropriation and thus encourage information sharing); Robert Mazzoleni & Richard R. Nelson, *Economic Theories About the Benefits and Costs of Patents*, 32 J. ECON. ISSUES 1039–40 (1998) (describing different strands of disclosure theory generally); *Merges, supra* note 16 (describing how property rights, including patent rights, may induce parties to disclose information before, during, and after contract formation that they otherwise may withhold for fear that the value of their property will be lost).

68. *Merges, supra* note 16; *see also* Nancy T. Gallini, *The Economics of Patents: Lessons from Recent U.S. Patent Reform*, 16 J. ECON. PERSP. 131, 132 (2002) (suggesting that patents may induce parties to disclose information that they may otherwise withhold).

69. *Merges, supra* note 16.

70. Ouellette, *supra* note 19, at 555–62 (discussing this “quid pro quo” strand of disclosure theory and its critiques).

informational value from an inventor's patent,⁷¹ the inventor herself loses informational value in exchange for exclusive rights.

Of course, it is certainly true that under coordination and disclosure theories, patents have informational value to inventors in the sense that exclusive rights allow inventors to better appropriate the value of their inventions. As discussed, these theories generally posit that patents allow inventors to derive value from their technical information by protecting it against appropriation.⁷² But patents are still a net negative to inventors in terms of informational impact. This is so because, all else being equal, these theories suggest that inventors would rather keep the information to themselves. While the economic prospect of exclusive rights may motivate them to disclose their information as part of a patent application,⁷³ these theories indicate that if inventors could simultaneously obtain these exclusive rights while keeping their inventions secret, they would.⁷⁴ Hence, while exclusive rights may provide inventors with significant economic compensation in return for their disclosures, the disclosures nonetheless remain a loss to inventors.⁷⁵

A remaining theoretical camp—"natural rights" theorists⁷⁶—is, at first blush, somewhat less straightforward in how it incorporates the informational value of patents—or lack thereof—within its reasoning. For instance, these theories might be viewed as less instrumental than other theoretical camps, because they surmise that intellectual property rights, including patent rights, arise based in large part on the effort and/or personhood with which an inventor has imbued her invention.⁷⁷ And if they are less instrumental because they are not focused on inducing

71. *Id.* at 559–62 (providing evidence that patent disclosures do in fact provide value to parts of the public, contrary to some earlier critiques).

72. *See supra* note 67 and accompanying text.

73. Andrew A. Schwartz, *The Corporate Preference for Trade Secret*, 74 OHIO ST. L.J. 623, 630–33 (2013) (discussing the traditional understanding of how parties choose between trade secrecy and patents for their inventions and innovations).

74. Long, *supra* note 23, at 635 (indicating that inventors are assumed to suffer losses when disclosing information and that exclusive rights are meant to balance that loss).

75. *Id.*

76. *See* Mark A. Lemley, *The Rise of Faith-Based IP*, 62 UCLA L. REV. 1328, 1337 (2015) (describing and critiquing this group of scholars on the basis of their purported lack of reliance on evidence); ROBERT P. MERGES, JUSTIFYING INTELLECTUAL PROPERTY (2011).

77. *See* Wendy J. Gordon, *A Property Right in Self-Expression: Equality and Individualism in the Natural Law of Intellectual Property*, 102 YALE L.J. 1533 (1993); Merges, *supra* note 76; David Opderbeck, *Beyond Bits, Memes, and Utility Machines: A Theory of Intellectual Property as Social Relations*, 10 U. ST. THOMAS L.J. 738 (2013); Margaret Jane Radin, *Property and Personhood*, 34 STAN. L. REV. 957 (1982).

invention, disclosure, commercialization, or some other similar goal, this may mean that the role that information disclosures and exclusive rights play in these theories is less relevant—or less clear—than in other theoretical camps.

Yet in vital respects, exclusive rights seem to play a similar role under natural rights theories as they do in other theoretical camps. That is, even if the basis for such rights under natural rights theories differs, exclusive rights remain key because they help protect the effort and/or personhood of the inventor.⁷⁸ And to the extent that information disclosures made as part of the patenting process limit those rights, they would have a negative impact on the natural rights of inventors. Or if the information disclosures are properly calibrated to what the inventor actually hopes to protect, they may be viewed in some sense as a positive. But as with other theoretical schools, exclusive rights remain a patent's key value, even if the information disclosures ultimately support that value. The informational value of patents, in other words, may subsist, if at all, in apprising the public of the inventor's exclusive rights. But those rights are ultimately the key value of the patent in the natural rights camp, too, because they protect an inventor's purported inherent interests.

III. THE INFORMATIONAL VALUE OF PATENTS

Under predominant patent law theories, then, patents have little if any informational value to the patent holder. The information disclosures required as part of the patenting process may certainly have some value to the public at large, which remains a topic of vigorous scholarly debate.⁷⁹ But the real prize for the patent owner remains a patent's exclusive rights because of the economic possibilities that those rights represent. And to the extent that patent law's disclosure requirements detract from those rights, these information disclosures may have a negative impact on the patent owner.

This Article argues, however, that patents can provide significant informational value to the patent owner, not just the public. And the information disclosures that patent law requires, as well as exclusive rights, are key to creating that value, as the following sections explore.

78. Gordon, *supra* note 77; Radin, *supra* note 77.

79. See, e.g., Fromer, *supra* note 19, at 547–96 (concluding that current patent law disclosure requirements do not adequately stimulate innovation); Ouellette, *supra* note 19, at 552–566 (concluding that nanotech researchers often do rely on patented information in their innovative efforts).

A. SIGNALING THEORY AND PATENTS

As briefly mentioned in the Introduction, one of this Article's primary arguments is that patent holders use patents in patent pledging contexts in order to credibly signal information to product, labor, and capital markets, and that these signals can result in the sender and recipient of such signals realizing significant informational value. In order to better show how this may be so, this Part now turns to a brief discussion of signaling theory.

Signaling theory has a long history in a variety of disciplines.⁸⁰ This Article does not attempt to exhaustively canvass that extensive literature. Nor does it suggest that any particular strand of signaling theory precisely matches the behaviors of patent pledgers in signaling information to product, labor, and capital markets in each instance. Instead, it argues that some basic tenets of signaling theory provide a useful conceptual lens by which to better understand the previously underappreciated informational functions of patents in patent pledging contexts.

A signal can be understood as a costly behavior meant to communicate information about the sender to recipients.⁸¹ And the cost of the behavior is a crucial piece of the informational communication.⁸² That is, recipients of the signal appreciate that "only senders with a particular characteristic can afford, or are willing, to send the signal."⁸³ Appreciation of this characteristic based on the signal may then result in some mutually beneficial outcome between the parties.⁸⁴ The signal thus communicates information to recipients that, but for the signal, they may otherwise have difficulty discerning or trusting.⁸⁵

An example helps better illustrate this theory in practice. For instance, educational attainment may function as a signal to potential employers that a candidate will be productive.⁸⁶ This is so, in part, because of the opportunity, monetary, and psychological costs involved in obtaining an education.⁸⁷ Without these costs, the signal may be a poor proxy for future productivity, since everyone could obtain the same credentials at no cost.⁸⁸

80. See sources cited *supra* note 29 and accompanying text.

81. See Moore, *supra* note 29.

82. *Id.*

83. *Id.* at 882.

84. *Id.*

85. *Id.*

86. See Michael Spence, *Job Market Signaling*, 87 Q.J. ECON. 355, 358 (1973).

87. See *id.*

88. See *id.*

This signal may thus help employers make more rational hiring choices, as well as help potential employees obtain well-paying jobs.⁸⁹

In patent pledging contexts, patents can play similar roles. Some previous scholarship has applied signaling theory to patent law. For instance, Clarisa Long has argued that obtaining patents is valuable to parties because doing so credibly signals to capital markets that the party obtaining the patent is a firm worth investing in, which is a characteristic that may be otherwise difficult to verify.⁹⁰ That signal is credible in part because of the costs that obtaining a patent may entail, as well as the ability of third parties to verify some of the information that patents convey.⁹¹ Such signals may then translate into market support for the party obtaining the patent.⁹² Long points to this function of patents in attempting to explain why parties continue to spend time and resources amassing patents, when the evidence shows that very few patents are ever enforced or translate into direct economic benefits in the form of royalties or other economic remuneration.⁹³ Other studies have provided some empirical evidence in support of Long's arguments.⁹⁴

This Article builds on these previous studies, while arguing that the informational potencies of patents go beyond what they propose. Patents provide informational value to inventors and the public, not only within

89. *Id.*

90. Long, *supra* note 23.

91. *Id.*

92. Ted Sichelman & Stuart J.H. Graham, *Patenting by Entrepreneurs: An Empirical Study*, 17 MICH. TELECOMM. & TECH. L. REV. 111, 115 (2010) (finding evidence that many startups rely "heavily on patents as signals to the market to improve their chances of raising financing, being acquired, and going public," and that this evidence lends support to signaling patent law theories).

93. Long, *supra* note 23, at 626–27; *see also* Edmund W. Kitch, *Property Rights in Inventions, Writings, and Marks*, 13 HARV. J.L. & PUB. POL'Y 119, 122–23 (1990) (concluding that most patents are so narrow that they are relatively worthless); Mark A. Lemley, *Rational Ignorance at the Patent Office*, 95 NW. U. L. REV. 1495, 1503–04 (2001) (showing that many issued patents are abandoned, presumably because of their weak economic prospects); Robert P. Merges, *As Many as Six Impossible Patents Before Breakfast: Property Rights for Business Concepts and Patent System Reform*, 14 BERKELEY TECH L.J. 577, 603 (1999) (concluding that most patented technologies will fail commercially and/or present few economic advantages).

94. *See, e.g.*, Stuart J.H. Graham et al., *High Technology Entrepreneurs and the Patent System: Results of the 2008 Berkeley Patent Survey*, 24 BERKELEY TECH. L.J. 1255, 1307 (2009) (suggesting that patents are important to startups as a signal to investors of discipline and quality); Darian M. Ibrahim, *Financing the Next Silicon Valley*, 87 WASH. U. L. REV. 717, 750–51 (2010) (discussing that patents have value as signals to investors); Sichelman & Graham, *supra* note 92 (same).

capital markets, but also within labor and product markets.⁹⁵ The phenomenon of patent pledging, as will be discussed more fully in Part IV *infra*, provides significant evidence of patentees using patents as informational tools in this broader range of scenarios.

Furthermore, Long's work seems to largely treat features of patents, such as rights of exclusion and information disclosures, as somewhat irrelevant to a patent's signaling functions. For instance, Long does not appear to view rights of exclusion as important to her signaling theory of patents.⁹⁶ Indeed, she notes that her article is meant to show that exclusive rights represent only one particular function of a patent.⁹⁷ Even patent law's information disclosures may be irrelevant to her brand of signaling theory because, as other scholars suggest, the actual "disclosure of any given patent is not terribly relevant to the signal."⁹⁸ Instead, what is important for purposes of previous brands of signaling theory as applied to patent law is, simply, that a patent exists.⁹⁹ If the market had to actually review the contents of the patent in order to assess the technical merits of the invention, for instance, then the informational advantages of the patent purportedly diminish.¹⁰⁰

In contrast, this Article argues that, in patent pledging scenarios, both exclusive rights and information disclosures play a key role in creating a patent's informational value, as a signal and otherwise. And both patentees and the relevant public realize this value. The next Sections assess how.

B. HOW EXCLUSIVE RIGHTS CAN BOLSTER THE INFORMATIONAL FUNCTIONS OF PATENTS

As discussed, traditional patent law theories typically treat a patent's rights of exclusion as the key value of a patent.¹⁰¹ Exclusive rights are valuable under these views because they represent an economic prize to inventors that incentivizes them to develop, disclose, or further develop

95. Long does note the possibility of patents being useful in signaling information to labor markets. *See* Long, *supra* note 23. But her article primarily focuses on what signals patents convey to capital markets.

96. *Id.* at 627 (challenging "the traditional assumption that exclusivity is the alpha and the omega of the private value of patent rights" and instead arguing that patents have value as a tool for "credibly publicizing information").

97. *Id.*

98. Holbrook, *supra* note 18, at 137–38.

99. *Id.*

100. Long, *supra* note 23, at 665 (discussing how such assessments may offset whatever informational gains a patent may otherwise represent because, for instance, determining whether a patent is valid entails significant costs).

101. *See supra* Section II.B.

their inventions for the benefit of society.¹⁰² Or, in the natural rights camp, exclusive rights are valuable because they help protect an inventor's inherent rights in their inventions.¹⁰³

Rights of exclusion are also important in this Article's informational account of patents, but for different reasons. For instance, if a party wishes to signal to capital, labor, or product markets information about that party through a patent pledge, a patent's rights of exclusion make that signal much more credible. This is so because, as with traditional signaling theory, sacrificing these otherwise valuable economic rights as part of a pledge is costly behavior that may aid the signal's recipients in inferring attributes of the signaler that are otherwise difficult to verify or trust.¹⁰⁴ These signals may thus lead the signaler and recipient of the signal to pursue economic opportunities in accordance with the signal's apparent message.¹⁰⁵

To illustrate: a party wishing to signal its intention to collaborate with others around core technologies, to pursue an open model of innovation, or to provide a certain type of working environment to employees may not convey much of a message through mere lip service.¹⁰⁶ In other words, it is difficult to give credence to such intentions if the signaler does not back up those intentions with something more than simply a public announcement.¹⁰⁷ Purporting to sacrifice a government-sanctioned property right,¹⁰⁸ on the other hand, helps underscore the seriousness and credibility of the message, as well as the future intentions of the pledger. And this is so, in part, because such a sacrifice entails costs. Indeed, such costs may also aid the signal recipient in inferring characteristics about the signaler that are otherwise difficult to observe, such as the culture of the company more generally.

Relatedly, rights of exclusion help establish a key baseline against which an informational signal can be assessed. For instance, if a party

102. *Id.*

103. *Id.*

104. Spence, *supra* note 29.

105. *Id.*

106. See Jack L. Goldsmith & Eric A. Posner, *Moral and Legal Rhetoric in International Relations: A Rational Choice Perspective*, 31 J. LEGAL STUD. 115, 125–29 (2002) (referring to costless communication intended to facilitate coordination in international relations as cheap talk).

107. *See id.*

108. Under the Patent Act, patents are explicitly defined as a property right. See 35 U.S.C. § 261 (2012) (indicating that “patents shall have the attributes of personal property”).

possesses a trove of key patented inventions in a particular technological area and publicly pledges that the public may use them, the surrendered right of exclusion in such a case more ably conveys the magnitude and credibility of the commitment.¹⁰⁹ If, on the other hand, a party commits to license to the relevant public a few relatively insubstantial patents covering non-essential technologies, the signal might be of enduring exclusion rather than clear inclusion.¹¹⁰ Indeed, a party's failure to pledge patents at all, particularly as the patent pledging phenomenon spreads in IT and other industries, would also appear to signal to markets information about the non-pledger's business strategies and company culture more generally. But in each of these cases, a patent's initial endowment of exclusion is essential to effectively signaling to various markets the patent holder's intentions, as well as in some cases characteristics of the signaler that may otherwise be difficult to observe. And these signals, in turn, may ultimately lead parties to alter their behavior in reliance on the signal, such as two parties collaborating together or labor market participants pursuing employment opportunities with the signaler.¹¹¹

Of course, rights of exclusion can also potentially disserve informational purposes. Perhaps the most obvious reason is that enduring rights of exclusion may make relying on a supposed informational signal risky.¹¹² For instance, in the patent pledging context, scholars have wrestled with whether such pledges are legally enforceable, and whether they should be.¹¹³ Given such legal uncertainties, the informational takeaway for recipients of patent pledges may be simply one of avoidance.¹¹⁴ In other words, despite whatever informational signals a

109. In contrast, some have complained when parties have pledged to the public patents that they perceive as insignificant. *See, e.g.*, Florian Mueller, *Google's Promise Not to Assert 10 Patents Against Open Source Software: Just a PR Stunt*, FOSS PATENTS (Mar. 28, 2013), <http://www.fosspatents.com/2013/03/googles-promise-not-to-assert-10.html> [<http://perma.cc/VK99-J7F2>].

110. *Id.*

111. Contreras, *supra* note 4 (arguing that such inducement should be legally enforceable).

112. For this reason, some scholars have advocated for a more robust licensing means of pledging patents to a commons. *See, e.g.*, Jason Schultz & Jennifer M. Urban, *Protecting Open Innovation: The Defensive Patent License as a New Approach to Patent Threats, Transaction Costs, and Tactical Disarmament*, 26 HARV. J.L. & TECH. 1 (2012) (proposing the Defensive Patent License as a surer way to protect open innovation against patent risks).

113. *See, e.g.*, Contreras, *supra* note 4 (arguing that patent pledges, to the extent that they induce market reliance, should be enforceable).

114. *See* Eric Blattberg, *Here's What Tesla's 'Good Faith' Patent Stance Actually Means*, VENTUREBEAT (June 14, 2014), <http://venturebeat.com/2014/06/14/heres-what-teslas>

patent pledger may have otherwise intended, the enduring presence of exclusive rights may carry a powerful message of its own: the ongoing possibility of those rights being asserted.¹¹⁵

But the key to resolving such possible conflicts would seem to lie with the patent holder. That is, a patent holder should be able to eliminate a great deal of informational uncertainty if they so desire by crafting a patent pledge in a clear and unequivocal manner.¹¹⁶ Of course, a patent's basic exclusionary nature can certainly trip up clumsy patent owners who, despite a desire to signal a purpose of inclusion through a permissive patent pledge, simply fail to communicate their intentions accordingly. But such informational blunders remain the fault of the patent owners, not patents themselves.

Furthermore, the risk of a patent pledger ultimately enforcing its rights against a party relying on the pledge may be negligible for other reasons. For instance, a public patent pledge, even if crafted somewhat ambiguously, may at least mean that the pledger is more willing to enter into a contractual relationship with third parties for clear access to the patented technologies. Indeed, in many cases third parties may prefer a contract over simple reliance on the pledge. Hence, even if a pledge in any given situation is less than perfect as a signal, it may still signal some information to the public that the pledger and recipients ultimately act upon, in this example in the form of a contractual relationship.

Of course, as briefly discussed above, many patent owners may not, in fact, desire to completely eliminate uncertainty as to their intentions through a public patent pledge, and that enduring uncertainty may be an important part of their informational purpose.¹¹⁷ In other words, a patent signal's primary informational purpose may be one of equivocation, which can serve the purpose of keeping competitors guessing as to the pledger's intentions.¹¹⁸ Or in other cases, the informational purpose behind a patent

-good-faith-patent-stance-actually-means [http://perma.cc/V7N6-55KC] (indicating that few if any major car manufacturers will simply rely on Tesla's pledge without more concrete contractual terms in place).

115. Cf. Schultz & Urban, *supra* note 112 (arguing, in part because of such possibilities, for adoption of a clear license that allows for use of patented technologies).

116. Some have criticized Tesla's pledge on this basis because, for instance, it equivocates by indicating that only those who use its patented technologies in "good faith" may benefit from the pledge. In other words, parties will be loathe to simply rely on the pledge absent a clearer license to the patented technologies. See Blattberg, *supra* note 114.

117. See, e.g., *supra* Section III.B (discussing Microsoft, whose patent pledges appear to fit this mold in certain respects).

118. *Id.*

pledge may be of signaling enduring exclusion, such as when a patent pledger carefully selects and pledges only patents relating to non-essential technologies.¹¹⁹ And such intentions are perfectly compatible with using patents as informational tools, since the varied values and uses of patents are not mutually exclusive.

In sum, the rights of exclusion that come with a patent are certainly valuable to some parties because they allow them to procure direct economic benefits through their exercise. But the same rights are also valuable as a key means by which patent pledgers credibly signal to the public a party's intentions of inclusion, exclusion, or a mix of both.

C. THE INFORMATIONAL VALUE OF PATENT LAW'S DISCLOSURE REQUIREMENTS

Patent law's disclosure requirements, as outlined in Section II.A above, are also important in facilitating the role and value of patents as informational tools in the patent pledging context. For instance, they provide pledgers with a ready means by which to signal information. That is, because patent applicants are required to make these disclosures as part of the application process,¹²⁰ and those disclosures become publicly available in most cases,¹²¹ the patent document itself becomes a standardized means by which to relay information to the public.¹²² And this standardization makes the costs of verifying certain aspects of the signal quite low in important respects.¹²³

Thinking of alternatives for credibly sending signals and facilitating information disclosure more generally helps illustrate the importance of the patent system in these regards. For instance, imagine that a company did not obtain patents on many of its innovations and instead, to the

119. *Id.*

120. 35 U.S.C. § 112 (2012) (specifying patent law's disclosure requirements).

121. § 122 (providing for publication of a patent application, either eighteen months after filing of the application or upon issuance of the patent).

122. ESTEBAN BURRONE & GURIOBAL SINGH JAIYA, WORLD INTELL. PROP. ORG., INTELLECTUAL PROPERTY (IP) RIGHTS AND INNOVATION IN SMALL AND MEDIUM-SIZED ENTERPRISES 3 (2004), http://www.wipo.int/sme/en/documents/pdf/iprs_innovation.pdf [<http://perma.cc/N9MG-ÅCT5>] (estimating that nearly 70% of the world's technical knowledge is contained in patents, and concluding that most of such information is never published anywhere else); Fromer, *supra* note 19, at 554–60 (discussing the primacy of the patent document in disclosing technical information to the public).

123. Long, *supra* note 23, at 665–66 (discussing costs associated with verifying patented information, noting that some aspects of such verification present low costs, while some present high costs).

extent possible, kept its innovations secret.¹²⁴ If it later decided to pledge these secrets to the public based on changing innovation preferences, the lack of patents in facilitating that pledge may prove to be a significant obstacle. For instance, the company would be forced to spend significant time and resources collecting and documenting the information that it wished to share with the public.¹²⁵ And even if it did so, the form and substance of the information may be difficult for the public to navigate.¹²⁶

Patented information, on the other hand, provides several significant advantages. First, as mentioned, the patent document discloses technical information in a formalized, well-understood format.¹²⁷ Hence, even where inadequacies remain in the document,¹²⁸ its standardized nature will mean that, in most cases, third parties should have a greater ability to comprehend it in comparison to the alternative discussed above.¹²⁹ Second and relatedly, patented information allows patent pledgers to be more precise in their messaging. In other words, because the patent includes formal patent claims that identify the boundaries of what the patent covers, the patent provides pledgers with a ready means by which to identify more specifically what the pledge does and does not cover.¹³⁰

Third, as discussed in the preceding Section, a patent's exclusive rights can also aid the patent pledger in using patents as informational tools. Hence, combining the informational potencies of such rights with the

124. This may be an option even in cases where companies, such as Tesla, distribute commercially available products containing the inventions. *See* Fromer, *supra* note 19, at 558 (“Inventors also will not necessarily learn about the most useful innovations from the mere existence and reverse-engineering of commercially available products because commercial success is not that well-correlated with the quality or usefulness of an inventive leap.”).

125. This is not an insurmountable hurdle, however, as many companies engage in this type of behavior and may see advantages in doing so. *See, e.g.*, Oren Bar-Gill & Gideon Parchomovsky, *The Value of Giving Away Secrets*, 89 VA. L. REV. 1857 (2003) (discussing why parties in many cases choose to publish their inventions, rather than patent them). Nonetheless, the patent document remains the primary means by which to convey technical information to the public. *See* Fromer, *supra* note 19, at 560 (“By process of elimination, the patent document is the principal way for an interested technologist to locate useful information about a patented invention.”).

126. *Id.*

127. BURRONE & JAIYA, *supra* note 122.

128. Fromer, *supra* note 19, at 563 (providing an overview of systematic deficiencies in patent disclosures).

129. BURRONE & JAIYA, *supra* note 122.

130. 35 U.S.C. § 112(b) (2012) (setting forth patent law's claiming requirement); *see also* Fromer, *supra* note 45 (discussing generally how patent law's claiming system might be improved).

informational advantages of patent law's disclosure requirements makes the patent system difficult to match in terms of informational advantages.

Last, a homegrown solution would fail to provide companies with non-informational value that patents may otherwise confer. For instance, even when pledging patents, companies may want to retain the ability to use those pledged patents defensively as "sticks" against those who may assert their patents against the pledger. Hence, patents may provide patent pledgers not only with informational value, but with defensive value as well. This final point again underscores a key point: patents can confer on their owners a variety of different types of value simultaneously. While the informational value of a patent may be significant to a patent holder, that value need not be the only value they derive from their patent.

Thus, though scholars have long debated whether patents are effective at communicating information, these scholars have typically focused almost exclusively on whether those communications benefit the public.¹³¹ And they have conducted their analysis through the lens of traditional patent law theories, which treat information disclosures as a sacrifice to the patent owner, and exclusive rights as the key value to her.¹³² In contrast, this Article argues that these information disclosures can be a key source of value to patent holders, as well as the public, when viewing patents as informational tools.

Indeed, when parties use patents as informational tools, such as in the patent pledging context, many of the purported deficiencies of patent law's disclosure requirements wane. For instance, commentators have often argued that innovators largely ignore patents when pursuing technological innovation,¹³³ either because patents do not disclose useful technical information,¹³⁴ there are too many patents to read,¹³⁵ or because inventors fear claims of willful infringement.¹³⁶

131. See Devlin, *supra* note 33; Fromer, *supra* note 19; Ouellette, *supra* note 19; Note, *supra* note 33.

132. See *supra* Section II.A.

133. Iain M. Cockburn & Rebecca Henderson, *Survey Results from the 2003 Intellectual Property Owners Association Survey on Strategic Management of Intellectual Property* (2003), http://www.ipo.org/wp-content/uploads/2013/04/survey_results_revised.pdf [<http://perma.cc/VPS6-K58G>] (reporting that 65% of the surveyed intellectual property rights owners do not always read patents before pursuing innovation); Fromer, *supra* note 19, at 561 (reviewing evidence that suggests that most inventors do not know of related patents until after their own invention is completed).

134. See Devlin, *supra* note 33, at 403 (arguing that information disclosed in patents often fails to satisfy patent law's disclosure requirements and thus fails to be useful to others); Lemley, *supra* note 21, at 746 (arguing that several factors lead to patent documents that are often opaque and, therefore, of dubious value to those reading them);

But when parties use patents as informational tools, the likelihood that parties will read and rely on technical information in these patents increases. First, parties will almost certainly have fewer concerns about willful infringement in inspecting the patents, especially if the pledge is permissive in nature and thus signals an inclusionary intent on the part of the patent holder.¹³⁷ Second, specifically identifying patents as part of a patent pledge may increase the likelihood that others will assess such patents, because the publicity surrounding the pledge naturally increases interest in the patented technologies themselves.¹³⁸ Though the oft-cited opaqueness of patents remains a concern,¹³⁹ other recent evidence suggests that patents are a more valuable source of technical information than others have argued.¹⁴⁰ In short, when parties use patents as informational tools, they create informational value for both themselves and the relevant public, while simultaneously addressing many of the perceived problems with patents as effective “teachers” of information.¹⁴¹

In sum, patent law’s information disclosure requirements and rights of exclusion play surprising roles when viewing patents as informational tools. Rights of exclusion can become a key means of inclusion for patent pledgers and others using patents as a signaling mechanism. And information disclosed as part of the patent application process, rather than

Note, *supra* note 33 at 2025–26 (arguing that patent documents often fail to enable others to reproduce the patented technology and that patents are often drafted in a way that reduces their value to third parties).

