

BERKELEY TECHNOLOGY LAW JOURNAL

VOLUME 38, ISSUE 4

ANNUAL REVIEW OF LAW AND TECHNOLOGY

2023

Pages

1089–1504

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 © 2023. Regents of the University of California.
All Rights Reserved.



Berkeley Technology Law Journal
University of California
School of Law
3 Law Building
Berkeley, California 94720-7200
editor@btlj.org
<https://www.btlj.org>

BERKELEY TECHNOLOGY LAW JOURNAL

VOLUME 38

ISSUE 4

2023

ARTICLES

FOREWORD	1089
<i>Shih-wei Chao & Rebecca Ho</i>	
THE UTILITY OF PATENT ELIGIBILITY	1093
<i>Caressa N. Tsai</i>	
DRUGS, DECEPTION, AND DISCLOSURE.....	1131
<i>Garreth W. McCrudden</i>	
TAKE A PICTURE: COPYRIGHT AND STATE SOVEREIGN IMMUNITY	1169
<i>Sarah Davidson</i>	
RESOLVING COPYRIGHT’S DISTORTIONARY EFFECTS	1207
<i>Akshat Agrawal</i>	
UNDERSTANDING UNICOLORS: MISTAKES OF LAW DON’T NECESSARILY INVALIDATE COPYRIGHT REGISTRATION CERTIFICATES.....	1249
<i>Samantha Cox-Parra</i>	
DANGLING THE CARROT OF TRADEMARK REGISTRATION.....	1273
<i>Brigitte Desnoes</i>	
WHY IT’S TIME TO BAN GEOFENCE SEARCHES IN LIGHT OF <i>UNITED STATES V. CHATRIE</i>	1307
<i>Danny Drane</i>	
DATA SURVEILLANCE AND ABORTION BANS AFTER <i>DOBBS</i>	1341
<i>Leila Nasrolahi</i>	
CLEARLY REPUGNANT: CORRECTING THE COURT’S FAILED APPROACH TO ANTITRUST ENFORCEMENT	1373
<i>M. A. Katz</i>	
PROTECTING THE COMPETITIVE PROCESS IN VERTICAL MERGER.....	1405
<i>Zhudi Huang</i>	

ARTICLES (CONTINUED)

GOTTA CATCH 'EM ALL: LEGISLATIVE OVERREACH IN FLORIDA AND
TEXAS ANTI-MODERATION LAWS.....1437
Utkarsh Srivastava

THE FORGOTTEN PUBLIC INTEREST STANDARD.....1469
Bogdan Belei

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 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, Law Library, LL123 South Addition, Berkeley Law, University of California, Berkeley, 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, Law Library, LL123 South Addition, Berkeley Law, University of California, Berkeley, Berkeley, CA 94720-7210; (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 CHICAGO MANUAL OF STYLE (16th ed. 2010) and to THE BLUEBOOK: A UNIFORM SYSTEM OF CITATION (Columbia Law Review Ass'n et al. eds., 21st ed. 2020). Please cite this issue of the *Berkeley Technology Law Journal* as 38 BERKELEY TECH. L.J. ____ (2023).

BTLJ ONLINE

The full text and abstracts of many previously published *Berkeley Technology Law Journal* articles can be found at <https://www.btlj.org>. Our site also contains a cumulative index; general information about the *Journal*; the *BTLJ Blog*, 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 Scholastica 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 Scholastica submission website can be found at <https://btlj.scholasticahq.com/for-authors>.

Citations. All citations should conform to THE BLUEBOOK: A UNIFORM SYSTEM OF CITATION (Columbia Law Review Ass'n et al. eds., 21st ed. 2020).

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.

SPONSORS

2022–23

The *Berkeley Technology Law Journal* and the Berkeley Center for Law & Technology acknowledge the following generous sponsors of Berkeley Law’s Law and Technology Program:

ALLEN & OVERY LLP	AXINN, VELTROP & HARKRIDER LLP
BAKER MCKENZIE	BAKER BOTTS L.L.P.
BOIES SCHILLER FLEXNER LLP	CHARLES RIVER ASSOCIATES
COOLEY LLP	CORNERSTONE RESEARCH
COVINGTON & BURLING LLP	DESMARAIS LLP
DLA PIPER	DURIE TANGRI LLP
FENWICK & WEST LLP	FISH & RICHARDSON P.C.
GENENTECH, INC.	GEN LAW FIRM
GIBSON, DUNN & CRUTCHER LLP	GILEAD SCIENCES, INC.
GREENBERG TRAURIG, LLP	GTC LAW GROUP PC
HAYNES AND BOONE, LLP	HOGAN LOVELLS
IRELL & MANELLA LLP	JINGTIAN & GONGCHENG
JONES DAY	KEKER, VAN NEST & PETERS LLP
KILPATRICK TOWNSEND & STOCKTON LLP	KING & SPALDING LLP

SPONSORS

2022–23

KING & WOOD MALLESONS	KIRKLAND & ELLIS LLP
KNOBBE MARTENS	LATHAM & WATKINS LLP
MARKS & CLERK	MCDERMOTT WILL & EMERY
MICROSOFT	MORGAN, LEWIS & BOCKIUS LLP
MORRISON & FOERSTER LLP	OCEAN TOMO
ORRICK HERRINGTON & SUTCLIFFE LLP	PAUL HASTINGS, LLP
QUALCOMM TECHNOLOGIES, INC.	QUINN EMANUEL URQUHART & SULLIVAN, LLP
ROBINS KAPLAN LLP	ROPES & GRAY LLP
SIDLEY AUSTIN LLP	TENSEGRITY LAW GROUP LLP
VAN PELT, YI & JAMES LLP	VIA LICENSING CORPORATION
WANHUIDA INTELLECTUAL PROPERTY	WEIL, GOTSHAL & MANGES LLP
WESTERN DIGITAL CORPORATION	WHITE & CASE LLP
WILMER CUTLER PICKERING HALE AND DORR LLP	WILSON SONSINI GOODRICH & ROSATI
WINSTON & STRAWN LLP	WOMBLE BOND DICKINSON

BOARD OF EDITORS

2022–23

Executive Board

Editors-in-Chief
DYLAN HOULE
JESSICA LI

Managing Editor
THOMAS MATTES

Sr. Executive Editor
RYAN CAMPBELL

Sr. Articles Editors
LUCILLE DAI-HE
MARIAN LEE
ROSS MOODY

Sr. Online Content Editors
AL MALECHA
KATHERINE WANG

Sr. Student Publications Editors
SHIH-WEI CHAO
REBECCA HO

Sr. Scholarship Editor
WILL KASPER

Sr. Production Editor
SARAH DAVIDSON

Editorial Board

Articles Editors
KEATON BLAZER
BRANDON DAILEY
BRIGITTE DESNOES
MATTHEW LUEVANO
LAWRENCE MYUNG
ELIZABETH OH
NIKKI SEICHEPINE
YUHAN WU
ROBERT ZHU

Notes & Comments Editor
NOAH COHEN

Student Publication Editor
KATE BULGER

Submissions Editors
ABRAHAM (ABIE)
BRAUNER
HYO JIN JEONG
ANNIE KIM
KERMIT RODRIGUEZ

Technical Editors
CHIAGOZIEM
MARK ANEKE
ZHUDI HUANG
CARESSA TSAI
SOPHIE ZOLLMAN

Alumni Relations Editor
EVA SPITZEN

External Relations Editor
PRIMAVERA MARTINEZ

Jr. Production Editors
SEUNGHAN BAE
ALEX LE

Podcast Editor
ISABEL JONES

Jr. Podcast Editor
ERIC AHERN

BHAVYA SUKHAVASI
ZOE SUN

Member Relations Editors
JAEYOUNG CHOI
PAYTON FONG

Symposium Editors
MEAGHAN KATZ
HUNTER KOLON

Web & Technology Editors
LINDA CHANG
PETER JOHNSTON

MEMBERSHIP

2022–23

Members

AKSHAT AGRAWAL	NORA AL HENAKI	NOOR ALANIZI
BRIAN ANDERSON	LAURAN BARLAND	HARVINDER (HARRY) BASSI
ANVITA BHANDARI	MARIT BJORN LUND	NICOLE BOUCHER
TIFFANEY BOYD	CARLISLE BOYLE	KATHY BUI
EMMA BURKE	ELLEN CAEN	ANGELICA CAMELO
ADRIANA CRISTINA CASADO BARRAL	JENNIFER CHENG	NIDHI CHIKKERUR
ALEX CHOI	TANISHA CHOUDHARY	WILLIAM CLARK
OWEN COOPER	DAISY CORONA	MAYA DARROW
ADAM DAVID	MAX DAVIS	UDIT DEDHIYA
HEATHER DEMOCKER	MICHELLE D'SOUZA	VERNON ESPINOZA
ASHLEY FAN	JOE FITZPATRICK	LAURA FORERO
BRIONNE FRAZIER	ALEXANDER GONZALEZ-TORRES	DYLAN GOOD
ANAN HAFEZ	LYNN HE	JISOO HONG
ELLEN JIN	CLEA JULLIEN	BRITTA KAJIMURA
SHREYA KAPOOR	PHILIP KATZ	JOHN KAUFFMAN
SHABRINA DEFI KHANSA	JUNG KIM	BEN KING
GULBABIL KOKVER	THELMA CHRISTEL KRISTJÁNSDÓTTIR	CYRUS KUSHA
JENNA KWAK	JESSICA KWOK	KRISTY LAM
JOY LEE	CRYSTAL LEE	MADISON ANN LEE

MEMBERSHIP

2022–23

Members (continued)

YAOQI LI	YU-JEN LIU	ANGELA MA
MARLEY MACAREWICH	MAHIMA MADAN	HARSHINI MALLI
BAILEY MALVEAUX	ANYA MANSOOR	ROBIN MARSHALL
GARRETH MCCRUDDEN	EDLENE MIGUEL	JOSH MIMURA
SUMAN MISHRA	JOHN MOORE	KEVIN NAVI
KAYLA NGUYEN	ANDREW OLIVEI	OLUWAPELUMI (PELUMI) OMONIYI
DAMILOLA OYEBAYO	HYEMI PARK	BEN PEARCE
ALESSANDRA POTINI	RUOYIN QI	JAANVI RATHI
KRITI RATHI	SAYAKA REED	SAMUEL REISS
JOE ROSS	CHRIS ROZOLIS	BANI SAPRA
RIDHIMA SAXENA	SIMONE SCHÜRCH	MASON SEDEN- HANSEN
SAKSHI SHARM	PAVANA SHIVAKUMAR UDASI	AISHWARYA SIHAG
ELIZABETH BASANA	IAN SMITH	ANDREA SOTELO GASPERI
VALERIE SIMANJUNTAK		
UTKARSH SRIVASTAVA	SANDEEP STANLEY	ANDREW STONE
CEM TEBER	SANSERN TEJSIHADHANANANDA	YAVUZ USAKLIOGLU
PEINING WANG	EMILY WELSCH	DOMINICK WILLIAMS
YIPING XU	YING (ELECTRA) YAO	JJ YOUNG
ENID YU	NICOLE ZEINSTR	YICHEN ZHANG
NINA ZHANG	BOJUN ZHAO	

BERKELEY CENTER FOR
LAW & TECHNOLOGY
2022–23

Executive Director

WAYNE STACY

Staff

MARK COHEN
*Sr. Fellow & Director,
BCLT Asia IP Project*

ALLISON SCHMITT
*Fellow & Director,
BCLT Life Sciences
Project*

JANN DUDLEY
Associate Director

RICHARD FISK
*Assistant Director,
Events & Communications*

IRYS SCHENKER
Office Administrator

JUSTIN TRI DO
Media Coordinator

ABRIL DELGADO
Events Specialist

Fellow

YUAN HAO
Senior Fellow

KATHRYN
HASHIMOTO
Copyright Law Fellow

RAMYA
CHANDRASEKHAR
*Biometric Regulatory
Fellow*

ROBERT BARR
*Executive Director
Emeritus*

BERKELEY CENTER FOR LAW & TECHNOLOGY

2022–23

Faculty Directors

KENNETH A.
BAMBERGER
*The Rosalinde and
Arthur Gilbert Foundation
Professor of Law*

CATHERINE CRUMP
*Robert Glushko Clinical
Professor of Practice in
Technology Law and
Director, Samuelson Law,
Technology and Public Policy
Clinic*

CATHERINE FISK
*Barbara Nachtrieb
Armstrong Professor of Law*

CHRIS JAY HOOFNAGLE
Professor of Law in Residence

SONIA KATYAL
Roger J. Traynor
*Distinguished Professor of
Law and Associate Dean,
Faculty Development and
Research*

ORIN S. KERR
*William G. Simon
Professor of Law*

PETER S. MENELL
Koret Professor of Law

ROBERT P. MERGES
*Wilson Sonsini Goodrich &
Rosati Professor of Law*

DEIRDRE K. MULLIGAN
*Professor in the
School of Information*

TEJAS N. NARECHANIA
*Robert and Nanci Corson
Assistant Professor of Law*

OSAGIE K. OBASOGIE
*Haas Distinguished Chair,
Professor of Law, and
Professor of Bioethics*

ANDREA ROTH
Professor of Law

PAMELA SAMUELSON
*Richard M. Sherman
Distinguished Professor of
Law*

PAUL SCHWARTZ
*Jefferson E. Peysers
Professor of Law*

ERIK STALLMAN
*Assistant Clinical Professor
and Associate Director,
Samuelson Law, Technology
& Public Policy Clinic*

JENNIFER M. URBAN
*Clinical Professor of Law
and Director, Samuelson
Law, Technology &
Public Policy Clinic*

MOLLY VAN
HOUWELING
*Harold C. Hobbach
Distinguished Professor of
Patent Law and
Intellectual Property*

REBECCA WEXLER
Assistant Professor of Law

FOREWORD

Shih-wei Chao[†] & *Rebecca Ho*^{††}

The Annual Review is a yearly publication of the *Berkeley Technology Law Journal* that provides a summary of many of the year's major developments at the intersection of law and technology. Our aim is to provide a valuable resource for judges, policymakers, practitioners, students, and scholars. Each Note provides a primer on a particular area of law, a development in that area of law, and commentary on that development.

The twelve Notes in this Issue continue a tradition of covering a wide range of topics. The Notes address developments in patent, copyright, trademark, privacy, antitrust, and content and media regulation.

I. PATENT LAW

The first Note¹ in this Section examines the patent-eligible subject matter doctrine after the Federal Circuit's decision in *American Axle & Manufacturing, Inc. v. Neapco Holdings LLC*.² In *American Axle*, the Federal Circuit applied the *Mayo/Alice* framework to determine whether an invention was eligible for patent protection. This Note asserts that the *Mayo/Alice* framework is problematic, as it overlaps with other patentability tests and disregards how all inventions inevitably include some level of abstraction. This Note proposes a revised notion of patent eligibility that is anchored in utility doctrine.

The second Note³ in this Section addresses a conflict in the pharmaceutical industry. To obtain approval from the Federal Drug Administration (FDA), a pharmaceutical company must present its product as being similar to already-approved drugs. However, when the pharmaceutical company seeks patent protection from the United States Patent and Trademark Office (USPTO) for

DOI: <https://doi.org/10.15779/Z38HM52M16>

© 2023 Shih-wei Chao & Rebecca Ho.

† Senior Student Production Editor, *Berkeley Technology Law Journal*; J.S.D. Candidate, University of California, Berkeley, School of Law.

†† Senior Student Production Editor, *Berkeley Technology Law Journal*; J.D., 2023, University of California, Berkeley, School of Law.

1. Caressa N. Tsai, Note, *The Utility of Patent Eligibility*, 38 BERKELEY TECH. L.J. 1093 (2024).

2. 967 F.3d 1285 (Fed. Cir. 2020), *cert. denied*, 142 S. Ct. 2902 (2022).

3. Garreth W. McCrudden, Note, *Drugs, Deception, and Disclosure*, 38 BERKELEY TECH. L.J. 1131 (2024).

the same product, it must distinguish the product from existing prior art. To resolve this conflict, this Note recommends a new system for USPTO-FDA interaction during patent prosecution and a complementary post-patent-issuance solution.

II. COPYRIGHT LAW

The first Note⁴ in this Section examines the interplay between copyright, state sovereign immunity, and the federal Takings Clause. After the Texas Supreme Court's decision in *Jim Olive Photography v. University of Houston*,⁵ state actors can essentially appropriate copyrighted material with impunity. To hold state actors liable for the unauthorized taking of an individual's copyrighted material, this Note explores the feasibility of bringing lawsuits under the Takings Clause and proposes a revised, narrowly tailored version of the Copyright Reform Clarification Act.

The second Note⁶ in this Section asserts that copyright law inherently favors certain genres of cultural expression over others, because certain genres of music, such as Indian classical music, inevitably either infringe upon another's copyright or contain a high amount of *scenes a faire* elements. To encourage investment in these genres of music, this Note proposes narrowing the scope of copyright's derivative and reproductive rights.

The third Note⁷ in this Section examines the Supreme Court's decision in *Unicolors v. H&M*.⁸ While the legal issue in the case—whether good faith mistakes of fact or law render copyright registration applications invalid—was a narrow one, this Note asserts that the underlying facts of the case suggest there is an issue with “copyright trolling.” This Note recommends several policy solutions to deter copyright trolls from abusing the copyright system.

4. Sarah Davidson, Note, *Take A Picture: Copyright and State Sovereign Immunity*, 38 BERKELEY TECH. L.J. 1169 (2024).

5. 624 S.W.3d 764, 768 (Tex. 2021), *cert. denied*, 142 S. Ct. 1361 (2022).

6. Akshat Agrawal, Note, *Resolving Copyright's Distortionary Effects*, 38 BERKELEY TECH. L.J. 1207 (2024).

7. Samantha Cox-Parra, Note, *Understanding Unicolors: Mistakes of Law Don't Necessarily Invalidate Copyright Registration Certificates*, 38 BERKELEY TECH. L.J. 1249 (2024).

8. *Unicolors, Inc. v. H&M Hennes & Mauritz, L.P.*, 142 S. Ct. 941 (2022).

III. TRADEMARK LAW

The Note⁹ in this Section critiques the Federal Circuit’s decision in *In re Elster*.¹⁰ In *In re Elster*, the Federal Circuit held that the USPTO’s refusal to register “TRUMP TOO SMALL” as a trademark violated the First Amendment’s protection of political speech. This Note asserts that the Federal Circuit in *In re Elster* failed to properly consider the context of the speech at issue and that the Federal Circuit should have analyzed the registrability of the trademark under the First Amendment’s limited public forum framework.

IV. PRIVACY

The first Note¹¹ in this Section asserts that law enforcement’s use of the geofence search warrant endangers individuals’ rights to privacy and political speech. In light of these dangers, this Note recommends that Congress enact a blanket prohibition on all law enforcement use of the geofence search warrant.

The second Note¹² in this Section examines the Supreme Court’s decision in *Dobbs v. Jackson Women’s Health Organization*, where the Court held that there is not a constitutional right to receive an abortion. This Note posits that, post-*Dobbs*, digital data surveillance will become the primary mode of enforcing abortion bans.¹³ This Note argues that the use of data surveillance will chill desirable, legal activities and that Congress should enact privacy legislation to limit the use of data surveillance.

V. ANTITRUST

The first Note¹⁴ in this Section examines the relationship between antitrust law and regulation in the context of Personal Social Networks. This Note argues that the relationship between antitrust law and regulation is symbiotic, and not adversarial. To properly balance the relationship between antitrust law and regulation in the context of Personal Social Networks, this Note asserts that the Personal Social Networks must first be broken down to manageable

9. Brigitte Desnoes, Note, *Dangling the Carrot of Trademark Registration*, 38 BERKELEY TECH. L.J. 1273 (2024).

10. 26 F.4th 1328 (Fed. Cir. 2022).

11. Danny Drane, Note, *Why It’s Time to Ban Geofence Searches in Light of United States v. Chatrie*, 38 BERKELEY TECH. L.J. 1307 (2024).

12. Leila Nasrolahi, Note, *Externalities of Maximally Enforcing Abortion Bans*, 38 BERKELEY TECH. L.J. 1341 (2024).

13. 142 S. Ct. 2228 (2022).

14. M. A. Katz, Note, *Clearly Repugnant: Correcting the Court’s Failed Approach to Antitrust Enforcement*, 38 BERKELEY TECH. L.J. 1373 (2024).

sizes before regulation can be crafted to protect consumers from harms like hate speech and privacy invasions.

The second Note¹⁵ in this Section examines the goals of antitrust law in the context of vertical mergers. This Note contends that the purpose of antitrust law is to protect competition. As vertical mergers do not directly eliminate competitors, this Note asserts that vertical mergers present a unique situation to test the meaning of competition.

VI. MEDIA AND CONTENT REGULATION

The first Note¹⁶ in this Section scrutinizes regulations by Texas and Florida that restrict the ability of social media platforms to moderate content.¹⁷ This Note concludes that the regulations are unconstitutional, as they are not narrowly tailored, and that they cannot be modified to become constitutional.

The second Note¹⁸ in this Section examines the Supreme Court's decision in *FCC v. Prometheus Radio Project*.¹⁹ In that case, Court held that the Federal Communication Commission (FCC) had the authority to revoke media cross-ownership rules. This Note asserts that the Court largely sidestepped the normative issues concerning the public interest standard, which requires broadcast licensees to operate in the "public interest, convenience and necessity." This Note argues that the FCC should revitalize the public interest standard in light of an alarming trend of deregulation in the media broadcasting industry.

15. Zhudi Huang, Note, *Protecting the Competitive Process in Vertical Merger*, 38 BERKELEY TECH. L.J. 1405 (2024).

16. Utkarsh Srivastava, Note, *Gotta Catch 'Em All: Legislative Overreach in Florida and Texas Anti-Moderation Laws*, 38 BERKELEY TECH. L.J. 1437 (2024).

17. FLA. STAT. § 501.2041 (2022); TEX. CIV. PRAC. & REM. CODE ANN. § 143A.002 (West 2021).

18. Bogdan Belej, Note, *The Forgotten Public Interest Standard*, 38 BERKELEY TECH. L.J. 1469 (2024).

19. 141 S. Ct. 1150 (2021).