135. Lemley, *supra* note 21, at 746 (arguing that the vast number of patents applied for and issued each year make reading all the relevant patents a “Herculean task”).

136. *Id.* (explaining that “lawyers often advise engineers not to read competitor patents for fear of becoming a willful infringer”); Doug Lichtman, *Substitutes for the Doctrine of Equivalents: A Response to Meurer and Nard*, 93 GEO. L.J. 2013, 2023 (2005) (arguing that very few innovators read patents for fear of willful infringement).

137. This may be even truer if innovators are less concerned with willful infringement than is often supposed, as some recent scholarship suggests. Ouellette, *supra* note 19, at 578–80.

138. Indeed, Tesla’s pledge led to several public analyses of its patent portfolio, an outcome that had not happened prior to the pledge. See Brian Fung, *How to Build a Tesla, According to Tesla*, WASH. POST (June 23, 2014), <https://www.washingtonpost.com/blogs/the-switch/wp/2014/06/23/how-to-build-a-tesla-according-to-tesla> [<http://perma.cc/HBR4-GC8J>].

139. See *supra* note 134 and accompanying text.

140. Ouellette, *supra* note 19, at 561–65 (analyzing previous surveys, as well as conducting an original survey on nanotechnology researchers, in coming to the conclusion that many parties do, in fact, rely on patented information in their innovative activities).

141. Holbrook, *supra* note 18, at 146 (concluding that patents are “ineffectual teachers” of technical information).

a necessary sacrifice, becomes a key benefit to inventors and the public alike in the informational account of patents.

Part IV below now examines specific examples of patent pledges in order to (1) more clearly highlight the informational value that parties may derive from patents, and (2) set the stage for Part V's examination of recent Supreme Court decisions and their likely informational impacts.

IV. THE RISE OF PATENT PLEDGES

Patent pledges are an increasingly important part of the patent landscape.¹⁴² Forms of patent pledging have existed for some time. For instance, in the standard setting context, parties have agreed to “FRAND” licensing commitments for decades.¹⁴³ But in recent years the types and frequency of pledges have grown, so much so that scholars have increasingly devoted attention to tracking and analyzing their various permutations.¹⁴⁴

This Part builds on the theoretical discussion of Part II and Part III by arguing that the patent pledging phenomenon is a manifestation of patents being used as informational tools. In other words, patent pledgers appear to use patents to efficiently and credibly signal information to product, labor, and capital markets about their innovation activities and goals. Transmission of this information may then facilitate a variety of economic purposes, depending on the pledge. Thus, through their private ordering efforts, patent pledgers increasingly manifest that patents have informational value beyond what typical accounts of the patent system posit. And this value has implications for patent law and policy more generally, as Part VI *infra* will more fully argue.

In support of these arguments, the following sections examine several recent patent pledges made by different innovators. It is beyond the scope of this Article to examine in detail all patent pledges, something others have in part attempted.¹⁴⁵ But the informational themes identified in the representative examples discussed below can help explain patent pledging scenarios in general.

142. Contreras, *supra* note 4 (examining this growing phenomenon in detail).

143. See generally Mark A. Lemley, *Intellectual Property Rights and Standard-Setting Organizations*, 90 CALIF. L. REV. 1889 (2002) (providing an early analysis of this form of patent pledging).

144. Contreras, *supra* note 4.

145. *Id.*; see Program on Info. Justice, *supra* note 3 (cataloguing the known existing non-SDO patent pledges).

A. TESLA

As mentioned in the Introduction, in June 2014 Tesla's CEO, Elon Musk, indicated on the company's blog that "Tesla will not initiate patent lawsuits against anyone who, in good faith, wants to use our technology."¹⁴⁶ Musk went on to provide several reasons behind this surprising move. According to the post, Tesla was "created to accelerate the advent of sustainable transport."¹⁴⁷ Yet, Musk admits, Tesla is unable to address the "carbon crisis" on its own because it is unable to produce enough electronic cars fast enough.¹⁴⁸ Annual new vehicle production is nearly 100 million per year, and approximately 2 billion cars are on the road today.¹⁴⁹ According to Musk, Tesla simply cannot satisfy such demand with electronic car options, especially as most major car manufacturers continue to flood the market with carbon-emitting vehicles instead.¹⁵⁰ Thus, Musk and Tesla, in the spirit of the open source software movement, argue that collaboration is the key to making electronic car technology a greater success than it already is.¹⁵¹ And to that end, the company publicly pledged its patents in support thereof.¹⁵²

Tesla's use of its patents as part of its pledge manifests the informational value of Tesla's patents in a number of ways. First, the exclusive rights that come with patents played a key informational role for both Tesla and the public. For instance, without patents, Tesla merely announcing a general desire to collaborate, while perhaps praiseworthy, is not particularly noteworthy. Parties look to collaborate with third parties all the time.¹⁵³ But without more, little if any fanfare will accompany such intentions, even if a company publicly announces them.¹⁵⁴ Of course, once a collaborative effort between parties is formed, such a collaboration may generate significant press, depending on the parties, products, and

146. Musk, *supra* note 1.

147. *Id.*

148. *Id.*

149. *Id.*

150. *Id.*

151. *Id.*

152. *Id.*

153. See Gary Hamel, Yves Doz & C.K. Prahalad, *Collaborate with Your Competitors—and Win*, HARV. BUS. REV., Jan–Feb. 1989, at 133 (discussing the advantages of competitive collaboration).

154. Indeed, it is difficult to even find examples of such scenarios, most likely because engaging in this type of behavior seems to present few if any advantages. And this is so despite the fact that collaboration between companies has long been a key competitive advantage. See *id.*

dynamics involved.¹⁵⁵ But a mere desire to collaborate, even if publicly announced, is hardly newsworthy.

Yet Tesla's announcement immediately made headlines and continues to do so, despite constituting essentially a publicly announced wish with no concrete collaborations in place.¹⁵⁶ Why? Because Tesla publicly disavowed its intention to exercise economically valuable rights associated with its patents.¹⁵⁷ With rights of exclusion in place, for instance, Musk's purported sacrifice of those rights lends Tesla's collaborative intentions greater significance and credibility.¹⁵⁸ After all, if Tesla is sincere about sacrificing those economically valuable rights in pursuit of collaborative efforts—and the evidence¹⁵⁹ and pledge itself¹⁶⁰ suggests that it is—then it seems logical to conclude that the company must genuinely wish to collaborate with third parties in the electronic vehicle industry.¹⁶¹ Indeed, as discussed above, signals are generally more credible if the signaler incurs costs in making the signal, since these costs enable recipients to infer certain characteristics about the signaler.¹⁶² These conditions seem satisfied with Tesla's pledge.

Hence, in Tesla's case, this reasoning suggests that the credibility of Tesla's signal was quite high. After all, though Tesla is a relatively new company, the company already boasts a significant patent portfolio with

155. See, e.g., Laura Entis, *Target to Collaborate With Top Pinterest Users on Party Collections*, ENTREPRENEUR (Feb. 11, 2014), <http://www.entrepreneur.com/article/231432> [<http://perma.cc/W5HH-KGRB>].

156. Ryan Davis, *Devil's In the Details of Tesla's Open Patent Pledge*, LAW360 (June 13, 2014), <http://www.law360.com/articles/547910/devil-s-in-the-details-of-tesla-s-open-patent-pledge> [<https://perma.cc/9SJ2-NTDN>] (providing one of countless immediate analyses of Musk's announcement); Mike Masnick, *Elon Musk Clarifies That Tesla's Patents Really Are Free; Investor Absolutely Freaks Out*, TECHDIRT (Feb. 18, 2015), <https://www.techdirt.com/articles/20150217/06182930052/elon-musk-clarifies-that-teslas-patents-really-are-free-investor-absolutely-freaks-out.shtml> [<https://perma.cc/G5ZL-EBKB>] (providing a more recent analysis of the pledge relating to an investor group's argument that Tesla investors should be worried).

157. Mike Masnick, *Elon Musk Destroys the Rationale for Patents, Opens Up All of Tesla's*, TECHDIRT (June 12, 2014), <https://www.techdirt.com/articles/20140612/11253427557/elon-musk-destroys-rationale-patents-opens-up-all-teslas.shtml> [<https://perma.cc/A99J-D2H8>] (lauding Tesla for pledging all of its patents, rather than some subset thereof, and contrasting Tesla's approach with that of most other companies, which continue to rely on patents in ways that, according to the author, hurt innovation).

158. See *id.*

159. Masnick, *supra* note 156 (detailing Musk reasserting that anyone is free to use Tesla's patents).

160. Musk, *supra* note 1.

161. Masnick, *supra* note 156.

162. See *supra* Part III.

worldwide coverage.¹⁶³ And Musk's pledge covers all of them.¹⁶⁴ Thus, rather than carefully pledging a few trivial patents, Tesla implicated its entire portfolio with its pledge, which helped underscore the seriousness of Tesla's message of collaboration.¹⁶⁵ This sacrifice may have also aided outsiders in inferring characteristics about the company that may be otherwise difficult to observe, such as a certain type of company culture that may be attractive to potential employees, collaborators, and investors. The rights of exclusion that come with a patent thus greatly enhanced Tesla's ability to use its patents as a tool for communicating its innovation preferences to product, labor, and capital markets, even if its approach caused concern in some corners of the capital market.¹⁶⁶

Patent law's disclosure requirements also facilitated Tesla's collaboration signal. For instance, the patent system in general provides a ready means by which to signal information, and the information disclosed in Tesla's patents provides a means by which third parties can assess Tesla's technology for purposes of possible collaboration and use.¹⁶⁷ Indeed, following Tesla's announcement, some parties have performed such assessments.¹⁶⁸

Without patents in place, on the other hand, it would be more difficult for third parties to make such assessments, even with some sort of "patent-less" pledge, for a number of reasons.¹⁶⁹ First, while third parties could simply inspect Tesla's commercially available products in order to better understand the company's technology, a number of factors make such inspections less effective at yielding the same type and quality of information that a patent provides.¹⁷⁰ For instance, patent law includes no requirement that an invention must be commercialized,¹⁷¹ and some

163. Fung, *supra* note 138.

164. See Mike Lloyd, *How Good Are the Tesla Patents? Who Might Want a Free License?* AMBERBLOG (July 24, 2014) <http://www.ambercite.com/index.php/amberblog/entry/how-good-are-the-tesla-patents-who-might-want-a-free-license> [<http://perma.cc/UMY5-QX4X>].

165. See Musk, *supra* note 1.

166. Masnick, *supra* note 156 (providing an account of one such investor group). Of course, the fact that some parties are concerned about Tesla's pledge and its effects on investors is simply another indicator of the pledge's credibility.

167. Fung, *supra* note 138 (detailing an analysis of Tesla's patent portfolio, which followed Tesla's patent pledge); Lloyd, *supra* note 164 (same).

168. See Fung, *supra* note 138.

169. *But see generally* Bar-Gill & Parchomovsky, *supra* note 125 (detailing how parties sometimes simply publish information relating to their innovations, rather than seeking a patent on them, based on certain economic motives).

170. Fromer, *supra* note 19, at 557–58.

171. *Id.*

evidence suggests that most patented inventions are not, in fact, ever put into commercial use.¹⁷² In some industries, even those patented inventions that do find their way into commercial products typically do so long after a patent has been published.¹⁷³ And finally, patentees often make commercially available products difficult to reverse-engineer, thereby making attempts to glean information about a patented invention prohibitively expensive in many cases.¹⁷⁴

Hence, patent law's enablement and written description requirements, combined with patent law's early publication requirement, mean that in many cases Tesla's patents provide technological insights that go beyond what simply inspecting the vehicles would yield.¹⁷⁵ Indeed, patents often disclose technical information that has not yet been commercialized (or which may never be commercialized by the inventor), which means that many aspects of Tesla's patented technology would not be available through inspecting commercially available products.¹⁷⁶ And even in cases where a skilled artisan could eventually obtain much of the same information through inspection that is provided in the patent, patent law's disclosure requirements speed up the process for obtaining such information.¹⁷⁷

Of course, patents are certainly not the only means by which to publicly share technical information relating to Tesla's inventions.¹⁷⁸ Absent patents, Tesla could have devised some other means to collect and systematically document its technical accomplishments.¹⁷⁹ And perhaps in some ways such a system could improve upon what the patent system offers in terms of informational value, to both the public and the patent holder;¹⁸⁰ after all, complaints abound about the low informational value of the patent system.¹⁸¹

But despite these possible advantages, the patent system offers several key benefits in comparison to a possible homegrown solution, as

172. Kurt M. Saunders, *Patent Nonuse and the Role of Public Interest as a Deterrent to Technology Suppression*, 15 HARV. J.L. & TECH. 389, 391 n.10 (2002).

173. Note, *supra* note 33, at 2016.

174. *Id.* at 2017.

175. See Fromer, *supra* note 19, at 557–58.

176. See *id.*

177. *Id.*

178. See generally Bar-Gill & Parchomovsky, *supra* note 125 (detailing how parties sometimes simply publish information relating to their innovations, rather than seeking a patent on them, based on certain economic rationales).

179. See *id.*

180. See *id.*

181. See sources cited *supra* note 33 and accompanying text.

mentioned above. First and foremost, the patent system already exists as a systematic way to publicly document technical accomplishments.¹⁸² Without relying on the patent system, for instance, Tesla may have to invest significant time and resources in building a systematic means for documenting and publishing its technical achievements.¹⁸³ And without the vetting that occurs as part of the patenting process, third parties would have to simply take Tesla at its word as to whether the documented technologies actually represent significant technical accomplishments worthy of their consideration.¹⁸⁴ Without that vetting process in place, therefore, the signal in general may thus become less credible.

Third parties may also have to learn to navigate and understand the homegrown documentation, whereas in most cases they are already familiar with patent documents, despite their potential ambiguities.¹⁸⁵ Last but importantly, as discussed above, exclusive rights that come with a patent can also facilitate the informational purposes of parties.¹⁸⁶ In other words, the patent system's information disclosure requirements present significant informational advantages over a homegrown solution, particularly when combining these advantages with the informational potencies of exclusive rights.

Patents thus aided Tesla in signaling information to several distinct markets. One obvious target is other car manufacturers, that is, the product market.¹⁸⁷ In other words, Tesla seems to have clearly intended to signal to other car manufacturers and innovators its innovation preferences

182. Ouellette, *supra* note 19, at 542 n.61 (noting that the patent system is well-entrenched and required to remain so based on international law obligations that the United States has taken upon itself); BURRONE & JAIYA, *supra* note 122 (estimating that around 70% of the world's technical information is contained in patents and is not published elsewhere).

183. This is not an insurmountable hurdle, as parties publish technical information absent patents all the time. *See, e.g.*, DEFENSIVE PUBLICATIONS, <http://www.defensivepublications.org> [<http://perma.cc/7QXU-XUQQ>] (providing tools and forms for publishing technical information in order to help prevent patents from issuing on prior art technical accomplishments). Nonetheless, the patent system remains a more entrenched, understood, and systematic means of doing so.

184. Sean Tu, *Luck/Unluck of the Draw: An Empirical Study of Examiner Allowance Rates*, 2012 STAN. TECH. L. REV. 10, 12–17 (2012) (reviewing the patent examination process).

185. Lemley, *supra* note 21, at 746–47 (noting the opaqueness of many patents).

186. *See supra* Section III.B.

187. Musk, *supra* note 1 (“We believe that Tesla, *other companies making electric cars*, and the world would all benefit from a common, rapidly-evolving technology platform.” (emphasis added)).

and activities, and to encourage such parties to adopt a similar approach to innovation within the electronic vehicle industry.¹⁸⁸

Tesla also appears to have intended to signal information to the labor market.¹⁸⁹ Indeed, a collaborative approach to innovation may signal an open and collaborative company culture, which may be attractive to many potential employees.¹⁹⁰ In the software industry, for example, the norms of the free and open source software movement have become so entrenched that a company's commitment to that movement often proves to be a useful recruiting tool.¹⁹¹

Last, Tesla's pledge also may have signaled important information to capital markets. While some investors decried Tesla's move because it appeared to sacrifice Tesla's economic rights in exchange for little,¹⁹² others lauded it as economically advantageous.¹⁹³ Although Tesla's pledge thus carries some risk that the economic benefits the company derives from the informational value of its patents will fall short of the direct economic benefits that it sacrificed, Tesla placed its bet on collaborative innovation. And its patents played a significant role in communicating that bet to each of product, labor, and capital markets.

Thus, the potential informational benefits of patents manifest in the Tesla example turn on their head many of the primary informational concerns that scholars have expressed about patents in general. For instance, as mentioned, some studies show that very few parties rely on

188. *Id.*

189. *Id.* ("Technology leadership is not defined by patents, which history has repeatedly shown to be small protection indeed against a determined competitor, but rather by the ability of a company to attract and motivate the world's most talented engineers. We believe that applying the open source philosophy to our patents will strengthen rather than diminish Tesla's position in this regard.")

190. Orly Lobel, *Why Noncompetes May Give You the Least Desirable Employees*, WALL ST. J. (Jan. 22, 2014), <http://blogs.wsj.com/accelerators/2014/01/22/orly-lobel-why-non-competes-may-give-you-the-least-desirable-employees> [<http://perma.cc/WA82-2Y3J>] (describing how restrictive non-competes may actually disincentivize employees); Musk, *supra* note 1.

191. Matt Asay, *Why Open Source Is Becoming a Big Developer-Recruiting Tool*, READWRITE (Oct. 24, 2014), <http://readwrite.com/2014/10/24/open-source-recruiting-facebook-netflix-twitter> [<http://perma.cc/26BC-D3SG>] (describing how Facebook has used its commitment to free and open source software as a recruiting tool).

192. See Masnick, *supra* note 156.

193. Walter Frick, *What Tesla Knows That Other Patent-Holders Don't*, HARV. BUS. REV. (June 12, 2014), <https://hbr.org/2014/06/what-tesla-knows-that-other-patent-holders-dont> [<https://perma.cc/2TBN-AJRP>] (lauding Tesla's move because it enables competitors to work with it in advancing electronic car technology while also preserving distinct advantages in the form of tacit, uncodified knowledge).

technical information gleaned from patents in pursuing innovation.¹⁹⁴ As others have noted, willful infringement standards under patent law may lead many parties to shield themselves from patents.¹⁹⁵ Others complain that patents too often disclose few useful technical details, and that patents are often purposely crafted ambiguously.¹⁹⁶ When using patents as informational tools, however, many of these concerns lessen. Tesla's competitors, for instance, now have greater incentives to inspect Tesla's patents without fear of that inspection later resulting in a willful infringement allegation. Indeed, the pledge helped pique interest in Tesla's patented technologies and encouraged parties to review the company's patents.¹⁹⁷

Furthermore, Tesla also has greater incentives to include significant technical details in its patents in hopes of encouraging and enabling others to collaborate on electronic vehicle innovation. This is in contrast to what some scholars have argued. For instance, some scholars suggest that when parties wish to encourage collaboration in order to increase the collective pie, they may publish, rather than patent, their inventions in hopes of encouraging others to adopt their technologies and build upon them.¹⁹⁸ A patent, according to these scholars, is less conducive to such purposes because of the presence of exclusive rights.¹⁹⁹

But as this Article has argued, patents, coupled with a pledge, may be even more useful with such economic motives in mind because of the informational potencies of exclusive rights discussed above. Indeed, since patents can also be used to effectively disclose technical information (and may even be more advantageous than an unpatented alternative), the advantages of patents in these regards seem even more pronounced. Some of the aforementioned informational concerns will certainly persist in the patent pledging context. For instance, the imperfect nature of language means that patent disclosures and claims will always lack perfect precision. And third parties may still avoid even pledged patents due to concerns about willful infringement, particularly if the patent pledge is largely about signaling enduring exclusionary purposes. Nonetheless, using patents for informational purposes generally means that the informational value of patents is more apparent and in many cases easier to harness.

194. *See supra* notes 133–136 and accompanying text.

195. *Id.*

196. *See id.*

197. *See* Fung, *supra* note 138; Lloyd, *supra* note 164.

198. *See* Bar-Gill & Parchomovsky, *supra* note 125.

199. *See id.*

B. MICROSOFT

Microsoft has frequently been maligned as the enemy of open innovation in the software world, particularly because of the widespread perception that the company maintains an aggressive patent attitude towards users of free and open source software.²⁰⁰ Indeed, at one time top executives at the company infamously referred to the free and open source software movement as a “cancer”²⁰¹ and akin to “communism.”²⁰²

Yet in the last several years, Microsoft has changed its stance.²⁰³ The company has come to provide open support to the free and open source software movement.²⁰⁴ In recent years, for instance, it has been a significant contributor to the Linux kernel project, the famous open source software operating system that helps power much of the computing world today. Microsoft has offered this support despite owning significant numbers of patents relating to Linux.²⁰⁵ It has also established an open

200. See, e.g., Cade Metz, *Meet Bill Gates, The Man Who Changed Open Source Software*, WIRED (Jan. 30, 2012), <http://www.wired.com/2012/01/meet-bill-gates> [<http://perma.cc/JRS6-VVRE>] (describing some of this cantankerous history as well as Microsoft’s softening towards FOSS in recent years).

201. Thomas C. Greene, *Ballmer: “Linux Is a Cancer”*, THE REGISTER (June 2, 2001), http://www.theregister.co.uk/2001/06/02/ballmer_linux_is_a_cancer [<http://perma.cc/S8YV-N8VW>].

202. Graham Lea, *MS’ Ballmer: Linux Is Communism*, THE REGISTER (Jul. 31, 2000), http://www.theregister.co.uk/2000/07/31/ms_ballmer_linux_is_communism [<http://perma.cc/TA67-3CMU>] (likening Linux, the most successful FOSS project at the time, to communism).

203. Steven J. Vaughan-Nichols, *Microsoft: The Open-Source Company*, ZDNET (Jan. 26, 2015), <http://www.zdnet.com/article/microsoft-the-open-source-company> [<http://perma.cc/8BKB-VF46>] (describing Microsoft’s relative embrace of open source software).

204. See, e.g., Maria Deutscher, *Microsoft’s Open-Source Makeover a Developer-Only Campaign—So Far*, SILICON ANGLE (Mar. 23, 2015), <http://siliconangle.com/blog/2015/03/23/microsofts-open-source-makeover-a-developer-only-campaign-so-far> [<http://perma.cc/45HM-UBK3>] (describing recent changes at Microsoft aimed at more fully embracing open innovation); Pinsent Masons, *Microsoft Patent Pledge ‘Worse Than Useless,’ Say Open Source Lawyers*, OUT-LAW.COM (Nov. 15, 2006), <http://www.out-law.com/page-7480> [<http://perma.cc/YT53-A64H>] (describing a deal in which Microsoft would work with Novell in making Windows more compatible with Linux); Microsoft, *Openness*, MICROSOFT OPENNESS BLOG, <http://openness.microsoft.com/blog> [<http://perma.cc/S4JJ-4ZAW>] (describing Microsoft’s approach to and adoption of open innovation).

205. Todd Bishop, *Surprise: Microsoft Makes List of Top 20 Linux Kernel Contributors—First Time Ever*, GEEKWIRE (Apr. 3, 2012), <http://www.geekwire.com/2012/surprise-microsoft-list-top-linux-kernel-contributors> [<http://perma.cc/533W-YNVB>]; Steven J. Vaughan-Nichols, *While You Shouldn’t Expect Windows to Be Open-Sourced in Your Lifetime, Microsoft—Yes, Microsoft—Is the Fifth Largest Code Contributor to Linux 3.0*, ZDNET (July 17, 2011), <http://www.zdnet.com/article/top-five-linux-contributor-microsoft> [<http://perma.cc/A5JX-CQKV>].

source software foundation aimed at providing support to open source software projects.²⁰⁶ It has also released a growing number of technologies under permissive open source software licensing terms.²⁰⁷

In the same vein, starting in 2006 Microsoft began pledging not to sue open source software developers who create, use, and distribute non-commercial software.²⁰⁸ Since then, Microsoft has made a variety of other patent pledges and promises relating to certain of its technologies and open source software users and developers.²⁰⁹ Many of these programs explicitly aim at encouraging interoperability between Microsoft products and those of third parties.²¹⁰ In other words, Microsoft makes some of its technologies more readily available through its interoperability programs and patent pledges in order to encourage third parties to develop technologies that complement its own products.²¹¹

206. Kelly Fiveash, *MS-Backed CodePlex Foundation Morphs into Outercurve*, THE REGISTER (Sept. 28, 2010), http://www.theregister.co.uk/2010/09/28/codeplex_outercurve_microsoft [<http://perma.cc/V3C8E-T667>].

207. Neil McAllister, *Microsoft in OPEN-SOURCE .Net Love-In with New Foundation*, THE REGISTER (Apr. 4, 2014), http://www.theregister.co.uk/2014/04/04/microsoft_launches_open_source_net_lovein_with_new_foundation [<http://perma.cc/V67C-32HU>] (describing Microsoft's surprising move of releasing some of its core technologies under permissive FOSS licenses).

208. MICROSOFT, THE MICROSOFT PATENT PLEDGES FOR IMPLEMENTATIONS OF MICROSOFT COMMUNICATIONS PROTOCOL PROGRAM TECHNICAL SPECIFICATIONS, <https://msdn.microsoft.com/en-us/openspecifications/dn865017> [<https://perma.cc/XL7E-NQ6H>]; Program on Info. Justice, *supra* note 3; see Matthew Broersma, *Promise from Microsoft Not to Sue over Standards*, ZDNET (Sep. 13, 2006), <http://www.zdnet.com/article/promise-from-microsoft-not-to-sue-over-standards> [<http://perma.cc/45HS-JA4X>] (OSP); Jason Matusow, *Your Input Requested*, MATUSOW'S BLOG (Nov. 11, 2006), <http://blogs.msdn.com/b/jasonmatusow/archive/2006/11/11/your-input-requested.aspx> [<http://perma.cc/4VY5-RW8M>] (describing Microsoft executive's request to the FOSS community for input on how to improve Microsoft's patent pledge to FOSS developers that it issued as part of a Novell-Microsoft business arrangement).

209. See Julie Bort, *Microsoft Legally Promises Not to Sue Open Source Developers*, NETWORK WORLD (Dec. 16, 2009), <http://www.networkworld.com/article/2232861/microsoft-subnet/microsoft-legally-promises-not-to-sue-open-source-developers.html> [<http://perma.cc/XWN7-WGXE>]; Ina Fried, *Microsoft Pledges Not to Sue over Open Source*, CNET (Feb. 21, 2008), <http://www.cnet.com/news/microsoft-pledges-not-to-sue-over-open-source> [<http://perma.cc/SB88-AWQK>].

210. See *Microsoft Interoperability Program (MIP)*, MICROSOFT, <https://msdn.microsoft.com/en-us/openspecifications/dn646762> [<https://perma.cc/6YAN-AQUU>]; *Interoperability Principles Program*, MICROSOFT, <https://msdn.microsoft.com/en-us/openspecifications/dn646764> [<https://perma.cc/V34C-835Q>].

211. Matt Asay, *What if Microsoft Really Did Open-Source Windows?*, INFOWORLD (Apr. 17, 2015), <http://www.infoworld.com/article/2911159/microsoft-windows/what-if-microsoft-really-did-open-source-windows.html> [<http://perma.cc/WK7P-55WA>].

But Microsoft's apparent embrace of open innovation has been accompanied by increased patent monetization efforts.²¹² Indeed, though Microsoft engaged in a fair amount of patent rattling early on with respect to open source software,²¹³ it has been only more recently that the company has actually utilized its vast patent portfolio against commercial users of Android and Linux.²¹⁴ And in some of these cases, the license deals Microsoft has struck have been quite large—one deal alone is said to be worth \$1 billion per year.²¹⁵ Others estimate that Microsoft receives multiple billions of dollars in annual licensing revenues through its patent licensing programs with respect to open source software technologies.²¹⁶

Viewing patents as informational tools helps explain this apparent contradiction between Microsoft's growing openness and expanding patent monetization efforts. For instance, Microsoft's interoperability programs and related patent pledges are carefully circumscribed.²¹⁷ Unlike Tesla, which simply pledged its entire patent portfolio to the public,²¹⁸

(discussing Microsoft's recent possible openness to releasing Windows as FOSS due to the same economic logic).

212. Alan Shimel, *Does Android Make Microsoft More Money Than Windows Mobile?*, NETWORK WORLD (Oct. 24, 2011), <http://www.networkworld.com/article/2220945/opensource-subnet/does-android-make-microsoft-more-money-than-windows-mobile-.html> [<http://perma.cc/CX7Z-4KUT>] (discussing Microsoft's growing patent monetization efforts and successes with respect to Android and Linux technologies).

213. Joel Hruska, *Microsoft: 235 Specific Patent Infringements in Linux*, ARS TECHNICA (May 14, 2007), <http://arstechnica.com/tech-policy/2007/05/microsoft-235-patent-specific-patent-infringements-in-linux> [<http://perma.cc/3DEQ-M362>] (discussing Microsoft's claim that Linux and other open source technologies infringe at least 235 of its patents).

214. Simon Phipps, *Microsoft 'Loves' Linux? Then Stop Attacking Open Source*, INFOWORLD (Oct. 31, 2014), <http://www.infoworld.com/article/2841412/open-source-software/microsoft-attacks-open-source-linux.html> [<http://perma.cc/DUY3-QEAW>] (reviewing this trend).

215. *Id.*

216. Steven J. Vaughan-Nichols, *Biggest Patent Win Ever? Microsoft's Billion Dollar a Year Samsung Deal*, ZDNET (Oct. 5, 2014), <http://www.zdnet.com/article/biggest-patent-win-ever-microsofts-billion-dollar-a-year-samsung-deal> [<http://perma.cc/B7MV-TF9E>] (estimating that Microsoft made \$3.4 billion on Android-related licensing alone).

217. See Martin LaMonica, *Free-Software Lawyers: Don't Trust Microsoft's Open XML Patent Pledge*, CNET (Mar. 13, 2008), <http://www.cnet.com/news/free-software-lawyers-dont-trust-microsofts-open-xml-patent-pledge> [<http://perma.cc/BJ4F-7QBN>] (discussing one group's concerns with limitations in a Microsoft patent pledge relating to its XML technologies); *Programs*, OPEN SPECIFICATIONS DEV CENTER, <https://msdn.microsoft.com/en-us/openspecifications/dn646761> [<https://perma.cc/44WK-2E5C>] (providing an overview of Microsoft's specific patent licensing and pledge programs).

218. Musk, *supra* note 1.

Microsoft limits its pledges to specific Microsoft technologies²¹⁹ and, in some cases, certain types of developers (i.e., non-commercial).²²⁰ In other words, the informational takeaway from Microsoft's patent pledging programs seems clear: if a party uses the delineated technologies and patents in the specific ways described in the pledges, that party has Microsoft's blessing. But everyone else is fair game. Microsoft's lucrative licensing programs relating to Android and Linux technologies give substance to this message.²²¹

As with Tesla, particular features of the patent system aided Microsoft in signaling to the various markets its intentions. Exclusive rights make Microsoft's informational signals to the product market more credible, both its exclusionary gestures as well as its more inclusionary ones.²²² With respect to exclusionary signals, rights of exclusion by their nature carry with them a message of possible exclusion,²²³ and Microsoft implicitly reasserted an exclusionary intent when it carefully assessed its patent portfolio and specifically excluded many of its patents from its interoperability programs.²²⁴ Indeed, exclusion from such a program, which undoubtedly underwent significant internal review and deliberation, may actually mean that such rights are more likely to be asserted in the future, since such internal deliberations seem to have identified them as important to leave outside of the interoperability programs and associated pledges.²²⁵ The rise of Microsoft's patent monetization efforts following

219. *Microsoft Community Promise*, OPEN SPECIFICATIONS DEV CENTER, <https://msdn.microsoft.com/en-us/openspecifications/dn646766> [<http://perma.cc/4P78-52SS>] (listing specific technological specifications to which the company's "Community Promise" relating to patent non-assertion applies).

220. *The Microsoft Patent Pledges for Implementations of Microsoft Communications Protocol Program Technical Specifications*, OPEN SPECIFICATIONS DEV CENTER, <https://msdn.microsoft.com/en-us/openspecifications/dn865017> [<http://perma.cc/D6NK-C8LK>] (providing a patent non-assertion pledge with respect to open source software developers).

221. See Vaughan-Nichols, *supra* note 216.

222. See *supra* Section III.B.

223. See Lorelei Ritchie, *Reconciling Contract Doctrine with Intellectual Property Law: An Interdisciplinary Solution*, 25 SANTA CLARA COMPUTER & HIGH TECH. L.J. 105, 118 (noting the exclusionary nature of patents).

224. See *Microsoft Community Promise*, *supra* note 219.

225. Indeed, the fact that Microsoft's increased patent monetization efforts seem to have followed some of its early pledges provides some confirmation to this, as forecast by some. See Peter Judge, *Microsoft Opens Up: Everything You Need to Know*, ZDNET (Feb. 24, 2008), <http://www.zdnet.com/article/microsoft-opens-up-everything-you-need-to-know> [<http://perma.cc/K8WS-WEUQ>] ("Interoperability that is safe only for non-commercial software excludes Microsoft's number one competitor, Linux. . . . So, right there it tells you that this is a promise to do nothing that matters.").

such deliberations appears to provide some confirmation of this intuition.²²⁶

Conversely, the patents and associated technologies that were included in such programs helped bolster the credibility of Microsoft's message of partial openness.²²⁷ After all, if Microsoft were not serious about this partial openness, why would it sacrifice its exclusive rights, which it chose not to do in so many other cases?²²⁸ This is not to say that more generous pledges, such as Tesla's, are less in earnest. But it is to say that the circumscribed nature of Microsoft's pledge indicates some amount of deliberation, which in turn may function as a proxy for credibility.

It is true that antitrust concerns may have in part motivated Microsoft's move towards more openness, since the pledges may have helped satisfy the relevant government authorities that the company was moving away from certain anti-competitive behaviors.²²⁹ But antitrust concerns certainly do not provide a complete explanation. Indeed, as others have noted, Microsoft seems to have begun to realize what other companies have known for some time: openness can yield significant economic returns that exceed a more proprietary approach, particularly where significant economic value lies not in the ceded technologies, but in complementary goods and services to them.²³⁰ Of course, the simplest way to signal openness may be to never pursue rights of exclusion in the first place.²³¹ But in a world of rapid technological change and strategy, obtaining a patent may be a safer approach since it allows the patent owner to signal exclusion or inclusion (or a mix of both).²³² Furthermore, as discussed above, sacrificing rights of exclusion lends greater credence to the signal, as well as enabling recipients thereof to potentially infer information about Microsoft that may be otherwise difficult to verify.²³³

The information disclosures that are part of the patenting process also facilitated Microsoft's signals of both exclusion and inclusion. Microsoft,

226. See Vaughan-Nichols, *supra* note 216.

227. See *supra* Section III.B.

228. See *Microsoft Community Promise*, *supra* note 219.

229. Judge, *supra* note 225 (discussing antitrust actions in the European Union that played a role in pushing Microsoft towards greater openness).

230. Asay, *supra* note 211 (discussing the economic logic of opening up Windows technologies).

231. Bar-Gill & Parchomovsky, *supra* note 125.

232. See generally Stuart J.H. Graham & Ted Sichelman, *Why Do Start-Ups Patent?*, 23 BERKELEY TECH. L.J. 1063 (2008) (discussing the many roles that a patent can play for start-up companies).

233. See *supra* Part III.

for instance, was able to more specifically delineate patents and the associated technologies included in its pledges, as well as those without.²³⁴ Third parties could thus inspect specific patents, their claimed scope, as well as their technical disclosures in assessing Microsoft's vows.²³⁵ While other regimes might also facilitate such informational purposes, the patent regime provides a ready, well-entrenched means by which to realize them.²³⁶

In addition to signaling information to Microsoft's competitors in the product market, it also seems likely that Microsoft sought to signal information to the labor market through its pledges. As mentioned, Microsoft has long been viewed with intense suspicion in certain parts of software developer communities.²³⁷ But good relationships with developer communities are important, both in terms of hiring²³⁸ and employee morale.²³⁹ By purporting to sacrifice some of its valuable patent rights in support of open source software development, Microsoft signaled to developer communities a new kind of Microsoft (at least in some respects).²⁴⁰ Indeed, the company has engaged in a variety of public relations efforts over the years in hopes of improving its public image, particularly within developer communities.²⁴¹ And again, this relatively new company culture may be difficult for outsiders to observe and trust but for the costly signal associated with Microsoft's patent pledge.²⁴²

Last, the pledge also likely signaled information to capital markets. For instance, investors might be more confident in a company adapting to the times than one stubbornly clinging to its past.²⁴³ And sacrificing exclusive rights helped credibly signal to investors that the company was,

234. See *Programs*, *supra* note 217.

235. 35 U.S.C. § 122 (2012) (requiring in most cases that patent applications be published eighteen months after filing or upon issuance of the patent).

236. Ouellette, *supra* note 19, at 556 (noting that the patent system is "entrenched" as part of the international patent system).

237. Metz, *supra* note 200.

238. Asay, *supra* note 191 (detailing Facebook's successes in using open source as a recruiting tool).

239. See Lobel, *supra* note 190 (arguing that restrictive clauses in employee agreements can actually decrease employee morale and thus impact future recruiting efforts).

240. Vaughan-Nichols, *supra* note 203.

241. Deutscher, *supra* note 204 (detailing some of these efforts focused on developer communities).

242. See *supra* Section III.A.

243. See, e.g., Leo Sun, *An Open-Source Microsoft Windows? I Don't Think So*, THE MOTLEY FOOL (Apr. 8, 2015), <http://www.fool.com/investing/general/2015/04/08/an-open-source-microsoft-windows-i-dont-think-so.aspx> [<http://perma.cc/JQE7-VWD9>] (providing analysis from an investor website discussing some of Microsoft's moves towards more openness).

in fact, adapting to the times, while also indicating the particular ways it was doing so. In other words, this manifestation of patent restraint pointed to a more collaborative, open company generally. Yet the confined nature of Microsoft's pledges also signaled to capital markets the likelihood of ongoing patent monetization efforts and business as usual, which later patent licensing activities confirmed.²⁴⁴

In sum, Microsoft used patents and its accompanying pledges to signal to product, labor, and capital products a variety of informational messages. One of a patent's key virtues in such scenarios, therefore, is its multi-dimensional nature. In other words, Microsoft is able to simultaneously use its patents and accompanying pledges in support of its patent monetization efforts as well as its more collaborative initiatives. And it is able to do so in significant part because of the informational potencies of patents.

C. TWITTER

Twitter, the social networking service, has also recently joined the patent pledging trend.²⁴⁵ Twitter's patent pledge is in form and substance distinct from the other pledges already discussed. Twitter, like Tesla, introduced its patent pledge on the company's blog.²⁴⁶ The essence of Twitter's "Innovator's Patent Agreement" (IPA) is that the company agrees not to offensively assert its patents against third parties unless the employees responsible for developing the patented technology give the company permission to do so.²⁴⁷ And this employee control persists even if the patent is sold to a third party.²⁴⁸ Twitter thus explicitly curtailed its ability to assert its patents against third parties, though the company retained the right to assert its patents defensively against those that attack it first.²⁴⁹ Hence, though the IPA is not an outright abandonment of Twitter's patent rights, it significantly handicapped the company in using its patents in traditionally exclusionary ways.

244. See Vaughan-Nichols, *supra* note 216.

245. See Adam Messinger, *Introducing the Innovator's Patent Agreement*, TWITTER BLOGS (Apr. 17, 2012), <https://blog.twitter.com/2012/introducing-innovators-patent-agreement> [<http://perma.cc/VB4G-5N3S>].

246. *Id.*

247. *Id.*

248. *Id.*

249. William Carleton, *Reality Check: Twitter's Patent Pledge Not What It Seems*, GEEKWIRE (Apr. 18, 2012), <http://www.geekwire.com/2012/reality-check-twitters-patent-pledge-not-what-it-seems> [<http://perma.cc/DM35-MZCL>] (discussing some of the limitations of Twitter's patent pledge).

As with the other pledges already reviewed, the exclusive rights that come with a patent played a key role in facilitating Twitter's informational purposes in making the pledge. For instance, the company explicitly expressed hope that other technology companies would join it in similarly curtailing their own patent rights in order to foster an environment of innovation rather than litigation.²⁵⁰ But this message might have rung rather hollow if Twitter did not own exclusive rights in the first place.²⁵¹ In other words, asking other parties to give up their economic rights while doing the same is one thing. Making such a request of third parties, without undertaking the same sacrifices, is quite another. At the time of the pledge, Twitter owned very few patents outright,²⁵² though, as others have noted, the company even then was in the process of acquiring significant numbers of patents.²⁵³ Indeed, the company has recently significantly bolstered its portfolio, thereby augmenting its message, since the IPA applies to those patents as well.²⁵⁴ Exclusive rights thus helped Twitter communicate its activities and goals to third parties, and put the company in a better position to foster those preferences more generally.²⁵⁵

Patent law's information disclosure requirements also aided Twitter in using its patents as an informational tool, since third parties could look to Twitter's patents and accompanying disclosures to assess the magnitude and credibility of its commitment.²⁵⁶ While patent disclosures are far from

250. Messinger, *supra* note 245 ("We . . . have just started to reach out to other companies to discuss the IPA and whether it might make sense for them too.").

251. *See supra* Part III.

252. George Anders, *Twitter's Odd Patent Portfolio: No Sign of 140-Character Gold*, FORBES (Sept. 14, 2013), <http://www.forbes.com/sites/georgeanders/2013/09/14/twitters-slim-patent-portfolio-lacks-any-140-character-gold> [<http://perma.cc/SH63-VRKL>] (discussing the relative smallness of Twitter's patent portfolio a year after the IPA was pledged).

253. Gene Quinn, *Patents are Important: Bursting the Twitter Patent Mythology*, IPWATCHDOG (Sept. 29, 2014), <http://www.ipwatchdog.com/2014/09/29/patents-are-important-bursting-the-twitter-patent-mythology> [<http://perma.cc/U6M8-WBZ3>] (indicating that Twitter has been active in pursuing patent applications since its founding).

254. Jacqueline Sahagian, *Twitter Fleshes Out Its Patent Portfolio*, THE CHEAT SHEET (Jan. 31, 2014), <http://www.cheatsheet.com/technology/twitter-fleshes-out-its-patent-portfolio.html> [<http://perma.cc/32HL-KCWA>] (detailing Twitter's acquisition of over 900 patents from IBM).

255. Kal Raustiala & Chris Sprigman, *The Twitter I.P.A.*, FREAKONOMICS (May 3, 2012), <http://www.freakonomics.com/2012/05/03/the-twitter-i-p-a> [<http://perma.cc/5S4K-9RN5>] (describing the IPA as a form of "norms entrepreneurship").

256. *See* Anders, *supra* note 252 (analyzing Twitter's patent portfolio and noting its paucity, especially with respect to technologies key to the Twitter service); Quinn, *supra*

perfect, they nonetheless provide third parties a significant amount of technical information that those parties can then take into account in assessing Twitter's (and others') pledge.²⁵⁷ It is true that reviewing patent documents will not reveal the intentions of the inventors themselves, which may lessen the credibility of Twitter's signal since some of the relevant inventors may be more than willing to allow the company to use their patents offensively. Nevertheless, the self-imposed restrictions still evince significant sacrifice, particularly when considering that no other companies have adopted a similar policy of self-restraint. These restrictions thus signal some credibility, even if the sacrifice could have been more far-reaching.

Other aspects of patent law's disclosure requirements also likely proved helpful to Twitter in using its pledge as a sales pitch to the labor market. For example, patent law requires patent applicants to list the specific inventors that developed the patented technology.²⁵⁸ This means that, although most patents ultimately belong to companies, the actual inventors listed on any given patent document are typically the employees that invented the patented technology.²⁵⁹ The employer is then normally listed on the patent document as the assignee of the employee's rights in the inventions, since most companies require their employees to assign their rights to anything that the employees invent while with the company.²⁶⁰

Traditionally, employee inventors thus have little say in what happens to their inventions after they assign ownership rights to the company.²⁶¹ But Twitter went to great lengths in its blog post to make clear that the company believes that its employees deserve better, and that Twitter, with its pledge, will do better by whomever the company employs.²⁶² In addition to empowering its current employees, the pledge may thus also serve as an important tool for recruiting potential employees. This may be particularly so since most companies are routinely taking the opposite

note 253 (analyzing Twitter's patent portfolio and concluding that the company has obtained several patents relating to key technologies, contrary to popular belief).

257. BURRONE & JAIYA, *supra* note 122 (indicating that the vast majority of the world's technical information is contained in patents and is not published elsewhere).

258. 35 U.S.C. § 115 (2012).

259. Michael Risch, *Patent Portfolios as Securities*, 63 DUKE L.J. 89, 99 (2013).

260. *Id.*

261. *See generally* ORLY LOBEL, *TALENT WANTS TO BE FREE: WHY WE SHOULD LEARN TO LOVE LEAKS, RAIDS, AND FREE RIDING* (2013) (detailing the many ways in which employers assert rights in the intellectual contributions of their employees).

262. Messinger, *supra* note 245.

approach to employer-employee relationships in terms of intellectual asset ownership.²⁶³

Without a patent document listing the inventors and describing their technical contributions, it is more difficult to imagine exactly how Twitter would have conveyed and brought to pass its message of employee empowerment. This is not to say that Twitter could not have devised some other means of doing so. But it is to say that patents and their accompanying information disclosures provided the company with a ready means for delivering its message. A patent's exclusive rights also helped bolster Twitter's purpose in this respect, since returning in part the exclusive rights that come with a patent back to the employee is a powerful informational signal in its own right.²⁶⁴ Indeed, that signal may allow recipients thereof to infer characteristics of the company that are otherwise difficult to confirm, such as an attractive, employee-first environment.²⁶⁵

These same labor market signals might have also been intended for capital markets. For instance, investors may look favorably upon Twitter's attempt to improve employee morale by empowering its employees through the patent pledge, since plenty of evidence suggests that companies benefit significantly when employees are happy.²⁶⁶ And the innovation preferences and activities that the pledge signals may also prove attractive to investors generally,²⁶⁷ though some may take the opposite view.²⁶⁸ But patents played a key role in helping Twitter communicate information to each of product, labor, and capital markets, regardless of how those markets ultimately respond to such signals.

263. See LOBEL, *supra* note 261.

264. See Joe Brockmeier, *Why Every Company Should Adopt Twitter's Innovator's Patent Agreement*, READWRITE (Apr. 17, 2012), <http://readwrite.com/2012/04/17/why-every-company-should-adopt> [<http://perma.cc/7SSU-8MZJ>] (listing the many recruiting advantages that an agreement such as the IPA gives a company, and arguing that other companies should follow Twitter's lead).

265. See *supra* Part III.

266. Meghan M. Biro, *Happy Employees = Hefty Profits*, FORBES (Jan. 19, 2014), <http://www.forbes.com/sites/meghanbiro/2014/01/19/happy-employees-hefty-profits> [<http://perma.cc/X26B-4S2K>].

267. See Jeff Roberts, *Why Twitter's Patent Policy Is an Asset, Not a Problem*, GIGAOM (Oct. 21, 2013), <https://gigaom.com/2013/10/21/why-twitters-patent-policy-is-an-asset-not-a-problem> [<http://perma.cc/DWD6-LET4>] (providing a favorable investor perspective on Twitter's approach to patents).

268. Michael Kanellos, *Twitter's Patently Absurd Patent Policy*, FORBES (Apr. 17, 2012), <http://www.forbes.com/sites/michaelkanellos/2012/04/17/twitters-patently-absurd-patent-policy> [<http://perma.cc/58B4-LF9X>] (expressing cynicism regarding Twitter's patent policy); Leslie Picker, *Twitter's Lack of Patents Seen as a Risk to Investors*, BLOOMBERG (Oct. 17, 2013) (describing Twitter's relative paucity of patents as a risk).

D. IBM

IBM made waves in 2005 when it pledged 500 of its patents to the free and open source software movement.²⁶⁹ As part of this pledge, IBM committed not to assert these specific patents against “any individual, community, or company working on or using software that meets the Open Source Initiative (OSI) definition of open source software.”²⁷⁰ At the time, IBM claimed the pledge was the largest of its kind.²⁷¹

In several respects, IBM’s pledge is different than some of the others discussed above. First, IBM’s pledge targeted a specific type of developer, namely, those that develop, use, and distribute open source software.²⁷² No other party benefits from the pledge.²⁷³ And even parties engaged in open source software development, use, and distribution remain possible targets of IBM’s tens of thousands of non-pledged patents.²⁷⁴

Second and relatedly, the pledge covered 500 specific patents, rather than IBM’s entire patent portfolio.²⁷⁵ While 500 patents may seem like a large number, IBM is regularly granted thousands of patents per year, and has filed for more U.S. patents than any other company for twenty-two straight years.²⁷⁶ Thus, instead of implicating its entire portfolio, IBM

269. See, e.g., Steve Lohr, *I.B.M. to Give Free Access to 500 Patents*, N.Y. TIMES (Jan. 11, 2005), <http://www.nytimes.com/2005/01/11/technology/ibm-to-give-free-access-to-500-patents.html> [<http://perma.cc/GXN2-TYNV>] (discussing IBM’s move as unconventional); Lucy Sherriff, *IBM Pledges 500 Patents to OS Developers*, THE REGISTER (Jan. 11, 2005), http://www.theregister.co.uk/2005/01/11/ibm_patent_donation [<http://perma.cc/8CKQ-EHTZ>]; Cynthia L. Webb, *Big Blue Opens the Patent Vault*, WASH. POST (Jan. 11, 2005), <http://www.washingtonpost.com/wp-dyn/articles/A376-2005Jan11.html> [<http://perma.cc/VQT7-JDJ8>].

270. *IBM Pledges 500 U.S. Patents to Open Source in Support of Innovation and Open Standards*, IBM (Jan. 11, 2005), <https://www-03.ibm.com/press/us/en/pressrelease/7473.wss> [<http://perma.cc/2EYH-LGCT>].

271. *Id.*

272. Wen Wen, Marco Ceccagnoli & Chris Forman, *Opening Up IP Strategy: Implications for Open Source Software Entry by Start-Up Firms*, MANAGEMENT SCIENCE 9 (Articles In Advance, 2015), <http://dx.doi.org/10.1287/mnsc.2015.2247> [<https://perma.cc/8Z44-RSB5>].

273. See *id.*

274. Ann Bednarz, *IBM Retains Patent Crown, Frees 500 Patents for Open Source Use*, NETWORK WORLD (Jan. 11, 2005), <http://www.networkworld.com/article/2328329/software/ibm-retains-patent-crown--frees-500-patents-for-open-source-use.html> [<http://perma.cc/3JEN-ZUJP>] (discussing the disparity between the 500 patents pledged and the ever-growing patent portfolio of the company).

275. *Id.*

276. Clayton Browne, *IBM Filed Most US Patents for 22nd Straight Year*, VALUEWALK (Jan. 12, 2015), <http://www.valuewalk.com/2015/01/ibm-filed-us-patents-22nd-straight-year> [<http://perma.cc/7ZEA-MQEF>].

carefully selected and pledged 500 specific patents relating to operating system and other technologies, where IBM believed that a more open model of innovation presented the company with greater economic prospects than an exclusionary approach.²⁷⁷ Some studies suggest IBM's patent pledge proved successful in encouraging complementary technological innovation to the technological areas that the pledged patents covered.²⁷⁸

How did patents facilitate this economic move? In important respects, they functioned as an informational conduit to third parties. First, by sacrificing a large block of patents and their accompanying exclusive rights, IBM signaled to the product market the seriousness of its altered approach to patent rights.²⁷⁹ Indeed, such a sacrifice may have aided many market participants in inferring certain characteristics of the company, including perceptions of IBM as a willing collaborator, an attractive possible employer, and a company with economic vision worth investing in. By the same token, the circumscribed nature of the pledge, in light of IBM's tens of thousands of unpledged patents, also signaled that this altered approach did not apply across the board.²⁸⁰ For those tens of thousands of non-pledged patents, IBM signaled that it intended business as usual.²⁸¹ And even for the pledged patents, some commentators later questioned IBM's intentions based on purported violations of the pledge.²⁸²

Patent law's disclosure and claiming requirements aided IBM in achieving these informational objectives. For instance, these requirements, once fulfilled during the patenting process, provided IBM with a ready means by which to pledge with specificity. Indeed, with its 500-patent pledge, the company precisely signaled the areas of technology in which it hoped to encourage complementary innovation, and which areas remained

277. YOCHAI BENKLER, *THE WEALTH OF NETWORKS: HOW SOCIAL PRODUCTION TRANSFORMS MARKETS AND FREEDOM* 46–47 (2006); *see also* Wen Wen et al., *supra* note 272.

278. *Supra* note 277.

279. Lohr, *supra* note 269 (labeling IBM's approach to patents unconventional and contrasting it with the typically exclusionary approach of most patent holders).

280. Brockmeier, *supra* note 264 (expressing some skepticism about the pledge because of the large number of patents that were left out of it).

281. *Id.*

282. *See, e.g.*, Florian Mueller, *IBM Breaks the Taboo and Betrays Its Promises to the FOSS Community*, FOSS PATENTS (Apr. 6, 2010), <http://www.fosspatents.com/2010/04/ibm-breaks-taboo-and-betrays-its.html> [<http://perma.cc/RJN9-87AC>] (detailing an alleged violation of the pledge).

on the outside.²⁸³ And patent law's disclosure and claiming requirements provided the company with a systematic, well-documented means of providing those specifics.²⁸⁴

IBM's pledge likely also functioned as an informational signal to capital markets. IBM had long lost ground to Microsoft and others in competitively licensing operating system and related technologies to third parties.²⁸⁵ With its pledge, IBM signaled a shift in focus: rather than attempting to directly monetize operating system and related technologies, the company would focus on generating revenues from complementary products and services while encouraging innovation in these ceded technologies.²⁸⁶ This hoped-for innovation, in turn, would aid IBM in realizing gains from its core revenue-generating products and services.²⁸⁷

Last, the pledge may have also been intended as a recruiting tool, like with the other pledges reviewed above.²⁸⁸ Especially given the sour feelings of many in developer communities vis-à-vis Microsoft,²⁸⁹ siding with open

283. See Jim Wagner, *IBM's Patent Pledge Ripples Open Sourcers*, INTERNETNEWS.COM (Jan. 13, 2005), <http://www.internetnews.com/bus-news/article.php/3458551/IBMs-Patent-Pledge-Ripples-Open-Sourcers.htm> [<http://perma.cc/79TH-D3JD>] (quoting some commentators who believe that IBM pledged some crucial patents for FOSS development, while also quoting other commentators who believe that a significant number of relevant patents remain unpledged).