THE UTILITY OF PATENT ELIGIBILITY

Caressa N. Tsai[†]

TABLE OF CONTENTS

I.	INTRODUCTION	1094
II.	HISTORY OF PATENT ELIGIBILITY JURISPRUDENCE	1096
A.	EARLY NOTIONS OF INELIGIBLE “PRINCIPLES”	1097
B.	THE EMERGENCE OF TWO INELIGIBILITY TRACKS	1098
1.	<i>Track One: “Natural” Ineligibility</i>	1099
2.	<i>Track Two: “Formulaic” Ineligibility</i>	1101
C.	EVOLUTION OF THE <i>MAYO/ALICE</i> TWO-STEP TEST	1103
D.	IMPACT OF THE <i>MAYO/ALICE</i> TEST ON “NATURAL” AND “DIAGNOSTIC” INVENTIONS	1107
III.	<i>AMERICAN AXLE</i>	1110
A.	CASE SUMMARY	1110
B.	ANALYSIS OF THE <i>MAYO/ALICE</i> TEST IN <i>AMERICAN AXLE</i>	1112
1.	<i>Step One of Mayo/Alice and the Ineligibility Bars</i>	1112
2.	<i>Step Two of Mayo/Alice and the Redundant Inquiries</i>	1114
a)	From § 101 to § 103	1114
b)	From § 101 to § 112	1116
IV.	A REVISED FRAMEWORK FOR PATENT ELIGIBILITY	1118
A.	EXISTING PERSPECTIVES ON § 101 UTILITY	1119
1.	<i>Social and Moral Utility</i>	1119
2.	<i>Operable Utility</i>	1120
3.	<i>Specific, Substantial, and Credible Utility</i>	1121
B.	THE PROPOSED UTILITY-ELIGIBILITY FRAMEWORK	1123
1.	<i>Scope-Limiting Utility</i>	1123
2.	<i>Innovation Policy Justifications</i>	1125
3.	<i>Implications of and Potential Improvements to the Proposal</i>	1126

DOI: <https://doi.org/10.15779/Z38JH3D39Z>

© 2023 Caressa N. Tsai.

[†] Ph.D., McMaster University, Department of Biochemistry and Biomedical Sciences, 2021; J.D. Candidate, University of California, Berkeley, School of Law, Class of 2024. I am grateful to Talha Syed for his guidance in preparing this Note. I would also like to thank Peter Menell, Allison Schmitt, Garreth McCrudden, Shih-wei Chao, Hunter Kolon, and the participants of the Fall 2022 Law & Technology Writing Workshop, for feedback, editorial assistance, and many helpful discussions.

V. CONCLUSION1128

I. INTRODUCTION

Patent eligibility doctrine is in a state of disarray. Section 101 of the Patent Act defines the scope of patent-eligible subject matter in simple, broad language: “[w]hoever invents or discovers any new and useful process, machine, manufacture, or composition of matter, or any new and useful improvement thereof” is eligible for patent protection.¹ But after several years of the Federal Circuit and Supreme Court expansively interpreting this statute,² both courts began to slowly establish a set of *ineligible* subject matter areas, adding a gloss over the text of § 101.³ Over time, “laws of nature, natural phenomena, and abstract ideas” became the judicial exceptions to patent eligibility.⁴ Then, beginning in 2012, *Mayo Collaborative Services v. Prometheus Laboratories, Inc.*,⁵ *Association for Molecular Pathology v. Myriad Genetics, Inc.*,⁶ and *Alice Corp. v. CLS Bank International*⁷ launched the *Mayo/Alice* test for evaluating the § 101 eligibility of inventions directed to one of the judicial exceptions.⁸

There is no shortage of writing on the problems with the *Mayo/Alice* test.⁹ Several patent examiners, practitioners, scholars, and jurists have agreed that patent eligibility doctrine is in urgent need of clarification from either the Court or Congress.¹⁰ For many, a key pressure point is that the *Mayo/Alice* test

1. 35 U.S.C. § 101.

2. 1 PETER S. MENELL, MARK A. LEMLEY, ROBERT P. MERGES & SHYAMKRISHNA BALGANESH, *INTELLECTUAL PROPERTY IN THE NEW TECHNOLOGICAL AGE: 2021* 177 (2021) (“[B]y the early 2000s, the Federal Circuit had pretty much lowered the patentable subject matter hurdle to a chalk line on the track.”).

3. *Berkheimer v. HP Inc.*, 890 F.3d 1369, 1374 (Fed. Cir. 2018) (Lourie, J., concurring).

4. *Id.*

5. 566 U.S. 66 (2012).

6. 569 U.S. 576 (2013).

7. 573 U.S. 208 (2014).

8. Talha Syed, *Reconstructing Patent Eligibility*, 70 AM. U. L. REV. 1937, 1959 (2021).

9. Robin C. Feldman, *Rewarding Failure with Patents*, YALE J. HEALTH POL’Y L. & ETHICS 2 n.2 (forthcoming 2023), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4223347 (compiling a long list of scholarly proposals relating to patent eligibility doctrine).

10. *See, e.g.*, David O. Taylor, *Amending Patent Eligibility*, 50 U.C. DAVIS L. REV. 2149 (2017) (arguing in favor of new legislation from Congress to address the lack of clarity for § 101); *Sequenom, Inc. v. Ariosa Diagnostics, Inc.*, SCOTUSBLOG (June 27, 2016), <https://www.scotusblog.com/case-files/cases/sequenom-inc-v-ariosa-diagnostics-inc/> (compiling amicus briefs urging the Court to grant certiorari in a controversial § 101 case based on a biotechnology invention). *But see, e.g.*, Jason D. Reinecke, *Is the Supreme Court’s Patentable Subject Matter Test Overly Ambiguous? An Empirical Test*, 2019 UTAH L. REV. 581 (2019) (analyzing the

appears to have specifically narrowed the eligibility of biological and software-based inventions¹¹—two critical areas of innovation, at the heart of what the patent system is designed to promote.¹² Indeed, many litigants are now leveraging the murky § 101 standard to challenge patent validity in these fields.¹³

Since *Alice*, the Court has bowed out of the eligibility problem, leaving district courts and the Federal Circuit to wrestle with § 101 on their own.¹⁴ In 2022, the Court denied yet another petition for certiorari in *American Axle & Manufacturing, Inc. v. Neapco Holdings LLC*.¹⁵ *American Axle* saw a district court invalidate a patent claim under § 101,¹⁶ and the Federal Circuit affirm that decision.¹⁷ And while § 101 jurisprudence has wreaked havoc on inventions in the life sciences and computational spaces for the past several years,¹⁸ the asserted patent in *American Axle* uniquely encompassed a *mechanical* device—something canonically thought to be unambiguously patent-eligible subject matter, and somewhat above the § 101 fray.

This Note uses *American Axle* to illustrate the critical problem with the current § 101 eligibility standard, beyond its contemporary restriction of patent eligibility in specific technological areas: the *Mayo/Alice* test asks the wrong questions of patent examiners and courts by (1) placing an emphasis on judicial

application of the *Mayo/Alice* test in several patent eligibility cases, to illustrate that § 101 jurisprudence may not be as unpredictable and dire as others have expressed).

11. See Syed, *supra* note 8, at 1940–42.

12. As per the Intellectual Property Clause of the Constitution, patents are awarded to inventors to “promote the progress of science and the useful Arts.” U.S. CONST. art. 1, § 8, cl. 8.

13. See Jeffrey A. Lefstin, Peter S. Menell & David O. Taylor, *Final Report of the Berkeley Center for Law & Technology Section 101 Workshop: Addressing Patent Eligibility Challenges*, 33 BERKELEY TECH. L.J. 551, 555–56, 576 (2018). An increasing number of patents are routinely invalidated as ineligible in district courts, with these holdings consistently affirmed by the Federal Circuit. *Id.* at 570, 576.

14. See, e.g., *Sequenom, Inc. v. Ariosa Diagnostics, Inc.*, 579 U.S. 928 (2016) (denying petition for writ of certiorari); *Athena Diagnostics, Inc. v. Mayo Collaborative Servs., LLC*, 140 S. Ct. 855 (2019) (same); *HP Inc. v. Berkheimer*, 140 S. Ct. 911 (2020) (same); *Hikma Pharms. USA Inc. v. Vanda Pharms Inc.*, 140 S. Ct. 911 (2020) (same).

15. 142 S. Ct. 2902 (2022).

16. *Am. Axle & Mfg., Inc. v. Neapco Holdings LLC*, 309 F. Supp. 3d 218 (D. Del. 2018).

17. *Am. Axle & Mfg., Inc. v. Neapco Holdings LLC*, 967 F.3d 1285 (Fed. Cir. 2020).

18. See U.S. PAT. & TRADEMARK OFF., PATENT ELIGIBILITY SUBJECT MATTER: REPORT ON VIEWS AND RECOMMENDATIONS FROM THE PUBLIC 34–38 (July 2017), https://www.uspto.gov/sites/default/files/documents/101-Report_FINAL.pdf [hereinafter USPTO Eligibility Report 2017] (describing criticism of recent § 101 jurisprudence from members of the life sciences and computational communities).

exceptions that are embedded into *all* inventions at some level of abstraction,¹⁹ and (2) inherently overlapping with the other substantive patentability doctrines.²⁰ As an alternative, urgently warranted framework to assess patent eligibility, this Note proposes a revised notion of patent eligibility anchored in *utility* doctrine, tethered to the word “useful” already present in § 101. To this end, Part II provides a history of patent eligibility jurisprudence to contextualize the evolution of the judicial exceptions. Part III summarizes *American Axle* and uses it as a paradigmatic example to analyze the problems with the *Mayo/Alice* test. Part IV proposes a method of assessing patent eligibility under § 101 to supplant the *Mayo/Alice* test, arguing that the word “useful” is sufficient to serve a scope-limiting function within all patent claims.

II. HISTORY OF PATENT ELIGIBILITY JURISPRUDENCE

Section 101 of the Patent Act was viewed originally as a minimum, low-bar threshold to patentability,²¹ as the text of the statute does not explicitly exclude any areas of subject matter from patent protection.²² However, over time, three judicially-added exceptions to § 101 came to be recognized as patent-ineligible: “laws of nature, natural phenomena, and abstract ideas.”²³ Many scholars trace the origins of these ineligible concepts to a set of nineteenth century cases that first disavowed the eligibility of “principles.”²⁴

19. See U.S. PAT. & TRADEMARK OFF., PATENT ELIGIBLE SUBJECT MATTER: PUBLIC VIEWS ON THE CURRENT JURISPRUDENCE IN THE UNITED STATES 16 n.130 (June 2022), <https://www.uspto.gov/sites/default/files/documents/USPTO-SubjectMatterEligibility-PublicViews.pdf> [hereinafter USPTO Eligibility Report 2022].

20. Many refer to § 102, § 103, and § 112 as the “substantive” patentability doctrines, drawing a line between these concepts and the § 101 standard. See, e.g., Syed, *supra* note 8, at 1960.

21. 35 U.S.C. § 101. Section 101 has explicitly read on a version of the four, present-day *eligible* categories (processes, machines, manufactures, and compositions of matter) since 1793. Syed, *supra* note 8, at 2030–31 (compiling the present-day text of § 101 and all previous versions of the statute in Table 1).

22. Rodney Swartz, *Separating Preemption from the Subject Matter Analysis of 35 U.S.C. § 101*, 61 SANTA CLARA L. REV. 903, 917 (2021). Indeed, the Court has interpreted the use of the word “any” to indicate the intent of Congress for an expansive, liberal approach to patentability, with the other substantive patentability doctrines functioning to more rigorously assess the extent of innovation. See *Diamond v. Chakrabarty*, 447 U.S. 303, 308 (1980).

23. *Berkheimer v. HP Inc.*, 890 F.3d 1369, 1374 (2018) (Lourie, J., concurring). These three ineligible subject matter areas are, in their very terms, well-known and regurgitated in countless opinions. Assessing whether (and why) a claim is “directed to” one or more of these “ineligible concepts” is much more elusive.

24. See, e.g., Jeffrey A. Lefstin, *Inventive Application: A History*, 67 FLA. L. REV. 565, 569–70 (2016); Peter S. Menell, *Forty Years of Wondering in the Wilderness and No Closer to the Promised Land: Bilski’s Superficial Textualism and the Missed Opportunity to Return Patent Law to Its Technology Mooring*, 63 STAN. L. REV. 1289, 1295–96 (2011); OREN BRACHA, OWNING IDEAS: THE

But in the decades since, these formative cases were misinterpreted, stretched, and applied beyond what courts previously anticipated.

A. EARLY NOTIONS OF INELIGIBLE “PRINCIPLES”

The Court first struggled with the notion of patent ineligibility in the context of so-called “principles.”²⁵ Questions as to the patentability of “principles” lurked in the background of early English patent law,²⁶ but came to a head in an 1841 case from the Court of Exchequer, *Neilson v. Harford*.²⁷ Neilson’s patent was directed to “the improved application of air to produce heat in fires, forges, and furnaces.”²⁸ Wrestling with a tension between (1) the “principle” that hot air more efficiently promotes ignition than cold air and (2) the “application” of injecting that hot air into a furnace, the court articulated that the patent was valid for claiming “not merely . . . a principle, but a machine embodying a principle.”²⁹ This became *Neilson*’s legacy—a principle may be eligible for patent protection to the extent that it is “embodied” or applied in some form.³⁰ But—as others have noted—the true dispute of *Neilson* was, surprisingly, related to the adequacy of *disclosure*, rather than patent eligibility.³¹

In parallel, early American patent jurisprudence had already suggested that patents could only claim “the contrivance or production of something which did not exist before,” which one might interpret as excluding “principles” as ineligible subject matter.³² In 1852, the Court solidified this idea in *Le Roy v. Tatham*,³³ using *Neilson* to more explicitly draw a line of ineligibility.³⁴ The patent in *Le Roy* involved improvements to the manufacture of wrought pipe.³⁵

INTELLECTUAL ORIGINS OF AMERICAN INTELLECTUAL PROPERTY, 1790–1909 261–72 (2016).

25. See BRACHA, *supra* note 24, at 265; Syed, *supra* note 8, at 1961.

26. Lefstin, *supra* note 24, at 578 (describing James Watt’s patent relating to steam engine improvements, and the associated debate over the notion of “patentable manufacture[s] or . . . ‘unorganized principles’” (citing *Boulton & Watt v. Bull*, 126 Eng. Rep. 651, 655 (1795))).

27. 151 Eng. Rep. 1266 (1841).

28. THOMAS WEBSTER, REPORTS AND NOTES OF CASES ON LETTERS PATENT FOR INVENTIONS 273 (1844) (reprinting *Neilson*’s patent). Patents at this primordial stage lacked formal claims, leaving litigants and courts to infer the scope of patented subject matter based on “a holistic examination of the specification and . . . the actual embodiments or experiments carried out by the patentee.” Lefstin, *supra* note 24, at 580.

29. *Neilson*, 151 Eng. Rep. at 1273.

30. See Lefstin, *supra* note 24, at 570.

31. See, e.g., Lefstin, *supra* note 24, at 581–82 (describing how the bulk of the argument in *Neilson* was directed to whether the actual heating apparatus of the invention was sufficiently enabled, a modern § 112 question).

32. *In re Kemper*, 14 F. Cas. 286, 288 (C.C.D.C. 1841); see Lefstin, *supra* note 24, at 594.

33. 55 U.S. 156, 175 (1852).

34. Lefstin, *supra* note 24, at 594.

35. 55 U.S. at 172–73.

As in *Neilson*, the case was much less directed to the fundamental eligibility of “principles,” but rather, what would today be defined as a claim construction issue.³⁶ Again, the Court articulated a philosophical view that would later anchor all of patent eligibility doctrine: “[a] principle is not patentable,” given that principles are “fundamental truth[s] . . . [that] no one can claim . . . [as] an exclusive right,” while “[a] new property . . . *when practically applied* . . . is patentable.”³⁷ And in 1854, in *O’Reilly v. Morse*,³⁸ the Court used *Neilson* to assert that “a principle [is] not patentable,”³⁹ but a “new application of a known principle” is.⁴⁰ This case, like *Neilson*, was not about patent eligibility, but rather, enablement.⁴¹

Thus, *Neilson*, *Le Roy*, and *Morse* came to stand for the notion that some form of “embodiment” or “practical application” could restore eligibility to an otherwise unpatentable “principle.” But three key issues lingered in the background: (1) no court had defined what a “principle” was;⁴² (2) no court had articulated a degree of “embodiment” or “practical application” required to restore eligibility to an otherwise unpatentable principle;⁴³ and (3) every court had proffered a notion of patent eligibility in cases that were truly about something else⁴⁴—enablement (*Neilson*, *Morse*) or claim construction (*Le Roy*). Given this precarious foundation—and potential anchoring in *other* patentability doctrines—courts were bound to later struggle with the meaning of patent ineligibility.

B. THE EMERGENCE OF TWO INELIGIBILITY TRACKS

In the first half of the twentieth century, “principles” remained ineligible, and “practical applications” of those principles remained eligible.⁴⁵ Professor

36. See Lefstin, *supra* note 24, at 595. The claim construction issue was whether the object of the patent was the application of a newly discovered principle or the machinery used in the pipe manufacture. *Id.*

37. *Le Roy*, 55 U.S. at 175 (emphasis added).

38. 56 U.S. 62 (1854).

39. *Id.* at 115.

40. *Id.* at 131.

41. Lefstin, *supra* note 24, at 596–97.

42. See Syed, *supra* note 8, at 1964 (“[W]ithout any clear rationale for *why* unapplied principles are ineligible, it becomes harder to know *what* falls within or outside the restriction.”).

43. See *id.* at 1963–64 (noting the ambiguity in three possible interpretations of *Neilson*, distinctly applied across the patent eligibility cases that followed).

44. See Michael Risch, *America’s First Patents*, 64 FLA. L. REV. 1279, 1283 (2012) (noting the flaws in the reliance on English law to construct U.S. patent law, and that “judicial discussion of principles almost never related to attempts to patent natural phenomena, but instead related to patent construction”).

45. See Lefstin, *supra* note 24, at 609.

Lefstin provides a comprehensive summary of eligibility-oriented case law and commentary during this time period, noting that the basic, minimal eligibility standard established by *Neilson* was left undisturbed.⁴⁶ But these years also watched the elusive concept of a “principle” evolve,⁴⁷ tracking the evolution of the “useful Arts”—the presumed object of the patent system.⁴⁸ The twentieth century welcomed radical progress in science and technology, which spawned areas of subject matter that intersected with “principles” in ways that muddled *Neilson*’s standard beyond comprehension. This history led to the emergence of two tracks of ineligibility: inventions that are (1) too “natural” or (2) too “formulaic.”

1. *Track One: “Natural” Ineligibility*

The “principles” of the nineteenth century trilogy were, at their core, mere correlations that reflected the relationships between factors: *Neilson*, between heated air and ignition efficiency;⁴⁹ *Le Roy*, between heated lead and wrought pipe continuity;⁵⁰ and *Morse*, between galvanic current and distanced character printing.⁵¹ These correlations represent pivotal discoveries and developments of the Industrial Revolution, which spawned inventions that we would now view as highly “mechanical”- and “materials”-oriented,⁵² and thus, for an unknown reason, de facto eligible for patent protection—assuming some degree of “practical application.” But these new industrial processes and machines soon launched an era of unprecedented scientific and technological discovery in newer, different areas, such as agriculture, biotechnology, chemistry, and medicine.⁵³ And in lockstep, interest in patent protection began

46. *See id.* at 609–23.

47. *See id.* at 609 (“pure scientific explanation”) (internal quotation marks omitted), 611 (“scientific or mathematical truth”), 612 (“mental steps or processes” and “purified natural products”), 616 (“process of nature”), 617 (“natural phenomena”), 619 (“natural law”).

48. MENELL ET AL., *supra* note 2, at 36 (noting that “useful Arts,” as written into the Constitution’s Intellectual Property Clause, encompass “what we would today call technology and scientific discovery”).

49. *See* 151 Eng. Rep. at 1273.

50. *See* 55 U.S. at 164.

51. *See* 56 U.S. at 77–78, 114–16.

52. *See* Risch, *supra* note 44, at 1308–10 (listing categories of “historical” patents on mills, steam, plows, pumps, leather, brick, wood, and more, falling into the “mechanical”- and “materials”-oriented category).

53. *See generally* Thomas Philbeck & Nicholas Davis, *The Fourth Industrial Revolution: Shaping a New Era*, COLUMBIA SIPA J. INT’L AFFAIRS (Jan. 22, 2019), <https://jia.sipa.columbia.edu/fourth-industrial-revolution-shaping-new-era> (recounting the impact of the iterative eras of industrial revolution on science and technology).

to grow for a set of resulting inventions that were intertwined with not just “principles,” but now, “nature.”⁵⁴

In 1948, *Funk Brothers Seed Co. v. Kato Inoculant Co.*⁵⁵ brought a “natural” invention before the Court, and generated a new version of eligibility doctrine.⁵⁶ The patentee claimed a composition of matter, comprising a favorable combination of bacterial strains.⁵⁷ This invention involved the exploitation of certain “qualities” of bacterial species; that is, their lack of mutual inhibition.⁵⁸ As in *Neilson*, *Le Roy*, and *Morse*, the patent in *Funk Brothers* intersected with a mere correlation—here, between species-specificity and mutual inhibition. And just like the “principles” of the Industrial Revolution, this correlation is one that might simply be defined as a “fundamental truth.”⁵⁹ But unlike the nineteenth century case trilogy, the Court in *Funk Brothers* fixated on the “natural” element of the correlation. Justice Douglas referred to the bacterial “qualities” of mutual non-inhibition as “natural” in almost every possible permutation: “the work of nature,” a “law of nature,” a “phenomenon of nature,” “nature’s secret,” a “natural principle,” “natural functioning,” and “perform[ance] in [a] natural way.”⁶⁰ He then asserted that such “natural” qualities were “part of the storehouse of knowledge of all men . . . free to all men and reserved exclusively to none.”⁶¹ Against this nature-oriented backdrop, the Court noted that while the claimed invention was an “application” of a “natural principle”—in fact, one that was also “new and useful”—it did not “satisfy the requirements of *invention or discovery*.”⁶² At first blush, one might think Justice Douglas was reading in “invent[ion] or

54. See MENELL ET AL., *supra* note 2, at 169.

55. 333 U.S. 127 (1948).

56. See Syed, *supra* note 8, at 1965 n.95. While “natural” inventions had previously been discussed, it was only *Funk Brothers* that rose to the level of disturbing *Neilson*’s eligibility standard. See Lefstin, *supra* note 24, at 609–23.

57. *Funk Brothers*, 333 U.S. at 138 n.1 (presenting a representative claim from the asserted patent). At the time of the invention, farmers often inoculated crops with bacterial strains to support plant growth, leveraging the symbiotic, nitrogen-fixing properties of *Rhizobium* species. *Id.* at 128–29. However, a single *Rhizobium* species was typically symbiotic with only certain types of crops, meaning that those growing multiple crop types needed to customize inoculants with different bacterial species. *Id.* Combining multiple *Rhizobium* species into a single inoculant product failed because different bacterial species were mutually inhibitive. *Id.* at 129–30. Uniquely, the claimed inoculant combined a specific set of strains, across bacterial species, that were mutually *non-inhibitive*, and therefore favorable for use in agricultural applications. *Id.* at 130.

58. *Id.*

59. *Le Roy v. Tatham*, 55 U.S. 156, 175 (1852).

60. *Funk Brothers*, 333 U.S. at 130–32.

61. *Id.*

62. *Id.* at 131 (emphasis added).

discover[y]” from the eligibility statute itself.⁶³ But to support this assertion, he cited *Cuno Engineering Corp. v. Automatic Devices Corp.*,⁶⁴ a 1941 case that established a substantive “inventive step” or “non-obvious” requirement for patentability—now codified as § 103.⁶⁵ This grafted an aggressively elevated standard on top of *Neilson’s* original articulation.⁶⁶ The mere application of a “principle”—now, cast as a “natural” concept—would only be patent-eligible if also “inventive.”⁶⁷

Scholars have since noted the drastic impact of *Funk Brothers’* new “inventive application” standard on eligibility doctrine.⁶⁸ And in parallel, a fledgling policy rationale for patent *ineligibility* had also emerged, swirling around the fear of patenting “nature.”⁶⁹ Tellingly, in the *Funk Brothers* concurrence, Justice Frankfurter expressed concerns with the introduction of “vague and malleable terms” such as “‘the work of nature’ and the ‘laws of nature,’” because “[e]verything that happens may be deemed ‘the work of nature,’ and any patentable composite exemplifies in its properties ‘the laws of nature.’”⁷⁰ Still, however, “nature” was irreversibly introduced into the vocabulary of patent ineligibility, where it would remain.

2. *Track Two: “Formulaic” Ineligibility*

As inventions continued to circulate around “nature” at the precipice of the biotechnology revolution, simultaneous developments carried another category of patents into the digital age, introducing a new swath of patent eligibility problems.⁷¹ So, the Court articulated a new type of ineligible subject matter in 1972, in *Gottschalk v. Benson*.⁷² The claims in *Benson* were directed to methods of converting binary coded decimal numbers into pure binary numerals.⁷³ This invention was related to a concept that the Court marked as “ineligible”: the “mathematical procedure,” or “algorithm,” used for the

63. 35 U.S.C. § 101 (“Whoever *invents or discovers* any new and useful process, machine, manufacture, or composition of matter, or any new and useful improvement thereof, may obtain a patent therefor.”) (emphasis added).

64. 314 U.S. 84 (1941).

65. Syed, *supra* note 8, at 1968–69.

66. Lefstin, *supra* note 24, at 623.

67. *See id.* at 629.

68. *See id.* at 631 (“[C]ommentary in the immediate wake of *Funk Brothers* recognized its true nature . . . [having] demanded inventive application as a condition of patentability.”).

69. Syed, *supra* note 8, at 1965, 1967–68.

70. *Funk Brothers*, 333 U.S. at 134–35 (Frankfurter, J., concurring).

71. *See* MENELL ET AL., *supra* note 2, at 268.

72. 409 U.S. 63, 70–72 (1972).

73. *Id.* at 73–74 (excerpting representative claims from the asserted patent).

conversion.⁷⁴ Citing *Le Roy*, *Morse*, and *Funk Brothers*, the Court invalidated the patent, treating the “algorithm” as an “abstract intellectual concept[]” that was a “basic tool[] of scientific and technological work.”⁷⁵ Herein was a new, more explicit articulation of what the Court perhaps feared most for patents directed to ineligible subject matter: a patent with “no substantial practical application except in connection with a digital computer . . . would *wholly pre-empt* the mathematical formula.”⁷⁶ Interestingly, the Court proposed a method of restoring eligibility to the offending “algorithm” that felt lighter than that of *Funk Brothers*:⁷⁷ “[t]ransformation and reduction of an article ‘to a different state or thing’” might lend eligibility to claims not directed to “particular machines,” e.g., non-physicalized “algorithms.”⁷⁸ This case was the first time that an “abstract idea” was explicitly excluded as ineligible subject matter.⁷⁹

Then, in 1978, the Court expanded *Benson* and the exclusion of algorithms in *Parker v. Flook*.⁸⁰ The claims in *Flook* were directed to methods of updating an alarm limit based on present values.⁸¹ Again, this invention involved a freshly “ineligible” concept: the “mathematical algorithm or formula” used to calculate the updated alarm limit value.⁸² But instead of offering *Benson*’s relaxed “transformation” suggestion to restore eligibility to such an “algorithm,” the Court in *Flook* proposed an “inventive application” standard, much like in *Funk Brothers*.⁸³ And, incorrectly, the Court linked this proposal

74. *Id.* at 65, 67.

75. *Id.* at 67.

76. *Id.* at 71–72 (emphasis added); see also Mark A. Lemley, Michael Risch, Ted Sichelman & R. Polk Wagner, *Life After Bilski*, 63 STAN. L. REV. 1315, 1343, 1339–41 (emphasizing pre-emption considerations in § 101 analyses); Joyce C. Li, *Preemption, Diagnostics, and the Machine-or-Transformation Test: Federal Circuit Refinement of Biotech Method Eligibility*, 32 BERKELEY TECH. L.J. 379, 408 n.194, 408–11 (2017) (same).

77. See Syed, *supra* note 8, at 1971.

78. *Benson*, 409 U.S. at 70. Professor Syed details how the Court’s perception of patent-ineligible subject matter developed against a backdrop of overly “physicalist” concerns, rooted in a fundamental misconception as to the true “object” of a patent right. See Syed, *supra* note 8, 1958–61, 1977–80. This problem is *particularly* salient for digital age inventions that heavily intersected with mathematical formulas and algorithms, given their inherently “intangible” nature. See *id.* at 1969–72.

79. Lefstin, Menell & Taylor, *supra* note 13, at 559 n.29.

80. 437 U.S. 584 (1978).

81. *Id.* at 585.

82. *Id.* at 586.

83. *Id.* at 594; Lefstin, *supra* note 24, at 641 (“[T]he core of *Flook*[] [is] the difference between an unpatentable principle and a patent-eligible invention is invention in the application. But *Flook* took this core from *Funk Brothers*.”).

to *Neilson*.⁸⁴ Together, *Benson* and *Flook* laid the foundation for a second track of ineligible subject matter: “algorithms” or “formulas.”⁸⁵

C. EVOLUTION OF THE *MAYO/ALICE* TWO-STEP TEST

In the early 1980s, the Court redrew the patent eligibility standard yet again, first with a “natural,” then a “formulaic” invention. In *Diamond v. Chakrabarty*, the asserted patent claimed a *Pseudomonas* bacterium expressing at least two types of hydrocarbon-degrading plasmids.⁸⁶ As in *Funk Brothers*, a “natural” concept lurked within: the idea that bacteria are “products of nature.”⁸⁷ Here, for the first time, the Court expressed a version of the modern-day judicial exceptions—“laws of nature, physical phenomena, and abstract ideas”—apparently distilling these three areas out of *Flook*, *Benson*, *Funk Brothers*, *Morse*, and *Le Roy*.⁸⁸ But unlike the elevated standard of *Funk Brothers* and *Flook*, the Court in *Chakrabarty* adopted a relaxed approach, holding that because the patentee’s bacterium had “markedly different characteristics from [bacteria] found in nature”—that is, the carriage of plasmids—the invention was directed to patentable subject matter.⁸⁹ Indeed, Chief Justice Burger went so far as to assert that Congress intended for “anything under the sun that is made by man” to be patentable.⁹⁰ This decision opened up patent eligibility for genetically modified organisms,⁹¹ which for many years would lie at the heart of the biotechnology revolution.

The following year, the Court applied a similarly lowered eligibility standard to a “formulaic” invention. In *Diamond v. Diehr*, the asserted patent claimed methods for molding synthetic rubber compounds.⁹² As in *Benson* and

84. Others have provided far more comprehensive accounts of the *Flook* Court’s mistaken reading of *Neilson*. See, e.g., MENELL ET AL., *supra* note 2, at 286–87; Lefstin, *supra* note 24, at 581–87.

85. See Syed, *supra* note 8, at 1969–72. Notably, the Court also did not clarify how these concepts were distinct from the “natural” correlations of *Neilson*, *Le Roy*, *Morse*, and *Funk Brothers*. Professor Syed notes the confusion in the canonical listing of the three judicial exceptions (laws of nature, natural phenomena, and abstract ideas) in that both “natural phenomena” and “abstract ideas” include concepts within the “laws of nature” category, and instead suggests that these ineligibility categories should be termed “laws of nature, products of nature, and abstract formulas.” *Id.* at 1977–78.

86. 447 U.S. 303, 305 (1980).

87. *Id.* at 306.

88. *Id.* at 309 (citing this exact set of cases).

89. *Id.* at 310 (emphasis added).

90. *Id.* at 309 (relying on the legislative history of the 1952 Patent Act).

91. MENELL ET AL., *supra* note 2, at 287.

92. 450 U.S. 175, 220 n.5 (1981) (excerpting representative claims from the asserted patent).

Flook, the methods involved a “formula.”⁹³ But the Court seemed to slightly step back from the heightened standards of either case, although not quite to the leniency of *Neilson*, *Le Roy*, and *Morse*. First, Justice Rehnquist framed the current version of the three judicial exceptions to patent eligibility—“laws of nature, natural phenomena, and abstract ideas”—citing to *Flook*, *Benson*, and *Funk Brothers*.⁹⁴ He also noted that “novelty” . . . is of no relevance in determining whether the subject matter of a claim falls within the § 101 categories of possibly patentable subject matter,⁹⁵ apparently overruling the “inventive application” concept of *Funk Brothers* or *Flook*.⁹⁶ And finally, he affirmed a *Benson*-esque “transformation” requirement for “formulaic” inventions, asserting that “[limit[ing]] the use of a formula to a particular technological environment” could not overcome the presumption of ineligibility.⁹⁷ Undoubtedly, *Diehr* introduced immense confusion as to the status of *Flook*.⁹⁸ But just as *Chakerabarty* expanded the eligibility of “natural,” biotechnology-type inventions, *Diehr* did the same for “formulaic,” software-type inventions.

Predictably, biotechnology and software patents exploded in the years following *Chakerabarty* and *Diehr*.⁹⁹ Unfortunately, however, the newly relaxed standards for patent eligibility also inspired an influx of internet-related and business method patents, many of which were met with disdain.¹⁰⁰ This anchored the later enactment of the *Mayo/Alice* test, as the Court attempted to reel in patent issuance with yet another redefined eligibility standard.

After three decades of silence as to patent eligibility, in 2010, the Court granted review of a Federal Circuit case on a “formulaic” business method patent. In *Bilski v. Kappos*,¹⁰¹ the asserted patent claimed a business method for

93. *Id.* at 187.

94. *Id.* at 185.

95. *Id.* at 188–89.

96. MENELL ET AL., *supra* note 2, at 288.

97. *Diehr*, 450 U.S. at 191–92 (stating that an unpatentable principle cannot be transformed into a patentable process without significant “postsolution activity”).

98. *See id.* at 213–16 (Stevens, J., dissenting).

99. *See* MENELL ET AL., *supra* note 2, at 289. This movement arguably spurred the development of startup culture, as expanding patent portfolios lent credibility to early-stage business endeavors, helping many small companies attract investors. *Id.*

100. *Id.*; Syed, *supra* note 8, at 1974. Many have argued that both software and business method inventions should be viewed as categorically incapable of meeting a threshold of patentability. *See, e.g.*, Ognjen Zivojnovic, *Patentable Subject Matter After Alice—Distinguishing Narrow Software Patents from Overly Broad Business Method Patents*, 30 BERKELEY TECH. L.J. 807, 808–09 (2015); Issie Lapowsky, *EFF: If You Want to Fix Software Patents, Eliminate Software Patents*, WIRED (Feb. 25, 2015, 9:00 PM), <https://www.wired.com/2015/02/eff-eliminate-software-patents/>.

101. 561 U.S. 593 (2010).

“managing the consumption risk costs of a commodity,” that is, hedging risks during trading of commodities. The Court invalidated the patent as ineligible subject matter, stating that the claims merely reduced “the basic concept of hedging,” into “a mathematical formula . . . just like the algorithms at issue in *Benson* and *Flook*.”¹⁰² The Court further held that granting such a patent “would pre-empt use of th[e] approach in all fields.”¹⁰³ With this, the Court reilluminated *Benson*’s “pre-emption” fears, and resurrected a higher, pre-*Diebr* standard for assessing the eligibility of “formulaic” inventions. The next four years saw three decisions that solidified this version of the eligibility standard for two “natural” inventions and one “formulaic” invention.

First, in *Mayo*, the asserted “natural” patent was directed to drug dose optimization, specifically, determining the appropriate dosage level of a nucleoside analog to treat inflammatory bowel disease.¹⁰⁴ The Court asserted that the claims recited a “law of nature”—the correlation between the thiopurine drug dosage administered to a patient, the resulting toxic metabolites produced in their body, and the overarching toxicity and/or therapeutic efficacy of treatment.¹⁰⁵ Then, the Court asked “whether the claims do significantly more than simply describe these natural relations,” suggesting a search for a *Funk Brothers*-esque “inventive concept” to “transform an unpatentable law of nature into a patent-eligible application of such a law.”¹⁰⁶ Failing to find one, the Court held that the asserted claims were ineligible under § 101.¹⁰⁷

Then, in *Myriad*, the asserted “natural” patent was directed to BRCA1 genomic DNA (claim 1) and cDNA (claim 2).¹⁰⁸ Based on the Court’s subjective interpretation of what “naturally occurring” meant for the exons and introns of DNA, it held that claim 2—involving cDNA—did “not present the same obstacles to patentability as *naturally* occurring, isolated DNA segments” and was thus eligible under § 101, falling outside of the scope of

102. *Id.* at 611.

103. *Id.* at 612.

104. *Mayo Collaborative Servs. v. Prometheus Lab’s, Inc.*, 566 U.S. 66, 72, 77, 82 (2012).

105. *See id.* at 76.

106. *See id.* at 77, 82.

107. *See id.*

108. *Ass’n for Molecular Pathology v. Myriad Genetics, Inc.*, 569 U.S. 576, 584 (2013). BRCA1 is a gene that encodes for a protein that researchers have linked to the development of breast cancer. *Id.* The claims of the asserted patent in *Myriad* were not formally construed, so this interpretation is based on the district court’s presumption that claim 1 was directed to “naturally occurring” DNA. *Ass’n for Molecular Pathology v. U.S. Pat. & Trademark Off.*, 702 F. Supp. 2d 181, 217 (S.D.N.Y. 2010).

the judicial exceptions.¹⁰⁹ The “naturally occurring” DNA of claim 1, on the other hand, was deemed to be an ineligible product of nature (encompassed within the “natural phenomena” judicial exception).¹¹⁰ As in *Mayo*, the Court held that the “isolation” of BRCA1 genomic DNA involved the mere *non-inventive* separation of the DNA from its flanking genomic regions, and therefore invalidated claim 1 under § 101.¹¹¹

Finally, in *Alice*, the asserted “formulaic” patent was directed to a business method of mitigating settlement risks by using third-party intermediaries.¹¹² Drawing a parallel to the formulaic business method of *Bilski*, the Court held that the invention was “drawn to the abstract idea of intermediate settlement,” and that there was no “inventive concept” present in its application.¹¹³ Invalidating the claims under § 101, the Court emphasized the foreboding concept of a “fundamental truth”¹¹⁴ and the attendant “pre-emption concern”¹¹⁵ for the final time, and then crystallized the current two-step *Mayo/Alice* test.¹¹⁶

The test is as follows: first, one must determine if the asserted claim is directed to a patent-ineligible concept, i.e., a law of nature, natural phenomenon, or abstract idea (step one).¹¹⁷ If so, the claim is presumptively ineligible for patent protection, unless one can identify some “inventive concept” embodied within the claimed invention that is “sufficient to ensure that the patent in practice amounts to *significantly more* than a patent upon the ineligible concept itself,” beyond “well-understood, routine, conventional activity” (step two).¹¹⁸

After *Alice* in 2014, the Court stepped away from patent eligibility, leaving district courts and the Federal Circuit to apply the *Mayo/Alice* test on their own.¹¹⁹ This has not gone well. Many view the new *Mayo/Alice* test as

109. *Myriad*, 569 U.S. at 594–95 (emphasis added). The *Myriad* Court has received intense criticism of its scientific outlook; however, it was *technically* correct at least in noting that the “natural” genomic DNA of claim 1 is measurably distinct from cDNA. The former contains both coding (exon) and non-coding (intron) regions, and the latter contains only coding (exon) regions and is typically produced by laboratory technicians.

110. *Id.* at 579.

111. *Id.* at 591.

112. *Alice Corp. v. CLS Bank Int’l*, 573 U.S. 208 (2014).

113. *Id.* at 212.

114. *Id.* at 220.

115. *Id.* at 223.

116. *See id.* at 217–18.

117. *Id.* at 217 (citing *Mayo Collaborative Servs. v. Prometheus Lab’y, Inc.*, 566 U.S. 66, 77 (2012)).

118. *Id.* at 217–18 (quoting *Mayo*, 566 U.S. at 72–73).

119. Lefstin, Menell & Taylor, *supra* note 13, at 555–56.

antagonistic towards innovation, given the cabined scope of eligibility for certain technologies.¹²⁰ To summarize the problems with the *Mayo/Alice* test: on the one hand, it appears that the test stems from a protracted misunderstanding of the nineteenth century case law on patenting “principles.” On the other, the test seems to capture the Court’s sensible ambivalence towards “natural” or “formulaic” inventions that threaten to “wholly preempt”¹²¹ something that is a “fundamental truth”¹²² or “part of the storehouse of knowledge of all men.”¹²³ Such inventions run the risk of limiting, rather than promoting innovation, and are thus incompatible with the fundamental purpose of the patent system.¹²⁴ Setting this venerable policy rationale aside, however, the *Mayo/Alice* test left the three issues¹²⁵ from *Neilson*, *Le Roy*, and *Morse* unresolved: (1) the meaning of an invention “directed to” a law of nature, natural phenomenon, or abstract idea is unclear (expanding upon the previous, elusive concept of a “principle”); (2) the degree of “inventive application” required to restore eligibility to one of these otherwise unpatentable judicial exceptions is confusing (replacing the previous, undefined degree of “embodiment” or “practical application”); and (3) the entire eligibility standard remains anchored in a “markedly ahistorical reading” of foundational case law (maintaining the nineteenth century confusion of enablement and claim construction case law into eligibility doctrine).¹²⁶

D. IMPACT OF THE *MAYO/ALICE* TEST ON “NATURAL” AND “DIAGNOSTIC” INVENTIONS

“Diagnostic” inventions provide a particularly useful case study into the problems of the *Mayo/Alice* test, with *American Axle* emerging as a variant at the end of this grouping. The Court’s § 101 jurisprudence has had a striking impact on patent eligibility in this arena because diagnostic inventions are often perceived as “natural”—a biological correlation is typically leveraged to infer

120. See USPTO Eligibility Report 2022, *supra* note 19, at 21, 25, 28. The now-routine ineligibility of life sciences and computational inventions has spurred patent applicants in both areas to express great frustration with the Court’s § 101 jurisprudence. See Lefstin, Menell & Taylor, *supra* note 13, at 555; USPTO Eligibility Report 2017, *supra* note 18, at 34–38. Beyond the disparate impact on these technological areas, the *Mayo/Alice* test would also arguably render some of the “most famous[ly] patented inventions” ineligible today—that is, “historic” patents mostly comprising mechanical inventions embodied in tangible, physical instruments. See Michael Risch, *Nothing is Patentable*, 67 FLA. L. REV. F. 45, 51–53 (2015) (providing examples in Table 3, such as the electric motor, Morse code, or the light bulb).

121. *Gottschalk v. Benson*, 409 U.S. 63, 71–72 (1972).

122. *Le Roy v. Tatham*, 55 U.S. 156, 175 (1852).

123. *Funk Bros. Seed Co. v. Kalo Inoculant Co.*, 333 U.S. 127, 130 (1948).