284. See generally Colleen V. Chien, *Exclusionary and Diffusionary Levers in Patent Law*, S. CAL. L. REV. (forthcoming 2016), http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2624692 [<https://perma.cc/JN4A-H8PZ>] (discussing the traditional focus on the exclusionary value of patents, and arguing that patents also have value as tools in diffusing information to the public).

285. Andrew Orłowski, *OS/2 a Quarter Century On: Why IBM Lost Out and How Microsoft Won*, THE REGISTER (Nov. 27, 2012), http://www.theregister.co.uk/2012/11/27/the_os_wars_os2_25years_old [<http://perma.cc/7TZK-3HFD>] (describing in detail how Microsoft's operating system technologies won out over IBM's over time); see also Robert Prentice, *Vaporware: Imaginary High-Tech Products and Real Antitrust Liability in a Post-Chicago World*, 57 OHIO ST. L.J. 1163, 1229 n.270 (1996).

286. Amy L. Landers, *The Antipatent: A Proposal for Startup Immunity*, 93 NEB. L. REV. 950, 992–93 (2015) (suggesting that the pledge was motivated in part to encourage development of the Linux operating system as a viable alternative to Microsoft's Windows technologies); Wen Wen et al., *supra* note 272, at 8.

287. Landers, *supra* note 286; Wen Wen et al., *supra* note 272, at 8.

288. Brockmeier, *supra* note 264.

289. Julie Bort, *Microsoft Exec: If You Hate Microsoft, You Don't Really Know Us*, BUSINESS INSIDER (Feb. 21, 2012), <http://www.businessinsider.com/microsoft-exec-if-you-hate-microsoft-you-dont-really-know-us-2012-2> [<http://perma.cc/LY3N-2ZKR>] (starting the article with the question, in reference to Microsoft, “[h]ow do you convince a generation of people who have grown up thinking you are evil to change their minds and love you instead?”).

source software developers via its pledge may have been a particularly savvy public relations move on the part of IBM at the time.²⁹⁰

In sum, IBM appears to have used its pledge to signal information to capital, labor, and product markets. There was, of course, no guarantee that investors, competitors, or developers would respond favorably to such signals.²⁹¹ Regardless of the efficacy of the signals, however, patents played a critical role in helping convey them.

E. THE PATENT IMPOVERISHED

The above patent pledging examples all concern well-known, larger companies that often boast significant patent portfolios. The question naturally arises, then, what to make of the purported informational value of patents for those possessing few patents. Can and do smaller companies that possess a more limited number of patents realize the same types of informational value from their patents as, say, an IBM or a Tesla?

As discussed briefly above, some accounts suggest that smaller companies realize informational value from patents primarily as a signal to investor communities.²⁹² In other words, start-ups and the like may acquire patents in order to convey to capital markets that they are serious innovators and thus worthy of investment. But the informational values of patents as described in this Article go beyond this narrative, and the question remains whether these smaller patent holders are able to derive these broader types of informational value from their patents.

The short answer is that such informational uses are certainly a possibility for smaller companies, but other considerations may lead them to adopt a different strategy with respect to the few patents that they possess. For instance, smaller companies heavily dependent on investor funding in the early stages of their lifecycles may be more risk-averse to engaging in behavior that may put off investors, which may include patent pledging. In other words, the safer approach may be to acquire patents in hopes of signaling innovativeness to investors, but to do little else with the patents unless investors otherwise advise. Hence, larger companies with significant portfolios appear to have greater leeway in how they use their

290. Scott Merrill, *Is IBM Splitting Hairs with Open Source?*, TECHCRUNCH (Apr. 8, 2010), <http://techcrunch.com/2010/04/08/is-ibm-splitting-hairs-with-open-source> [<http://perma.cc/EW49-BMF8>] (calling IBM's pledge "an obvious PR stunt to get buddy-buddy with the open source community").

291. Wagner, *supra* note 283 (providing a variety of viewpoints in response to IBM's pledge).

292. Long, *supra* note 23.

patents, since their relative stability means that they are not dependent on early-stage investment in the same way that a start-up company may be.

Furthermore, the businesses of larger companies are likely to be more diverse than those of smaller entities and thus demand different patent strategies. Hence, because of this diversification, larger companies with significant patent portfolios may be more likely to utilize patents in ways that yield the types of informational value discussed in this Article. Smaller companies, on the other hand, are more likely to still be seeking their commercial footing, and pledging away patent rights may be a risky behavior in light of that commercial status.

Last, though no absolute reasons bar smaller companies from realizing some of the informational value of patents as described in this Article, their smaller numbers of patents may mean that any signals that they do seek to convey are limited. Put differently, a smaller company that pledges its few patents may send a less significant signal to markets than IBM pledging 500 patents, simply because the sacrifice of the smaller company, in absolute terms, is less. Of course, the opposite conclusion may also be reasonable because, in relative terms, the smaller company may have sacrificed more than IBM, given IBM's tens of thousands of patents. But the uncertainty of how markets will perceive the smaller patent holder's pledge, combined with the uncertainty of how the smaller company's investors may receive its actions, may cumulatively mean that smaller companies with few patents are less likely to engage in patent pledging than the IBMs of the world. Or, even when they do engage in patent pledging, the informational effects of their activities are fewer. The informational value of patents described in this Article may thus largely be a big firm phenomenon, at least for now, even in cases where smaller companies with few patents do engage in patent pledging.²⁹³

V. SCOTUS AND THE INFORMATIONAL ROLE OF PATENTS

Parts II–IV of this Article have provided an informational account of patents. Unlike traditional theories of patent law, which view any possible informational value of patents from the perspective of the public, this Article argues that patents can have informational value to patent owners

293. Indeed, a database of patent pledgers lists a number of smaller pledgers. *See* Program on Info. Justice, *supra* note 3. But the pledges of such parties do not appear to garner the same type of attention as those of larger companies such as Tesla, Microsoft, and others. This all suggests that the size of the company and its numbers of patents do affect what informational value pledgers may derive from use of their patents in this way.

as well. The phenomenon of patent pledging, several examples of which Part IV reviewed, provide evidence in support of this claim.

Part V now turns to recent Supreme Court case law in order to assess what informational impacts these decisions may have. This review is particularly relevant now, as the U.S. Supreme Court has been quite active in the last several years in rendering key patent law decisions.²⁹⁴ While the informational value of patents has not been explicitly considered in these decisions, many of these decisions nonetheless may have significant consequences for patents as informational tools, as the below sections demonstrate. Going forward, this Article argues that courts and Congress should consider the informational value of patents in making decisions regarding patent law, particularly in industries where patent pledging is more common.

A. THE INFORMATIONAL IMPACT OF RECENT PATENTABLE SUBJECT MATTER CASES

The following sections first summarize recent Supreme Court decisions regarding patentable subject matter and then assess what those cases may mean for the informational value of patents.

1. *Patentable Subject Matter at the Supreme Court*

In recent years, the U.S. Supreme Court has spent considerable time assessing what constitutes “patentable subject matter.”²⁹⁵ Section 101 of the Patent Act sets forth the patentable subject matter requirement generally.²⁹⁶ This section indicates that anyone may obtain a patent on an invention so long as that party “invents or discovers any new and useful process, machine, manufacture, or composition of matter, or any new and useful improvement thereof.”²⁹⁷ On its face, § 101 thus lacks express subject matter limitations beyond requiring that the item is within one of the enumerated categories and proves “new and useful.” As others have noted, the categorical limitations do not actually function as significant limitations, since almost anything that a party could patent easily fits

294. Ouellette, *supra* note 37 (noting this trend, and listing all recent Supreme Court cases dealing with patent law).

295. John M. Golden, *Flook Says One Thing, Diehr Says Another: A Need for Housecleaning in the Law of Patentable Subject Matter*, 82 GEO. WASH. L. REV. 1765, 1766–75 (2014) (reviewing this trend and the relevant cases).

296. 35 U.S.C. § 101 (2012).

297. *Id.*

within one of them.²⁹⁸ Indeed, those favoring an expansive view of patentable subject matter often point to language from the legislative history of the 1952 Patent Act, which indicates, in part, that “anything under the sun made by man” is meant to be patentable subject matter.²⁹⁹ The America Invents Act, passed in 2011 and representing the most significant patent law reform since the 1952 Patent Act, did nothing to alter this understanding.³⁰⁰

Yet over time courts have developed a number of exceptions to patentable subject matter that are not explicitly listed in the statute.³⁰¹ These common law exceptions generally preclude a patent that purports to claim a “law of nature,” an “abstract idea,” or “natural phenomena.”³⁰² Though some scholars have questioned the value of these exceptions,³⁰³ others argue that the Supreme Court has made clear that it does not intend to abandon them.³⁰⁴

Indeed, the recent slew of Supreme Court patentable subject matter cases has largely dealt with determining when these exceptions apply. In *Bilski v. Kappos*, for example, the Court held that a patent claiming the concept of risk-hedging as applied to energy markets constituted an abstract idea ineligible for patent protection.³⁰⁵ In *Mayo Collaborative Services v. Prometheus Laboratories, Inc.*, the Court reviewed the “law of nature” exception and ultimately held that the patent at issue did not claim patentable subject matter because it attempted to claim a law of nature relating to drug dosages and metabolite levels in the blood.³⁰⁶ And in *Association for Molecular Pathology v. Myriad Genetics, Inc.*, the Court

298. Mark A. Lemley, Michael Risch, Ted Sichelman & R. Polk Wagner, *Life After Bilski*, 63 STAN. L. REV. 1315, 1328 (2011).

299. S. Rep. No. 82-1979, at 5 (1952).

300. See, e.g., Robert A. Armitage, *Understanding the America Invents Act and Its Implications for Patenting*, 40 AIPLA QUART. J. 1 (2012) (failing to mention patentable subject matter even once in over 133 pages of summary of the America Invents Act’s implications).

301. See *Bilski v. Kappos*, 130 S. Ct. 3218, 3225 (2010).

302. *Id.*

303. Michael Risch, *Everything Is Patentable*, 75 TENN. L. REV. 591, 595 (2008) (arguing for jettisoning the patentable subject matter requirement and instead relying on patent law’s other requirements in assessing patentability).

304. Lemley et al., *supra* note 298, at 1326.

305. *Bilski*, 130 S. Ct. at 3229–30.

306. *Mayo Collaborative Servs. v. Prometheus Labs., Inc.*, 132 S. Ct. 1289, 1297–98 (2012) (holding that method claims requiring the administration of a drug, followed by assessing the metabolite levels in the blood and then adjusting the drug level in order to change the metabolite levels claimed a law of nature and thus did not qualify for patent protection).

addressed the “natural phenomena” or “product of nature” exception, ultimately holding that isolated DNA sequences are not patentable subject matter because the DNA segment occurs in nature (even if not in isolated form).³⁰⁷

Most recently, the Court revisited the “abstract idea” exception in *Alice Corp. Proprietary Ltd. v. CLS Bank International*.³⁰⁸ The Court, building on the earlier patentable subject matter cases mentioned directly above, formalized a two-step approach to assessing whether one of the patentable subject matter exceptions applies.³⁰⁹ First, a court is to assess whether the patent claims are directed to a law of nature, abstract idea, or natural phenomenon.³¹⁰ If so, the court then moves to the second step, which asks whether the patent claims involve an “inventive concept.”³¹¹ In other words, if the patent includes elements or a combination of elements “sufficient to ensure that the patent in practice amounts to significantly more than a patent upon [a law of nature, abstract idea, or natural phenomena],” then the patent claims patentable subject matter.³¹² Otherwise, it does not.³¹³

According to the Court, the patents in question in *Alice* claimed mitigating settlement risk in financial transactions by using a computer system as a third-party intermediary.³¹⁴ The Court determined that this idea of intermediated settlement was an abstract idea not eligible for patent protection, and that requiring computer implementation did not transform this abstract idea into a patent-eligible invention because generic computer implementation did not involve an “inventive concept.”³¹⁵

2. *The Possible Informational Effects of the Alice Decision*

Cumulatively, these patentable subject matter cases have several important possible implications relating to the informational value of patents. First, they may boost the informational value of patents by requiring more detailed disclosures within the patent document in order to

307. *Ass’n for Molecular Pathology v. Myriad Genetics, Inc.*, 133 S. Ct. 2107, 2111 (2013).

308. *Alice Corp. Proprietary Ltd. v. CLS Bank Int’l*, 134 S. Ct. 2347 (2014).

309. *Id.*

310. *Id.*

311. *Id.*

312. *Id.*

313. *Id.*

314. *Id.*

315. *Id.*

avoid the applicability of one of these exceptions. For instance, though some early evidence following *Alice* suggests that very few current business method and software patents are valid when applying *Alice's* standards,³¹⁶ the post-*Alice* cases in which patents have survived the patentable subject matter question indicate that the greater the level of specificity in the patent document, the more likely that the patent claims will survive.³¹⁷ Thus, while cases since *Alice* may not be entirely consistent,³¹⁸ the trend for now indicates that a greater amount of specificity and detail in the patent may save the claims from invalidation.³¹⁹

Hence, this greater level of disclosure and specificity in patent applications may in some respects increase the informational value of patents. For instance, from the public's perspective, more specific and detailed patent disclosures will increase the value of patents by providing clearer notice of the patent's scope as well as greater insight into the technical accomplishments involved.³²⁰ This greater value to the public could, in turn, increase the informational value of patents to patentees that use patents as informational tools. For instance, more specific and detailed

316. Steven Seidenberg, *Business-Method and Software Patents May Go Through the Looking Glass After Alice Decision*, A.B.A. J. (Feb. 1, 2015), http://www.abajournal.com/magazine/article/business_method_and_software_patents_may_go_through_the_looking_glass_after [<http://perma.cc/C5VU-YWVP>] (reviewing evidence of court and administrative agency action in invalidating business method and software patents based on the *Alice* holding).

317. See, e.g., Gene Quinn, *Federal Circuit Finds Software Patent Claim Patent Eligible*, IPWATCHDOG (Dec. 5, 2014), <http://www.ipwatchdog.com/2014/12/05/federal-circuit-finds-software-patent-claim-patent-eligible> [<http://perma.cc/2R3K-8CMQ>] (discussing one such case). Some scholars have argued for standards akin to this standard previously. See, e.g., Mark A. Lemley, *Software Patents and the Return of Functional Claiming*, 2013 WIS. L. REV. 905, 962–63 (2012).

318. Quinn, *supra* note 317.

319. Ryan Davis, *5 Tips for Getting Software Patents Approved Post-Alice*, LAW360 (Sept. 2, 2014), <http://www.law360.com/articles/572491/5-tips-for-getting-software-patents-approved-post-alice> [<http://perma.cc/5WFH-AV4R>] (advising that narrowing claims and including more specific technical details in a patent application may well help save a patent from invalidation); Gene Quinn, *The Road Forward for Software Patents Post-Alice*, IPWATCHDOG (Feb. 25, 2015), <http://www.ipwatchdog.com/2015/02/25/the-road-forward-for-software-patents-post-alice> [<http://perma.cc/2BS8-YPDL>] (“[B]ased on the few decision points we have so far it seems that the courts will place disproportionate weight on the sufficiency of the specification when making decisions about patent eligibility . . . which means patent applications and issued patents must have robust technical disclosure to even stand a chance of getting through the patent eligibility threshold.”).

320. Fromer, *supra* note 19 (arguing in favor of the benefits of disclosure generally). *But see* Holbrook, *supra* note 18 (arguing generally that information disclosures under patent law have little if any value).

disclosures in the patent document will enable patent holders to use patents more effectively in signaling to product, labor, and capital markets information about themselves, their technologies, and their innovation preferences.³²¹

But this “sea-change” in patentable subject matter standards could have adverse effects on the informational role of patents as well. First, if fewer parties pursue or maintain patents because of these decisions,³²² or if fewer patents are valid based on the § 101 question more generally,³²³ then fewer parties may have exclusive rights with which to signal information to the relevant markets. Instead, such parties may opt for trade secrecy given the uncertain state of patentable subject matter following the Supreme Court’s decisions.³²⁴ And a greater preference for trade secrecy may mean that fewer inventions are disclosed to the public in general.³²⁵

Second, more specific and detailed disclosures may mean that, for the patents that are granted, the scope of those rights is severely limited. For instance, as mentioned, though courts are not supposed to read limitations from the rest of the patent document into the patent claims as a general matter, they sometimes do based on a variety of patent law doctrines.³²⁶ Indeed, core patent law requires courts to read patent claims “in light of the specification,”³²⁷ which includes the inventor’s technical disclosures. Furthermore, since parties are likely to narrow claims in hopes of overcoming the patentable subject matter hurdle, patent rights may be further limited.³²⁸ Hence, by including more specific and detailed disclosures and claims in their patent applications in response to the Supreme Court’s patentable subject matter decisions, inventors may

321. See *supra* Parts I & III.

322. See Seidenberg, *supra* note 316 (indicating that pursuing patents on software and business methods may not be worth it in light of the Supreme Court’s new standards and the high invalidation rate of such patents post the Court’s decisions).

323. See Jasper L. Tran, *Software Patent: One-Year Review of Alice v. CLS Bank*, 97 J. PAT. & TRADEMARK OFF. SOC. 532, 534 (2015) (noting that in the year since the *Alice* decision, around 82% of software patents have been invalidated when applying the *Alice* standards).

324. Schwartz, *supra* note 73 (describing the conventional wisdom for selecting between trade secrecy and patent protection).

325. For a counterargument to this contention, see Mark A. Lemley, *The Surprising Virtues of Treating Trade Secrets as IP Rights*, 61 STAN. L. REV. 311, 312–15 (2008) (arguing that trade secret protection actually facilitates the disclosure of inventions).

326. *Phillips v. AWH Corp.*, 415 F.3d 1303, 1312–23 (Fed. Cir. 2005) (en banc) (describing sources to be used in construing patent claims, including “intrinsic evidence” such as the technical disclosures included as part of the patent application).

327. *Id.*

328. Davis, *supra* note 319.

inadvertently (or advertently) limit their patent rights. And restricting rights in such a manner may mean that the informational potency of those rights also decreases, since parties purporting to sacrifice those rights may not be sacrificing much.³²⁹

As to the first concern—that parties will either stop pursuing patents or those that they do obtain will simply be invalid—a few responses are in order. First, it seems unlikely that parties serious about innovation will simply choose to forego patent protection for their innovations, including with respect to software and business method innovations, which some claim have been most impacted. Patents provide significant advantages, including strict liability for infringers and unique rights of exclusion, over other forms of protection.³³⁰ This likelihood of ongoing patenting may be particularly high since information technology is so important in today's global economy.³³¹

Second, the Court was clear in *Alice* and its other decisions that it was not precluding any specific type of technology from patent eligibility. *Alice*, for instance, does not mention the word “software” or explicitly discuss “business methods.”³³² Thus, the more likely effect of *Alice* and the other Supreme Court decisions is that patent drafters will adjust their practices to adapt to these cases.³³³ And the primary adjustment appears to be including more specific details in the patent document itself.³³⁴

Third, patents obtained post-*Alice* will likely have greater informational value as a result of the more detailed and specific disclosures contained therein, which may provide another reason for some parties to pursue such patents.³³⁵ Indeed, since most patents are never litigated (and

329. See *supra* Section III.A (discussing signaling theories that postulate that the lower the cost of the signal, the less credible it is likely to be).

330. But for recent scholarship challenging the assertion that patent law is a strict liability regime see Saurabh Vishnubhakat, *An Intentional Tort Theory of Patents*, 68 FLA. L. REV. (forthcoming 2016), http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2492200 [<https://perma.cc/S7DG-F3VF>].

331. See generally WILLIAM J. KRAMER, BETH JENKINS & ROBERT S. KATZ, *THE ROLE OF THE INFORMATION AND COMMUNICATIONS TECHNOLOGY SECTOR IN EXPANDING ECONOMIC OPPORTUNITY* (2007).

332. Robert Sachs, *Alice, the Illusory Death of Software Patents*, IPWATCHDOG (June 27, 2014), <http://www.ipwatchdog.com/2014/06/27/alice-the-illusory-death-of-software-patents> [<http://perma.cc/6L7J-N8HH>].

333. Quinn, *supra* note 317 (discussing what adaptations make sense in light of *Alice*).

334. Davis, *supra* note 319 (same); Quinn, *supra* note 317.

335. See Seidenberg, *supra* note 316; *supra* text accompanying note 321.

thus never invalidated)³³⁶ the many uses of patents—including the informational functions discussed herein—may have greater relevance in a post-*Alice* world.³³⁷

Of course, certain parties may opt out of the patent system as a result of these decisions. For instance, so-called “patent trolls”—which assert patents against others, but do not produce products or services themselves—may opt out of the system because, as some contend, their models depend in many cases on asserting overbroad and vague patents and then relying on the high costs of litigation to force a settlement.³³⁸ *Alice* and other patent law changes discussed *infra* may make these models more difficult to sustain.³³⁹ And if patent trolls stop providing a ready market for those wishing to sell some stake in their patents, it may mean that fewer parties ultimately pursue patents.³⁴⁰

But to the extent that such parties actually do rely on vague and ambiguous patents in their business models, the effects of the Supreme Court’s decisions would seem to be positive as an informational matter. After all, excessive ambiguity and vagueness in patents make informational

336. For reasons why alleged infringers may not focus on invalidity defenses even when patent litigation is commenced, see generally Roger Allan Ford, *Patent Invalidity Versus Noninfringement*, 99 CORNELL L. REV. 71 (2013).

337. Tran, *supra* note 323.

338. Mark A. Lemley & A. Douglas Melamed, *Missing the Forest for the Trolls*, 113 COLUM. L. REV. 2117, 2173 (2013) (indicating that in the IT industry, overbroad patents are the norm and facilitate patent trolling); Lisa Allen, *The Problem With Patent Trolls*, SAND HILL (Jan. 26, 2015), <http://sandhill.com/article/the-problem-with-patent-trolls> [<http://perma.cc/5C9V-LK3E>] (indicating that patent trolls often rely on vague and overbroad software patents, and that the *Alice* and other recent Supreme Court decisions may make the patent trolling business model more difficult to sustain because these decisions may make such patents even more suspect than they already were); Claire Bushey, *Why This Lawyer Is Rethinking Patent Lawsuits*, CRAIN’S CHICAGO BUSINESS (June 6, 2015), <http://www.chicagobusiness.com/article/20150606/ISSUE01/306069991/tripping-up-the-trolls> [<http://perma.cc/7ACX-HM4S>] (indicating that some prominent patent trolls are considering abandoning the business because of *Alice*, among other patent law changes).

339. For recent evidence suggesting this type of business model is alive and well, however, see Mike Masnick, *Patent Trolls Strike Back: Trolling Rebounds After Brief Supreme Court-Enabled Dip*, TECHDIRT (July 14, 2015), <https://www.techdirt.com/articles/20150714/10344431637/patent-trolls-strike-back-trolling-rebounds-after-brief-supreme-court-enabled-dip.shtml> [<http://perma.cc/ZPK6-ZSTU>] (noting a recent spike in patent troll activity).

340. Paul Schneck, *Not So Scary, After All: In Defense of Patent Trolls*, FORBES (Feb. 1, 2013), <http://www.forbes.com/sites/ciocentral/2013/02/01/not-so-scary-after-all-in-defense-of-patent-trolls> [<http://perma.cc/P2MG-VXAH>] (arguing that patent trolls are allies with independent inventors and others versus large corporations).

content more difficult to decipher.³⁴¹ So to the extent that post-*Alice* such uncertainty is more difficult to foment, all the better.

But it may still be the case that some parties cease pursuing and patenting their inventions absent the patent troll ally. In other words, some inventors may be motivated to invent by the possibility of monetizing those inventions by patenting them and then licensing or selling their patents to a patent troll or similar entity.³⁴² Rather than boosting the informational value of patents, therefore, the Supreme Court's patentable subject matter cases may mean that there's simply less information created.³⁴³

There are reasons to doubt this outcome, however. For instance, for inventive ideas that do have societal value, it seems likely that market opportunities will continue to exist or develop over time, including by way of patent sales or licensing.³⁴⁴ In other words, while business models that rely on vague and overbroad patents may dry up, others will likely step in to harness otherwise valuable inventive ideas, which in turn should provide parties with incentives to invent these valuable things.³⁴⁵

The second possible negative implication of the Supreme Court's patentable subject matter decisions mentioned above also merits a few responses. To recapitulate that concern: because the Supreme Court cases appear to require more specific and detailed disclosures to save patent claims from invalidation, the result may be excessively narrow patents. And such narrowness may mean that patents have less informational impact because (1) the narrow rights, when sacrificed, are simply less

341. See Fromer, *supra* note 19.

342. See generally James F. McDonough III, *The Myth of the Patent Troll: An Alternative View of the Function of Patent Dealers in an Idea Economy*, 56 EMORY L.J. 189 (2006) (arguing that patent trolls benefit the market by providing liquidity, market clearing, and increased efficiency to the patent markets).

343. *Id.*

344. There is some recent evidence, however, that patent licensing and the patent troll model in general does not actually lead to innovation. See, e.g., Robin Feldman & Mark A. Lemley, *Does Patent Licensing Mean Innovation?* (Stanford Law & Econ., Olin Working Paper No. 473, 2015), http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2565292 [<https://perma.cc/6CKA-UCHX>] (concluding that it does not). But see Gene Quinn, *Flawed Survey Erroneously Concludes Patent Licensing Does Not Contribute to Innovation*, IPWATCHDOG (Feb. 22, 2015), <http://www.ipwatchdog.com/2015/02/22/flawed-survey-erroneously-concludes-patent-licensing-does-not-contribute-to-innovation> [<http://perma.cc/K8PD-LDPA>] (critiquing Feldman & Lemley).

345. But as noted above, recent evidence suggests that patent trolling is alive and well, despite *Alice*. So the worry that these patent dealers are set to vanish, and with them the good that they do, seems exaggerated at the least. See Masnick, *supra* note 339.

meaningful as signals, and (2) narrowed patents may mean less technical disclosure in some respects.