124. U.S. CONST. art. 1, § 8, cl. 8.

125. See discussion *supra* Section II.A.

126. Lefstin, Menell & Taylor, *supra* note 13, at 560.

a physiological “state” (e.g., disease, resistance, responsiveness) based on the presence of a biological “marker” (e.g., a macromolecule such as DNA, RNA, or protein, or a by-product such as a metabolite derived from an administered drug).¹²⁷

Here, several cases follow a similar pattern. An inventor holds a patent on some form of medical diagnostic strategy, which links an underlying “marker” with the identification of some health-relevant “state.” The patent is then challenged for validity under § 101, and like *Mayo* or *Myriad*, it is invalidated due to its intersection with a “natural” concept. And inventors, patent examiners, scientists, and even judges at the Federal Circuit bemoan the resulting invalidity and emphasize that the invention is remarkably innovative. In *Ariosa Diagnostics, Inc. v. Sequenom, Inc.*, the invalidated patent claimed “methods of using cffDNA,” including “making a diagnosis of certain fetal characteristics based on the detection of paternally inherited cffDNA”—the marker here is cffDNA; the state is the fetal characteristic.¹²⁸ Many viewed this as an extraordinary invention that merited patent protection.¹²⁹ In *Cleveland Clinic Foundation v. True Health Diagnostics LLC*, the invalidated patent claimed “methods for characterizing a test subject’s risk for cardiovascular disease” by measuring endogenous myeloperoxidase levels—the marker, myeloperoxidase; the state, cardiovascular disease.¹³⁰ Again, many felt that this invention should have been patent-eligible; in fact, the PTO had published a hypothetical example—strikingly similar to the claims at issue in *Cleveland Clinic*—that it deemed patent-eligible.¹³¹ And in *Athena Diagnostics, Inc. v. Mayo Collaborative Services, LLC*, the invalidated patent claimed “methods for diagnosing neurological disorders” by detecting anti-muscle-specific tyrosine kinase (MuSK) antibodies—the marker, anti-MuSK antibodies; the state, the

127. Robert M. Califf, *Biomarker Definitions and their Applications*, 243 EXPERIMENTAL BIOLOGY & MED. 213, 213–15. For example, the invasiveness (state) of breast cancer can often be inferred from increased *HER2* oncogene expression (marker). Cristina Grávalos & Amaya Jimeno, *HER2 in Gastric Cancer: A New Prognostic Factor and Novel Therapeutic Target*, 19 ANNALS ONCOLOGY 1523, 1523 (2008); see also N. Lynn Henry & Daniel F. Hayes, *Cancer Biomarkers*, 6 MOLECULAR ONCOLOGY 140 (2012) (describing the identification and use of biomarkers for cancer diagnostics).

128. 788 F.3d 1371, 1373–74 (Fed. Cir. 2015). “cffDNA” refers to cell-free fetal DNA.

129. See, e.g., *id.* at 1380–81 (Linn, J., concurring) (deeming the invention “truly meritorious”).

130. 859 F.3d 1352, 1356 (Fed. Cir. 2017). Myeloperoxidase is an enzyme associated with inflammatory immune responses. Amjad A. Khan, Mohammed A. Alsahli & Arshad H. Rahmani, *Myeloperoxidase as an Active Disease Biomarker: Recent Biochemical and Pathological Perspectives*, 6 MED. SCIS. (BASEL) 1, 1–3 (2018).

131. *Federal Circuit Clarifies PTO Guidance Regarding Subject Matter Eligibility*, MORGAN LEWIS (Apr. 3, 2019), https://www.morganlewis.com/pubs/2019/04/federal-circuit-clarifies-pto-guidance-regarding-subject-matter-eligibility#_ftn4.

neurological disorder.¹³² The “cry for help”¹³³ after *Athena* was even further exaggerated, with eleven amici briefs filed supporting the ultimately unsuccessful certiorari petition.¹³⁴

In all of these cases,¹³⁵ district courts and the Federal Circuit offered similar reasoning in finding ineligibility. Diagnostic methods are necessarily “directed to” laws that judges view as “natural.” To be sure, the concept of a biological marker—the target of medical diagnostics—is theoretically a “natural” one, typically being a macromolecule that is endogenous to the human body and reflective of internal physiology. Thus, all marker-state relationships can be judicially interpreted as “natural” laws under step one of the *Mayo/Alice* test. So, the test proceeds to step two, requiring an “inventive concept” within the invention—where the diagnostics fail. Most diagnostic approaches apply the marker-state relationship to simply assess the state, using routine techniques in the art. And this does not meet the “inventive” standard of the *Mayo/Alice* test.

Of course, a measure of “inventiveness” is already embedded within diagnostic inventions—just not in a format suited to the *Mayo/Alice* test. It is the mere use of the marker-state relationship *to assess the state* that many would deem “inventive”—swallowed entirely by *Mayo/Alice*’s step two.¹³⁶ Thus, medical diagnostic inventions have become presumptively ineligible for patent protection, to the concern of many.¹³⁷ But worse yet is that the medical

132. 915 F.3d 743, 746–47 (Fed. Cir. 2019). Anti-MuSK antibodies are produced in people with conditions such as myasthenia gravis, and are part of an autoimmune response against important endogenous proteins.

133. Petition for Writ of Certiorari, *Athena*, 140 S. Ct. 855 (2020).

134. Dennis Crouch, *Athena v. Mayo: Strong Amicus Support*, PATENTLYO (Nov. 6, 2019), <https://patentlyo.com/patent/2019/11/athena-strong-support.html>.

135. *But see* *Vanda Pharms Inc. v. West-Ward Pharms Int’l Ltd.*, 887 F.3d 1117, 1121 (Fed. Cir. 2018) (providing an exception to the aforementioned set of cases, where the asserted patent claimed “method[s] for treating schizophrenia patients with iloperidone” using patient genotype and cytochrome P450 2D6 metabolic activity—the marker, cytochrome P450 2D6; the state, iloperidone responsiveness). Unlike the other diagnostic cases, the Federal Circuit held that the claims in *Vanda* were patent-eligible, despite a striking parallel between *Vanda* and *Mayo*. *Id.* at 1136.

136. Of course, assessing inventiveness in this way would necessarily involve an examination of the relevant prior art. MPEP § 2141 (9th ed. Rev. 10, June 2020) (providing the examination guidelines for assessing patentability under § 103).

137. *See, e.g.*, Rebecca S. Eisenberg, *Prometheus Rebound: Diagnostics, Nature, and Mathematical Algorithms*, 122 *YALE L.J. ONLINE* 341 (2013) (explaining the uncertain impact of *Mayo* on the subject matter eligibility of molecular diagnostic inventions); Philip Merksamer, *Ariosa Diagnostics v. Sequenom: Metastasis of Mayo and Myriad and the Evisceration of Patent Eligibility for Molecular Diagnostics*, 31 *BERKELEY TECH. L.J.* 495 (2015) (arguing that together, *Mayo* and *Myriad* eviscerated eligibility for molecular diagnostics); Lauren Matlock-Colangelo, *Broadly Unpatentable: How Broad Method Claims Have Limited Patentability of Diagnostic Inventions*,

diagnostics and “natural” correlations of *Ariosa*, *Cleveland Clinic*, and *Athena* are not far from those from the nineteenth century era case law in *Neilson*, *Le Roy*, and *Morse*. The correlations of cffDNA with fetal characteristics,¹³⁸ myeloperoxidase with cardiovascular disease,¹³⁹ and anti-MuSK antibodies with neurological disorders¹⁴⁰ are of the same “natural” quality as the correlations of heated air with ignition efficiency,¹⁴¹ heated lead and wrought pipe continuity,¹⁴² and galvanic character with distanced character printing.¹⁴³ *American Axle* shows that such a correlation can also be drawn out of a “mechanical” invention, suggesting that the wrath of § 101 is not as specific to the life sciences as previously framed.

III. *AMERICAN AXLE*

A. CASE SUMMARY

As in *Ariosa*, *Cleveland Clinic*, and *Athena*, the Federal Circuit invalidated yet another “natural” correlation-based invention under § 101 in *American Axle*.¹⁴⁴

119 COLUM. L. REV. 797 (2019) (addressing the issue of overbroad claims in the diagnostic invention space); Li, *supra* note 76 (summarizing Federal Circuit case law relating to biotechnology inventions with a focus on medical diagnostic inventions). Fundamentally, the complaint is that patent protection is critical for diagnostic inventions, albeit not as conclusively as for pharmaceutical products. See Lefstin, Menell & Taylor, *supra* note 13, at 582–83.

138. *Ariosa Diagnostics, Inc. v. Sequenom, Inc.*, 788 F.3d 1371, 1373–74 (Fed. Cir. 2015).

139. *Cleveland Clinic Found. v. True Health Diagnostics LLC*, 859 F.3d 1352, 1356 (Fed. Cir. 2017).

140. *Athena Diagnostics, Inc. v. Mayo Collaborative Servs., LLC*, 915 F.3d 743, 746–47 (Fed. Cir. 2019).

141. *Neilson v. Harford*, 151 Eng. Rep. 1266, 1273 (1841).

142. *Le Roy v. Tatham*, 55 U.S. 156, 164 (1852).

143. *O’Reilly v. Morse*, 56 U.S. 62, 77–78, 114–16 (1854).

144. 967 F.3d 1285 (Fed. Cir. 2020). There is a complex procedural history leading up to the final Federal Circuit opinion in *American Axle*. American Axle & Manufacturing, Inc. (American Axle) first sued Neapco Holdings LLC (“Neapco”) in 2015, for infringement of U.S. Patent No. 7,774,911. *Am. Axle & Mfg., Inc. v. Neapco Holdings LLC*, 309 F. Supp. 3d 218 (D. Del. 2018). The district court invalidated the American Axle patent. *Id.* Then, American Axle appealed to the Federal Circuit, which affirmed the district court’s holding. *Am. Axle & Mfg., Inc. v. Neapco Holdings LLC*, 939 F.3d 1355 (Fed. Cir. 2019). American Axle then petitioned for both a panel rehearing and a rehearing en banc. The Federal Circuit granted the panel rehearing and withdrew the previous opinion. *Am. Axle & Mfg., Inc. v. Neapco Holdings LLC*, 966 F.3d 1294 (Fed. Cir. 2020). Issuing a modified opinion after rehearing, the Federal Circuit affirmed the district court decision again, with similar reasoning to the initial opinion and a few key analytical changes. *Am. Axle & Mfg., Inc. v. Neapco Holdings LLC*, 967 F.3d 1285 (Fed. Cir. 2020). This is the opinion that this Note will focus on. The same day that the refreshed *American Axle* opinion was issued, the Federal Circuit denied the petition for rehearing en banc in an evenly split, 6-6 vote—yielding 5 different

The asserted patent was directed to “an improved method for damping various types of vibrations in a hollow [drive]shaft.”¹⁴⁵ In an automobile, the driveshaft is the part of the vehicle that connects the transmission to the axle shaft and transmits drive torque (rotary power) from the engine to the wheels.¹⁴⁶ This positioning makes driveshafts vulnerable to vibrations that produce undesirable, disruptive noise for passengers.¹⁴⁷ American Axle’s invention addressed this problem with the application of liners to driveshafts to attenuate those vibrations.¹⁴⁸ Liners are susceptible to vibration—just as driveshafts are—but may vibrate at different frequencies, allowing for dampening of the vibration of the driveshafts that they might embrace.¹⁴⁹ Leveraging this concept, the claimed method of manufacture in the asserted patent included a step where variables of the liners (e.g., mass, stiffness) are tuned¹⁵⁰ to alter the ability for the liner to dampen vibration in the driveshaft.¹⁵¹

The Federal Circuit invalidated American Axle’s patent using the *Mayo/Alice* test.¹⁵² Under step one, Judge Dyk identified the claims to be directed to “use of a natural law of relating frequency to mass and stiffness—

opinions written by Judges Dyk, Chen, Newman, Stoll, and O’Malley. *Am. Axle & Mfg., Inc. v. Neapco Holdings LLC*, 966 F.3d 1347 (Fed. Cir. 2020). American Axle then filed a petition for certiorari to the Supreme Court, but as in every patent eligibility case since *Alice*, the petition was denied. *Am. Axle & Mfg., Inc. v. Neapco Holdings LLC*, 142 S.Ct. 2902 (2022).

145. U.S. Patent No. 7,774,911 col. 2 ll. 40–41. The specification describes three types of vibration that can occur: bending (lateral) mode, torsion mode, and shell mode. *Id.* at col. 1 ll. 42–43.

146. *See id.* at col. 4 ll. 20–31.

147. *See id.* at col. 1 ll. 8–27, 39–43 (noting that driveshafts are typically made of a “relatively thin-walled steel or aluminum tubing and as such, can be receptive to various driveline excitation sources . . . [that] cause the [drive]shaft to vibrate”).

148. Brad M. Scheller, Andrew H. DeVoogd & Matthew A. Karambelas, *SCOTUS Declines to Answer Calls for Clarification in American Axle v. Neapco*, NAT’L L. REV. (July 13, 2022), <https://www.natlawreview.com/article/scotus-declines-to-answer-calls-clarification-american-axle-v-neapco>.

149. *See* U.S. Patent No. 7,774,911 col. 7–8.

150. In the initial claim construction hearing, the court construed the “tuning” terms of U.S. Patent No. 7,774,911 (hereinafter, “the ’911 patent”) to mean “controlling characteristics [of a liner, e.g., mass or stiffness] . . . to match a relevant frequency or frequencies to reduce at least two types of vibration.” Put differently, here, to “tune” a liner refers to altering certain properties of it (e.g., mass, stiffness) for an expected output (e.g., frequency of vibration). *Am. Axle & Mfg., Inc. v. Neapco Holdings LLC*, No. 15-1168-LPS, 2017 WL 1334733, at *5 (D. Del. Apr. 7, 2017).

151. U.S. Patent No. 7,774,911 col. 10 ll. 10–27. Independent claims 1, 22, and 36 of the ’911 patent were all directed to attenuating two forms of vibration: shell mode and bending mode. *Id.* at col. 10, ll. 10–27, col. 11, ll. 24–36, col. 12, ll. 32–47.

152. *Am. Axle & Mfg., Inc. v. Neapco Holdings LLC*, 967 F.3d 1285 (Fed. Cir. 2020).

i.e., Hooke’s law.”¹⁵³ The court held that representative claim 22 was “directed to a natural law because it clearly invokes a natural law, and *nothing more*, to accomplish a desired result.”¹⁵⁴ Moving to step two, the court held that the claims failed to identify an “inventive concept,” highlighting again that they did “nothing more” than instruct one to apply Hooke’s law when designing a driveshaft liner to reduce vibration.¹⁵⁵

Judge Moore wrote a fervent dissent, arguing that the claims were in fact not directed to Hooke’s law under step one, nor did they lack an inventive concept under step two.¹⁵⁶ She described the majority’s new explanation for why claim 22 was directed to Hooke’s law as the freshly introduced “nothing more” test, which inappropriately requires appellate judges to “resolve questions of science *de novo* on appeal,” playing the role of scientific experts.¹⁵⁷

B. ANALYSIS OF THE *MAYO/ALICE* TEST IN *AMERICAN AXLE*

American Axle illustrates that current § 101 eligibility doctrine imposes an elevated hurdle for even mechanical devices that are perceived to be “natural.” Is this outcome acceptable? Should *American Axle*’s driveshaft invention—and others like it—be deemed ineligible for patent protection? This Note submits that the *Mayo/Alice* test has narrowed patent eligibility for the wrong types of inventions, by asking the wrong questions of patent examiners and courts. *American Axle* is a useful case to illustrate the flaws in both steps of the *Mayo/Alice* test, justifying its removal and replacement with an alternative standard under § 101. This Section will first present the problems with the *Mayo/Alice* test, as demonstrated by *American Axle*, to advocate for its removal from eligibility doctrine. Part IV will then propose a replacement standard under § 101.

1. Step One of *Mayo/Alice* and the Ineligibility Bars

Step one of the *Mayo/Alice* test sets out an inescapable trap in asking if the asserted claim is directed to a law of nature, natural phenomenon, or abstract

153. *Id.* at 1294. Hooke’s law is a formula used to calculate the force created by a spring (e.g., a driveshaft) that has been displaced. Michael Oliver, *Greasing the Wheels of Patent Law: Clarifying the Judicial Exceptions via American Axle & Manufacturing, Inc. v. Neapco Holdings LLC*, 29 J. INTELL. PROP. L. 370, 379 (2022). Specifically, it mathematically relates the mass and/or stiffness of an object to its vibration frequency.

154. *American Axle*, 967 F.3d at 1297 (emphasis added).

155. *Id.* at 1298–99.

156. *Id.* at 1305 (Moore, J., dissenting).

157. *Id.* at 1309, 1311 (Moore, J., dissenting).

idea.¹⁵⁸ This “directed to” question can almost always be answered affirmatively. Professor Risch aptly proposes that at some level of abstraction, “every invention will look like an abstract idea or natural phenomenon.”¹⁵⁹ This bears similarity to Judge Newman’s dissent from the denied en banc petition in *American Axle*, stating that so many inventions can simply be “reduc[ed] to mathematical abstractions and algorithms,” or Justice Frankfurter’s concurrence in *Funk Brothers*, asserting that “[e]verything that happens may be deemed ‘the work of nature.’”¹⁶⁰ As detailed in Section II.B, centuries of eligibility jurisprudence led the Court to select laws of nature, natural phenomena, and abstract ideas as the three ineligible concepts, which this Note reclassified as either natural or formulaic. Regardless of the nomenclature for these nebulous ideas, they are interwoven inextricably into most inventions. The invention in *American Axle* simply intersected with yet another “natural” correlation (Hooke’s law), just like the inventions in *Neilson*, *Le Roy*, *Morse*, *Ariosa*, *Cleveland Clinic*, and *Athena*. Like the medical diagnostics, the *marker* in *American Axle* was the liner variable (mass and/or stiffness), correlated with the *state* of vibration intensity. At this level of abstraction, it is hard to imagine an invention without such a correlation.

This Note argues that this problem is enough to warrant the complete removal of the ineligibility bars from the § 101 standard.¹⁶¹ In theory, the Court identified laws of nature, natural phenomena, and abstract ideas to be presumptively ineligible because their standalone patentability runs a greater risk of undue pre-emption, precluding too much valuable follow-on innovation.¹⁶² This, of course, is a very strong policy rationale. But if one can find a law of nature, natural phenomenon, or abstract idea in *any* invention if sufficiently abstracted, what purpose does step one of the *Mayo/Alice* test serve? Notably, the Court never even articulated a persuasive answer to the question of *why* the three ineligible concepts were predisposed to the pre-emption concern.¹⁶³ And while many scholars have proposed alternative, slightly clearer ideological lines along which to partition eligibility¹⁶⁴—*all* the

158. *Alice Corp. v. CLS Bank Int’l*, 573 U.S. 208, 217 (2014) (citing *Mayo Collaborative Servs. v. Prometheus Lab’ys, Inc.*, 566 U.S. 66, 77 (2012)).

159. Risch, *supra* note 120, at 53 (emphasis added).

160. 333 U.S. 127, 134–35 (1948) (Frankfurter, J., concurring).

161. Others have previously proposed and embraced the idea of overruling the ineligibility bars. *See, e.g.*, Risch, *supra* note 120.

162. *Gottschalk v. Benson*, 409 U.S. 63, 71–72 (1972); *Bilski v. Kappos*, 561 U.S. 593, 611–12 (2010); *Alice*, 573 U.S. at 216 (2014).

163. Syed, *supra* note 8, at 1983 n.175.

164. Several scholars have proposed alternative ways of partitioning eligibility. *See, e.g.*, Lefstin, Menell & Taylor, *supra* note 13, at 563 n.50 (with “the result of human effort”); *id.* at 563–64 n.53 (with “physical or tangible form”); *id.* at 564 (with “practical application”); *id.*

existing suggestions are susceptible to the same fundamental problem. Defining a *concept* or a *category* of concepts as presumptively ineligible forces patent examiners and courts to perform an abstraction analysis without the guiderails of experience in that field (i.e., not from the standard perspective of “one of ordinary skill in the art”).¹⁶⁵ There is no avoiding the inherent subjectivity of such an inquiry; therefore, this Note contends that a line of ineligibility cannot and should not ever be drawn.

2. *Step Two of Mayo/Alice and the Redundant Inquiries*

American Axle also illustrates the futility of step two of the *Mayo/Alice* test, which inevitably bleeds into the substantive patentability doctrines of § 103 and § 112 by requiring an assessment of the prior art (§ 103) or an interrogation into the extent of disclosure (§ 112). It is certainly possible that the invention in *American Axle* was truly undeserving of a patent grant under § 103 or § 112. But in carrying out step two of the *Mayo/Alice* test, the Federal Circuit performed covert § 103 and § 112 analyses without the backbone of either statute to inappropriately invalidate *American Axle*’s patent on § 101 grounds.

a) From § 101 to § 103

As discussed in Section II.B.1, the Court in *Funk Brothers* grafted an “inventive application” standard into § 101, which the *Mayo/Alice* test solidified into the step two search for an “inventive concept.”¹⁶⁶ But the vague notion of inventiveness is also seen in § 103 of the Patent Act, which provides that an invention must not “have been *obvious* . . . to a person having ordinary skill in the art to which the claimed invention pertains.”¹⁶⁷ To assess patents

(with a specified list of categories, e.g., discoveries, scientific theories, mathematical methods, aesthetic creations, schemes, etc.); Patent Eligibility Restoration Act of 2022, S. 4734, 117th Cong. (2022) (with an evaluation of “technological” qualities, specifically proposed by Senator Tillis as a possible legislative reform to § 101); Syed, *supra* note 8, at 1981 (with an assessment of an “applied” rather than “basic” quality of the claimed invention). In contrast, the Court partitioned eligibility across a murky ideological line: things that are “fundamental truth[s]” or “part of the storehouse of knowledge of all men” are ineligible; thus, things that are “non-fundamental” or perhaps even “anything under the sun made by man” are eligible. *Le Roy v. Tatham*, 55 U.S. 156, 175 (1852); *Funk Bros. Seed Co. v. Kalo Inoculant Co.*, 333 U.S. 127, 130 (1948); *Diamond v. Chakrabarty*, 447 U.S. 303, 309 (1980).

165. MPEP § 2141 (9th ed. Rev. 10, June 2020).

166. *Alice*, 573 U.S. at 217–18 (quoting *Mayo*, 566 U.S. at 72–73).

167. 35 U.S.C. § 103 (emphasis added). Initially—and controversially—the non-obvious standard was viewed as a requirement for inventive, creative genius. *Cuno Eng’g Corp. v. Automatic Devices Corp.*, 314 U.S. 84, 90–91 (1941). The 1952 codification of the modern § 103 statute took a step back from the “genius” reference, but left the notion of non-obviousness intertwined with that of inventiveness. *Ryko Mfg. v. Nu-Star, Inc.*, 950 F.2d 714, 718 (Fed. Cir. 1991).

under § 103, courts compare claimed inventions against “the scope and content of the prior art” and theorize what one of skill in the pertinent art would obviously or non-obviously dream up.¹⁶⁸ And under step two of the *Mayo/Alice* test, the *American Axle* court did just this. In its very terms, step two of the *Mayo/Alice* test invites a prior art inquiry, asking whether the “inventive concept” is beyond “well-understood, routine, conventional activity already engaged in by the scientific community.”¹⁶⁹ It is hard to see how one could attempt to answer this question without looking to the prior art, mirroring the exact analysis that § 103 prescribes. Indeed, Judge Dyk focused on the extent to which liner manipulation relied upon techniques that were well-known in the prior art, describing the claimed advance as “simply controlling various *known* characteristics of the liner so as to achieve attenuation of two vibration modes . . . [using methods that were] *well known* in the automotive industry.”¹⁷⁰

In other words, the Federal Circuit affirmed that there was no inventive concept in *American Axle* because the implementation of the driveshaft invention was well-known in the field. There are two issues here. First, carrying out a canonical § 103 analysis under the guise of a § 101 challenge guts the structure of the patent statutes. Second, it forces judges to carry out a § 103 analysis without robust access to the relevant prior art (namely, extrinsic evidence to support the science).¹⁷¹ Rather than carrying out a comprehensive prior art analysis—one that would be supported by litigants defending against a § 103 invalidity challenge—the Federal Circuit judges were left with only the shreds of § 101 arguments.¹⁷² It may well be that the driveshaft invention in *American Axle* did not merit patent protection for a lack of inventiveness over

168. *Graham v. John Deere Co.*, 383 U.S. 1, 17 (1966). Secondary considerations such as “commercial success, long felt but unsolved needs, [and] failure of others” are also factored into the § 103 analysis. *Id.*

169. *Mayo*, 566 U.S. at 79–80 (emphasis added).

170. *See American Axle*, 967 F.3d at 1290.

171. *See id.* at 1311 (Moore, J., dissenting). Worse yet, as litigants often bring § 101 challenges at early phases of litigation, this pseudo-§ 103 analysis might occur without claim construction or discovery.

172. To leave open the option of a pseudo-§ 103 analysis under § 101 is to disincentivize litigants from performing their own comprehensive prior art analyses as part of § 103 challenges. And litigants are much better positioned than the court to perform such an analysis. *See Parke-Davis & Co. v. H.K. Mulford Co.*, 189 F. 95, 115 (C.C.S.D.N.Y. 1991) (noting “the extraordinary condition of law which makes it possible for a man *without any knowledge of even the rudiments of chemistry* to pass upon such questions as these,” in reference to the court needing to rely heavily on expert evidence to understand the technical details of the asserted patent) (emphasis added). If the majority and dissent had been given equivalent access to the prior art, perhaps their respective interpretations of the “inventive concept” (or lack thereof) in *American Axle* would have converged.

the prior art in the automobile manufacturing industry. But if so, the asserted claims should have been invalid under § 103, and not § 101.¹⁷³

b) From § 101 to § 112

As discussed in Section II.A, § 101 and § 112 have been blurred from the start. *Neilson* and *Morse*, two key nineteenth century “eligibility” cases, both hinged on disputes about the adequacy of disclosure—a modern § 112 inquiry.¹⁷⁴ *Neilson* set forth that embodiments or applications of “principles” were patent-eligible, but the patent at issue was truly being challenged for failing to adequately enable the invention.¹⁷⁵ *Morse* then used *Neilson* to reaffirm the importance of applying a “principle” to achieve eligibility, but again, the patent at issue suffered from inadequate disclosure.¹⁷⁶ The core elements of *Neilson* and *Morse* would today be strictly defined as § 112 problems.

To meet the § 112 standard, the specification of a patent must adequately disclose the invention, such that “any person skilled in the art to which it pertains . . . [could] make and use” it in the same way.¹⁷⁷ This statute is now understood to require two distinct elements within the concept of disclosure: written description and enablement.¹⁷⁸ Briefly, the written description doctrine requires that the specification indicate that the inventor was “in possession” of the claimed invention as of the application filing date.¹⁷⁹ The enablement doctrine requires that the patent owner adequately teach and support the entire range of embodiments set out in the patent claims.¹⁸⁰

173. See Brief of Professors Jeffrey A. Lefstin & Peter S. Menell as Amici Curiae in Support of Petitioner for a Writ of Certiorari at 6–7, *Am. Axle & Mfg., Inc. v. Neapco Holdings LLC*, 142 S.Ct. 2902 (2022) (No. 20-891), 2021 WL 859724 [hereinafter Lefstin & Menell Brief].

174. Lefstin, *supra* note 24, at 581–82, 596–97.

175. See *id.* at 580; Lefstin & Menell Brief, *supra* note 173, at 11.

176. See Lefstin, *supra* note 24, at 596–97.

177. 35 U.S.C. § 112(a).

178. These two requirements were deemed to be distinct in 2010. *Ariad Pharms. v. Eli Lilly & Co.*, 598 F.3d 1336 (Fed. Cir. 2010) (en banc). *Ariad* set out a *quid pro quo* rationale of patent protection: that inventors should only be awarded the exclusionary, negative rights to their inventions if they give the public something in return—more than the mere existence of their invention in the world. *Id.* at 1345; Jacob Adam Schroeder, *Written Description: Protecting the Quid Pro Quo Since 1793*, 21 *FORDHAM INTELL. PROP. MEDIA & ENT. L.J.* 63, 66–67 (2010). A third requirement (best mode) exists within the statute but is now deemed irrelevant for modern patent practice. MENELL ET AL., *supra* note 2, at 262.

179. MENELL ET AL., *supra* note 2, at 273.

180. *Id.* at 263.

The invention in *American Axle* seemed to pose a conventional § 112 problem.¹⁸¹ The court focused on a lack of information as to how to craft the liner and drive shaft, and the absence of “any physical structure or steps for achieving the claimed result.”¹⁸² Bizarrely, Judge Dyk used these deficiencies to conclude that the driveshaft invention lacked an “inventive concept” under step two of the *Mayo/Alice* test, rather than that it more plausibly failed to meet the § 112 enablement standard.

Here, there are another two underlying issues: the implementation of a canonical § 112 analysis shoehorned into a § 101 challenge (1) guts the structure of the patent statutes, and (2) forces judges to carry out a § 112 analysis without the boundaries of the actual statute.¹⁸³ As Judge Newman described in her dissent, the majority’s analysis required the patent claims to go beyond mere definiteness, “inject[ing] a heightened enablement requirement into the § 101 analysis” that failed to adequately refer back to the specification.¹⁸⁴ Section 112—on its own—is better positioned to do the analysis that the *American Axle* court grasped at. And recent decisions even suggest a trend towards intensifying the § 112 requirement,¹⁸⁵ which will more robustly police the “pre-emption” concerns that anchor much of eligibility

181. See Lefstin & Menell Brief, *supra* note 173, at 3 (arguing that *American Axle* “presents an excellent vehicle for clarifying the interplay of § 101 and § 112 of the Patent Act”).

182. 967 F.3d 1285, 1295 (Fed. Cir. 2020).

183. For § 112 written description, the court should have explicitly defined the genus being claimed (liners designed with mass and/or stiffness factored in) and contemplated the entire range of contained species that must be adequately disclosed. For § 112 enablement, the court should have looked to the *specification* to assess the adequacy of disclosure in the driveshaft invention. Instead, the judges all fixated on the plain text of the *claims*, and whether they explained *how* to apply Hooke’s law on their own. See discussion *supra* Section III.A.

184. *American Axle*, 967 F.3d at 1317 (Moore, J., dissenting).

185. See, e.g., *Juno Therapeutics, Inc. v. Kite Pharma, Inc.*, 10 F.4th 1330 (Fed. Cir. 2021) (tightening the written description requirement of § 112 to require more comprehensive disclosure, more rigorously enforcing that an inventor must have possessed the full scope of the claim at the time of filing); *Amgen Inc. v. Sanofi*, 987 F.3d 1080 (Fed. Cir. 2021) (tightening the enablement requirement of § 112 to require disclosure of all possible embodiments (species) within a genus and more rigorously enforcing that the scope of the claims at issue must be commensurate with the scope of disclosure). To be sure, some have expressed disappointment in what has become of the § 112 standard in recent years, and its potential impact on genus claims. Dmitry Karshedt, Mark A. Lemley & Sean B. Seymore, *The Death of the Genus Claim*, 35 HARV. J.L. & TECH. 1 (2021). Others have expressed that this is, perhaps, an overstatement. Christopher M. Holman, *After Granting Certiorari in Enablement Case, Supreme Court Declines Opportunity to Address Written Description*, PATENTLYO (Nov. 11, 2022), <https://patentlyo.com/patent/2022/11/certiorari-opportunity-description.html>. Regardless of outlook, recent case law has clearly pushed for a § 112 standard that more rigorously polices claims with functional breadth.

doctrine.¹⁸⁶ So, just as for § 103, perhaps *American Axle*'s invention did not merit patent protection—but under § 112, not § 101.¹⁸⁷

Overall, collapsing the § 103 and § 112 patentability doctrines into § 101, as shown by *American Axle*, creates redundancy between the intentionally partitioned patent statutes and lessens the value of each substantive doctrine for challenging invalidity or defending validity. Twisting § 103 and § 112 questions into § 101 allows litigants and courts to dilute or concentrate the impact of either statute. Removing the *Mayo/Alice* test as an option for covert § 103 or § 112-type inquiries would reinvigorate both doctrines, eliminating the needless and harmful redundancy between the statutes.

IV. A REVISED FRAMEWORK FOR PATENT ELIGIBILITY

Part III uses *American Axle* as a vehicle to argue that steps one and two of the *Mayo/Alice* test are fundamentally flawed. The first step forces a subjective evaluation of “ineligibility” that can capture almost any invention, and the second step forces a betrayal of the partitioned structure of the Patent Act. But if the *Mayo/Alice* test were to be set aside, what could stand in its place? A persuasive solution is to find a sweet spot for the § 101 standard: one that is low enough to avoid an administratively frustrating overlap with § 103 or § 112, but high enough to remain mindful of policy concerns such as whole field pre-emption. That is, the vestiges of the substantive patentability doctrines should be filtered out, but some baseline level of eligibility must be maintained.

To this end, this Note suggests that the existing eligibility inquiry under § 101 should be reoriented to simply focus on utility, stemming from the word “useful” in § 101. That is, the *Mayo/Alice* test ought to be replaced with a utility-oriented eligibility framework, distinct from the current understandings

186. One way of dealing with the pre-emption concern is to treat “laws of nature” as presumptively ineligible for patent protection because claiming such a law might pre-empt its use in all contexts. But a proper evaluation of patent applications under § 112 might render this unnecessary. The enablement requirement of § 112, for example, should theoretically require an inventor to disclose *all possible means* of using a law of nature. That is, if an inventor has such broad claims so as to risk “whole pre-emption” of a law of nature, then they will have to understand it—and all its potential—well enough to describe those pre-empted uses in depth. This, for most laws of nature, is a remarkably high standard to meet—possibly, high enough to disqualify many inventions that are currently dealt with on eligibility grounds. The inherent similarity between the function of § 112 disclosure (in policing claim breadth) and § 101 eligibility (in guarding against pre-emption) is, perhaps, the reason for the inevitable overlap between these doctrines, as seen in *American Axle*. This is distinct from the overlap between § 103 and § 101, which might be better characterized as an artifact of the doctrinal language (the sharing of the word “inventive”).

187. Lefstin & Menell Brief, *supra* note 173, at 7.

of both § 101 eligibility and § 101 utility. This Part will first contextualize this proposal among the existing interpretations of § 101 utility, and then detail the parameters of this Note’s heightened utility-eligibility § 101 standard.

A. EXISTING PERSPECTIVES ON § 101 UTILITY

Currently, patent examiners and courts interpret the word “useful” in the text of § 101¹⁸⁸ to require simply that an invention have some extraordinarily minimal form of “utility.” This is not collapsed into § 101 eligibility, but rather, treated as its own requirement. By many accounts, the modern-day § 101 utility standard is a very low threshold that most inventions easily satisfy.¹⁸⁹ But despite its now-defunct status, the § 101 utility standard initially had much more significance—possibly, in fact, entirely anchoring patent protection in its early days.¹⁹⁰ And in more recent years, specific concerns have spawned efforts to heighten the utility requirement for certain types of patents.¹⁹¹ Together, the history of patent utility and its attendant policy considerations indicate that three existing perspectives on utility are available.¹⁹² This Section will briefly summarize each perspective.

1. *Social and Moral Utility*

An early interpretation of “useful” in § 101 appeared to require “social utility” of inventions.¹⁹³ In the formative years trailing the codification of the 1790 patent regime, patents were treated as a *privilege* that the Patent Board¹⁹⁴ had the power to bestow.¹⁹⁵ In keeping with this privilege-oriented framework, the Board would “weigh the social costs and benefits underlying each grant,”¹⁹⁶ encouraging patent petitions that extolled the virtuous public benefits of

188. 35 U.S.C. § 101 (“Whoever invents or discovers any new and *useful* process, machine, manufacture, or composition of matter, or any new and useful improvement thereof, may obtain a patent therefor.”) (emphasis added). The word “useful” has been part of the patent statutes since their first form in the 1790 Patent Act. Syed, *supra* note 8, at 2030–31 (compiling the present-day text of § 101 and all previous versions of the statute in Table 1). It is even woven into the Intellectual Property Clause itself. U.S. CONST. art. 1, § 8, cl. 8 (referring to the “*useful* Arts”) (emphasis added).

189. *See, e.g.,* Juicy Whip, Inc. v. Orange Bang, Inc., 185 F.3d 1364, 1366 (Fed. Cir. 1999) (“The threshold for utility is not high.”).

190. *See* discussion *infra* Section IV.A.2.

191. *See* discussion *infra* Section IV.A.3.

192. *See* Syed, *supra* note 8, at 2028 (articulating three available perspectives of utility).

193. *See* BRACHA, *supra* note 24, at 188–89, 202–03.

194. The Patent Board of the 1790 Patent Act is the historical analog to the modern PTO. *Id.* at 203–04.

195. *Id.* at 194.

196. *Id.* at 194–97.

inventions.¹⁹⁷ Utility, at this point, was enshrined as a serious patentability requirement.

Later, this “social utility” view shifted into one that was more moral in nature. As recounted by Professor Bracha, the ideological and practical perception of patents transformed from “economic privilege[s] . . . bestowed on inventors” into “inventors’ rights.”¹⁹⁸ Patents went from a discretionary privilege conferred by the Patent Board to an almost presumptive right to be checked by the courts, after issuance.¹⁹⁹ With this, the previous “useful” inquiry—asking, more or less, whether an invention did something *worthwhile for society*—morphed into a more subjective, morally tinged assessment to weed out “mischievous” inventions.²⁰⁰ While popular for some time, neither the social nor moral views of utility currently persist in U.S. patent law.²⁰¹

2. Operable Utility

The value-based assessments of utility were later replaced with an operability view, which remains the prevailing inquiry under present-day § 101 utility for most inventions. Under this perspective, the word “useful” in § 101 is read to suggest simply that an invention “works,” i.e., that it is operable for its disclosed purpose.²⁰² This operability view was ushered in by the statutory

197. *Id.* at 196–99.

198. *Id.* at 188. Some perspectives from this time suggest the embracing of a labor theory-type rationale to catalyze this ideological shift, for example, “a citizen has a right in the inventions he may make, and he considers the law but as the mode by which he is to enjoy the fruits.” *See id.* at 190.

199. *Id.* at 200–02.

200. *See id.* at 203–07 (discussing, among other things, two conflicting constructions of “utility” that emerged in these years: one focused on the “objective social value” of an invention, where a patentee must show a mere “public benefit”; the other focused on a “moralistic framing,” where a patentee must show “that the invention should not be frivolous or injurious to the well-being, good policy, or sound morals of society”). The “moral utility” perspective can be traced back to Justice Story’s statement deeming inventions that are “injurious to the well-being, good policy, or sound morals of society” to be unpatentable. *Lowell v. Lewis*, 15 F. Cas. 1018 (C.C.D. Mass. 1817).

201. *Juicy Whip, Inc. v. Orange Bang, Inc.*, 185 F.3d 1364, 1368 (Fed. Cir. 1999) (putting the moral requirement of utility, which had begun to rear its head again, to rest: “[t]he requirement of ‘utility’ in patent law is not a directive to the Patent and Trademark Office or the courts to serve as arbiters of deceptive trade practices”). There is, now, a strong opposition to the notion that questions of morality or general community welfare would permeate the patent laws. *Id.* *But see* Laura A. Keay, *Morality’s Move Within U.S. Patent Law, From Moral Utility to Subject Matter*, 40 AIPLA Q.J. 409 (2012) (noting that § 101 case law represents the infiltration of morality considerations into patent law, once again, without the backbone of historical “moral utility”).

202. *Syed*, *supra* note 8, at 2028. The canonical example of an invention that would not meet this standard is a perpetual motion machine, which is a physical impossibility. *Id.*

reform leading to the 1836 Patent Act and the establishment of the PTO. The new “utility” of the 1836 Patent Act was not the same as the privilege-anchored version from the 1790 regime.²⁰³ The PTO did not have the Patent Board’s “discretionary powers to grant privileges,” rather, its role was “to certify the satisfaction of standard patentability criteria.”²⁰⁴ With the privilege view put to rest, the social and moral questions were replaced with a quest for uniform patentability.²⁰⁵ Patents, strictly, would be “rights” open to all, with the *market* serving as the arbiter of utility, rather than the PTO.²⁰⁶ So long as the patent met other criteria—now codified as the substantive doctrines of § 102, § 103, and § 112²⁰⁷—the utility of an invention would be naturally assessed based on its demand in the market, such that the inventor would derive value from their patent in proportion to its market utility.²⁰⁸ And with a market-oriented, rights-based view of patent protection, a *substantive* utility requirement was seemingly unwarranted.²⁰⁹ While previously “a central defining feature of patents,” utility became “the exotic periphery of patent law” by the end of the nineteenth century.²¹⁰

3. *Specific, Substantial, and Credible Utility*

A third perspective on utility comes from *Brenner v. Manson*, which added a gloss over the operability view just discussed: utility must be specific, substantial, and credible.²¹¹ These terms are loosely defined, with “specific” suggesting *not vague*, “substantial” suggesting *not throw-away*, and “credible” suggesting *believable for its purpose*,²¹² to be evaluated from the perspective of one of skill in the art.²¹³ The *Manson* Court rationalized this elevated utility standard by citing policy concerns as to the scope of patentability and the notion of pre-

203. BRACHA, *supra* note 24, at 209–13.

204. *Id.* at 209. Professor Bracha describes this shift in perspective as reflective of Jacksonian ideology, which eschewed a former paternalistic implication of the 1790 regime’s utility standard: that the government was appropriately positioned to identify an objective, social utility common to all, and then promote it with the privilege of a patent. *Id.* at 209–12.

205. *Id.* One possible view is that the original casting of the patent right—then, a privilege—only emphasized “use” as a placeholder, before the complete set of patent doctrines took shape.

206. *Id.*

207. *See id.* at 202–03.

208. *See id.* at 212–14 (explaining the market-oriented view of patent valuation).

209. *Id.* at 215. For some time, courts still scrutinized the utility requirement under the 1836 regime, but the utility requirement did eventually decline in importance entirely. *Id.* at 214.

210. *Id.* at 216.

211. 383 U.S. 519, 527–33 (1966); MENELL ET AL., *supra* note 2, at 248.

212. *See id.* at 245–48.

213. MPEP § 2107 (9th ed. Rev. 10, June 2020).

emption—not unlike the case for eligibility.²¹⁴ In evaluating the meaning of “useful” for a patent claiming a steroid,²¹⁵ Justice Fortas focused on the risk of monopolization with a tone mirroring that of *Benson*, *Bilski*, and *Alice*.²¹⁶ Without a showing of use, he opined that “the metes and bounds of [a] monopoly are not capable of precise delineation,” threatening to “confer power to block off whole areas of scientific development, without compensating benefit to the public.”²¹⁷

Manson remains the zenith of § 101 utility. Although all inventions are now formally required to show specific, substantial, and credible utility, these adjectives are rarely a bar to patentability, in perception or practice.²¹⁸ However, *Manson* has at least appeared to have an effect on chemical compositions (e.g., requiring perhaps a feasible therapeutic application for a claimed compound,²¹⁹ even if supported only from *in vitro* testing²²⁰) and biotechnology inventions (e.g., allowing claims directed to coding DNA fragments only if the resultant translated proteins had known functional uses that the inventor could articulate²²¹). The specific application of the *Manson* standard to biotechnology patents was precipitated by the attempted patenting of human genes in the 1990s,²²² which inspired the PTO to “mov[e] toward a

214. Professor Taylor has noted that, at the time of *Manson*, the Court used the utility requirement to address “the very concerns of the Supreme Court in its recent cases addressing patent eligibility.” Taylor, *supra* note 10, at 2189 (2017).

215. The patent in *Manson* claimed a steroid composition but did not specify any use for that composition. Instead, the patent specified the known use of other steroid compositions, which were similar in structure to the claimed composition. 383 U.S. at 533–34.

216. *See id.* at 534–35.

217. *Id.* The *Manson* opinion weighed heavily in favor of the *quid pro quo* rationale to patent protection, where inventors are required to not only adequately disclose their inventions under § 112, but also produce an invention that is “useful” *to some end*, to provide “a significant and presently available benefit to the public.” *See In re Fisher*, 421 F.3d 1365, 1371 (Fed. Cir. 2005).

218. *See MENELL ET AL.*, *supra* note 2, at 245–46, 248 (calling *Manson* the “high-water mark” of the utility doctrine and noting that “[m]ost applications of the doctrine have been limited in the hurdles they place before inventors”).