As mentioned, however, the greater detail and specificity required will likely increase the informational value of patents on the whole because that detail and specificity will allow parties to be even more precise in their messaging.³⁴⁶ Recipients of those informational signals will therefore benefit, since they will have greater ease deciphering the signal's actual meaning.³⁴⁷ Indeed, patent rights based on broad patent claims with minimal disclosures may generally have less informational value, even when used in an explicitly informational context, because their scope remains so uncertain.³⁴⁸

Of course, it remains true that narrower patents may have less informational impact in some cases, simply because sacrificing these narrower rights in a patent pledging context may be less meaningful.³⁴⁹ By way of analogy, a party claiming to support some cause by donating a few dollars to that cause may have less credibility than the party that donates thousands of dollars to the same cause. Similarly, a patent holder with a narrow (and thus less economically valuable) patent may obtain less signaling impact when pledging that patent than if they were able to pledge a broader, more economically valuable patent.³⁵⁰

But broad, economically valuable patents are still available, even after the Supreme Court's recent decisions.³⁵¹ Rather than prohibiting such patents, the Supreme Court's patentable subject matter decisions instead appear to require patentees making such claims to support them with

346. See Seidenberg *supra* note 316; *supra* text accompanying note 321.

347. Fromer, *supra* note 19, at 568.

348. *Id.*

349. See *supra* Section III.A (discussing signaling theories that postulate that the lower the cost of the signal, the less credible it is likely to be).

350. *Id.*

351. Bradley Knepper & Benjamin Lieb, *The New Methods for Drafting Patent Applications Post-Alice*, INSIDE COUNSEL (Jan. 23, 2015), <http://www.insidecounsel.com/2015/01/23/the-new-methods-for-drafting-patent-applications-p> (describing drafting methods to overcome the hurdles to patentability that *Alice* introduced); Robert Merges, *Symposium: Go Ask Alice—What Can You Patent After Alice v. CLS Bank?*, SCOTUSBLOG (June 20, 2014) [<http://www.scotusblog.com/2014/06/symposium-go-ask-alice-what-can-you-patent-after-alice-v-cls-bank>] [<http://perma.cc/L3NL-QARQ>] (noting that the *Alice* decision did not rule that software or business methods are ineligible for patent protection); Thomas D. Nguyen, *Patent Eligibility of Computer Software Inventions in a Post-Alice Era*, NAT'L L. REV. (May 15, 2015), <http://www.natlawreview.com/article/patent-eligibility-computer-software-inventions-post-alice-era> [<http://perma.cc/E6UM-L6NX>] (noting that software patentability is still possible following the *Alice* decision).

extensive technical disclosures that demonstrate that the inventor is not simply claiming an abstract idea, law of nature, or natural phenomenon (or is not claiming beyond what the inventor actually invented).³⁵² If an inventor is unable to satisfy this threshold, then one of the patentable subject matter exceptions may apply.³⁵³ And while the more detailed technical disclosures may narrow the rights significantly, such outcomes seem to better align patent law with its purpose: granting inventors rights to what they actually invented.³⁵⁴ These outcomes may also actually make the patents more valuable, since they are, on the basis of such increased specificity, more likely to be found valid.

Of course, if lower courts interpret the Supreme Court's standards post-*Alice* such that essentially all software or business method patents are framed as abstract ideas or laws of nature (and are thus invalid), then that may be a cause for concern for the informational value of patents, as well as the value of patents in general.³⁵⁵ But at least in theory (and some limited practice since the decisions),³⁵⁶ the Supreme Court's decisions do not require such a result, instead appearing to require detailed and specific disclosures and claims in order to avoid having one of the patentable subject matter exceptions apply.³⁵⁷ And that approach promises an increase of informational value for patents, rather than a diminution thereof.

In sum, the Supreme Court's recent patentable subject matter cases seem poised to increase the informational value of patents, both to the general public and to patentees using patents as informational tools. While the Court's standards will certainly undergo revision over time, for now the trend appears positive as an informational matter.

B. DEFINITE CLAIMS

The Supreme Court also recently implicated the informational value of patents with its decision in *Nautilus, Inc. v. Biosig Instruments, Inc.*³⁵⁸ At issue in the *Nautilus* case was the Patent Act's requirement that the patent document conclude with "one or more claims particularly pointing out and

352. Knepper & Lieb, *supra* note 351.

353. *Id.*

354. *See generally* Holbrook, *supra* note 18 (arguing that enablement that manifests possession is more theoretically consistent with the typical justifications for the patent system).

355. Tran, *supra* note 323 (detailing such a general trend).

356. *Id.* (detailing some cases where patents have not been invalidated based on application of the *Alice* standard).

357. Knepper & Lieb, *supra* note 351.

358. *Nautilus, Inc. v. Biosig Instruments, Inc.*, 134 S. Ct. 2120 (2014).

distinctly claiming the subject matter which the inventor or a joint inventor regards as his invention.”³⁵⁹ These formal patent claims define the scope of what a patent covers, yet uncertainty remained regarding how “definite” the claims needed to be, given that language will always remain an imperfect tool for delineating boundaries.³⁶⁰

The Court of Appeals for the Federal Circuit, which has nationwide jurisdiction for appeals that include a patent issue, had previously set a somewhat indefinite definiteness standard.³⁶¹ It had held that a claim is indefinite “only when it is ‘not amenable to construction’ or ‘insolubly ambiguous.’”³⁶² But the Supreme Court in *Nautilus* overruled the Federal Circuit, holding instead that the definiteness standard is met when a “patent’s claims, viewed in light of the specification and prosecution history, inform those skilled in the art about the scope of the invention with reasonable certainty.”³⁶³

This “reasonable certainty” standard should thus result in more concrete and clear patent claims than the Federal Circuit’s previous standard. Indeed, that former standard promised a rather indefinite future, since, so long as *some* construction could be applied to the claims, the patent claims satisfied the Federal Circuit’s standard.³⁶⁴

This greater level of definiteness, in turn, would appear to boost the informational value of patents, to both the public as well as patentees that use their patents as informational tools. It boosts the informational value of patents to the public by making it more likely that patent claims will be more specific and clear than they otherwise may have been. That this would be so is intuitive: patent drafters will almost certainly take the *Nautilus* standard into account when drafting patent claims, which dictates greater precision than the Federal Circuit’s previous standard required.³⁶⁵ With these more specific and definite claims, patent owners using patents as informational tools should also benefit because their signals will be that much clearer. Hence, while the *Nautilus* decision may appear at first blush as a loss of value to patent owners, taking into account the informational

359. 35 U.S.C. § 112(b) (2012).

360. Fromer, *supra* note 45, at 757–58 (discussing the imprecision of language).

361. *Biosig Instruments, Inc. v. Nautilus, Inc.*, 715 F.3d 891, 898 (Fed. Cir. 2013), *cert. granted*, 134 S. Ct. 896 (2014), *vacated*, 134 S. Ct. 2120 (2014).

362. *Id.*

363. *Nautilus*, 134 S. Ct. at 2128–29 (2014).

364. *Id.*

365. Harold C. Wegner, *Post-Nautilus Patent Drafting for Claim Definiteness and Proper Generic Scope*, 10TH ANN. ADVANCED PAT. L. INST. (Mar. 12–13, 2015) (providing drafting guidance to patent prosecutors in light of the *Nautilus* decision).

value of patents suggests it may also be a win in certain informational respects.

Of course, many of the same concerns raised in the preceding Section may equally apply here. That is, the reduced scope of patent claims that the Supreme Court's definiteness standard could bring about may dampen incentives to invent and pursue patents on those inventions. Weakening those incentives, therefore, may mean less inventive information created and publicly disclosed. Furthermore, the narrower scope of patent claims that the *Nautilus* decision may lead to may also inadvertently lessen the informational impact of patents in, for instance, patent pledging scenarios.

But for the same reasons discussed above, there is good reason to doubt such outcomes. In the year since the *Nautilus* case, for instance, there has been no discernible reduction in the number of patent applications filed.³⁶⁶ Parties are still filing a record number of patent applications, and are likely simply taking into account the Supreme Court's new standards when filing them. Furthermore, though these applications may be narrower in theory because they are more definitively scoped, it may also be the case that they are consequently more valuable because they are less likely to be found invalid. Indeed, as an informational matter, these more precise patent claims should mean that both the public and patentees benefit.

In sum, the *Nautilus* decision is another recent Supreme Court decision that promises informational benefits to patentees and the public alike. While the Court's standard will almost certainly undergo refinement in subsequent cases, for now the *Nautilus* holding points to a more promising informational future for patents.

VI. AN INDUSTRY-SPECIFIC PHENOMENON?

Much of the informational value of patents discussed in the preceding Parts may be industry-specific. For instance, other scholars have argued that patents play different roles depending on the industry.³⁶⁷ In the

366. See, e.g., U.S. Patent Trade Office, *U.S. Patent Activity Calendar Years 1790 to the Present, Table of Annual U.S. Patent Activity Since 1790*, USPTO, http://www.uspto.gov/web/offices/ac/ido/oeip/taf/h_counts.htm [<http://perma.cc/U2S8-XGGN>] (last modified Oct. 12, 2015) (showing that 2014 saw the most patent applications filed with the U.S. Patent Office in history, despite the *Nautilus* ruling issuing in July of that year).

367. See generally JAMES BESSEN & MICHAEL J. MEURER, *PATENT FAILURE: HOW JUDGES, BUREAUCRATS, AND LAWYERS PUT INNOVATORS AT RISK* (2008) (arguing that the patent system imposes more costs than benefits in most technology sectors); BURK & LEMLEY, *supra* note 36 (arguing that courts should treat different technology

pharmaceutical and biotechnology industries, for instance, a patent's rights of exclusion appear to be more critical in helping parties recoup their traditionally higher research and development costs.³⁶⁸ Conversely, in the IT industries generally, and the software industry in particular, patents may not be as critical in this regard because of the relatively lower costs of research and development and other factors that enable parties to thrive without having to rely on asserting exclusive rights against others.³⁶⁹

The patent pledging phenomenon provides some support to the broader point that the role of patents differs depending on the industry. For instance, though the phenomenon cannot be confined to any one industry, most of the known pledges have been made in the IT industry, while only a few relate to biotechnology, and none as of yet have come from pharmaceutical companies.³⁷⁰ Hence, though patents may have latent informational value in the biotechnology and pharmaceutical industries, this informational value is currently most clearly manifest in the IT sector.

But the more specific point that patents are unnecessary or even harmful in the IT industry, simply because participants in these industries do not rely on a patent's exclusionary rights the same way they do in the biotechnology and pharmaceutical industries, is incomplete. Indeed, as discussed in Parts III and IV *supra*, the informational value of patents manifest in the patent pledging context highlights previously unexamined informational roles of patents in promoting a variety of economic goals. While this informational value of patents may not always offset the costs that patents purportedly impose in industries, in some cases it may. This informational role of patents should thus become a greater consideration in developing patent law and policy, particularly in industries, such as IT, where that value is apparent.

Indeed, some scholars argue that courts already de facto tailor patent law according to industry, and that courts can address many other problems in patent law by increasing this trend.³⁷¹ To the extent that they do so, courts should take into account the informational value of patents in performing such tailoring. For instance, courts in applying patent law's disclosure requirements to IT-related patents would do well to interpret

sectors differently in terms of patent law in order to elide significant hindrances to innovation that the current patent system causes).

368. BURK & LEMLEY, *supra* note 36.

369. *Id.*

370. Program on Info. Justice, *supra* note 3 (listing only two out of 160 patent pledges relating to biotechnology, and none relating to the pharmaceutical industry).

371. BURK & LEMLEY, *supra* note 36.

those requirements expansively. Doing so would increase the informational value of patents by requiring patent applicants to include more specific technical details in their applications. While such requirements may at first blush seem like a loss of value to such patent holders, in reality they may actually make those patents more valuable as an informational matter. What is more, doing so would also tailor the scope of the patent claims to better fit how innovation typically occurs in the industry, as other scholars have argued.³⁷²

By neglecting to take into account the informational role of patents in their decisions, on the other hand, courts fail to consider an increasingly important use of patents in certain industries and may, consequently, craft decisions that harm innovation in those industries. For instance, while the Supreme Court cases reviewed above may have significant informational impacts, particularly in the software industry, it seems clear that the Court's reasoning did not explicitly take into account those possible impacts. This Article's findings suggest that it and other courts should.

This point applies to Congress and other legislative bodies as well. Indeed, in recent years Congress has implemented the most far-reaching patent reforms in decades.³⁷³ And it continues to consider a variety of additional patent law reforms, largely aimed at curbing the "patent troll" phenomenon.³⁷⁴ State legislatures have also become involved with a variety of litigation reforms aimed at addressing what is perceived as excessive patent litigation.³⁷⁵ Yet absent from these legislative discussions are the informational roles that patents may play. This is not to suggest that such informational roles should dominate such discussions. But it is to argue that these roles should become part of the conversation.

372. *Id.*

373. Quentin Palfrey, *Patent Reform: Celebrating the One Year Anniversary of the America Invents Act*, THE WHITE HOUSE BLOG (Sept. 17, 2012), <https://www.whitehouse.gov/blog/2012/09/17/patent-reform-celebrating-one-year-anniversary-america-invents-act> [<http://perma.cc/95N7-PSGU>] (identifying the America Invents Act of 2011 as "one of the most significant legislative reforms to the patent system in our Nation's history").

374. *Patent Progress's Guide to Federal Patent Reform Legislation*, PATENT PROGRESS, <http://www.patentprogress.org/patent-progress-legislation-guides/patent-progress-guide-patent-reform-legislation> [<http://perma.cc/GRB5-QVB2>] (reviewing six separate bills in Congress relating to patent law for the 2015–2016 term).

375. Jonathan Griffin, *States, Congress Continue to Fight Against Patent Trolls*, NCSL BLOG (Mar. 18, 2015), <http://www.ncsl.org/blog/2015/03/18/states-congress-continue-to-fight-against-trolls.aspx> [<https://perma.cc/Q7NB-9BPK>] (reviewing evidence showing that thirty-nine out of fifty states have introduced some form of legislation aimed at curbing so-called patent troll problems).

Indeed, better recognizing the informational value of patents in specific industries bolsters the case for more disclosure generally, as discussed in the judicial context above. But Congress need not leave the issue for courts to sort out; it could also take an active role in both studying the patent pledging phenomenon and altering patent law in order to better serve industries that increasingly rely on patents as informational instruments rather than solely as exclusionary ones.

VII. CONCLUSION

Traditionally, patents have been viewed as primarily exclusionary tools. Indeed, conventional patent law theories assume that the key benefit of a patent to the patent owner is the exclusive rights that come with it. Yet the patent pledging phenomenon reviewed in this Article shows that patents can also yield significant informational value for both patent holders and the public alike. As such, these informational roles of patents deserve greater consideration by courts and legislatures in formulating and implementing patent law and policy, particularly in industries, such as IT, where patent pledging is more prevalent.

This is not to say that the informational and exclusionary values of patents are mutually exclusive, or that one value need trump the other in policy discussions. Indeed, the same party may use patents in a variety of ways, sometimes as an informational tool, and other times as an exclusionary one. Such varied uses should not be viewed as conflicting or even contradictory. Instead, they simply reflect a variety of economic interests that a party may have, and which different uses of patents may facilitate. Better adapting patent law theory and policy to these realities would thus improve patent law's capacity to support innovators in pursuing their varied economic goals and thereby promote innovation more generally. And that, in the final analysis, is the purpose of having a patent system at all.

None of this is to say that patents could not be improved as informational tools. Some previous proposals aimed at improving disclosures under patent law, for instance, may also have merit when viewing patents as informational tools.³⁷⁶ And proposals focused on limiting patent holders' typically robust rights of exclusion under certain circumstances may also, if implemented, improve the informational value

³⁷⁶ Fromer, *supra* note 19, at 563–94 (discussing various means for improving disclosures under patent law).

and uses of patents.³⁷⁷ This Article does not explicitly review and assess these proposals. But it does suggest that such proposals—and patent law and policy in general—should better take into account the informational value of patents articulated in this Article.

377. See, e.g., Katherine J. Strandburg, *What Does the Public Get? Experimental Use and the Patent Bargain*, 2004 WIS. L. REV. 81 (2004) (proposing expanding the experimental use exception to patent infringement).

GAGS AS GUIDANCE: EXPANDING NOTICE OF NATIONAL SECURITY LETTER INVESTIGATIONS TO TARGETS AND THE PUBLIC

Rebecca Wexler[†]

ABSTRACT

National Security Letters (NSLs) are administrative subpoenas that the FBI uses to demand information from Internet service providers without prior judicial approval. They almost always include nondisclosure orders, commonly called “gags,” which prohibit the recipient from discussing the letter’s contents or even its mere existence. Courts and commentators have expressed concern that these gags may be overbroad prior restraints that violate the First Amendment and shroud government surveillance in undue secrecy. On November 30, 2015, an NSL gag order was lifted in full for the first time after a federal district judge found no “good reason” to retain it.

This Note considers the related rights of NSL targets. It argues that the FBI should provide notice of NSL investigations to targets and the public once government interests in secrecy abate. Specifically, once a nondisclosure order is lifted, thereby authorizing the recipient of the NSL to reveal any information about it that she desires, the government should disclose that same information. Enhancing transparency about government surveillance in this manner would not risk harm or cause undue administrative burden. It would harmonize with longstanding, closely related domestic criminal statutes. And it would advance core principles that underlie the Fourth Amendment. Moreover, the First Amendment offers ready balancing tests that can easily and reasonably be applied to guide government notice practices.

DOI: <http://dx.doi.org/10.15779/Z38DK0M>

© 2016 Rebecca Wexler.

[†] Rebecca Wexler is a member of the Yale Law School J.D. Class of 2016 and a Fellow of the Information Society Project. This Note received the 2015 Judge William E. Miller Prize for best paper concerning the Bill of Rights at Yale Law School. It also won first place in the 2015 *Berkeley Technology Law Journal* student writing competition. The author is deeply grateful to Akhil Reed Amar for his inspiration and generous guidance. Orin Kerr, Jonathan Manes, Peter Swire, and Lee Tien provided thoughtful comments on earlier drafts. Jack Balkin, Philip Bobbitt, Cindy Cohn, Mark M. Jaycox, Daniel Solove, and Patrick Toomey offered helpful feedback over email and conversations, as did Jane Ostrager and Bradley Silverman. The author especially thanks Waqas Akmal, Erica Fisher, and the editors of the *Berkeley Technology Law Journal* for their support and excellent editorial suggestions.

TABLE OF CONTENTS

I.	INTRODUCTION	327
II.	THE PROPOSAL: PROVIDING NOTICE OF NSL INVESTIGATIONS ONCE NONDISCLOSURE ORDERS END	331
III.	STATUTORY PRECEDENTS: BALANCING NOTICE WITH COMPETING INTERESTS IN SECRECY, SECURITY, AND CONFIDENTIALITY	337
IV.	CONSTITUTIONAL PRINCIPLES: FOURTH AMENDMENT SUPPORT FOR NOTICE AND FIRST AMENDMENT GUIDES FOR DISCLOSURE	342
	A. NOTICE TO THE TARGET PROTECTS THE INDIVIDUAL RIGHT TO CONTEST FOURTH AMENDMENT REASONABLENESS IN JURY TRIALS	345
	1. <i>Non-Governmental Actors Should Police Government Abuse</i>	345
	2. <i>The Right to Contest Implies an Individual Right to Notice</i>	349
	B. NOTICE TO THE PUBLIC PROTECTS THE REPUBLICAN RIGHT TO CONTEST FOURTH AMENDMENT REASONABLENESS IN THE POLITICAL ARENA	352
	1. <i>Republican Fourth Amendment Rights</i>	352
	2. <i>The Right to Contest Implies a Public Right to Notice</i>	355
	C. FIRST AMENDMENT STANDARDS FOR NOTICE.....	357
V.	CONCLUSION.....	360

I. INTRODUCTION

On November 30, 2015, a National Security Letter gag order was lifted in full for the first time.¹ National Security Letters (NSLs) are administrative subpoenas that the FBI uses to demand information from Internet service providers without prior judicial approval.² Over 300,000 NSLs have been issued over the past decade.³ They almost always include nondisclosure orders, commonly called “gags,” which prohibit the recipient from discussing the letter’s contents or even its mere existence.⁴ One such recipient, Nicholas Merrill, fought an eleven-year court battle to be able to speak freely about the NSL that he received.⁵ In November 2015, he finally won; a federal district court found no “good reason” to retain the gag in relation to any part of the NSL.⁶ Mr. Merrill may now tell anyone he wants anything about the NSL that he chooses to reveal, including precisely what kinds of information the FBI wanted to know—and about whom. Yet, despite the fact that this investigation ended without criminal charges years ago, the actual target of the investigation has no right to any of that information at all.

This Note proposes that the FBI provide notice to the targets of NSL investigations once government interests in secrecy abate. The government should also make public more information about its actual NSL practices than is currently required. The following discussion offers standards for

1. *Merrill v. Lynch*, No. 14-CV-9763, 2015 WL 9450650, at *3 (S.D.N.Y., Aug. 28, 2015). The author served as counsel for Mr. Merrill in this case, along with supervising attorneys Jonathan Manes and David Schulz and fellow students Amanda Lynch and Lulu Pantin, as part of the Media Freedom and Information Access Clinic at Yale Law School. *E.g.*, Press Release, Info. Soc’y Project, *Following Major First Amendment Victory, National Security Letter Recipient Nicholas Merrill Able to Reveal Previously Undisclosed Scope of FBI Warrantless Surveillance Tool* (Nov. 30, 2015), <http://isp.yale.edu/node/6037> [<https://perma.cc/H9AT-TEEX>] [hereinafter Info. Soc’y Project, *First Amendment Victory*]. The ACLU represented Mr. Merrill in a prior case that lifted his NSL gag in part. *Doe v. Holder*, 703 F. Supp. 2d 313 (S.D.N.Y. 2010).

2. 18 U.S.C. § 2709 (2012); *see also* Charles Doyle, *National Security Letters: Proposals in the 113th Congress*, CONG. RESEARCH SERV., R43322, at 1 (2015), <https://www.fas.org/sgp/crs/intel/R43322.pdf> [<https://perma.cc/HAD9-ZRZN>].

3. *E.g.*, Elec. Frontier Found., *National Security Letters: FAQ*, DEFENDING YOUR RIGHTS IN THE DIGITAL WORLD, <https://www.eff.org/issues/national-security-letters/faq> [<https://perma.cc/4RTR-229K>]; Elec. Privacy Info. Ctr., *Foreign Intelligence Surveillance Act Court Orders 1979-2014*, EPIQ.ORG, http://epic.org/privacy/wiretap/stats/fisa_stats.html#background [<https://perma.cc/UD9X-GBE2>].

4. *See Reauthorizing the USA Patriot Act: Hearing Before the S. Comm. on the Judiciary*, 111th Cong. 6 (2009) (statement of Glenn Fine, Inspector Gen., U.S. Dep’t of Just.), www.justice.gov/oig/testimony/t0909.pdf.

5. Info. Soc’y Project, *First Amendment Victory*, *supra* note 1.

6. *Merrill*, 2015 WL 9450650 at *3.

when and how much notice should be provided. Specifically, once an NSL nondisclosure order is lifted, thereby authorizing the recipient of the letter to disclose information about it to whomever she desires, the government should provide that same information to the target of the investigation. Further, when the government permits the recipients of NSL nondisclosure orders to publish general information about the letters they receive, the government should make similar materials available for public review. Either Congress should amend the relevant NSL statutes to require such notice, or the Executive should adopt a policy to provide it unilaterally.

Expanding notice of NSL investigations to targets and the public will enhance transparency about government surveillance. NSLs have existed since the 1980s, but originally targeted only information pertaining to foreign powers or their agents.⁷ In 2001, the USA PATRIOT ACT added domestic surveillance of U.S. persons to the FBI's NSL authority and authorized the agency to collect a far broader category of information relevant to "protect against international terrorism or clandestine intelligence activities."⁸ Congress also altered the NSL procedural requirements, permitting the FBI to deploy more NSLs more quickly, enabling an exponential increase in their issuance.⁹ Currently, the FBI issues almost sixty NSLs each day.¹⁰ After the leak of classified information by Edward Snowden in 2013, the President's Surveillance Review Group proposed mandatory prior judicial approval for all NSLs and nondisclosure orders.¹¹ To date, this proposal has not been adopted. Instead, the USA FREEDOM Act of 2015 expanded existing procedures

7. *E.g.*, Elec. Frontier Found., *In re: National Security Letter*, DEFENDING YOUR RIGHTS IN THE DIGITAL WORLD, <https://www.eff.org/cases/re-matter-2011-national-security-letter> [<https://perma.cc/34DP-G5XL>]; *see also*, Charles Doyle, *National Security Letters in Foreign Intelligence Investigations: A Glimpse at the Legal Background* 2, CONG. RESEARCH SERV., RS22406, at 1–2 (2014) [hereinafter Doyle, *A Glimpse*], <http://fas.org/sgp/crs/intel/RS22406.pdf> [<https://perma.cc/RK29-Y8NW>] (noting that Congress passed the first NSL statute as an amendment to RFRA).

8. Doyle, *A Glimpse*, *supra* note 7, at 1–2; § 2709(b).

9. *Id.* at 2.

10. RICHARD A. CLARK ET. AL, LIBERTY AND SECURITY IN A CHANGING WORLD: REPORT AND RECOMMENDATIONS OF THE PRESIDENT'S REVIEW GROUP ON INTELLIGENCE AND COMMUNICATIONS TECHNOLOGIES 93 (2013), https://www.whitehouse.gov/sites/default/files/docs/2013-12-12_rg_final_report.pdf [<https://perma.cc/7EQR-KKLP>].

11. *E.g.*, Julian Sanchez, *Can We Do Without National Security Letters*, JUST SECURITY (Jan. 9, 2014), <http://justsecurity.org/5351/national-security-letters> [<https://perma.cc/CK64-HMAE>].

that can be used to limit or terminate NSL nondisclosure orders,¹² and also created new ones.¹³

Notice for NSLs would reduce the risk of government overreach for this surveillance authority. NSLs have been abused in the past. Two reports by the Department of Justice's Inspector General found that "the FBI had used NSLs in violation of applicable NSL statutes, Attorney General Guidelines, and internal FBI policies."¹⁴ Wrongful behavior may have been even more common than was documented because possible additional violations were not "identified or reported" as required.¹⁵ There is also evidence that some NSL investigations may have been constitutionally unreasonable. An Inspector General report on Section 215 Orders, a related surveillance statute, confirmed that the FBI has used NSLs to conduct investigative actions after authorization for those same actions was denied by the Foreign Intelligence Surveillance Court (FISC).¹⁶ The FISC cited concern that individuals were being targeted "solely on the basis of activities protected by the First Amendment."¹⁷ No further review addressed that concern before the FBI conducted identical actions under its NSL authority.¹⁸ If the FISC's concerns were accurate, these investigations may have been both statutory violations and also

12. USA FREEDOM Act of 2015, H.R. 2048, 114th Cong. § 3511(b)(1)(A)–(B) (2015) [hereinafter USA Freedom Act]. This law codified a prior procedure for recipient initiated judicial review of NSL nondisclosure orders, which was originally developed by the Second Circuit. *See, e.g., Doe v. Mukasey*, 549 F.3d 861, 879 (2008) ("The Government could inform each NSL recipient that it should give the Government prompt notice, perhaps within ten days, in the event that the recipient wishes to contest the nondisclosure requirement. Upon receipt of such notice, the Government could be accorded a limited time, perhaps 30 days, to initiate a judicial review proceeding to maintain the nondisclosure requirement . . .").

13. The USA FREEDOM Act also added new procedures to limit the gags. The Attorney General is now required to review all NSL nondisclosure orders and lift them when the facts no longer support a need for secrecy. USA Freedom Act § 502(f). The Act also permits NSL recipients themselves to disclose general information about the letters they receive. USA Freedom Act § 604.