219. *See In re Brana*, 51 F.3d 1560, 1565 (Fed. Cir. 1995).

220. *See Cross v. Iizuka*, 753 F.2d 1040, 1048 (Fed. Cir. 1985).

221. Timothy A. Worrall, *The 2001 PTO Utility Examination Guidelines and DNA Patents*, 16 BERKELEY TECH. L.J. 123, 133 (2001).

222. *Chakrabarty*’s permissiveness towards living organism patentability led scientists to begin seeking patent protection on expressed sequence tags (ESTs)—fragments of cDNA, not whole genes—in the early 1990s. Daniel J. Kevles & Ari Berkowitz, *The Gene Patenting Controversy: A Convergence of Law, Economic Interests, and Ethics*, 67 BROOK. L. REV. 233, 235–37 (2001). Many disapproved of this trajectory, given the concern that an EST patent landscape would foreclose considerable future research—the pre-emption rationale, again. *Id.* at 237–39. These years also saw the attempted patenting of the *BRC A1* and *BRC A2* genes, which was also met with “overwhelmingly negative” public perception. Robert Cook-Deegan & Christopher Heaney, *Patents in Genomics and Human Genetics*, 11 ANN. REV. GENOMICS & HUM.

stronger requirement for utility.”²²³ And the 2001 Utility Guidelines issued by the PTO effectively pushed the timing of chemical and biotechnology patenting further downstream in the discovery process.²²⁴ In these contexts, specific, substantial, and credible were interpreted to mean that inventors could only claim what they truly had in hand and truly understood on a functional level.²²⁵

B. THE PROPOSED UTILITY-ELIGIBILITY FRAMEWORK

The previous Section describes three currently available perspectives on utility. This Note suggests a fourth construction of utility to replace the existing § 101 eligibility inquiry. The next three Sections will describe the parameters of this fourth utility perspective, clarify how it supports the policy justifications for current eligibility doctrine, and explain its implications for overall patentability.

1. *Scope-Limiting Utility*

There are three layered requirements embedded in this Note’s proposal. The first requirement is that in assessing whether an “invent[ion] or discover[y]” of a “process, machine, manufacture, or composition of matter” is “useful” under § 101,²²⁶ patent examiners and courts should require that *some* finite use or set of uses is delimited by the claims. Specifically, the claimed use or set of uses must not be *infinite* or left out of the claims and only suggested in the specification. This is distinct from the existing standard of operable

GENETICS 383, 389–90 (2010). The most common concern was another depiction of the pre-emption rationale: that gene patents would specifically preclude the development of future sequencing technology and diagnostic tests. Christopher M. Holman, *Debunking the Myth that Whole-Genome Sequencing Infringes Thousands of Gene Patents*, 30 NATURE BIOTECHNOLOGY 240, 240–41 (2012).

223. *In the Crossfire: Collins on Genomes, Patents, and ‘Rivalry,’* 287 SCI. 2396, 2397 (2000) (transcribing an interview with Francis Collins, the head of the National Human Genome Research Institute and leader of the Human Genome Project). Collins referred to the gene patents feared by many as “generation one” applications, where inventors had “just a [DNA] sequence” but “no clue as to what it does.” The PTO’s Utility Guidelines deemed these applications to be insufficiently “specific” to meet the *Manson* standard. Purists may also include, as “generation one” applications, claims directed to sequences with a vague construction of function based on homology—not unlike the steroid composition claims in *Manson*. *Id.*

224. See Utility Examination Guidelines, 66 Fed. Reg. 1092, 1097–99 (Jan. 5, 2001).

225. See Timothy A. Worrall, *The 2001 PTO Utility Examination Guidelines and DNA Patents*, 16 BERKELEY TECH. L.J. 123, 132 (2001) (explaining the stringency of the newer utility examination guidelines on certain types of inventions).

226. 35 U.S.C. § 101 (“Whoever invents or discovers any new and *useful* process, machine, manufacture, or composition of matter, or any new and useful improvement thereof, may obtain a patent therefor.”) (emphasis added).

utility under § 101, which does require that inventions “work” for their disclosed purposes,²²⁷ but does not require setting out those disclosed purposes *in the claims*.²²⁸ That is, under the current § 101 utility standard, applicants can satisfy the low-bar utility requirement through a merely qualitative suggestion of use in the specification. The instant proposal, instead, is a quantitative requirement of use, embedded into the claims.

The second requirement of this proposal makes clear the implications of placing a set of uses in the claim language itself. In this Note’s framework, the scope of the exclusionary patent right would be limited to the uses laid out in the application’s claims (of course, also accommodating for equivalents and after-arising technology, to be interpreted from the specification).²²⁹ And in keeping with the current standards under § 112, these uses would need to be commensurate with the scope of the material disclosed in the patent specification, which would require both adequate disclosure and the contemplation of all embodiments. Currently, § 112 doctrine requires only claims containing means-plus-function language to be interpreted in this way.²³⁰

The third and final requirement answers a lingering question within this framework—what are the qualities of an adequate “use” for the purposes of limiting claim scope? Scope-limiting uses must be “specific, substantial, and credible,” as per *Manson*. On its own, this requirement does not signal a change in existing § 101 utility doctrine—patent examiners are already advised to apply *Manson*’s adjectives to all inventions.²³¹ However, when integrated with the other two requirements, this would push the *Manson* standard outside of merely chemical and biotechnology patents, where its application currently has the most force. This casting of “use” is closer in meaning to the “practical application” requirement of *Neilson*, *Le Roy*, and *Morse*, rather than the “inventive application” requirement of *Funk Brothers*, *Flook*, *Mayo*, *Myriad*, and

227. Syed, *supra* note 8, at 2028.

228. See *infra* note 247 and accompanying text (explaining the current standard of utility applied to claim language, with the example of standalone composition of matter claims).

229. Recall that the “right” conferred by a patent is a negative, exclusionary one, to exclude others from making, using, or selling the invention—specifically, the invention as “defined and limited by the language in that patent’s claims.” *Corning Glass Works v. Sumimoto Elec. U.S.A., Inc.*, 868 F.2d 1251, 1257 (Fed. Cir. 1989). Here, that language will specifically include the clearly articulated set of uses that the applicant lays out.

230. See MPEP § 2181 (9th ed. Rev. 10, June 2020). At least one Note has, similarly, proposed that claims directed to one of the judicial exceptions could be treated as means-plus-function claims (as an alternative to the *Mayo/Alice* test). See, e.g., Nicholas Stroger, *An Automatic Means-Plus-Function Limitation for Otherwise Unpatentable Subject Matter*, 22 WAKE FOREST J. BUS. & INTELL. PROP. L. 243 (2022).

231. See MPEP § 2107 (9th ed. Rev. 10, June 2020).

Alice.²³² This Note proposes to simply replace “practical application” with a requirement for an explicit use or set of uses, not required to be “practical” or “inventive,” but “specific, substantial, and credible.”

2. Innovation Policy Justifications

Why is this scope-limiting utility framework a viable substitute for the existing § 101 eligibility inquiry? Is this proposal enough to replace the ineligibility bars on laws of nature, natural phenomena, and abstract ideas? To answer these questions, it is useful to think back to the innovation policy rationale anchoring the existence of the § 101 eligibility standard itself. Part II of this Note traced the history of patent eligibility jurisprudence, revealing that the strongest justification for the ineligibility bars was expressed in *Benson*,²³³ *Bilski*,²³⁴ and *Alice*.²³⁵ the threat of pre-emption. The § 101 eligibility standard, ideally, is a bulwark against inventions that are novel, non-obvious, and adequately disclosed, yet also pre-empt the use of a critical, pseudo-universal concept in other inventions.²³⁶ In theory, these “pre-emptive” inventions threaten to impede more innovation than they would promote, betraying the fundamental purpose of the patent system.

This Note agrees with the importance of the pre-emption rationale, as patents that tie up “building blocks”²³⁷ and monopolize entire technological fields are a deterrent to innovation. But the *Mayo/Alice* test does not suffice to address the pre-emption rationale. The nuances of this argument aside, borne out in the preceding Parts—the *Mayo/Alice* test asks the wrong questions and is fundamentally unclear. The proposed scope-limiting framework, instead, addresses the pre-emption rationale with more clarity.²³⁸ The Court’s pre-emption fears were arguably tailored to disallowing claims that “substantially encompass” an ineligible concept. But, as others have noted, what really should be disallowed are claims that fail to “impose[] a meaningful limit” on

232. See discussion *supra* Part II.

233. 409 U.S. 63, 71–72 (1972).

234. 561 U.S. 593, 611–12 (2010).

235. 573 U.S. 208, 223 (2014).

236. *Le Roy v. Tatham*, 55 U.S. 156, 175 (1852); *Funk Bros. Seed Co. v. Kalo Inoculant Co.*, 333 U.S. 127, 130 (1948).

237. Syed, *supra* note 8, at 1967–68 & 1967 n.105 (describing the “building block” rationale that emerged from the Court’s eligibility jurisprudence, getting to the heart of the pre-emption concern).

238. Again, arguably, the *Manson* opinion swirled around this same policy rationale. Justice Fortas justified the heightened utility standard applied to the invention in *Manson* against pre-emption concerns. 383 U.S. at 534–35.

the concept.²³⁹ Requiring an explicit articulation of scope-limiting uses to be set out in the patent claims themselves does just this—it imposes a meaningful limit. Thus, this casting of utility-eligibility improves upon the *Mayo/Alice* test by shifting the burden of assessing that “meaningful limit” away from patent examiners and courts and toward applicants, who themselves must write their uses into their claims with precision. And in doing so, the proposed model avoids the ineligibility bar question entirely by uniformly imposing this standard across all inventions.²⁴⁰

3. *Implications of and Potential Improvements to the Proposal*

It bears emphasizing that this Note is not the first to call attention to the utility requirement as a vehicle for patent reform. Many have specifically looked to § 101 utility in the context of patent eligibility, albeit not with the exact boundaries of the instant framework.²⁴¹ Compared to these other

239. See David V. Sanker & Jillynne Quinn, *A Quantitative Approach to Overcoming § 101 Rejections*, PAT. LAW. 17 (2021), <https://www.morganlewis.com/-/media/files/publication/outside-publication/article/2021/a-quantitative-approach-to-overcoming-101-rejections-the-patent-lawyer.pdf?rev=bf8e7a96c0e44e058c09c24a04855099&hash=BED99736E739B6E7F1E590386001A633>. David Sanker and Jillynne Quinn have proposed a “quantitative approach” to evaluating § 101, painting the eligibility question as one that should compare: (1) the realm of options included in the ineligible concept; with (2) the realm of options included in the *claim* relating to the ineligible concept. *Id.* This, in effect, is a search for a meaningful limit—a patent-eligible invention should simply be one in which the realm of options covered by the claim is “meaningful[ly] limit[ed],” compared to the realm of imaginable options covered by the ineligible concept itself. *Id.* The “meaningful limit” language was proposed by the PTO as a useful way of thinking about patent eligibility. 2019 Revised Patent Subject Matter Eligibility Guidance, 84 Fed. Reg. 4 (Jan. 7, 2019), <https://www.govinfo.gov/content/pkg/FR-2019-01-07/pdf/2018-28282.pdf> (“A claim that integrates a judicial exception into a practical application will apply, rely on, or use the judicial exception in a manner that imposes a *meaningful limit* on the judicial exception, such that the claim is more than a drafting effort designed to monopolize the judicial exception.”) (emphasis added).

240. For a possible exception to this standard, see discussion *infra* Section IV.B.3.

241. See, e.g., Sean B. Seymore, *Making Patents Useful*, 98 MINN. L. REV. 1046 (2014) (proposing the entire deletion of the utility requirement, with an argument resembling that of this Note regarding the current eligibility standard—that it is redundant with the other patentability doctrines); Robin C. Feldman, David A. Hyman, W. Nicholson Price II & Mark J. Ratain, *Negative Innovation: When Patents Are Bad for Patients*, 39 NATURE BIOTECHNOLOGY 914, 914–15 (2021) (highlighting the vulnerabilities of pharmaceutical patents to negative innovation—where patent law incentivizes innovation into directions that are net harmful to the public—and suggesting that the utility standard should be heightened for pharmaceutical patents to require that they “actually improve social welfare relative to the prior art,” for example, by “requir[ing] certification of likely improvement, followed by a demonstration that the improvement had materialized, on pain of losing the patent”); Michael Risch, *A Surprisingly Useful Requirement*, 19 GEO. MASON L. REV. 17 (2011) (arguing that the utility standard should be more harmoniously woven into the other substantive patentability doctrines); Sean M. O’Connor, *The Lost “Art” of the Patent System*, 2015 U. ILLINOIS L. REV. 1397, 1476 (2015)

proposals, the instant scope-limiting utility standard does not clearly expand or contract the overall stringency of § 101.²⁴² Instead, it treats a different set of inventions as patent-eligible. For some inventions—those that are vulnerable to characterization as “natural” or “formulaic,” which currently succumb to the *Mayo/Alice* test—this proposal makes § 101 a lower bar. As discussed in Section IV.B.1, requiring applicants to claim a specific, substantial, and credible set of uses is a lower standard than the “inventive application” requirement of *Mayo/Alice*, closer to the “practical application” requirement of nineteenth century eligibility case law. This Note has already advanced several arguments to support this outcome. Briefly, again, § 103 and § 112 can adequately police most inventions challenged under § 101, and the imposition of the proposed framework would still impose a meaningful limit to guard against pre-emption concerns. All the “natural” and “diagnostic” inventions discussed in Section II.D—deemed ineligible under § 101—would likely satisfy this Note’s version of § 101 utility-eligibility, given their specific, substantial, and credible uses²⁴³ in the context of fetal characteristic analysis,²⁴⁴ cardiovascular disease risk assessment,²⁴⁵ and neurological disorder diagnosis.²⁴⁶ The same is true for the invention in *American Axle*, which articulated a use for the invention in the specific, substantial, and credible context of driveshaft assembly.

For other inventions—those that do not intersect with the so-called ineligibility concepts, do not traditionally claim a set of uses, and have not yet been discussed by this Note—this proposal would radically elevate the § 101 hurdle. Replacing the *Mayo/Alice* test with the instant framework means that

(describing that limiting the scope of patent-eligible subject matter those inventions that “progress or advance the use of natural materials or forces for *practical (useful)* ends” would remove the need for judicial exceptions, in an argument based on a historical reconstruction of the target “art” of the patent system) (emphasis added); Taylor, *supra* note 10, at 2188–89 (proposing, among other types of amendments to repair eligibility doctrine, that the utility requirement be elevated to require that inventors describe the use of an invention in the claims, rather than merely the specification); Syed, *supra* note 8, at 2029, 2033–35 (suggesting that the word “useful” in § 101 might illustrate a “subject-matter delimiting” role for patent utility doctrine, wherein “useful” connotes “applied,” and that only practical, downstream applications of knowledge should be patent-eligible subject matter, rather than basic, unapplied knowledge).

242. See *supra* note 241.

243. Of course, one of skill in the art would need to perceive the data in each patent as credible to satisfy this analytical standard.

244. *Ariosa Diagnostics, Inc. v. Sequenom, Inc.*, 788 F.3d 1371 (Fed. Cir. 2015).

245. *Cleveland Clinic Found. v. True Health Diagnostics LLC*, 859 F.3d 1352 (Fed. Cir. 2017).

246. *Athena Diagnostics, Inc. v. Mayo Collaborative Servs., LLC*, 915 F.3d 743, 746–47 (Fed. Cir. 2019).

all inventions would be subjected to the same level of scrutiny. Even inventions that are not “directed to” ineligible concepts would require an explicit articulation of scope-limiting, specific, substantial, and credible uses. In this sense, the proposed utility-eligibility framework needs further refinement. On the one hand, the current § 112 disclosure requirements are, as is, tailored to serve this function—inventors should theoretically not be able to claim uses of an invention that are not yet understood or even anticipated. But on the other hand, many patentees benefit from claims untethered to uses—for example, claims directed to entirely novel, synthetic small molecules.²⁴⁷ Thus, there are perhaps some technological areas that need to be somehow exempted from this Note’s proposal, suggesting a degree of unavoidable exceptionalism in eligibility doctrine that merits further research and discussion.²⁴⁸

V. CONCLUSION

This Note proposes a revised framework for § 101, replacing the current eligibility standard with a heightened utility requirement: that patents must lay out a finite set of specific, substantial, and credible uses in their claims,

247. See generally Benjamin N. Roin, *Unpatentable Drugs and the Standards of Patentability*, 87 TEXAS L. REV. 503 (2009) (explaining how pharmaceutical firms view strong composition of matter patents as essential in the drug development process). The current § 101 eligibility and utility standards, even with the guidelines of *Manson*, would allow a composition claim directed to that small molecule itself, with no express articulation of its use(s) (i.e., “A compound of formula X, or a pharmaceutically acceptable salt thereof,” with nothing more). That is, an applicant’s discovery of new chemical matter is currently sufficient under § 101 for them to tie up all future uses of it. Under this Note’s framework, such a claim would be invalid—the applicant would be entitled to claim only the uses of that small molecule that they understand enough to satisfy *Manson*’s specific, substantial, and credible standard (i.e., an allowed claim would need to recite “A compound of formula X, or a pharmaceutically acceptable salt thereof, for [one or more specific, substantial, and credible uses].”). To be sure, viewed against the reward-incentivization framework of the patent system, it is unfair and perhaps even entirely unnecessary to entitle a patent applicant to yet-uncontemplated uses of a discovery, simply because they were the one to discover it. Further, the PTO’s notion of a “meaningful limit” and the Court’s consistent articulation of the pre-emption fear indicates a concern for claims of this nature. See generally Sanker & Quinn, *supra* note 239 (describing the “meaningful limit” issue).

248. See, e.g., Dan L. Burk & Mark A. Lemley, *Biotechnology’s Uncertainty Principle*, 54 CASE W. RES. L. REV. 691 (2004) (arguing that patent law is unavoidably mired in technology specificity, perhaps warranting distinct legal standards for distinct technological areas, e.g., “a consciously designed . . . patent policy” that would be specific to biotechnology). *But see*, e.g., R. Polk Wagner, *Exactly Backwards: Exceptionalism and the Federal Circuit*, 54 CASE W. RES. L. REV. 749 (2004) (arguing that Federal Circuit case law does not indicate as pervasive a degree of technology exceptionalism as argued by Professors Burk and Lemley, and that as a matter of policy, allowing or encouraging such exceptionalism is prone to several issues).

commensurate with the scope of disclosure in the patent specification. But this proposal leaves much room for further refinement, in having dealt primarily with inventions of the “natural” and not “formulaic” variety, not analyzing the implications of the utility-eligibility framework for software or business method patents, and not fine-tuning *Manson’s* “specific, substantial, and credible” standard outside of the chemical and biotechnology contexts.

Earlier, this Note laid out three issues that were left open by *Neilson*, *Le Roy*, and *Morse*, and then aggravated by *Mayo*, *Myriad*, and *Alice*. The revised framework for § 101 addresses each one, as follows: (1) it is no longer relevant what principles, laws of nature, phenomena, or abstract ideas are, nor what it means for an invention to be directed to them; (2) the degree of “application” required of an invention is simply a finite set of specific, substantial, and credible uses, explicitly laid out in the claims of a patent; and (3) some deference has been given to both the early nineteenth century eligibility case law and the § 101 jurisprudence that followed it, retaining the original “application” standard but also paying respect to the later pre-emption concerns. Perhaps a heightened form of utility is the antidote to patent eligibility doctrine’s disarray.

DRUGS, DECEPTION, AND DISCLOSURE

Garreth W. McCrudden[†]

TABLE OF CONTENTS

I.	INTRODUCTION	1132
II.	INCONSISTENT REPRESENTATIONS AT THE USPTO AND FDA	1133
A.	PATENT EXCLUSIVITY, REGULATORY APPROVAL, AND MISALIGNED INCENTIVES	1133
1.	<i>Patenting</i>	1137
2.	<i>Obtaining Regulatory Approval</i>	1137
3.	<i>Misaligned Incentives</i>	1138
B.	THE NEED FOR REFORM.....	1140
III.	TACKLING INCONSISTENT REPRESENTATIONS DURING PATENT PROSECUTION.....	1143
A.	A NEW SYSTEM OF USPTO-FDA INTERACTION.....	1143
B.	WHAT THIS NEW SYSTEM ACHIEVES—AND WHAT IT DOES NOT	1145
IV.	TACKLING INCONSISTENT REPRESENTATIONS AFTER PATENT ISSUANCE.....	1150
A.	INEQUITABLE CONDUCT DOCTRINE UNDER <i>THERASENSE</i>	1151
B.	THREE CASE STUDIES IN INCONSISTENT REPRESENTATION: <i>BRUNO</i> , <i>BELCHER</i> , AND <i>BAXTER</i>	1156
1.	<i>Bruno Independent Living Aids, Inc. v. Acorn Mobility Services</i> , <i>Ltd.</i>	1157
2.	<i>Belcher Pharmaceuticals, LLC v. Hospira, Inc.</i>	1158
3.	<i>Baxter International, Inc. v. CareFusion Corp.</i>	1159
C.	A NEW “PHARMA EXCEPTION” TO <i>THERASENSE</i>	1161
V.	CONCLUSION.....	1167

DOI: <https://doi.org/10.15779/Z38N87314D>

© 2023 Garreth W. McCrudden.

[†] (he/him); Doctor of Philosophy (D.Phil.), University of Oxford, Department of Chemistry, 2017; Juris Doctor (J.D.) Candidate, University of California, Berkeley, School of Law, Class of 2024. I am especially grateful to Talha Syed for his encouragement, guidance, and support in preparing this Note. I would also like to thank Allison Schmitt, Caressa Tsai, Rebecca Ho, Samantha Cox-Parra, Brigitte Desnoes, Nicolas Altemose, and the editors of the *Berkeley Technology Law Journal* for their invaluable feedback and assistance.

I. INTRODUCTION

How can we address the problem of pharmaceutical companies making inconsistent representations to the United States Patent and Trademark Office (USPTO or PTO) and the United States Food and Drug Administration (FDA)? Pharmaceutical innovators seeking both patent protection and regulatory approval of their drug products experience fundamentally misaligned incentives when they engage with the USPTO and FDA. On the one hand, FDA approval is often faster and cheaper for pharmaceuticals that bear significant similarities to already-approved drugs. On the other, successfully patenting a new small molecule requires an inventor to distinguish their product from the existing prior art—which often includes those very same drugs. As a consequence, pharmaceutical innovators are at once motivated to disclose to the FDA information about existing drug products and to hide that same information from the USPTO, even (or perhaps especially) when that information may be material to patentability. And, if innovators submit to that temptation, they can end up with patents that, in reality, should never have been issued in the first place.

Permitting pharmaceutical companies to make inconsistent representations to the USPTO and FDA is harmful both to the integrity of the patent system and to the public good. President Biden recognized as much in a July 2021 Executive Order, calling on the USPTO and FDA to work together “to help ensure that the patent system, while incentivizing innovation, does not also unjustifiably delay generic drug and biosimilar competition.”¹ Since President Biden issued his Executive Order, the heads of the USPTO and FDA have reiterated time and again that meaningful change is needed in the pharmaceutical industry both to uphold foundational patent-law doctrines and to provide public access to affordable drug products. But, to date, the agencies have provided very little indication of what that change will look like in practice.

This Note proposes two solutions to the problem of inconsistent representation at the USPTO and FDA. Part II outlines the nature of the problem, as well as the growing demand for reform. Part III proposes a first solution: a new system of USPTO-FDA interaction during patent prosecution. Because the overall effectiveness of such a system could be somewhat limited by issues of confidentiality, timing, and noncompliance, Part IV offers a complementary post-patent-issuance solution to the problem of inconsistent representation. Specifically, Part IV argues that the Federal Circuit should revise its inequitable conduct doctrine to create a “pharma exception” to the

1. *See* Exec. Order No. 14306, 86 Fed. Reg. 36,987, 36,997 (July 9, 2021).

otherwise exceedingly high legal standards outlined in *Therasense*.² Finally, Part V summarizes the key takeaways from Parts II, III, and IV.

II. INCONSISTENT REPRESENTATIONS AT THE USPTO AND FDA

This Part begins by providing a basic overview of (1) patent exclusivity, which is granted by the USPTO; and (2) regulatory approval, which is required in the pharmaceutical context by the FDA.³ Though these two systems often work in tandem, they are separate and distinct. Unique hurdles to patentability and regulatory approval create misaligned information-disclosure (or information-nondisclosure) incentives for pharmaceutical innovators seeking to both patent their new drug products and sell those products in interstate commerce. Thus, this Part argues, pharmaceutical innovators are at once motivated to disclose to the FDA information that may be material to patentability and to hide (or at least recharacterize) that same information when prosecuting a patent application at the USPTO.⁴ Finally, this Part summarizes the growing demand for meaningful change to the patenting and regulatory systems—both from within the USPTO and FDA as well as further afield.

A. PATENT EXCLUSIVITY, REGULATORY APPROVAL, AND MISALIGNED INCENTIVES

A fundamental—but often misunderstood—characteristic of any patent issued by the USPTO is that it does not grant its owner any affirmative rights.⁵

2. *Therasense, Inc. v. Becton, Dickinson & Co.*, 649 F.3d 1276 (Fed. Cir. 2011) (en banc).

3. *Compare* 35 U.S.C. § 101 (“Whoever invents or discovers any new and useful process, machine, manufacture, or composition of matter, or any new and useful improvement thereof, may obtain a patent therefor.”) (emphasis added), *with* 21 U.S.C. § 355(a) (“No person shall introduce or deliver for introduction into interstate commerce any new drug, unless an approval of an application filed pursuant to [this section] is effective with respect to such drug.”).

4. *See* Letter from Patrick Leahy, U.S. Senator, and Thom Tillis, U.S. Senator, to Andrew Hirshfeld, Performing Functions & Duties Under Sec’y Com. for Intell. Prop. & Dir. USPTO 1 (Sept. 9, 2021) (<https://www.leahy.senate.gov/imo/media/doc/20210909%20Letter%20to%20PTO%20on%20FDA%20submissions.pdf>) [hereinafter *Leahy & Tillis Letter*] (“[I]nconsistent statements submitted to the Food and Drug Administration (FDA) to secure approval of a product—asserting that the product is the same as a prior product that is already on the market—can then be directly contradicted by statements made to the PTO to secure a patent on the product.”).

5. *See Patentability Versus Freedom-To-Operate*, BUCKINGHAM, DOOLITTLE & BURROUGHS, LLC (May 10, 2021), <https://www.bdblaw.com/patentability-versus-freedom-to-operate/> (“Most often people mistakenly believe that a patent gives them the right to make,

In reality, patent rights are negative rights.⁶ The patent owner has the ability to exclude another from making, using, selling, offering for sale, or importing the claimed invention for twenty years from the filing date of the earliest nonprovisional application to which priority is claimed.⁷ However, the patent does not automatically confer to its owner the right to make, use, sell, offer for sale, or import that same invention.⁸ In fact, in many instances where the patent owner wishes to take any of those affirmative steps, they first need to obtain some sort of regulatory approval from an administrative body.⁹

In the pharmaceutical context, regulatory approval for new small-molecule drug products—patented or otherwise—must be obtained from the FDA before the drug can be sold in interstate commerce.¹⁰ Obtaining patent protection from the USPTO and market approval from the FDA are separate endeavors. But the incentive to obtain patent exclusivity from the USPTO is sustained in part by the FDA’s informationally demanding regulatory approval standards.¹¹ In fact, there are two distinct FDA-created informational costs

use, and sell an invention. Not so. A patent does not confer the right to do anything but sue others for patent infringement. This is perhaps the single most misunderstood feature of patents, and at the same time one of the most expensive mistakes an innovator can make.”).

6. See PETER S. MENELL, MARK A. LEMLEY, ROBERT P. MERGES & SHYAMKRISHNA BALGANESH, *INTELLECTUAL PROPERTY IN THE NEW TECHNOLOGICAL AGE: 2022* 177 (2022) (“The exclusionary right is in a sense a negative right . . . [A] patent does not automatically grant an affirmative right to do anything.”); *Patterson v. Kentucky*, 97 U.S. 501, 506 (1879) (“All which [patents] primarily secure is the *exclusive* right in the discovery.”) (emphasis added).

7. 35 U.S.C. § 154(a)(1) (“Every patent shall . . . grant to the patentee, his heirs or assigns, of the right to exclude others from making, using, offering for sale, or selling the invention throughout the United States or importing the invention into the United States.”); 35 U.S.C. § 154(a)(2) (“Subject to the payment of fees under this title, such grant shall be for a term beginning on the date on which the patent issues and ending 20 years from the date on which the application for the patent was filed in the United States.”).

8. See *Bio-Tech. Gen. Corp. v. Genentech, Inc.*, 80 F.3d 1553, 1559 (Fed. Cir. 1996) (“[T]he existence of one’s own patent does not constitute a defense to infringement of someone else’s patent. It is elementary that a patent grants only the right to *exclude others* and confers no right on its holder to make, use, or sell.”) (citation omitted); *Herman v. Youngstown Car Mfg. Co.*, 191 F. 579, 584 (6th Cir. 1911) (“A patent is not the grant of a right to make or use or sell. It does not, directly or indirectly, imply any such right. It grants only the right to exclude others.”).

9. For example, before entering interstate commerce, an insecticide manufacturer likely needs to obtain regulatory approval from the Environmental Protection Agency (EPA), and a radio broadcaster likely needs to obtain regulatory approval from the Federal Communications Commission (FCC).

10. See 21 U.S.C. § 355(a).

11. See *id.* §§ 355(b), (j) (describing the statutory requirements for obtaining FDA approval for new drugs).

that encourage patenting.¹² First, FDA approval standards vastly increase the cost of pharmaceutical innovation by requiring innovators to generate significant quantities of clinical (safety and effectiveness) data.¹³ Second, FDA approval standards massively decrease the cost of pharmaceutical imitation by permitting imitators (primarily, generics manufacturers) to reap the benefits of those same clinical data without having to generate them *de novo*.¹⁴ To make generating costly clinical data worthwhile, pharmaceutical innovators thus need market exclusivity—and patents help them get it.¹⁵

For a pharmaceutical innovator, then, patents are incredibly important—so long as the financial benefits of exclusivity are unlocked by FDA approval. Without approval, the innovator cannot produce, market, or sell their patented drug—and, as a result, they lose out on the highly supramarginal profits that are commonplace in the pharmaceutical industry.¹⁶ Of course, the exclusionary property rights that attach to drug patents can help delay the entry of generic competitors into markets for which FDA approval is part of the price of admission.¹⁷ But without FDA approval, the innovator company is *also* barred

12. See Talha Syed, *Should a Prize System for Pharmaceuticals Require Patent Protection for Eligibility?*, INCENTIVES FOR GLOB. HEALTH DISCUSSION PAPER NO. 2, at 14 (2009), https://healthimpactfund.org/pdf/DP2_Syed.pdf (suggesting that “the case for strong patent protection for pharmaceuticals may be largely based on the combination of regulatorily-mandated clinical trials for innovators and regulatorily-enabled piggybacking for imitators”).

13. See 21 U.S.C. § 355(b) (describing the statutory requirements for obtaining FDA approval for new drugs); Joseph A. DiMasi, Henry G. Grabowski & Ronald W. Hansen, *Innovation in the Pharmaceutical Industry: New Estimates of R&D Costs*, 47 J. HEALTH ECON. 20, 24 (2016) (showing, in Table 2, the average out-of-pocket clinical period costs for investigational compounds in Phases I, II, and III of regulatory approval).

14. See 21 U.S.C. § 355(j) (describing the statutory requirements for Abbreviated New Drug Applications or ANDAs).

15. See 35 U.S.C. § 154(a)(1) (“Every patent shall . . . grant to the patentee, his heirs or assigns, of the right to exclude others from making, using, offering for sale, or selling the invention throughout the United States or importing the invention into the United States.”); 35 U.S.C. § 154(a)(2) (“Subject to the payment of fees under this title, such grant shall be for a term beginning on the date on which the patent issues and ending 20 years from the date on which the application for the patent was filed in the United States.”).

16. See 21 U.S.C. § 355(a); Angus Liu, Eric Sagonowsky, Kevin Dunleavy, Fraiser Kansteiner & Zoey Becker, *The Top 10 Most Profitable Pharma Companies in 2021*, FIERCE PHARMA (June 14, 2022), <https://www.fiercepharma.com/special-reports/top-10-most-profitable-pharma-companies-2021> (ranking the top ten pharmaceutical companies by 2021 net income); ERIN H. WARD, KEVIN J. HICKEY & KEVIN T. RICHARDS, CONG. RESEARCH SERV., R46679, DRUG PRICING AND PHARMACEUTICAL PATENT PRACTICES 2 (2021), <https://crsreports.congress.gov/product/pdf/R/R46679> (“IP rights can deter or delay the market entry of generic drug or biosimilar competition, and thus may allow the rights holder to charge higher-than-competitive prices.”).

17. Rachel Sachs, *Drug Innovation: When Patents Work (And When They Don't)*, MILKEN INST. REV. (Oct. 23, 2020), <https://www.milkenreview.org/articles/drug-innovation-when->

from that same market.¹⁸ In other words, as much as patent exclusivity can help safeguard an innovator's market share once it is established, the innovator needs FDA approval to amass that market share in the first place.¹⁹

Few industries—if any—value patents as much as the pharmaceutical sector.²⁰ For proponents of strong patent rights, that can only be a good thing.²¹ Patent rights provide financial incentives to innovate—including, as noted above, compensation for the high cost of satisfying informationally demanding FDA approval standards.²² Patent rights also encourage early public disclosure of new inventions, which reduces duplicative research and development efforts between different inventors.²³ Since innovation and disclosure ultimately serve the public good, it is fair to reward the makers of new and useful drugs with patent exclusivity.²⁴ But what if our current system of granting patent exclusivity is not actually up to the task? What if pharmaceutical “inventors” are able to gain patent exclusivity for drug “inventions” that are, in reality, not inventive at all—is that still fair?

patents-work (“[R]egulatory exclusivities and patents function similarly, enabling innovators to block generic competitors from the market.”).

18. See 21 U.S.C. § 355(a) (“No person shall introduce or deliver for introduction into interstate commerce any new drug, unless an approval of an application filed pursuant to subsection (b) or (j) is effective with respect to such drug.”).

19. See Sachs, *supra* note 17 (“[R]egulatory exclusivities and patents function similarly, enabling innovators to block generic competitors from the market.”); 21 U.S.C. § 355(a) (“No person shall introduce or deliver for introduction into interstate commerce any new drug, unless an approval of an application filed pursuant to subsection (b) or (j) is effective with respect to such drug.”).

20. Sachs, *supra* note 17 (“[P]harmaceutical executives rate patents as far more important to innovation than do representatives of any other tech-driven business.”).

21. See generally MENELL ET AL., *supra* note 6, at 18–22 (describing the utilitarian justification for intellectual property rights).

22. See *id.* at 36 (“The public benefits directly [from patents] through the spur to innovation and disclosure of new technology.”); Amy Kapczynski & Talha Syed, *The Continuum of Excludability and the Limits of Patents*, 122 YALE L.J. 1900, 1908 (2013) (“Conventional economic actors will only produce a good when they can appropriate sufficient returns to recoup the capitalized costs of providing the good.”).

23. See Robert P. Merges & Richard R. Nelson, *On the Complex Economics of Patent Scope*, 90 COLUM. L. REV. 839, 871 (1990) (noting that granting patent rights after invention but before commercialization “allows the inventor to coordinate her activities with those of potential imitators to reduce ineffective duplication of inventive effort”).

24. See generally U.S. CONST. art. 1, § 8, cl. 8 (“To promote the Progress of Science and useful Arts, by securing for limited Times to Authors and Inventors the exclusive Right to their respective Writings and Discoveries.”); MENELL ET AL., *supra* note 6, at 19 (“The economic philosophy behind the [constitutional] clause empowering Congress to grant patents and copyrights is the conviction that it is the best way to advance public welfare through the talents of authors and inventors in ‘Science and useful Arts.’ Sacrificial days devoted to such creative activities deserve rewards commensurate with the services rendered.”).

1. *Patenting*

Patent applicants have several major hurdles to clear between filing a patent application and obtaining a valid patent. Chief among them is the need to show that their invention is both novel and nonobvious in view of any prior art. The novelty requirement²⁵ is relatively easy to understand. If a single prior art reference discloses each and every feature of the claimed invention, the reference anticipates the invention and renders it unpatentable.²⁶ The nonobviousness requirement²⁷ can seem somewhat less intuitive. However, at bottom, a patent applicant needs to demonstrate that their invention would not be obvious to a skilled artisan in view of the totality of the prior art's teachings (even if no single reference anticipates the invention).²⁸

2. *Obtaining Regulatory Approval*

To better understand how the FDA grants regulatory approval for new pharmaceutical products, consider one particular scenario: the approval of a small-molecule New Drug Application (NDA) under § 505 of the Federal Food, Drug, and Cosmetic Act.²⁹ Section 505 broadly describes the application process for legally marketing all new drugs in the United States.³⁰ If a drug product is entirely novel, in that it contains an active ingredient that has never been approved by the FDA, then the pharmaceutical innovator (known as “the sponsor”) will follow the regulatory pathway outlined in § 505(b)(1).³¹ To that end, the sponsor itself will generate (often from scratch) all the safety and efficacy data that are needed to support its case for regulatory approval.³²

If, however, the sponsor's new product is best described as a modification of a previously approved drug (e.g., a liquid formulation of an earlier-approved solid tablet), then the sponsor may be able to seek expedited approval under § 505(b)(2).³³ Even though, under a § 505(b)(2) framework, the sponsor still needs to provide full assurance of the drug's safety and efficacy to the FDA, they can satisfy some of those requirements by pointing to data that were

25. 35 U.S.C. § 102.

26. MPEP § 2131 (9th ed. Rev. 10.2019, June 2020).

27. 35 U.S.C. § 103.

28. MPEP § 2141 (9th ed. Rev. 10.2019, June 2020).

29. *See* 21 U.S.C. § 355.

30. *Id.*

31. *See* 21 U.S.C. § 355(b)(1).

32. *See* 21 U.S.C. § 355(b)(1)(A)(i).

33. *See* 21 U.S.C. § 355(b)(2).

submitted with the earlier-approved analog.³⁴ In practice, then, a § 505(b)(2) sponsor needs to first define the informational “bridge” between the already-approved product and their new formulation.³⁵ Then, they must provide whatever new data are needed to cross that bridge.³⁶ But because the informational bridge will, by its nature, always be shorter than the full regulatory pathway under § 505(b)(1), there are temporal and financial incentives for the sponsor to obtain approval under § 505(b)(2).³⁷

3. *Misaligned Incentives*

Briefly comparing these two processes—patenting and FDA approval via § 505(b)(2)—reveals parallel, though oppositely aligned, incentives.³⁸ Because of patent law’s novelty and nonobviousness requirements, a patent applicant will naturally want to create as much distance as possible between their invention and the prior art, which may include FDA-approved drug products that are already on the market. However, when that same inventor approaches the FDA to seek regulatory approval under § 505(b)(2), their prerogative is to emphasize the similarities between their new formulation and one or more of those same earlier-approved drug products. The more a sponsor can

34. FDA, DETERMINING WHETHER TO SUBMIT AN ANDA OR A 505(B)(2) APPLICATION 7–13 (May 2019), <https://www.fda.gov/media/124848/download> (outlining “Scientific Considerations for ANDAs and 505(b)(2) Applications”).

35. Ingrid Freije, Stéphane Lamouche & Mario Tanguay, *Review of Drugs Approved via the 505(b)(2) Pathway: Uncovering Drug Development Trends and Regulatory Requirements*, 54 THERAPEUTIC INNOVATION & REGUL. SCI. 128, 128 (2020) (“A drug submitted via 505(b)(2) can be approved based on data from studies not conducted by the sponsor, by relying on (1) Agency’s previous findings of safety and effectiveness (AFSE) of an approved drug; and/or (2) clinical and preclinical studies’ data from published literature without the right of reference. This requires not only a successful bridging to an RLD (reference listed drug), by the means of relative bioavailability (BA) or bioequivalence (BE) studies, but also some potential additional studies that may be needed to fully support efficacy and/or safety of the new product.”).

36. *Id.*

37. See Mitchell Katz, *Why Are 505(b)(2)s Gaining Increased Interest Among Midsize Biopharma Companies?*, LIFE SCI. LEADER (Feb. 7, 2018), <https://www.lifescienceleader.com/doc/why-are-b-s-gaining-increased-interest-among-midsize-biopharma-companies-0001> (explaining that approval under § 505(b)(2) “takes less time, cost, and risk to get product[s] onto the market because the active ingredient has been previously approved with data from a prior submission package”).

38. This misalignment of incentives could also occur, for example, when a pharmaceutical innovator submits a patent application with the USPTO and (1) an Investigational New Drug (IND) application, or (2) an NDA via § 505(b)(1) with the FDA in relation to the same small-molecule drug. Like § 505(b)(2) applications, IND applications and § 505(b)(1) applications may include information that is material to patentability. But the misalignment of incentives is particularly strong for § 505(b)(2) drug products because of the inherently comparative nature of the § 505(b)(2) pathway.

demonstrate a sameness between their formulation and its already-approved analogs, the shorter the § 505(b)(2) “bridge” to approval will be—and the cheaper and faster it will be for the sponsor to cross it.³⁹

Thus, there is a (potentially big) problem. The innovator is at once incentivized to share information about analogous competitor products with the FDA and to hide (or at least reframe) that same information when seeking patent exclusivity at the USPTO—even if the innovator suspects that the information speaks to the novelty or nonobviousness of their invention.⁴⁰ If they give into that temptation, they may well end up with a patent that does not meet the statutory requirements for patentability. Not only does such a patent offend the integrity of the patent system by undermining foundational principles of novelty and nonobviousness, but it also denies the public access to generics that are unfairly blocked by invalid patents.⁴¹

Unfortunately, it seems that at least some patent applicants do give into that temptation. In a 2021 decision, *Belcher Pharmaceuticals, LLC v. Hospira, Inc.* (“*Belcher IP*”), the Federal Circuit affirmed the District Court for the District of Delaware’s holding that a pharmaceutical patent assigned to Belcher Pharmaceuticals, LLC (“Belcher”) was unenforceable over a range of inconsistent representations that Belcher made when interacting with the USPTO and FDA.⁴² First, Belcher disclosed to the FDA information about similar third-party products that it later withheld from the USPTO.⁴³ Second, Belcher, when corresponding with the FDA, referred to the pH range of a competitor product as “old,” later asserting that that same pH range was “unexpectedly found to be critical” to its own invention when contesting an obviousness rejection at the USPTO.⁴⁴ As a result, the district court found—and the Federal Circuit agreed—that Belcher “did not merely withhold . . .

39. See Katz, *supra* note 37 (explaining why approval under § 505(b)(2) is faster, cheaper, and less risky).

40. See The Editorial Board, *Save America’s Patent System*, N.Y. TIMES (Apr. 16, 2022), <https://www.nytimes.com/2022/04/16/opinion/patents-reform-drug-prices.html> (“In 2014, for example, the E.P.A. discovered that some pesticide makers were routinely amplifying the novel effects of their latest products in patent applications, only to downplay the same effects to federal regulators. ‘They would tell the patent office that their pesticide deserved a patent because it was different than what was already out there,’ said Charles Duan, a public interest attorney and a member of the patent office’s public advisory committee . . . ‘Then they’d tell the E.P.A. that the same pesticide didn’t need extra regulatory clearance because it was no different than what was already out there.’ Experts have long warned that the same thing could easily be happening at the F.D.A.”).

41. See Sachs, *supra* note 17 (“[R]egulatory exclusivities and patents function similarly, enabling innovators to block generic competitors from the market.”).