14. Doyle, *National Security Letters: Proposals*, *supra* note 2, at 7.

15. *Id.*

16. U.S. DEPT OF JUSTICE, INSPECTOR GEN., REPORT ON 215 ORDERS 73 (2006) ("[T]he FISA Court twice refused to authorize Section 215 orders based on concerns that the investigation was premised on protected First Amendment activity, and the FBI subsequently issued NSLs to obtain information . . . based on the same factual predicate without first reviewing the underlying investigation to ensure it did not violate the First Amendment caveat.").

17. *Id.* at 72 n.67.

18. OFFICE OF THE INSPECTOR GEN., A REVIEW OF THE FBI'S USE OF SECTION 215 ORDERS FOR BUSINESS RECORDS IN 2006, 73 (Mar. 2008). The same prohibition applies to NSL searches. § 2709.

constitutionally problematic. NSL investigations could also be unlawful if they were not properly authorized, or if they were not relevant “to protect against international terrorism or clandestine intelligence activities.”¹⁹

Notice would help to ensure accountability. A notice requirement would enable the targets of unreasonable or abusive NSL investigations to sue the government. Congress could codify an express cause of action and specify a remedy for NSL violations, as it has already done for improper surveillance that violates the Wiretap Act or the Stored Communications Act.²⁰ Even without statutory amendment, targets who were investigated solely as a result of their First Amendment protected activities may have constitutional claims that would enable them to file *Bivens* suits for civil damages.²¹ Depending on the type of information that the FBI sought to obtain using an NSL, targets might also have Fourth Amendment *Bivens* claims.²² Admittedly, the Supreme Court’s current “third party doctrine,” which retracts Fourth Amendment safeguards for information that individuals share with others,²³ would pose a substantial hurdle to those claims. But courts have begun to restrict the scope of the third party doctrine when the government collects particularly sensitive electronic data,²⁴ some of which can ostensibly be obtained with an NSL.²⁵

19. § 2709(b); *see also*, *Doe v. Ashcroft*, 334 F. Supp. 2d 471, 496 (S.D.N.Y. 2004) (stating reasons why an NSL might be “unreasonable or otherwise unlawful”).

20. 18 U.S.C. § 2520; § 2707(a).

21. In *Bivens v. Six Unknown Named Agents*, the Court recognized a private right of action against federal officers for money damages in cases of alleged constitutional violations. 403 U.S. 388 (1971). The Third, Ninth, and D.C. Circuits have recognized *Bivens* claims for First Amendment violations. *E.g.*, *David Bloem v. Unknown Department of the Interior Employees*, 920 F. Supp. 2d 154, 159-160 (D.D.C. 2013). A rich array of secondary literature addresses First Amendment concerns about NSLs. *E.g.*, Jack Balkin, *Old-School/New-School Speech Regulation*, 127 HARV. L. REV. 2296, 2329-41 (2014); Niel M. Richards, *Intellectual Privacy*, 87 TEX. L. REV. 387 (2008) (locating intellectual privacy within First Amendment theory).

22. One criticism of *Bivens* as an enforcement mechanism is that monetary damages from abusive searches are often so small that potential plaintiffs will not pursue them. But this is not always true. For instance, the plaintiffs in *Clapper v. Amnesty International* spent significant sums in an effort to avoid government collection of confidential communications. *Clapper v. Amnesty International USA*, 133 S. Ct. 1138, 1146 (2013) (“[R]espondents declare that they have undertaken ‘costly and burdensome measures’ to protect the confidentiality of sensitive communications.”).

23. *See, e.g.*, Orin S. Kerr, *The Case for the Third Party Doctrine*, 107 MICH. L. REV. 561, 563 (2009) (“The ‘third-party doctrine’ is the Fourth Amendment rule that . . . [b]y disclosing to a third party, the subject gives up all of his Fourth Amendment rights in the information revealed.”).

24. For an excellent overview of recent Fourth Amendment case law as applied to sensitive electronic data, *see In re Application for Telephone Information Needed for a Criminal Investigation*, 2015 WL 4594558, at *7-9 (N.D. Cal., Jul. 29, 2015).

Finally, notice to the public through stronger reporting requirements would add another layer of oversight. As the Snowden revelations revealed, when intelligence agencies interpret their surveillance authority in secret, they can stretch the scope of that authority far beyond what regular observers would expect from reading the law on the books.²⁶ Greater public transparency about the FBI's actual NSL practices would counteract this risk.

Part II of this Note offers a detailed proposal for implementing such a notice policy without risking harm or causing undue administrative burden. Part III considers likely counterarguments and turns to longstanding statutory precedents for guidance on how to address them. It concludes that notice should outweigh competing interests in secrecy, security, and confidentiality for NSL investigations because, in this particular context, there is a substantial risk of government abuse. Possible constitutional support for the notice proposal is examined in Part IV. This Part builds on Akhil Reed Amar's theory of the original Fourth Amendment to develop a series of constitutional arguments in favor of notice to NSL targets and the public. It also offers a structural claim that the relationship between the First and Fourth Amendments substantiates using limits on nondisclosure orders as standards to guide the government's notice obligations.

II. THE PROPOSAL: PROVIDING NOTICE OF NSL INVESTIGATIONS ONCE NONDISCLOSURE ORDERS END

Policy shifts over the last two years suggest that a substantial number of nondisclosure orders will presumably be lifted soon. Until recently, almost all National Security Letter nondisclosure orders stayed in effect

25. See *infra* notes 72–77 and accompanying text. For instance, the FBI claims the authority to collect cell-site location information using an NSL. *E.g.*, Info. Soc'y Project, *First Amendment Victory*, *supra* note 1 (“The FBI also claims authority to obtain cell-site location information with an NSL, which effectively turns a cell phone into a location tracking device. In court filings, the FBI said that at some point it stopped gathering location data as a matter of policy, but that it could secretly chose to resume the practice under existing authority.”). Some courts have held that this information is protected by the Fourth Amendment. *E.g.*, *In re Application for Telephone Information*, 2015 WL 4594558, at *9 (“[T]he Court finds that individuals have an expectation of privacy in the historical CSLI associated with their cell phones, and that such an expectation is one that society is willing to recognize as reasonable.”).

26. *E.g.*, Jonathan Manes, *Online Service Providers and Surveillance Law Transparency*, 125 YALE L.J. F. 343 (2016).

indefinitely.²⁷ Only a handful of recipients managed to even challenge their gag orders in court, including Mr. Merrill,²⁸ four Connecticut librarians,²⁹ the Internet Archive,³⁰ two Internet service providers whose identity remains under seal because they are still fighting their case,³¹ and—most recently—an unidentified plaintiff who filed a challenge in September of 2015.³² But this state of affairs is changing. On January 17, 2014, President Obama instructed the Attorney General “to amend how we use National Security Letters so that [their] secrecy will not be indefinite, and will terminate within a fixed time unless the government demonstrates a real need for further secrecy.”³³ In response, the FBI announced a new policy to “presumptively terminate” all NSL gag orders after three years.³⁴ The USA FREEDOM Act of 2015 added yet another layer of change; the law directs the Attorney General to review all NSL gags at regular intervals and to rescind them when the facts no longer require secrecy.³⁵

This Part proposes extending these policy shifts with two easy reforms to reduce secrecy about the government’s use of NSLs without causing harm or undue administrative burden. First, when an NSL nondisclosure order ends, the government should notify the target of that NSL investigation. Second, the government should publish more information

27. *E.g.*, *Doe v. Ashcroft*, 334 F. Supp. 2d 471, 502 (S.D.N.Y. 2004) (concluding until 2004, no NSL had ever been challenged in court).

28. *E.g.*, *My National Security Letter Gag Order*, WASH. POST (Mar. 23, 2007), <http://www.washingtonpost.com/wp-dyn/content/article/2007/03/22/AR2007032201882.html> [<https://perma.cc/MB8D-RF9C>].

29. *E.g.*, Alison Leigh Cowan, *Four Librarians Finally Break Silence in Records Case*, N.Y. TIMES (May 31, 2006), <http://www.nytimes.com/2006/05/31/nyregion/31library.html> [<https://perma.cc/66DH-AAL3>].

30. *E.g.*, Ryan Singel, *FBI Targets Internet Archive with Secret National Security Letter, Loses*, WIRED.COM (May 7, 2008), <http://www.wired.com/2008/05/internet-archiv> [<https://perma.cc/5T4G-HHHY>].

31. *E.g.*, Cindy Cohn & Kurt Opsahl, *Justice Delayed: Ninth Circuit Sends EFF’s NSL Cases Back for Consideration Under USA FREEDOM*, DEFENDING YOUR RIGHTS IN THE DIGITAL WORLD (Aug. 31, 2015), <https://www.eff.org/deeplinks/2015/08/justice-delayed-ninth-circuit-sends-effs-nsl-cases-back-consideration-under-usa> [<https://perma.cc/QV34-CFD5>].

32. *E.g.*, Tim Cushing, *Another NSL Challenge is Made Public; Court Decides Government can Keep Gag Order in Place Indefinitely*, TECHDIRT (Dec. 29, 2015), <https://www.techdirt.com/articles/20151229/10144633190/another-nsl-challenge-is-made-public-court-decides-government-can-keep-gag-order-place-indefinitely.shtml> [<https://perma.cc/L7CU-SPGS>].

33. Office of the Dir. of Nat’l Intelligence, *Strengthening Privacy & Civil Liberties Protections*, IC ON THE RECORD, <http://icontherecord.tumblr.com/ppd-28/2015/privacy-civil-liberties> [<https://perma.cc/8P87-M7DP>].

34. *Id.*

35. USA Freedom Act § 502(f).

about NSLs than it currently makes available for public review. Congress could implement this proposal by amending 18 U.S.C. § 2709 to incorporate a notice requirement and expanding the government's current reporting obligations under 50 U.S.C. § 1873. Alternately, the Executive should adopt these policies unilaterally.

Expanding notice to NSL targets and the public in these ways would not be harmful. Before the government can rescind an NSL nondisclosure order in full, it must make a prior decision that revealing information about that NSL would not cause harm. This prior decision should function like a form of collateral estoppel; it should bar re-reconsidering the issue of harm. In other words, once a gag order is lifted in full, the government should not be able to claim, at a later time or in a different context, that it has an interest in the secrecy of the same information that the gag once restricted. Even if nondisclosure orders are only lifted in part, as long as the recipient is no longer gagged with respect to the identity of the target, that person should be notified of whatever portion of the letter is no longer restricted. Of course, once a gag order ends, the NSL recipient is authorized to notify the target directly. My proposal merely places the burden of notifying the target on the government, rather than leaving it to the ad hoc decision of an NSL recipient.

Note that there are a number of different reasons why third-party speech about an NSL investigation might pose an insignificant risk of harm. Perhaps the information at issue is granular and yet, even so, permitting the third party to notify the target of an investigation directly would not be harmful.³⁶ In this case, the government should notify the target directly. Alternately, perhaps the information at issue is general and so permitting the third party to speak about that information would not actually reveal the specifics of any particular investigation. In that case, the government should publish the general information for public review.

These forms of notice would require nominal administration. Indeed, the government should have little difficulty contacting the target of an NSL investigation because NSLs themselves collect the target's name, address, and billing records.³⁷ Further, standards to guide how much notice to provide, and when, already exist. The USA FREEDOM Act of 2015 recently established a new set of procedures to rescind NSL

36. Courts have made highly granular determinations of harm in the past, and partially lifted NSL gags as a result. *See, e.g.*, *Doe v. Holder*, 703 F. Supp. 2d 313, 318 (S.D.N.Y. 2010) (“ORDERED that motion . . . to lift the nondisclosure requirement . . . is GRANTED in part and DENIED in part.”).

37. § 2709(b)(1).

nondisclosure orders. It requires various decision-makers, including the Attorney General and courts, to determine whether disclosure would risk harm. The statute also provides tests for decision-makers to apply when making their determinations. The same procedures and tests could easily govern notice as well.

To implement my first proposal, then, the government could tie notice to NSL targets directly to existing statutory procedures for rescinding individual nondisclosure orders. For instance, Congress has directed the Attorney General to review NSL gags “at appropriate intervals” and to terminate them “if the facts no longer support nondisclosure.”³⁸ This same review could prompt government notice to NSL targets. Likewise, the government could use the tests that courts apply when evaluating nondisclosure orders.³⁹ Congress has stipulated that when an NSL recipient initiates ex post judicial review of a gag order, the government must provide “a statement of specific facts indicating” that disclosure “may result in” an enumerated harm.⁴⁰ Note that a risk of harm other than those enumerated will not justify maintaining the gag.⁴¹ If a district court then finds no “reason to believe” the government’s assertions, the court has discretion to order the gag terminated.⁴² This was the reasoning that a district court applied to lift Mr. Merrill’s eleven-year nondisclosure mandate in full. Under a parallel notice procedure, the court’s decision would oblige the government to provide all of the information contained in that NSL, including the categories of information the FBI had sought to obtain, to the target of the investigation.

Courts may also apply a constitutional analysis to rescind nondisclosure orders that violate the First Amendment. When a court undertakes a First Amendment analysis of an NSL gag, it must evaluate whether the government has a compelling interest in the speech prohibition. Precisely what level of First Amendment scrutiny should apply to NSL nondisclosure requirements is presently disputed.⁴³

38. USA Freedom Act § 502(f).

39. 18 U.S.C. § 3511 (b)(3).

40. § 3511(b)(1)–(2).

41. In *Doe v. Holder*, Judge Victor Marrero lifted part of Mr. Merrill’s gag order because he was “not persuaded that disclosure of these two categories of information would raise a substantial risk that any of the statutorily enumerated harms would occur.” *Doe v. Holder*, 703 F. Supp. 2d 313, 316 (S.D.N.Y. 2010).

42. § 3511(b)(3).

43. See, e.g., Brief of Amici Curiae Floyd Abrams Institute for Freedom of Expression and First Amendment Scholars in Support of the Parties Under Seal at 25–28, Nat’l Sec. Letter, *Under Seal v. Holder* (Sealed), Nos. 13-15957, 13-16732 (9th Cir. Mar. 31, 2014), <http://cdn.ca9.uscourts.gov/datastore/general/2014/05/23/13-15957,13>

Regardless, a holding that an NSL gag is unconstitutional necessarily establishes that speech about that NSL poses an insufficient risk of harm to outweigh an individual liberty enshrined in the Bill of Rights. In that case, the government could follow the court's holding and disclose the NSL to its target.

My second proposal concerns a related though significantly more abstract issue: to what extent should information about the government's NSL practices be publically available? I suggest that when the government authorizes private parties to reveal general information about NSL usage, it should publish that same material for public review. In other words, existing limits on the scope of NSL nondisclosure orders—which have already been negotiated and thoroughly reviewed—can guide the government's public disclosures with minimal administrative burden.

As of June 2015, the Director of National Intelligence must publish an annual report online every April.⁴⁴ The report needs to contain both “the total number” of NSLs issued over the preceding calendar year and also the “number of requests for information” that those letters contained.⁴⁵ This disclosure is inadequate, I argue, because it provides less information to the public than NSL recipients may reveal ad hoc. The USA FREEDOM Act allows even those NSL recipients who remain under gag to publish a semiannual report of the approximate number of letters they have received, and roughly how many of their customers were targeted, during the preceding 180 days.⁴⁶ This authorization to report certain controlled categories of information was not new. The Attorney General had previously permitted a select group of Internet service providers to publish “transparency reports” that described—in vastly general terms—the quantity of NSLs they had received and the number of their customer

-16731Floyd.pdf [https://perma.cc/KN8M-V4JT] (arguing that the more stringent tests in Pentagon Papers or Nebraska Press should apply); Balkin, *Old-School*, *supra* note 21, at 2334–35 (explaining that the Second Circuit applied the Freedman standard in one NSL case).

44. Charles Doyle, *National Security Letters in Foreign Intelligence Investigations: Legal Background*, CONG. RESEARCH SERV., RL33320, at 23 (July 30, 2015) [hereinafter Doyle, *Legal Background*] (noting that this reporting requirement was imposed by the USA FREEDOM Act).

45. 50 U.S.C. § 1876(b)(6), (c) (2012)

46. 50 U.S.C. § 1874(a)(b). Section 604 of USA Freedom codified the executive branch's authorization for recipients to report the number of NSLs they have received in large aggregated bands of one thousand, starting with 0–999. Because the authorized disclosure includes “0” in the lowest reporting band, it effectively bans smaller providers from disclosing that they have received at least one letter. *See* USA Freedom Act § 604.

accounts that were targeted.⁴⁷ The USA FREEDOM ACT codified the agreement and extended it to all NSL recipients.⁴⁸

To authorize these disclosures, then, both Congress and the Executive must have made a prior decision that revealing this particular information will not cause harm. Even so, the government has yet to make that same information publicly available. Private transparency reports—narrow as they may be—provide more information than the government’s annual disclosure. The bulk “transparency report” permissions ameliorate First Amendment concerns by loosening NSL nondisclosure orders, but don’t actually guarantee parallel transparency relief. In short, select private parties are currently able to make unilateral decisions about the secrecy of government surveillance practices. By negotiating these guidelines but failing to disclose the same, the government has empowered Internet service providers to act as discretionary filters on the public’s ability to know about government surveillance practices.

To resolve this incongruence, the government should increase the transparency level of its own reports to match that permitted private parties. There are various ways to accomplish this goal. The government could publish semiannual reports documenting in general terms how many NSLs it issued to various Internet service providers over the prior 180 days. These reports could use the same broad numeric ranges of 0–999 and 1000–1999 that NSL recipients may currently disclose.⁴⁹ Of course, one problem with this reporting scheme might be that when the government first adds an Internet service provider to its list of recipients, it would implicitly reveal that this particular entity received its first NSL during the prior 180 day period.⁵⁰ If the government were concerned that revealing this information would risk harm, it could redact the name of that entity until a subsequent reporting cycle. Alternately, the government could publish some information about each individual NSL every time a gag is

47. *Letter from James M. Cole, Deputy Attorney Gen., to Colin Stretch et al., Gen. Counsels of Tech. Cos.*, WASH. POST (Jan. 17, 2014), <http://www.washingtonpost.com/r/2010-2019/WashingtonPost/2014/10/07/National-Security/Graphics/dagletter.pdf> [<https://perma.cc/23Y7-MNXX>]. For an overview of advocacy efforts by NSL gag recipients, see Rebecca Wexler, *Warrant Canaries and Disclosure by Design: The Real Threat to National Security Letter Gag Orders*, 125 YALE L.J. F. 158 (2014).

48. 50 U.S.C. § 1874; USA Freedom Act § 604.

49. *Cf.* USA Freedom Act § 604.

50. Twitter is currently fighting a lawsuit that addresses a similar concern. *Twitter v. Lynch*, 2015 WL 5970295, at *1 (Oct. 14, 2015) (ordering Twitter to file an amended complaint in light of the USA Freedom Act).

lifted.⁵¹ Indeed, the government has disclosed similar information in the past that has proven useful. For instance, in response to an ACLU Freedom of Information Act request, the government released a list of every NSL it issued between 2001 and 2003. While the substance of the list was completely redacted, the number of entries proved to be a valuable piece of information and even helped a federal district court to evaluate the constitutionality of the NSL statute.⁵²

III. STATUTORY PRECEDENTS: BALANCING NOTICE WITH COMPETING INTERESTS IN SECRECY, SECURITY, AND CONFIDENTIALITY

Three primary issues complicate the goal of expanding notice to NSL targets and the public: concerns about interfering with covert surveillance or national security investigations; concerns about unanticipated exigent circumstances; and concerns about the reputational interests of third parties who aid an investigation. This Part draws on existing statutes that offer established solutions to balance these competing interests. Specifically, a number of statutes require the government to provide notice to the targets of some types of law enforcement investigations, regardless of whether they are ever charged with a crime.⁵³ Where these types of investigations are analogous to National Security Letters, the statutes that govern them offer useful guidance. When the government has no need to keep information about an NSL secret, I contend, the target's interest in knowing that specific information should outweigh competing concerns. And when the government has no manifest need for secrecy, the public's interest in knowing general information about the FBI's NSL practices should also prevail.

Providing notice of NSL investigations in the manner detailed in Part II would produce a negligible risk to covert surveillance or national security investigations; it only requires notice after the government

51. The author thanks Jonathan Manes for proposing that the government should publicly report information about each NSL once a gag order is lifted.

52. *Doe v. Ashcroft*, 334 F. Supp. 2d 471, 502 (S.D.N.Y. 2004) (“Although the entire substance of the document is redacted, it is apparent that hundreds of NSL requests were made during that period.”).

53. Patrick Toomey & Brett Max Kaufman, *The Notice Paradox: Secret Surveillance, Criminal Defendants, & the Right to Notice*, 54 SANTA CLARA L. REV. 843, 853 nn.29, 52 (2014) [hereinafter Toomey & Kaufman, *The Notice Paradox*], (citing Federal Rule of Criminal Procedure 41(f); its predecessors, Rule 41(d) and 18 U.S.C. § 622 (1928); 18 U.S.C. § 3103(a)(b) (2012); 18 U.S.C. § 2518(8)(d) (2012); 18 U.S.C. § 2518(8)(d) (2012); 18 U.S.C. § 3109 (2012); 18 U.S.C. § 3406(b)).

determines that it has no compelling interest in keeping that particular information secret.⁵⁴ Even so, it is worth noting that some existing statutes already enable notice to the targets of investigations in both surveillance and national security contexts. After the government uses certain covert monitoring techniques—including an electronic tracking device⁵⁵—it must notify the target whose person or property was tracked.⁵⁶ Judges reviewing applications for wiretaps also have discretion to require that the government provide notice to “the persons named in the order or the application.”⁵⁷ And the Foreign Intelligence Surveillance Act authorizes covert physical searches of U.S. residences for national security intelligence investigations,⁵⁸ but requires that once the government’s interests in the secrecy of those searches abate, the Attorney General must notify the person whose home was searched.⁵⁹

To be sure, not every harm can be anticipated. Particularly relevant then, a number of existing statutes permit delayed notice in exigent circumstances. Traditional search warrants generally require notice before the police physically enter a residence.⁶⁰ But this notice may be temporarily delayed if providing it immediately would result in one of five enumerated harms, including “flight from prosecution” or “intimidation of potential witnesses.”⁶¹ Notice may also be delayed if the warrant is particularly restrictive and prohibits “the seizure of any tangible property, any wire or electronic communication . . . [or] any stored wire or electronic information”⁶² The Stored Communications Act (SCA) requires notice when the government collects the contents of online communications without a warrant.⁶³ This notice too can be temporary

54. See *supra* Part II.

55. 18 U.S.C. § 3117(b).

56. FED. R. CRIM. P. 41(f)(2)(C) (requiring the law enforcement agent who executes the warrant to provide a copy of the warrant itself to the target).

57. 18 U.S.C. § 2518(8)(d).

58. 50 U.S.C. § 1824.

59. 50 U.S.C. § 1825(b).

60. See, e.g., Wayne R. LaFare, *Manner of Entry for Which Notice Ordinarily Required*, in 2 SEARCH & SEIZURE: A TREATISE ON THE FOURTH AMENDMENT § 4.8(b) (5th ed. 2015); see also, TELFORD TAYLOR, TWO STUDIES IN CONSTITUTIONAL INTERPRETATION: SEARCH, SEIZURE, AND SURVEILLANCE AND FAIR TRIAL AND FREE PRESS 80–81 (1969) (“a search warrant, although initially issued *ex parte*, becomes known to the individual, whose person or premises are the field of the search, as soon as the warrant is executed”).

61. 18 U.S.C. § 3103a(b)(1); 18 U.S.C. § 2705(B)(D).

62. § 3103a(b)(2).

63. 18 U.S.C. § 2703(b). NSLs are similar to SCA orders in that they all collect information from third party Internet service providers. They are also related via

delayed if—but only if—there is “reason to believe”⁶⁴ it might cause one of five enumerated harms, such as endangering “life or physical safety”⁶⁵ or “jeopardizing an investigation.”⁶⁶ An un-enumerated risk of harm does not justify delayed notice. NSL notice requirements could contain similar exigent circumstances provisions.

More complicated concerns arise from the potential reputational interests of third parties who cooperate with law enforcement investigations. For example, when the government demands customer information from Internet service providers, these service providers may prefer to keep that fact confidential. Even if they are legally compelled to assist law enforcement investigations—perhaps by an NSL—companies might not want their cooperation publicized. Though challenging, these types of concerns are also familiar. Some existing statutes require notice even when it might impinge on a third party’s business or confidentiality interests. For instance, when the IRS demands information from a third party record keeper as part of a tax liability investigation, it must notify the person whose information it collects.⁶⁷ There are also statutory precedents for favoring a target’s interest in notice over third party business interests when the government collects information from an Internet service provider. The SCA, mentioned above, governs these types of collections in domestic criminal investigations and requires notice to targets in certain circumstances.⁶⁸ Specifically, unless a judge reviews the government’s demand *ex ante*, the target usually has a right to notice under the SCA regardless of an Internet service provider’s confidentiality concerns.

The mere fact that NSL investigations involve third party service providers, then, should not prevent notice to NSL targets and the public. That said, current laws do not always require notice. One way to explain when they do and do not is to focus on the level of judicial oversight for the government’s activities. For the most part, the SCA obliges the government to choose between providing notice to the target and obtaining some form of prior judicial approval. For instance, unless the government obtains a full warrant, which requires probable cause, it must

legislative history. Indeed, the statute authorizing the FBI to issue NSLs to Internet service providers was enacted as part of the SCA. 18 U.S.C. § 2709.

64. 18 U.S.C. § 2705(a)(1).

65. § 2705(a)(2)(A).

66. § 2705(a)(2)(E).

67. 26 U.S.C. § 7602(c).

68. 18 U.S.C. §§ 2701–2711; *see also*, Orin S. Kerr, *A User’s Guide to the Stored Communications Act, and a Legislator’s Guide to Amending it*, 72 GEO. WASH. L. REV. 1208, 1223 (2004) (chart summarizing when the SCA requires notice).

notify the customer of an Internet service when it demands the contents of her communications.⁶⁹ Less process is required to collect non-content information, such as customer transactional records. Even so, unless the government obtains some type of court order, it usually needs the customer's consent before it can access his information.⁷⁰ The single exception is that collecting limited categories of basic subscriber information requires a mere administrative subpoena and no notice.⁷¹ Collecting non-content telephone logs under the Pen Register statute does not require notice either.⁷² Another way to parse these requirements, then, is to focus on the type of information the government collects: collecting the contents of communications generally require notices while collecting basic subscriber information generally does not.

Synthesizing these rules, a guiding principle for notice requirements seems to be reducing the risk of government abuse. When the risk of abuse is lower—as with warrants that expressly prohibit the seizure of any electronic communications—notice is less urgent. In those circumstances, a third party's confidentiality or business interests might prevail. When the risk of abuse is higher—as when the government uses an SCA order to collect sensitive transactional records without prior judicial oversight—notice is more urgent and should outweigh competing concerns.

How should these principles apply to NSLs, which currently require neither prior judicial approval nor notice? On the one hand, the statutes that govern NSLs supposedly restrict their use to the collection of lower-risk information. On its face, § 2709 vests the FBI with the authority to seek only limited materials about the customer of an Internet service provider, including her “name, address, length of service, and local and long distance toll billing records.”⁷³ Yet, on the other hand, the secrecy surrounding the letters has made it difficult to determine precisely what information the FBI collects—or believes it has authority to collect—under the statute. The FBI may secretly interpret its authority to have far broader scope than a plain reading of § 2709 at first suggests. For instance, the NSL that Mr. Merrill received demanded a surprising array of sensitive information about his customer, including URLs that could potentially expose the target's entire web browsing history, lists of every person the target had contacted over telephone, Skype, Facebook, instant

69. 18 U.S.C. § 2703(b).

70. § 2703(c)(1).

71. § 2703(c)(2)–(3).

72. 18 U.S.C. §§ 3121–3127.

73. 18 U.S.C. § 2709(b)(1).

messaging or email, and records of all the target's online purchases.⁷⁴ Especially alarming, the FBI demanded cell-site location information that could be used to monitor the target's physical location.⁷⁵ Cell-site location information can effectively transform a mobile phone into an electronic tracking device.⁷⁶ Courts across the country have considered what protections the Fourth Amendment requires before the government can collect precisely this information.⁷⁷ Many have held that a full warrant—or at least a court order—is needed.⁷⁸ More generally, changes in technology mean that the distinction between content and non-content may no longer identify which types of information are particularly sensitive for the government to collect.⁷⁹

The risk of government overreach or abuse of its NSL authority is arguably high. We know that such abuses have occurred in the recent past,⁸⁰ and ongoing secrecy around the FBI's actual NSL practices exacerbates this risk. As a result, interests in NSL notice should be afforded substantial weight. In contrast, third party business interests are hardly the sole confidentiality concerns associated with NSL investigations; the target has privacy interests—potentially constitutional ones—at stake as well. When balancing these interests, then, it is both reasonable and urgent to resolve the competing concerns in favor of notice to NSL targets and the public.

74. *E.g.*, Unredacted Attachment to 2004, NSL *Merrill*, 2015 WL 9450650, <https://yale.app.box.com/NSL-Attachment-Unredacted> [<https://perma.cc/JMH4-ELY8>].

75. *E.g.*, *Merrill*, 2015 WL 9450650, at *7 (noting a document authorizing the FBI to use an NSL to collect “cellular site and sector information”); *see also* Info. Soc’y Project, *First Amendment Victory*, *supra* note 1 (explaining that the FBI claimed the authority to collect cell-site location information in court filings in Mr. Merrill’s case).

76. *E.g.*, *In re Application for Telephone Information*, 2015 WL 4594558, at *1–3 (detailing CSLI technologies). Recall that the government must generally provide notice when it uses an electronic tracking device. FED. R. CRIM. P. 41(f)(2)(C).

77. Courts in California, Florida, Massachusetts, and New Jersey have found a “reasonable expectation of privacy” in cell-site location information. *In re Application for Telephone Information*, 2015 WL 4594558, at *11.

78. *Id.*

79. Chris Conley has provided an excellent overview of how changing technologies have blurred the line between content and non-content electronic information, and how these changes are affecting privacy laws. Chris Conley, *Non-Content is Not Non-Sensitive: Moving Beyond the Content/Non-Content Distinction*, 54 SANTA CLARA L. REV. 821, 829–39 (2014).

80. *See, e.g.*, Doyle, *A Glimpse*, *supra* note 7, at 3–4.

IV. CONSTITUTIONAL PRINCIPLES: FOURTH AMENDMENT SUPPORT FOR NOTICE AND FIRST AMENDMENT GUIDES FOR DISCLOSURE

This Part considers possible constitutional support for reducing NSL secrecy through notice. It explores how a policy to notify the targets of NSL investigations, and to increase the government's public reporting of its actual NSL practices, would further core principles that underlie the Fourth Amendment. To do so, it turns to established theoretical claims about the original understanding of the Fourth Amendment. Specifically, it draws on two key claims in Akhil Reed Amar's theory of the Fourth Amendment: First, a primary motive for drafting the Fourth Amendment was the agency problem of how to police government abuse. Second, the drafters sought to solve this problem by empowering non-governmental actors to enforce Fourth Amendment violations.⁸¹ Requiring notice to the targets of NSL investigations would further these principles by empowering targets to sue for civil damages. These suits—by private individuals not government agents—would mitigate a demonstrated risk of government abuse. Likewise, if the government were to report more information about the FBI's actual NSL practices, it would empower the public to oversee government surveillance more directly. Finally, I suggest that First Amendment balancing tests to weigh the government's interests in secrecy against an individual's right to free speech should provide standards to guide any Fourth Amendment notice obligations.

Some legal scholars and practitioners recently began to rehabilitate an under-theorized and largely marginalized idea that the Fourth Amendment implies a right to notice. Also building on Amar's originalist framework, Nola Breglio argued over a decade ago that to facilitate *Bivens* claims for foreign intelligence surveillance, the Fourth Amendment should "require law enforcement authorities to provide notice to every surveillance target once the investigation is complete."⁸² Similar arguments gained traction after *Clapper v. Amnesty International*, in which the Supreme Court denied standing to plaintiffs seeking to raise a Fourth Amendment challenge to a mass surveillance statute because they had "no actual knowledge" of whether the government had collected their particular

81. Akhil Reed Amar, *Fourth Amendment First Principles*, 107 HARV. L. REV. 757, 176 (1994).

82. Nola K. Breglio, Note, *Leaving FISA Behind: The Need to Return to Warrantless Foreign Intelligence Surveillance*, 113 YALE L.J. 179, 213 (2003).

communications.⁸³ These arguments for Fourth Amendment notice draw support from common law, longstanding practice codified in statutes, and recent case law.⁸⁴ Patrick Toomey and Brett Max Kaufman point out that notice is and has always been an “inevitable consequence” of most physical searches because the searches themselves are likely to disrupt a person or object in a readily apparent manner.⁸⁵ In other words, search targets notice when law enforcement officers stop and frisk them on the street, fingerprint them, force them to walk through a metal detector, or seize their belongings.⁸⁶ Toomey and Kaufman also observe that law enforcement officers at common law were required to “knock-and-announce” before conducting a search. Not only has the knock-and-announce rule been codified,⁸⁷ but the Court has also incorporated it into its Fourth Amendment reasonableness analysis.⁸⁸ Jonathan Witmer-Rich offers evidence that covert searches without simultaneous notice were either rare or non-existent when the Fourth Amendment was drafted.⁸⁹ As a result, he argues that courts have wrongly separated the Fourth Amendment knock-and-announce requirement from case law that governs delayed notice warrants.⁹⁰ Witmer-Rich claims that the Fourth

83. *Clapper*, 133 S. Ct. at 1143, 1147, 1148 (“[R]espondents cannot manufacture standing by choosing to make expenditures based on hypothetical future harm that is not certainly impending.”).

84. *See, e.g.*, Toomey & Kaufman, *The Notice Paradox*, *supra* note 53, at 851, 900 (arguing that “the right to notice is implicit in the right to privacy guaranteed by the Fourth Amendment,” and that courts must vigilantly police this right); Jonathan Witmer-Rich, *The Rapid Rise of Delayed Notice Searches, and the Fourth Amendment “Rule Requiring Notice,”* 41 PEPP. L. REV. 509, 571 (2014) (“[T]he argument that covert searching raises Fourth Amendment concerns does not rest only on a generalized analysis of privacy interests, but also on the common law—and constitutional—“rule requiring notice.”).

85. Toomey & Kaufman, *The Notice Paradox*, *supra* note 53, at 852 (citing *United States v. Chadwick*, 433 U.S. 1 (1977); *Camara v. Mun. Court of S.F.*, 387 U.S. 523 (1967); *Michigan v. Tyler*, 436 U.S. 499 (1978); *Berger v. New York*, 388 U.S. 41 (1967); *United States v. Freitas*, 800 F.2d 1451 (9th Cir. 1986)).

86. For examples of warrantless searches that take place throughout everyday life and that also trigger built-in structural notice, see Amar, *First Principles*, *supra* note 81, at 176 (“consider[ing] the vast number of real-life, unintrusive, non-discriminatory searches and seizures to which modern day Americans are routinely subjected: metal detectors at airports, annual auto emissions tests, inspections of closely regulated industries, public school regimens, border searches, and on and on”).

87. *See, e.g.*, 18 U.S.C. § 3109 (2012).

88. Toomey & Kaufman, *The Notice Paradox*, *supra* note 53, at 852–53 n.25 (citing *Wilson v. Arkansas*, 541 U.S. 927, 931–37 (1995)).

89. Witmer-Rich, *The Rapid Rise*, *supra* note 84, at 561–70.

90. *Id.*, at 571 (citing *Richards v. Wisconsin*, 520 U.S. 385, 387–88 (1997)). An example of the divide that Witmer-Rich identifies can be found in the Court’s

Amendment must require both practices, not just one, because these are logically continuous phenomena; a delayed notice search, he argues, is “simply a more extreme version of a no-knock search.”⁹¹

Nor is the idea of a Fourth Amendment notice right completely foreign to courts. Case law concerning the constitutional status of statutory notice practices is under-developed but has at times suggested that the right to notice might have constitutional underpinnings. Toomey and Kaufman offer an excellent collation and analysis of this doctrine.⁹² They point out that the Supreme Court has sometimes incorporated the adequacy of notice requirements into its Fourth Amendment analysis. In *Berger*, for instance, the Court invalidated an electronic eavesdropping statute under the Fourth Amendment in part because it provided “no requirement for notice.”⁹³ *Katz* found that delayed notice to accommodate exigent circumstances is constitutional.⁹⁴ The Sixth Circuit later recognized a “Fourth Amendment requirement of notice” for wiretaps.⁹⁵ And when the Supreme Court itself considered wiretaps in *Donovan*, it observed that Congress may have intended the notice provisions of the Wiretap Act to satisfy Fourth Amendment requirements. Justice Powell quoted Senator Hart, who introduced the notice provisions on the floor of the Senate: “notice of surveillance is a constitutional requirement of any surveillance statute.”⁹⁶

Yet, the scarcity of these theoretical and judicial gestures creates an opportunity to scrutinize the issue using other sources of authority. This Part turns to originalist understandings of the Fourth Amendment that can shed new light on a constitutional right to notice. It first considers

observation in *Katz* that the exigent circumstances exception to the notice requirement addresses considerations such as shock, fright or embarrassment that “are not relevant to the problems presented by judicially authorized electronic surveillance.” *Katz v. United States*, 389 U.S. 347, n.16 (1967).

91. Witmer-Rich, *The Rapid Rise*, *supra* note 84, at 576.

92. Toomey & Kaufman, *The Notice Paradox*, *supra* note 53, at 853, nn.31, 34, 36–39, 41 (citing *Berger v. New York*, 388 U.S. 41, 60 (1967); *United States v. Villegas*, 899 F.2d 1324, 1337 (2d Cir. 1990); *United States v. Freitas*, 800 F.2d 1451, 1456 (9th Cir. 1986); *Nordelli v. United States*, 24 F.2d 665, 666 (9th Cir. 1928); *United States v. Simons*, 206 F.3d 392 (4th Cir. 2000); *Dalia v. United States*, 441 U.S. 238 (1979); *United States v. Johns*, 948 F.2d 599 (9th Cir. 1991)).

93. *Berger v. New York*, 388 U.S. 41, 60 (1967); *see also*, Susan Freiwald, *Online Surveillance: Remembering the Lessons of the Wiretap Act*, 56 ALA. L. REV. 9, n.100 (2004).

94. *Katz*, 389 U.S. at n.16.

95. *United States v. Martinez*, 498 F.2d 464, 468 (1974).

96. Senator Hart in turn cited *Berger* and *Katz* as the authority for this Fourth Amendment interpretation. *United States v. Donovan*, 429 U.S. 413, 430 (1977) (quoting 114 CONG. REC. S14485–86 (1968) (statement of Sen. Hart)).

how the drafter's intent relates to the Article III standing issues in *Clapper*. Next, it examines how republican theories of collective Fourth Amendment rights can justify some level of public transparency for surveillance practices. Finally, it offers a novel structural claim that First Amendment limits on the government's power to censor third parties who are conscripted into a search—such as the Internet service providers who receive NSL nondisclosure orders—can guide what, when, and how much notice the Fourth Amendment may require.

A. NOTICE TO THE TARGET PROTECTS THE INDIVIDUAL RIGHT TO CONTEST FOURTH AMENDMENT REASONABLENESS IN JURY TRIALS

This Section draws on Amar's "civil-enforcement model"⁹⁷ of the Fourth Amendment to suggest that providing notice of NSL investigations to their individual targets would further the drafters' original intent. Amar argues that the agency issue of how to police government abuse was a key motivator for the drafters of the amendment. To solve this problem, the drafters deliberately sought to empower non-governmental actors to enforce Fourth Amendment violations, rather than entrust the task to any branch of the federal government.⁹⁸ This reading implies that establishing civil enforcement mechanisms, such as private causes of action, advances the drafters' original intent. Moreover, I would add, this is particularly true when there is a high risk that the government will abuse its search and seizure powers. Using prudential, textual, historical, and ethical modalities as well as executive precedent, I consider whether this original preference for private enforcement implies a right to notice for the targets of investigations.

1. *Non-Governmental Actors Should Police Government Abuse*

Amar's theory of the original Fourth Amendment contains two key claims for purposes of the NSL notice proposals discussed above: first, the drafters were concerned with how to police government abuse; and second, they sought to empower non-governmental actors to enforce Fourth Amendment violations. The following overview summarizes the textual and historical details that sustain this reading of original intent.⁹⁹

97. Amar, *First Principles*, *supra* note 81, at 811–14.

98. *Id.* at 176.

99. *E.g.*, AKHIL REED AMAR, THE BILL OF RIGHTS: CREATION AND RECONSTRUCTION 73, 96 (1998) [hereinafter AMAR, BILL OF RIGHTS] ("If we seek a paradigmatic image underlying the original Bill of Rights, we cannot go far wrong in picking the jury. Not only was it featured in three separate amendments (the Fifth, Sixth,

Amar joins Telford Taylor in theorizing that the drafters were particularly concerned over the due process limitations of warrant proceedings. The drafters sought to limit warrants, the argument goes, because warrants perfectly executed could indemnify law enforcement officers from post-intrusion civil suits.¹⁰⁰ At common law, judges or magistrates would issue warrants for stolen goods in secret *ex parte* proceedings that authorized surprise searches.¹⁰¹ To compensate for the due process limitations of these proceedings, the search target could adjudicate the legality of the intrusion after the fact in an adversarial proceeding.¹⁰² Warrants were a problem because they could reduce the target's capacity to mount a convincing case. As a result, any errors that resulted from the covert, non-adversarial nature of the initial warrant proceedings would become even more difficult to correct.¹⁰³

Amar finds textual support for these claims by reading the two clauses of the amendment in relation to one another. The second clause of the Fourth Amendment, which directs that “no Warrants shall issue, but upon probable cause,”¹⁰⁴ limits warrants by conditioning their issuance on *ex*

and Seventh), but its absence strongly influenced the judge-restricting doctrines underlying three other amendments (the First, Fourth, and Eighth.”); *see also* AKHIL REED AMAR, *AMERICA'S UNWRITTEN CONSTITUTION* [hereinafter *AMAR, AMERICA'S UNWRITTEN*] 174 (2012); AKHIL REED AMAR, *AMERICA'S CONSTITUTION: A BIOGRAPHY* 326 (2005) [hereinafter *AMAR, A BIOGRAPHY*].

100. Amar documents that, at the founding, the “aggrieved target” of an unreasonable search could sue the government official for damages under the common law of trespass. *AMAR, BILL OF RIGHTS*, *supra* note 99, at 69. If the official had obtained and complied with a legal warrant, whether particular or general, he could raise it in defense. Amar concludes that the warrant would compel a form of “declaratory judgment” or “directed verdict” for the defense. Not only would the warrant effectively have barred the trespass cause of action, but also and critically for this discussion, a directed verdict would rob the determination of reasonableness from the jury decision-maker. *Id.*

101. *Id.* at 72.

102. TAYLOR, *TWO STUDIES IN CONSTITUTIONAL INTERPRETATION*, *supra* note 60, at 82 (arguing that search warrants “are ancillary *ex parte* preludes to confrontation and controversy.”); *AMAR, BILL OF RIGHTS*, *supra* note 99, at 69–70 (arguing that warrants were disfavored by the drafters of the Fourth Amendment because “even when issued by a judge—and in some places executive magistrates also claimed authority to issue warrants—warrants lacked many traditional safeguards of judicial process: notice, adversarial presentation of the issues, publicity, and so on”).

103. *Cf.* AKHIL REED AMAR, *THE LAW OF THE LAND: A GRAND TOUR OF OUR CONSTITUTIONAL REPUBLIC* 336 (2015) (documenting how the framers “prefer[ed] open tort suits to secret warrants”).

104. U.S. CONST. amend. IV (“no warrants shall issue, but upon probable cause, supported by oath or affirmation, and particularly describing the place to be searched, and the persons or things to be seized”).

ante showings of probable cause and particularity. The result prohibits vague or general warrants in keeping with an antipathy towards them that many people recognize was both common at the time of the drafting and clear in the intent of the drafters. (State constitutions of the era incorporated similar restrictions on general warrants, and discussions of the First Congress also cautioned against general warrants.¹⁰⁵) But, Amar contends, the relative weight of the clauses shows that the drafters' hostility towards warrants reached beyond the general to condemn *all* warrants. To be sure, recent doctrine has expanded the probable cause and particularity requirements of the warrant clause to a wide variety of searches,¹⁰⁶ and deemed those without to be per se unreasonable unless they fall into one of a series of judge-made exceptions,¹⁰⁷ such as customs and border searches,¹⁰⁸ roadblock seizures,¹⁰⁹ and DNA cheek swabs.¹¹⁰ In short, twentieth century case law begins with the warrant clause and defines the reasonableness clause in relation to it. But the text that the drafters begot and bestowed begins with reasonableness and only subsequently restricts the subset of warranted searches. The reasonableness

105. An early precursor to the federal Fourth Amendment, Article XIV of the 1780 Massachusetts Constitution stated that all warrants lacking an oath, affirmation, or "special designation of the persons or objects of a search, arrest, or seizure" were per se unreasonable and impermissible. Similar prohibitions on general warrants appeared in the New Hampshire Constitution of 1784, and in proposals for federal amendments from constitutional conventions in Virginia, New York, and North Carolina. And James Madison cautioned against general warrants when introducing the Fourth Amendment to the First Congress. AMAR, LAW OF THE LAND, *supra* note 103, at 343–44.

106. *See, e.g.*, Illinois v. Gates, 462 U.S. 213, 230–31 (1983) (adopting a totality-of-the-circumstances test for probable cause and overruling the *Spinelli* and *Aguilar* tests); *Spinelli v. United States*, 393 U.S. 410, 587 (affirming the *Aguilar* test for probable cause); *Aguilar v. Texas*, 378 U.S. 108, 115 (1964) (requiring knowledge of underlying circumstances showing that information is "credible" or "reliable" in order to establish probable cause).

107. *See* Amar, *First Principles*, *supra* note 81, at 757 (Amar describes the Court's Twentieth Century doctrine as: "Warrants are not required—unless they are. All searches and seizures must be grounded in probable cause—but not on Tuesdays.")

108. *See, e.g.*, *United States v. Martinez-Fuerte*, 428 U.S. 543, 562 (1976) (relying on reasonableness, rather than a warrant and probable cause requirement, to permit suspicion-less stops of vehicles traveling over the Mexican border).

109. *See, e.g.*, *Michigan Dep't of State Police v. Sitz*, 496 U.S. 444, 450–52 (1990) (applying a mere reasonableness test to permit a twenty-five second delay at roadblocks).

110. *Maryland v. King*, 133 S. Ct. 1958, 1980 (2013) (applying a mere Fourth Amendment reasonableness requirement, and not a warrant requirement, to DNA cheek swabs, because the intrusion is negligible when weighed against countervailing government interests).

standard thus applies to all searches, while the second clause limits merely the quantum that satisfies the preconditions for a warrant.¹¹¹

This reading of the amendment incorporates an individualistic solution to the agency problem of policing government abuse: civil-jury enforcement. At common law, *ex post* adversarial proceedings to review the legality of a search often treated the reasonableness of that search as a factual issue for a civil jury. Thus another reason that the drafters worried about warrants, Amar argues, is that a lawful warrant defense could compel a directed verdict of reasonableness that robbed even the subsequent adversarial proceeding of jury review, and the jury in turn of its decision-making power.¹¹² This theft of decision-making authority from jury to judge mattered because it risked the excesses of a self-interested government self-policing. By ensuring jury supremacy, the drafters could also ensure that the final arbiters of Fourth Amendment reasonableness would not be on the government payroll.

The text of the amendment also suggests a republican justification for this “civil-enforcement model.” According to Amar, the drafters sought to empower “the people”—conceived as a collective entity—to oversee Fourth Amendment abuses. The first clause of the Fourth Amendment protects the “right of the people to be secure in their persons, houses, papers, and effects, against unreasonable searches and seizures”¹¹³ Amar points out that the words “persons, houses, papers, and effects” reference intimate materials of the most private sphere, a sphere apart from government that merits protection from its unreasonable intrusion.¹¹⁴ But, in seeming contradiction, the text does not bestow the right “to be secure” in these facets of the personal on the individuals who embody, inhabit, or produce them. Instead, it bestows this right on “the people,” a public entity in every sense of the term. The text thus entrusts the security of the most private sphere to its most public counterpart. How can a public collective entity possess rights to the security of an individual’s

111. Akhil Reed Amar, *Terry and Fourth Amendment First Principles*, 72 ST. JOHN’S L. REV. 1097, 1098 (1998) (“[T]he Fourth Amendment means what it says and says what it means: All searches and seizures must be reasonable. Reasonableness—not the warrant, not probable cause—thus emerged as the central Fourth Amendment mandate and touchstone.”).

112. Amar, BILL OF RIGHTS, *supra* note 99, at 69–71 (“A warrant issued by a judge or magistrate—a permanent government official, on the government payroll—has had the effect of taking a later trespass action away from a jury of ordinary citizens.”).

113. U.S. CONST. amend. IV.

114. AMAR, A BIOGRAPHY, *supra* note 99, at 326 (“Clearly, this amendment seemed to center on the domain of domesticity—on ‘persons’ in their private ‘houses’ as distinct from the people in the public square.”).

person or property? Amar resolves this puzzle by reading “the people” as an enforcement mechanism. More specifically, “the people” acting in symbolic collective as juries offer a means for enforcement other than by the paid officials of the very government whose intrusions the amendment protects against.¹¹⁵ Jury enforcement of the Fourth Amendment harmonizes both with the public entity of “the people” and with the republican nature of the Bill of Rights as a whole. Juries are populist, representing the sovereign electorate, and federalist, comprising local decision-makers.¹¹⁶ “[T]he central role of the jury in the Fourth Amendment should remind us,” Amar urges, “that the core rights of ‘the people’ were popular and populist rights—rights that the popular body of the jury was well suited to vindicate.”¹¹⁷

To summarize, the drafters sought to limit warrants in order to enable local juries to determine the reasonableness of searches. Civil juries representing the sovereign people—not salaried judges or executive officials who might act in the interests of government—would enforce the Fourth Amendment.¹¹⁸

2. *The Right to Contest Implies an Individual Right to Notice*

This Section draws on the “civil-enforcement model” of the Fourth Amendment to reason in support of an implied individual right to notice. To begin, I note that the drafters’ preference for civil juries can be read in combination with Article III to imply a private cause of action to contest the reasonableness of government searches. Because the Bill of Rights did not originally apply to the states, the drafters must have contemplated that

115. Thus, Amar writes, the phrase “right of the people” may “highlight the part that civil jurors, acting collectively and representing the electorate, were expected to play in deciding which searches and seizures were reasonable and how much to punish government officials who searched or seized improperly.” *Id.*

116. *Id.*

117. AMAR, BILL OF RIGHTS, *supra* note 99, at 73.

118. To be sure, the damages-as-remedy solution has practical limits. For instance, Anthony G. Amsterdam has questioned whether anti-criminal bias, professional conflicts of interest, and the difficulty of litigating against the police might improperly dissuade contingency-fee lawyers from filing damages claims. Anthony G. Amsterdam, *Perspectives on the Fourth Amendment*, 58 MINN. L. REV. 349, 430 (1974) (“Taking on the police in any tribunal involved a commitment to the most frustrating and thankless legal work I know.”); *see also id.* at 447–48 n.137. But the good news is that damages harmonize individual rights with individual remedies. And indeed, the tide may be shifting in current doctrine away from deterrence through exclusion of evidence in criminal trials and towards civil liability. *See, e.g., Virginia v. Moore*, 128 S. Ct. 1598, 1603 (2008) (“[F]ounding-era citizens were skeptical of using the rules for search and seizure set by government actors as the index of reasonableness.”).

civil juries would enforce the Fourth Amendment in federal courts.¹¹⁹ Article III grants federal court jurisdiction over nine categories of “cases” and “controversies.”¹²⁰ Jury enforcement thus presupposes a case or controversy, which in turn requires a private cause of action to file a complaint—or a “right to contest.” Without this right, the government’s search practices would never be subject to civil jury review. And indeed, the Supreme Court recognized such a cause of action for federal Fourth Amendment violations in *Bivens*.¹²¹

Arguing for an implied right to notice is a more challenging jump. Yet, I submit, the underlying motives that Amar identified for drafting the amendment offer a foundation to bridge this gap. Here is a possible prudential claim.¹²² Without notice, the peoples’ limited capacity to discover and dispute what they don’t know would emaciate civil jury enforcement of the Fourth Amendment. Alleged victims of unreasonable covert searches would either lack the knowledge and incentive to initiate a suit, or fail to satisfy the Article III case and controversy requirement to obtain standing (as happened in *Clapper*).¹²³

The alternative is to argue that searches are per se reasonable when no one notices them, so secret searches need not be litigated. But, as the heated public debates that followed the Snowden revelations show, this categorical determination is subject to popular dispute.¹²⁴ Establishing ex

119. The Supreme Court did not begin to apply the Bill of Rights to the states by incorporating it into the due process clause of the Fourteenth Amendment until the late Nineteenth Century. See, e.g., Jerold Israel, *Selective Incorporation Revisited*, 71 GEO. L.J. 253, 336–38 (1982); MICHAEL CURTIS, NO STATE SHALL ABRIDGE: THE FOURTEENTH AMENDMENT AND THE BILL OF RIGHTS (1986).

120. U.S. CONST. art. III, § 2 (“The judicial Power shall extend to all Cases, in Law and Equity, arising under this Constitution . . . to Controversies to which the United States shall be a party . . .”).

121. *Bivens v. Six Unknown Named Agents*, 403 U.S. 388, 391–93.

122. See, e.g., PHILIP BOBBITT, CONSTITUTIONAL INTERPRETATION 16–17 (1991) (describing the prudential modality as claims about the wisdom of certain actions and their political expediency that are “actuated by facts”).

123. U.S. CONST. art. III, § 2; TAYLOR, TWO STUDIES IN CONSTITUTIONAL INTERPRETATION, *supra* note 60, at 82 (arguing that search warrants “are ancillary ex parte preludes to confrontation and controversy”).

124. E.g., Editorial Board, *Mass Surveillance Isn't the Answer to Fighting Terrorism*, N.Y. TIMES (Nov. 17, 2015), <http://www.nytimes.com/2015/11/18/opinion/mass-surveillance-isnt-the-answer-to-fighting-terrorism.html> [<https://perma.cc/L5QB-RV45>]; *End Mass Surveillance Under the Patriot Act*, ACLU.ORG, <https://www.aclu.org/feature/end-mass-surveillance-under-patriot-act> [<https://perma.cc/8TXZ-6LBE>]; see also, Julie E. Cohen, *What Privacy Is For*, 126 HARV. L. REV. 1904 (2013); Neil M. Richards, *The Dangers of Surveillance*, 126 HARV. L. REV. 1934 (2013); Neil M. Richards, *Intellectual Privacy*, 87 TEX. L. REV. 387 (2008).

ante categorical immunity for any search that the government can keep secret from its target is thus analogous to a directed verdict compelled by a warrant; both rob the ultimate judgment of reasonableness from the sovereign people represented in juries. If the drafters of the Fourth Amendment sought specifically to empower non-governmental actors to police government abuses, then permitting the government to achieve categorical ex ante immunity by preventing anyone outside of it from even knowing about a search belies the amendment's original intent.

History lends additional support for a notice right. Built-in notice attended the types of searches most immediately contemplated by the drafters of the Fourth Amendment. Customs searches were of prime concern, generating complaints that lower-class government officers were entering elite homes without justification.¹²⁵ The two famous English cases that dominated colonial conversations about the government's investigative powers at the time were *Wilkes v. Wood* in 1763 and *Entick v. Carrington* in 1765.¹²⁶ In *Wood*, government officers operating under general warrants broke into the house of John Wilkes, and later arrested and detained his person. Wilkes was well aware of the happenings, a fact that enabled him successfully to sue in an ex post trespass action.¹²⁷ Entick likewise would have noticed when government officers broke into and damaged his home before seizing his papers.¹²⁸ And warrants at common law required government officers to provide search targets with a receipt of anything seized.¹²⁹

This is neither to say that the drafters must have intended only built-in notice nor that they would have had difficulty contemplating covert

125. See, for example, James Otis on Writs of Assistance as quoted in Telford Taylor's, *Two Studies*, *supra* note 60, at 37 ("Custom house officials may enter our houses when they please—we are commanded to permit their entry—their menial servants may enter . . ."). After Taylor published his book, Charles Warren showed that the pamphlet was actually written by Mercy Otis Warren. See AMAR, BILL OF RIGHTS, *supra* note 99, at 66.

126. *Wilkes v. Wood* (1763) 98 Eng. Rep. 489; Lofft 1 (K.B.); *Entick v. Carrington* (1799) 95 Eng. Rep. 807; 19 How. St. Tr. 1029 (K.B.); see also AMAR, BILL OF RIGHTS, *supra* note 99, at 66–67 ("[B]oth the target of the government search, John Wilkes, and the author of the opinion, Lord Chief Justice Charles Pratt (soon to become Lord Camden), became folk heroes in the colonies.").

127. Lord Camden deciding the case found that the warrants were illegal, ordered the government to pay substantial compensatory and punitive damages, and in the process transformed himself and Wilkes into popular heroes in the colonies whom the drafters would have known about. *Id.* at 65–68.

128. See AMAR, LAW OF THE LAND, *supra* note 99, at chs. 2, 7.

129. TAYLOR, TWO STUDIES IN CONSTITUTIONAL INTERPRETATION, *supra* note 60, at 82.

searches with the possibility of delayed notice. Rather, the primacy of notice in the searches discussed and experienced at the time of the drafting bolsters the idea that notice had some role in the original understanding of the amendment. For further support, these observations about the drafting era can also be combined with longstanding practice. Criminal procedure rules and statutes requiring notice¹³⁰ reveal at once popular acceptance, legislative enactment, and executive precedent to provide some form of notice.¹³¹ Because the constitutional reasonableness of Fourth Amendment searches is a subjective standard, these practices help to shape “the people’s” expectations in their image. If non-governmental actors are ultimately responsible for enforcing Fourth Amendment reasonableness, then these longstanding popular expectations should help to define constitutional meaning.

B. NOTICE TO THE PUBLIC PROTECTS THE REPUBLICAN RIGHT TO CONTEST FOURTH AMENDMENT REASONABLENESS IN THE POLITICAL ARENA

This Section shows how providing notice of NSL investigations directly to the public can also further the drafters’ original intent. First, it examines the republican, or collective, aspects of the Fourth Amendment. It then considers whether, if Fourth Amendment rights to notice exist, they belong merely to individual search targets or also vest in some form with the people as a whole. The following discussion presents textual, historical, prudential, and doctrinal analyses to show that the Fourth Amendment encompasses collective as well as individual rights. It then argues that to satisfy both of these elements, any right to notice should extend in some form directly to the public. Some level of public transparency about government search practices would thus advance the original understanding of Fourth Amendment.

1. *Republican Fourth Amendment Rights*

The Fourth Amendment is the only part of the Constitution to reference in a single clause both ‘the people’ and the individual ‘persons’ who compose that communal body. Certainly, the two concepts implicitly appear in relation to each other throughout the Bill of Rights. For instance, amendments five and six guarantee the individual rights of

130. See *supra* Part III.

131. For a helpful overview of executive precedent in this domain, see Jared Cole, Note, *Historical Gloss and Congressional Power: Control over Access to National Security Secrets*, 99 VA. L. REV. 1855 (2013).

‘persons’ and ‘the accused’ while also referencing the jury, an institution of popular sovereignty that embodies ‘the people’ in their collective identity.¹³² And the democratic process of ratification that gave rise to the Bill as a whole casts a populist sheen across the individual rights of amendments four through six.¹³³ But exceptional to the Fourth Amendment, the explicit reference and proximate juxtaposition of the terms ‘the people’ and ‘persons’ emphasize at once their uniqueness and their interdependence.

If the terms ‘the people’ and ‘persons’ were not unique, the first clause—protecting the “right of the people to be secure in their persons”¹³⁴—would be redundant. It would also clash with the use of these terms in other parts of the Constitution. Internal consistency demands that “the people” of Article I, § 2, cl. 1, “the people” whose rights Amendments I, II, and IV protect, and “the people” whose powers Amendments IX and X reserve, constitute a community capable of forming the “Union” of “We the People” that the Preamble envisions.¹³⁵ The phrase “the people” in Amendment XVII also necessarily refers to a union because, like Article I, § 2, cl. 1, it presupposes collective action through democratic elections.¹³⁶ In contrast, the term “person,” as it appears in Amendments V, XII, XIV and XXII, is individualistic and not collective; each has one and only one, while none can share. Persons are therefore incapable of union, leaving the two terms by necessity unique.

Yet the word “their” also highlights that “the people” and “*their* persons” inter-depend. In a common contemporary reading of the phrase, each individual must enjoy some basic security as a prerequisite to joining a union in the first place. Control over one’s person by necessity precedes any determination of what to do with it. As a result, Fourth Amendment rights to individual security support and enable “We the People” to form a union capable of choosing representatives to govern—the constitutional project writ large.¹³⁷ At the same time, “their” reads in the other direction to mean that the union as a whole possesses the persons it comprises.

132. See U.S. CONST. amend. II, V–VII; AMAR, A BIOGRAPHY, *supra* note 99, at 329 (“Of the five amendments in the Bill of Rights that did not directly invoke ‘the people,’ three explicitly referred to the closely related idea of the ‘jury.’”).

133. AMAR, A BIOGRAPHY, *supra* note 99, at 321 (“The text of the Bill itself poetically recapitulated its own populist enactment sage.”).

134. U.S. CONST. amend. IV.

135. See, for example, *United States v. Verdugo-Urquidez* for a discussion of the phrase “the people” as a term of art. 494 U.S. 250, 265–55 (1990).

136. U.S. CONST. amend. XVII.

137. Preamble; U.S. CONST. art. I, § 2, cl. 1.

Grammatically, this possessory reading of “their” requires an understanding of “the people” not merely as a single entity like “the union,” but also as a plural noun like “the shareholders.”

In an originalist reading of these terms, “their” could mean that the Fourth Amendment protects the rights of individuals in the electorate to the security of other persons in their keep. Indeed, the free male voters who constituted “the people” at the time of the drafting, in a strictly political interpretation of the concept,¹³⁸ enacted the amendment with the explicit intent to protect their own rights to the security of *their* women—persons other than themselves whom they nonetheless in some sense possessed under the laws of coverture.¹³⁹ Of primary concern were the male customs officers who searched for imported goods inside people’s homes,¹⁴⁰ and who might enter them by surprise before the women inside could dress.¹⁴¹ Hence, at the time, the drafters could reasonably have sought to ensure that “the people”—free and male—be secure in *their* persons, including their own bodies and also those of their wives and children. To be sure, the concept of virtual representation, whereby men participating in the electorate were to guarantee the interests of their female dependents, need not transform all individual rights into collective ones. Fourth Amendment rights may be unique in this sense because they focus so prominently on the domestic sphere of the home often associated with women.¹⁴²

138. See, e.g., AMAR, LAW OF THE LAND, *supra* note 99, at 295 (“Free women in 1787 had the rights of ‘persons’ . . . [but] did not as a rule vote in constitutional conventions or for state lawmakers or for Congress; nor did they serve on juries; nor were women part of the militia/people at the heart of the Second Amendment.”); see also, AMAR, A BIOGRAPHY, *supra* note 99, at 324 (“Elsewhere in the Bill of Rights, the phrase ‘the people’ generally gestured toward voters as the core rights-holders, even as the phrase in certain contexts plainly radiated beyond the core group.”).

139. See, e.g., Margaret Valentine Turano, *Jane Austen, Charlotte Brontë, and the Marital Property Law*, 21 HARV. WOMEN’S L.J. 179 (1998).

140. It thus makes particular sense that it was a woman, Mercy Otis Warren, who wrote the pamphlet about the Writ of Assistance decrying the ease with which lower class customs officers could enter domestic spaces. See AMAR, BILL OF RIGHTS, *supra* note 99, at 66.

141. For a contemporary example, see *Kyllo v. United States*, 522 U.S. 27, 30 (2001), expressing Fourth Amendment concerns that surveillance technologies might “disclose, for example, at what hour each night the lady of the house takes her daily sauna and bath.”

142. AMAR, A BIOGRAPHY, *supra* note 99, at 326 (“Clearly, this amendment seemed to center on the domain of domesticity—on ‘persons’ in their private ‘houses’ as distinct from the people in the public square.”).

While political realities have changed, the dual nature of individual and collective Fourth Amendment rights remains politically expedient. From a prudential perspective, collective security depends on individual security and vice versa. For “the people” as a whole to be secure, each must also be individually secure. Were some unfortunate few singled out for lesser treatment, their insecurity would radiate outward to degrade the rights of all.¹⁴³ The insecurity of some may cause the identity and unity of the collective as a whole to suffer and cause social unrest. An insecure union threatens the security of each individual within it.

Finally, from a doctrinal perspective, the Court’s exclusionary rule doctrine lends further support for a collective reading of Fourth Amendment rights. The exclusionary rule prohibits the state from introducing information obtained through unconstitutional searches as evidence in a criminal trial.¹⁴⁴ While often criticized and a divergence from jury enforcement,¹⁴⁵ the rule also manifests collective Fourth Amendment rights because it depends on deterrence for justification.¹⁴⁶ Deterrence is a collective good entirely inapplicable to the past injury of the party with standing. In sum, collective Fourth Amendment rights at once depend on and also encompass individual Fourth Amendment rights.

2. *The Right to Contest Implies a Public Right to Notice*

If the Fourth Amendment requires notice, the amendment’s republican elements intimate that notice should be given in some form

143. This is the sentiment memorialized so beautifully in Pastor Martin Niemöller’s iconic quote, “[i]n Germany they came first for the Communists, and I didn’t speak up because I wasn’t a Communist. Then they came for the Jews . . . Then they came for me, and by that time no one was left to speak up.” NAT’L WWII MUSEUM, EXPLORING PERSONAL AND COLLECTIVE RESPONSIBILITY IN WWII (2011), <http://www.nationalww2museum.org/learn/education/for-teachers/lesson-plans/pdfs/when-they-came-for-me.pdf> [<https://perma.cc/747S-9UHM>].

144. *E.g.*, *Mapp v. Ohio*, 367 U.S. 643, 649 (1961) (establishing the exclusionary rule doctrine).

145. The exclusionary rule extracts enforcement power from “the people” of the jury and hands it to the judge, once again expanding federal power in contradiction to the enactment history of the Bill as a whole. *See Amar, First Principles, supra* note 81, at 811 (“Fixated on the exclusionary rule, the twentieth-century Supreme Court has betrayed the traditional civil-enforcement model . . .”).

146. *See, e.g.*, Donald L. Doernberg, *The Right of the People: Reconciling Collective and Individual Interests Under the Fourth Amendment*, 58 N.Y.U. L. REV. 259, 283 (1983) (“[T]he remedy which has been created in response to law enforcement officials’ fourth amendment violations is said to be not personal but collective, designed to prevent future violations rather than to repair those that have already occurred. Its purpose is not to benefit the individual who invokes it but instead to protect society by deterring unlawful police behavior.”).

directly to the public. Notice to Congress alone would not fully address the amendment's populist concern with self-policing by a self-interested government. Amar's jury-enforcement model of the original Fourth Amendment is fundamentally populist;¹⁴⁷ the drafters, the argument goes, preferred juries over judges because they did not trust decision-makers who received a government salary.¹⁴⁸ The deeper enforcement logic at work is this: The drafters sought to solve the agency problem of relying on self-interested government actors to police government actions.¹⁴⁹ Their solution was to use actors who are completely outside of government to control for Executive overreach and abuse. Critically, the drafters did not choose a mere inter-branch checks and balances enforcement scheme. Indeed, they distrusted judges even though judges represent a different branch of government from the executive officials who might perform abusive searches. According to this logic, enforcement by any government agent—including legislators—would contravene original intent. After all, legislators receive government pay just like judges. And while legislators must eventually face the electorate, they need not do so until the end of a term. Any collective Fourth Amendment right to notice should thus vest directly with the public to reach non-government-agent ears.

To be sure, congressional review offers more protection than self-enforcement by the Executive alone.¹⁵⁰ And the Court's recent holdings on Fourth Amendment standing support a view that the political branches—not courts—should enforce collective Fourth Amendment rights. Arguments for the Court's decision in *Clapper*, for instance, focused on separation of powers issues:¹⁵¹ Courts have developed a prudential

147. See AMAR, A BIOGRAPHY, *supra* note 99, at 327 (arguing that the second and fourth amendments were designed “to protect popular rights via institutions (the military and the jury) that would embody ‘the people’ themselves”).

148. See AMAR, LAW OF THE LAND, *supra* note 99, at 296.

149. In Amar's originalist interpretation, “the Fourth Amendment evinces at least as much concern with the agency problem of protecting the people generally from self-interested government policy as with protecting minorities against majorities of fellow citizens.” AMAR, BILL OF RIGHTS, *supra* note 99, at 68.

150. For an interesting contrast, see Philip Bobbitt's argument that the notice to Congress mandated by the Spending Clause controls for the excesses of a self-policing, self-interested Executive Branch. Bobbitt, CONSTITUTIONAL INTERPRETATION, *supra* note 122, at 64–82 (analyzing the Iran–Contra affair).

151. As government counsel emphasized during oral argument in that case, “the basic, most fundamental point about the case or controversy requirement and the injury-in-fact requirement that is embedded in it is to preserve the separation of powers.” Transcript of Oral Argument at 58, *Clapper*, 133 S. Ct. 1138 (2013) (11-1025), http://www.supremecourt.gov/oral_arguments/argument_transcripts/11-1025.pdf [<https://perma.cc/DQT7-W5ZD>].

standing doctrine of nonjusticiability for generalized grievances in order to protect the balance of powers and prevent judicial overreach into political questions best addressed by the other branches.¹⁵²

Yet, more may sometimes be needed; the enactment history of the Bill of Rights suggests that something is missing from the prudential standing perspective. For example, the First Congress enacted the Bill as a whole not to establish, but rather to limit, the powers of the political branches.¹⁵³ It would thus have been illogical for the drafters of the Fourth Amendment to reference “the people” in order to identify a collective right that, through proper congressional enforcement, would actually engorge the power of the federal government. If today the political branches hold the sole key to Fourth Amendment remedies for unreasonable mass surveillance searches, they likely do so contrary to the drafter’s intent. Notice to Congress alone runs counter to the amendment’s original populist logic. Public notice, in contrast, permits political solutions for generalized grievances without risking unmitigated dependence on officials who receive government pay.

Applying this understanding of the Fourth Amendment to NSLs lends constitutional support for the proposal that the government should publish more information about its actual NSL practices for public review. Individual rights may be justiciable and collective rights political, but both may sometimes mandate notice to contest the reasonableness of searches—particularly when the risk of government abuse is high.

C. FIRST AMENDMENT STANDARDS FOR NOTICE

This Section posits that First Amendment standards should guide the government’s notice obligations. The previous Sections offered possible arguments to support the idea of a Fourth Amendment right to notice. They suggested that the government may have a constitutional duty to notify both the individual target of an investigation and also the public directly. However, as the exceptions to statutory notice requirements in

152. See also, Andrew Nolan, *Foreign Surveillance and the Future of Standing to Sue Post-Clapper*, CONG. RESEARCH SERV., R43107, at 4 (Jul. 10, 2013) (“[T]he rule barring adjudication of generalized grievances . . . ensure[s] that legal injuries that are shared in equal measure by all or a large class of citizens are properly a subject of the elected branches and not the judiciary.”). For instance, the Court in *Alderman v. United States* made clear that “Fourth Amendment rights are personal rights which . . . may not be vicariously asserted.” *Alderman v. United States*, 394 U.S. 165, 174 (1969).

153. AMAR, A BIOGRAPHY, *supra* note 99, at 315 (prefacing a discussion of the First Congress’s enactment of the Bill of Rights with the reflection, “[s]elf-denial is a wonderful thing to behold and an intriguing one to explain”).

criminal procedure show, the government may at times have compelling reasons to delay notice. For instance, disclosure might risk physical danger, threaten an investigation, or cause some other harm.¹⁵⁴ The difficult question, then, is when, what, and how much notice should the government provide? Fortunately, the First Amendment offers ready balancing tests that can easily and reasonably be applied to guide government notice practices.

Turning to the First Amendment to help resolve Fourth Amendment issues is not new. For example, Jack Balkin recently proposed that the First Amendment should offer less protection to “information fiduciaries,” or entities that enjoy certain relationships of trust and confidence with people who give them information, and thus that these entities may be subject to greater regulation than other speakers.¹⁵⁵ Kiel Brennan-Marquez then developed the “information fiduciary” idea to argue that, for similar reasons of trust and confidence, these same entities should have a different relationship to the Fourth Amendment.¹⁵⁶ Likewise, Daniel Solove has observed that Fourth Amendment limits on certain kinds of government information gathering have failed adequately to safeguard some First Amendment values, such as anonymity or the right to listen.¹⁵⁷ Solove argues that the First Amendment should fill the gaps and prevent the government from collecting information about certain types of expressive activities.¹⁵⁸ While my argument moves in a different direction—it seeks to draw on the First Amendment to safeguard Fourth Amendment values—similar rationales apply, such as the historical ties between the two amendments and the fact that they were drafted in response to some overlapping concerns.¹⁵⁹

There is also a particular, concrete justification for applying First Amendment tests to guide Fourth Amendment notice: as with policy

154. See *supra* text accompanying notes 60–66.

155. Jack Balkin, *Information Fiduciaries and the First Amendment*, 49 UC DAVIS L. REV. 1183 at 1185–86 (2016).

156. Kiel Brennan-Marquez, *Fourth Amendment Fiduciaries*, 84 FORDHAM L. REV. 611 (2015). Balkin also draws the connection between First and Fourth Amendment information fiduciaries. “The reasons why this information is not public discourse for purposes of the First Amendment also provide reasons why we should have a reasonable expectation of privacy for purposes of the Fourth Amendment.” Balkin, *Information Fiduciaries*, *supra* note 155, at 1231.

157. Daniel J. Solove, *The First Amendment as Criminal Procedure*, 82 N.Y.U. L. REV. 112, 121, 122, 132–49 (2007).

158. *Id.* at 152.

159. *Id.*, at 133 (“The First, Fourth, and Fifth Amendments share a common background in concerns about seditious libel.”).

arguments for notice about NSL investigations,¹⁶⁰ collateral estoppel should prevent the government from reconsidering a prior constitutional decision about whether revealing certain information would cause harm. This idea plays out in constitutional terms as follows.

The First Amendment bars the government from making any law “abridging the freedom of speech,”¹⁶¹ including the speech of private citizens whom the government conscripts into aiding a search (such as Internet service providers). But the rights of the people to life, liberty, and security, as enshrined in the Executive’s Commander-in-Chief powers, sometimes permit the government to intrude on an individual’s free speech rights (such as by issuing a nondisclosure order).¹⁶² Free speech rights can be outweighed when speech risks alerting targets to an *ongoing* covert investigation or revealing classified intelligence methods.¹⁶³ To evaluate whether an individual’s First Amendment rights should prevail even in the context of a national security investigation, courts apply a series of established tests to determine whether the speech would cause harm.¹⁶⁴

First Amendment protections against NSL nondisclosure orders thus establish instances of third-party speech that pose no national security harms, or that pose risks of harm too minimal to outweigh an individual liberty enshrined in the Bill of Rights. For example, judges, lawyers, and legal scholars have questioned whether NSL nondisclosure requirements are “overbroad.” In other words, if the gags restrict more speech than necessary, they violate the First Amendment.¹⁶⁵ These concerns persist despite recent statutory developments. United States District Judge Susan Illston of the Northern District of California is currently considering a challenge by two unnamed electronic communication service providers who argue that the post-amendment NSL statutes remain unconstitutional.¹⁶⁶

160. See *supra* Part II.

161. U.S. CONST. amend. I.

162. U.S. CONST. art. II, § 2.

163. See, e.g., *Doe v. Holder*, 703 F. Supp. 2d 313, 318 (S.D.N.Y. 2010) (lifting an NSL gag order in part).

164. See *supra* note 43.

165. See, e.g., *John Doe, Inc. v. Mukasey*, 549 F.3d 861 (2008); *In re Nat’l Sec. Letter*, 930 F. Supp. 2d 1064 (N.D. Cal. 2013); Clark et. al, LIBERTY AND SECURITY, *supra* note 10; Jack Balkin, *Old-School*, *supra* note 21, at 2329–41. The issue has also reached public debate. See, e.g., Maria Bustillos, *What It’s Like to Get a National-Security Letter*, NEW YORKER (June 27, 2013), <http://www.newyorker.com/tech/elements/what-its-like-to-get-a-national-security-letter> [<https://perma.cc/355X-37VM>].

166. Appellee/Appellant Under Seal’s Supplemental Brief Re USA Freedom Act at 1, *Under Seal v. Lynch*, Nos. 13-15957, 13-16731 (9th Cir. July 6, 2015), <http://cdn.ca9.uscourts.gov/datastore/general/2015/07/15/13-15957%20dkt%2099%20Supp%20Brief.pdf>

I contend that the same standards that require the government to rescind a nondisclosure order under the First Amendment should trigger any Fourth Amendment notice obligations. A prior determination that speech about specific information cannot be suppressed by a nondisclosure order should prevent the government from claiming in other contexts that disclosing that same information would be harmful. In other words, the government's national security interests should pose a parallel limitation on both First and Fourth Amendment rights. This parallel limitation has the following consequences. If the limitation is the same for both amendments, the moment that one is trumped by national security interests so must the other be; and the moment one overcomes these competing concerns, so does the other. The very same minute that government interests in secrecy fail to outweigh the First Amendment rights of a third party conscripted to aid an investigation, they should also fail to outweigh any Fourth Amendment interests in notice about that investigation.

Note that in situations where national security interests fail to outweigh the speech rights of a third party conscripted into an investigation, that third party's right to speak does not require her to speak.¹⁶⁷ If the government were to permit third-party speech that produces some probability of notice, but fail to provide notice itself, the government would toss the Fourth Amendment interests of targets and the public to the happenstance whims of their fellow citizens. The result is unbecoming to the Bill of Rights.

V. CONCLUSION

Expanding notice of NSL investigations to their targets and the public would enhance transparency about government surveillance; reduce the risk of government overreach and abuse; and help to ensure accountability for the government's actual NSL practices. There are good reasons to believe that it would also advance the core principles that underlie the Fourth Amendment. Moreover, the FBI could accomplish these benefits

[<https://perma.cc/5C96-Y5WT>]. This litigation is on remand after the USA Freedom Act of 2015 revises the statutes at issue. Order Remanding Cases, Under Seal v. Lynch, Nos. 13-15957, 13-16731, 13-16732 (9th Cir. Oct 8, 2014), <https://www.eff.org/document/9th-circuit-order-remanding-case> [<https://perma.cc/VQ7X-RPRC>].

167. *Wooley v. Maynard* held that the First Amendment protects "both the right to speak freely and the right to refrain from speaking at all." 430 U.S. 705, 714 (1977); *see also* *Agency for Int'l Dev. v. All. for Open Soc'y Int'l, Inc.*, 133 S. Ct. 2321, 2332 (2013) (striking a policy requirement as unconstitutional under the First Amendment because it "compels as a condition of federal funding the affirmation of a belief").

without risk of harm or an undue administrative burden. To be sure, the Supreme Court's current doctrine does not compel the FBI to adopt a notice practice for NSL investigations.¹⁶⁸ But, in addition to the prudential justifications detailed in Parts I–III, the Executive also has an independent obligation to interpret and enforce the Constitution in good faith.¹⁶⁹ Therefore, if Congress does not amend the NSL statutes to require notice to targets and the public, the Executive should adopt these policies of its own accord.

The Northern District of California is currently considering a challenge by two unnamed Internet service providers who argue that even with the USA FREEDOM Act amendments, NSL nondisclosure orders remain unconstitutional.¹⁷⁰ In a prior decision, the court agreed with the plaintiffs that the pre-amendment nondisclosure orders were insufficiently narrowly tailored—in both scope and duration—to the government's national security interests and thus facially unconstitutional.¹⁷¹ If the court again finds that the gags are unconstitutional under the First Amendment and lifts them in full, the court will have determined that the government's interests do not justify the scope and duration of its secrecy. If that happens, the government would similarly be unable to justify its lack of post-intrusion notice to the targets of those particular NSL searches. If the gag orders are lifted, perhaps the plaintiffs will notify the targets of these investigations and join them in a Fourth Amendment challenge.

168. See *supra* text accompanying notes 92–96.

169. See, e.g., Brian Galle, *The Justice of Administration: Judicial Responses to Executive Claims of Independent Authority to Interpret the Constitution*, 33 FLA. ST. U. L. REV. 157 (2005); John O. McGinnis & Charles W. Mulaney, *Judging Facts Like Law*, 25 CONST. COMMENT. 69, 110 (2008) (“Formally, no express clause of the Constitution singles out one branch or the other for exclusive responsibility of constitutional assessment. Indeed, members of all branches take an oath to uphold the Constitution.”).

170. See sources cited *supra* note 165.

171. *In re Nat'l Sec. Letter*, 930 F. Supp. 2d 1064, 1075–77 (N.D. Cal. 2013).