42. 11 F.4th 1345 (Fed. Cir. 2021).

43. *Id.* at 1354.

44. *Id.* at 1350–51.

information but also used emphatic language” to make inconsistent statements to the USPTO and FDA.⁴⁵

Belcher II is, of course, just one case.⁴⁶ But it represents an important tipping point in the wider recognition of the inconsistent-representation problem that is besmirching the pharmaceutical industry.⁴⁷ Within a week of the *Belcher II* decision, U.S. Senator Patrick Leahy of Vermont and U.S. Senator Thom Tillis of North Carolina penned a bipartisan letter to the USPTO requesting that it “take steps to reduce patent applicants’ [sic] making inappropriate conflicting statements in submissions to the PTO and other federal agencies,” including the FDA.⁴⁸ The senators’ letter echoed the demands for meaningful reform that nonprofit organizations, such as I-MAK, had been making for years.⁴⁹ After *Belcher II*, it seems like the USPTO and FDA are finally starting to listen.⁵⁰

B. THE NEED FOR REFORM

On July 9, 2021, President Biden issued Executive Order No. 14306, entitled “Executive Order on Promoting Competition in the American Economy.”⁵¹ Section 5(p)(vi) of the Executive Order stated that, “to help ensure that the patent system, while incentivizing innovation, does not also unjustifiably delay generic drug and biosimilar competition,” the FDA should send a letter to the USPTO “enumerating and describing any relevant concerns.”⁵² Biden’s Executive Order set in motion a series of communications between the FDA and the USPTO. Though the various communications differ in substance and scope, they all share a common message: There is an urgent need for change.

On September 10, 2021, Janet Woodcock, then-Acting Commissioner of Food and Drugs, sent a letter to the USPTO in accordance with Executive

45. *Id.* at 1352.

46. That said, there are relatively few cases that make it to the courts. *See infra* Section IV.C.

47. *See* The Editorial Board, *supra* note 40 (describing the problem of inconsistent representation at the USPTO and FDA).

48. *See* Leahy & Tillis Letter, *supra* note 4, at 1.

49. *See* I-MAK, STRENGTHENING COMPETITION FOR PRESCRIPTION DRUGS THROUGH PATENT AND DRUG REGULATORY REFORM 6 (2022), <https://www.i-mak.org/strengthening-competition-blueprint/> (describing proposals to “expand interagency collaboration, starting with partnership between the PTO and the FDA”).

50. USPTO Director Vidal later acknowledged the senators’ letter in a post on the USPTO “Director’s Blog.” *See* Kathi Vidal, *Duty of Disclosure and Duty of Reasonable Inquiry Promote Robust and Reliable Patents, Drive Competition and Economic Growth, and Bring Life-Saving Drugs to the American People*, DIRECTOR’S BLOG (July 28, 2022, 5:34 AM), <https://www.uspto.gov/blog/director/entry/duty-of-disclosure-and-duty>.

51. Exec. Order No. 14306, 86 Fed. Reg. 36,987, 36,997 (July 9, 2021).

52. *Id.*

Order No. 14306.⁵³ Woodcock wrote generally of her desire to increase “engagement between FDA and USPTO,” including, for example, offering USPTO Examiners “training on FDA’s public information and databases that may help USPTO locate pertinent references.”⁵⁴

Then, on July 6, 2022, USPTO Director Kathi Vidal, in response to Woodcock’s letter, asserted her desire to work with the FDA on “[e]xplor[ing] consistency in representations made to the USPTO and FDA,” such as “initiatives to require patent applicants to provide relevant information to the USPTO that has been submitted to other agencies.”⁵⁵ Further, on July 29, 2022, Director Vidal published a Notice in the Federal Register that broadly discussed the duties of disclosure and reasonable inquiry during patent prosecution.⁵⁶ Most notably, in Section V of the Notice, Director Vidal explained that:

“Each individual with a duty to disclose, or party with a duty of reasonable inquiry, should ensure that the statements made to the USPTO and other Government agencies, or any statements made on their behalf to other Government agencies regarding the claimed subject matter, are consistent Providing material information to other Government agencies, including the FDA, while simultaneously withholding the same information from the USPTO undermines both the intent and spirit of the duty of disclosure and violates those duties.”⁵⁷

Director Vidal specifically outlined several instances in which it may be incumbent upon patent applicants (or any other party involved in patent prosecution who has a duty to disclose) to share information with the USPTO that has arisen through dealings with other government agencies.⁵⁸ For example, a party with a duty to disclose should always review information they receive from other government agencies in relation to their invention to determine whether that information should be shared with the USPTO.⁵⁹ To

53. Letter from Janet Woodcock, Acting Comm’r Food & Drugs, to Andrew Hirshfeld, Performing Functions & Duties Under Sec’y Com. for Intell. Prop. & Dir. USPTO 4 (Sept. 10, 2021), <https://www.fda.gov/media/152086/download>.

54. *Id.*

55. Letter from Katherine K. Vidal, Under Sec’y Com. for Intell. Prop. & Dir. USPTO, to Robert M. Califf, Comm’r Food & Drugs 3–4 (July 6, 2022), <https://www.uspto.gov/sites/default/files/documents/PTO-FDA-nextsteps-7-6-2022.pdf>.

56. Duties of Disclosure and Reasonable Inquiry During Examination, Reexamination, and Reissue, and for Proceedings Before the Patent Trial and Appeal Board, 87 Fed. Reg. 45,764, 45,764–67 (July 29, 2022) [hereinafter July 2022 Notice].

57. *Id.* at 45,766.

58. *Id.* at 45,766–67.

59. *Id.*

illustrate this point, Director Vidal noted that pharmaceutical patentees who receive Paragraph IV certifications should review the certification to determine whether the factual and legal bases of the Paragraph IV challenge contain information that is material to the patentability of matters still pending before the USPTO (e.g., in a continuation application within the same family).⁶⁰ Likewise, patent practitioners violate their duty of good faith and candor under 37 C.F.R. § 1.56(a) when they devise deliberate schemes to prevent individuals with a duty to disclose from obtaining relevant information in the first place.⁶¹ As a consequence, the duty to disclose cannot be circumvented by “walling off the patent prosecution practitioners from the attorneys seeking FDA approval.”⁶²

On October 4, 2022, Director Vidal published a second Notice in the Federal Register requesting public comments on “proposed initiatives directed at bolstering the robustness and reliability of patents.”⁶³ The Notice described the letters previously exchanged between the USPTO and FDA, reiterating that the “USPTO could work with the FDA to ensure that our patent system properly and adequately protects innovation while not unnecessarily delaying generic and biosimilar competition.”⁶⁴ However, neither the specific USPTO initiatives described in the Notice nor the questions ultimately submitted for public comment referred to increased USPTO-FDA collaboration.⁶⁵

On November 7, 2022, Director Vidal published a third Notice in the Federal Register, in which she outlined a “Public Listening Session” to be jointly hosted by the USPTO and FDA on January 19, 2023.⁶⁶ In preparation

60. *Id.* Pursuant to 21 U.S.C. § 355(b)(2)(A)(iv), a “Paragraph IV” submission is made when a generic applicant includes “in its application a ‘certification’ that a patent submitted to FDA by the brand-name drug’s sponsor and listed in FDA’s [Orange Book] is, in the generic applicant’s opinion and to the best of its knowledge, invalid, unenforceable, or will not be infringed by the generic product.” FDA, PATENT CERTIFICATIONS AND SUITABILITY PETITIONS (2022), <https://www.fda.gov/drugs/abbreviated-new-drug-application-anda/patent-certifications-and-suitability-petitions>. Examples of “continuing applications” include continuation, divisional, and continuation-in-part applications. *See* MPEP § 201.02 (9th ed. Rev. 10.2019, June 2020); *see generally* Chen Chen, *Using Continuation Applications Strategically*, COOLEYGO, <https://www.cooleygo.com/using-continuation-applications-strategically/> (last visited Dec. 13, 2023) (describing continuation applications in the context of patent portfolio development).

61. July 2022 Notice, *supra* note 56, at 45,767.

62. *Id.*

63. Request for Comments on USPTO Initiatives to Ensure the Robustness and Reliability of Patent Rights, 87 Fed. Reg. 60,130, 60,130–34 (Oct. 4, 2022).

64. *Id.* at 60,130.

65. *See id.*

66. Joint USPTO–FDA Collaboration Initiatives; Notice of Public Listening Session and Request for Comments, 87 Fed. Reg. 67,019, 67,019–20 (Nov. 7, 2022) [hereinafter November

for the session, Director Vidal specifically requested written comments in response to the following question: “What mechanisms could assist patent examiners in determining whether patent applicants have submitted inconsistent statements to the USPTO and the FDA?”⁶⁷ Part III of this Note responds directly to Director Vidal’s request.

III. TACKLING INCONSISTENT REPRESENTATIONS DURING PATENT PROSECUTION

Director Vidal has, as outlined above, spoken repeatedly of her desire to increase USPTO-FDA interaction in ways that would help ensure the robustness of the patent system. However, the USPTO and FDA have yet to provide any detail on what this increased interaction would look like in practice. This Part, in response, proposes a new system of USPTO-FDA interaction that, to the extent possible, undercuts inconsistent representation before a patent issues—that is, during examination of a nonprovisional patent application.

A. A NEW SYSTEM OF USPTO-FDA INTERACTION

This Section argues that the USPTO should require that, when a patent applicant files an NDA relating to the same subject matter, they must provide the NDA application number to the USPTO. Then, USPTO Examiners must (1) search for the NDA in the FDA’s public databases, (2) review the information contained within the NDA submission, and (3) factor any relevant information into their patentability (in particular, novelty and nonobviousness) assessments during examination. The NDA information will then become part of the prosecution file for each patent application.

As an initial matter, the burden of creating and maintaining this (or any) new interagency system should lie primarily with the USPTO rather than the FDA. The FDA has made it clear that it is not—and has no desire to become—a patenting body.⁶⁸ In addition, the responsibility of overseeing interagency conduct that serves to reinforce duties of disclosure, good faith, and candor should fall on the agency that creates and perpetuates those

2022 Notice]; *see also* Joint USPTO-FDA Collaboration Initiatives; Notice of Public Listening Session and Request for Comments, 87 Fed. Reg. 11,902, 11,902–03 (Feb. 24, 2023) (explaining that the deadline for comment had been extended through March 10, 2023).

67. *Id.* at 67,021–22.

68. *See* Woodcock, *supra* note 53, at 2 (“FDA has an important but ministerial role with respect to patents.”).

duties—namely, the USPTO.⁶⁹ Because any new system of USPTO-FDA interaction will ultimately be intended to help the USPTO properly assess (or, as the case may be, reassess) patentability, it makes sense that USPTO Examiners will bear the burden of collecting and using FDA submissions to facilitate such assessments.⁷⁰

What will the new system of USPTO-FDA interaction look like in practice? When a patent applicant files a nonprovisional application, they will also be required to submit to the USPTO the application numbers of any relevant NDAs pending at or approved by the FDA.⁷¹ Then, during patent prosecution, the Examiner will use that NDA information to access and review publicly available FDA records—for example, using the Drugs@FDA database and the FDA's Orange Book.⁷² The Drugs@FDA database, in particular, contains (often redacted) correspondence between the FDA and the pharmaceutical sponsor, including approval letters, review letters, and general correspondence.⁷³ As a result, the Examiner will likely be able to note, for example, whether the sponsor claimed that their product was comparable to an already-approved Reference Listed Drug (RLD) as part of a § 505(b)(2) application.⁷⁴ Thus, even if substantive comments about the already-approved product are redacted, the Examiner may still gain baseline knowledge about

69. See MPEP § 2001 (9th ed. Rev. 10.2019, June 2020) (describing the Duty of Disclosure, Candor, and Good Faith).

70. Arguably, NDA information should still be submitted to the relevant prosecution files of relevant *issued* patents. In those cases, the Examiner can review the publicly available NDA records and, if they find information that raises new questions of patentability, the Examiner should be permitted to re-open prosecution. This process could, in many ways, mimic existing post-issuance proceedings such as *ex parte* reexamination. See 35 U.S.C. § 302.

71. To understand the differences between provisional and nonprovisional patent applications in the United States, see MPEP § 201 (9th ed. Rev. 10.2019, June 2020).

72. See Woodcock, *supra* note 53, at 4; *Drugs@FDA: FDA-Approved Drugs*, FDA, <https://www.accessdata.fda.gov/scripts/cder/daf/index.cfm>; FDA, APPROVED DRUG PRODUCTS WITH THERAPEUTIC EQUIVALENCE EVALUATIONS | ORANGE BOOK (Oct. 17, 2022), <https://www.fda.gov/drugs/drug-approvals-and-databases/approved-drug-products-therapeutic-equivalence-evaluations-orange-book> (“The publication *Approved Drug Products with Therapeutic Equivalence Evaluations* (commonly known as the Orange Book) identifies drug products approved on the basis of safety and effectiveness by the Food and Drug Administration (FDA) under the Federal Food, Drug, and Cosmetic Act (the Act) and related patent and exclusivity information.”).

73. See *Drugs@FDA: FDA-Approved Drugs*, FDA, <https://www.accessdata.fda.gov/scripts/cder/daf/index.cfm>.

74. See FDA, *supra* note 34, at 2 n.7 (“The RLD ‘is the listed drug identified by FDA as the drug product upon which an applicant relies in seeking approval of its ANDA.’ 21 CFR 314.3(b). Because an ANDA applicant is relying upon FDA’s finding that the RLD is safe and effective, FDA’s practice is to designate as RLDs drug products that have been approved for safety and effectiveness.”).

the existence of a comparable, already-approved product, which may well qualify as prior art that is material to patentability.

B. WHAT THIS NEW SYSTEM ACHIEVES—AND WHAT IT DOES NOT

Revisiting the *Belcher II* decision provides a helpful example of the potential usefulness of the new system of USPTO-FDA interaction proposed in this Part.⁷⁵ In that case, Belcher first submitted a § 505(b)(2) NDA for its epinephrine formulation in November 2012, which the FDA ultimately approved in July 2014.⁷⁶ The Drugs@FDA entry for Belcher’s formulation then became publicly accessible in March 2015.⁷⁷ Meanwhile, Belcher filed a nonprovisional patent application (claiming the same formulation) with the USPTO in August 2014.⁷⁸ The application later issued as a U.S. Patent in March 2016.⁷⁹ Thus, a full year lapsed between Belcher’s NDA being released on publicly available FDA databases (March 2015) and Belcher’s patent being issued by the USPTO (March 2016).

If the USPTO-FDA system proposed in this Part had been in place at that time, the USPTO Examiner would have had a full twelve months to review the publicly available components of Belcher’s NDA submissions on Drugs@FDA.⁸⁰ Had the Examiner undertaken such a review, they would have been made aware, for example, that Belcher listed Twinject, an already-approved epinephrine formulation, as an RLD in its § 505(b)(2) application.⁸¹ As explained by the district court in *Belcher I*, Twinject used the “old” pH range that Belcher later described as “critical” when trying to patent its own formulation at the USPTO.⁸² At minimum, then, notice of Belcher’s NDA would have made the Examiner aware of a patently material third-party product that Belcher did not disclose to the USPTO, despite Belcher’s belief

75. See *supra* Section II.A.3.

76. NDA NO. 205029, DRUGS@FDA: FDA-APPROVED DRUGS (Mar. 31, 2015), https://www.accessdata.fda.gov/drugsatfda_docs/nda/2014/205029Orig1s000TOC.cfm.

77. *Id.*

78. U.S. Patent Application No. 14/460,845 (filed Aug. 15, 2014).

79. U.S. Patent No. 9,283,197; *Belcher Pharms., LLC v. Hospira, Inc. (Belcher II)*, 11 F.4th 1345, 1349–50 (Fed. Cir. 2021).

80. Assuming, of course, that the patenting timeline remained otherwise unaltered.

81. See, e.g., FDA, PHARMACOLOGY REVIEW 46 (Jan. 30, 2013), https://www.accessdata.fda.gov/drugsatfda_docs/nda/2014/205029Orig1s000PharmR.pdf (“This submission is primarily based on published literature as a 505(b)(2) application and safety information from Twinject (NDA 020800, an approved drug) as the listed reference drug.”).

82. *Belcher Pharms., LLC v. Hospira, Inc. (Belcher I)*, 450 F. Supp. 3d 512, 522–23 (D. Del. 2020). Note that, while Belcher specifically described the pH range of another third-party formulation—that of Sintetica SA—as “old,” the *Belcher I* court explained that the Twinject and Sintetica SA formulations “had approximately the same pH.” *Id.*

that such a product was similar enough to its own to be listed as an RLD in its NDA.

But a new system of USPTO-FDA interaction is not just useful—it is necessary. First, creating a system of dual requirements—that patent applicants disclose NDA information and that USPTO Examiners review publicly available FDA records—increases the likelihood that the relevant information will be provided to and considered by the Examiner. To be sure, the duty of disclosure already encompasses the requirement to share with the USPTO material information submitted to other government agencies.⁸³ And USPTO Examiners have always been able to access FDA’s public databases without a formal system of USPTO-FDA interaction. But the status quo is clearly not working, at least with the effectiveness needed to tackle the problem of inconsistent representation.⁸⁴ Establishing a system of explicit disclosure and review requirements for patent applicants and Examiners, respectively, is thus necessary to uphold “both the intent and spirit of the duty of disclosure.”⁸⁵

Second, the FDA Orange Book—the closest existing analog of the proposed new system—only lists issued patents for approved drugs.⁸⁶ Consequently, by the time an Examiner is able to use the Orange Book to link patents and FDA records, prosecution is long over.⁸⁷ A system that instead connects FDA submissions to pending patent applications (at least some of the time) will help tackle inconsistent representation in the most effective way possible: *before* an invalid patent actually issues.⁸⁸ This will save critical USPTO resources in the long run by shortening the time spent by Examiners prosecuting ultimately unpatentable inventions.

The new system proposed herein will also have beneficial outcomes both for generics manufacturers and the general public. As things stand, an accused infringer has to wait for (in reality, invalid) patents to issue and appear in the

83. See July 2022 Notice, *supra* note 56, at 45,766.

84. See November 2022 Notice, *supra* note 66, at 67,021–22 (requesting public comment on possible mechanisms for tackling the inconsistent representation problem); Woodcock, *supra* note 53, at 4 (suggesting that USPTO Examiners could benefit from “training on FDA’s public information and databases that may help USPTO locate pertinent references”).

85. See July 2022 Notice, *supra* note 56, at 45,766.

86. See FDA, APPROVED DRUG PRODUCTS WITH THERAPEUTIC EQUIVALENCE EVALUATIONS | ORANGE BOOK (Oct. 17, 2022), <https://www.fda.gov/drugs/drug-approvals-and-databases/approved-drug-products-therapeutic-equivalence-evaluations-orange-book> (describing the Orange Book).

87. Of course, prosecution may be ongoing for other applications in the patent family. But there is currently no straightforward way for an Examiner to know that the patent application she is assessing is part of a family with issued patents listed in the Orange Book.

88. Again, that is not to say that patentees should not also be required to submit NDA information for *issued* patents. See *supra* note 70.

Orange Book before they can submit Paragraph IV unenforceability certifications based on inconsistent representation.⁸⁹ This delay has significant financial costs for generics manufacturers while they are frozen out of the market.⁹⁰ More importantly, it denies the public access to generic medicines that are unfairly blocked by invalid patents.⁹¹ It is thus imperative that the USPTO devise a system, such as the one proposed in this Part, that allows patent Examiners to access FDA submissions as early as possible in the patent prosecution timeline.

Admittedly, the proposed new system of USPTO-FDA interaction would be far from perfect. One significant problem is that our current patent prosecution and FDA approval processes suffer from a fundamental—if not fatal—incompatibility. Patent prosecution, unlike the FDA approval process, is inherently public.⁹² Indeed, USPTO Director Vidal, writing in her November 2022 Notice in the Federal Register, seemed to anticipate the problematic nature of this private-public dichotomy: As part of her request for mechanisms to tackle inconsistent representation, Director Vidal asked commenters to “explain whether such mechanisms present confidentiality concerns and, if so, how those concerns could be addressed.”⁹³

Likewise, the public-private problem has not gone unnoticed by nonagency advocates of greater USPTO-FDA interaction. For instance, I-MAK, a nonprofit organization, has suggested that pharmaceutical patent applicants should be required to submit copies of all FDA filings with the USPTO during prosecution.⁹⁴ Acknowledging the need to “avoid any issues relating to trade secrets,” I-MAK suggested that “[t]he sharing of information on drug products between FDA and PTO could be made through a

89. See FDA, *supra* note 60 (describing Paragraph IV certifications as governed by 21 U.S.C. § 355(b)(2)(A)(iv)).

90. See Sachs, *supra* note 17 (“Lower-priced generic versions of these drugs may not appear for decades—and may be delayed beyond the expected date by patent holders’ arcane strategies for extending their legal monopolies. In the meantime, patent holders may have no qualms about raising their prices year after year, putting their products even further out of reach.”).

91. See *id.*

92. Compare 37 C.F.R. § 1.11 (2012) (explaining the public components of patent prosecution), with 21 C.F.R. § 314.430 (2008) (discussing the “[a]vailability for public disclosure of data and information in an [NDA] or abbreviated application”). For this reason, this Part suggests that patent applicants should only be required to submit *NDA* information to the USPTO since those applications (or at least parts of them) eventually become accessible to the public. Other types of FDA submissions, e.g., INDs, generally do *not* become public. See 21 C.F.R. § 312.130 (2003) (discussing the “[a]vailability for public disclosure of data and information in an IND”).

93. See November 2022 Notice, *supra* note 66, at 67,022.

94. I-MAK, *supra* note 49, at 6.

memorandum of understanding.”⁹⁵ The problem with I-MAK’s recommendation is that the Examiner cannot keep confidential a patent applicant’s statements from an NDA and also use them as the basis of a novelty or obviousness rejection that becomes part of an entirely public patent prosecution record.⁹⁶ To be sure, the *Belcher II* timeline outlined at the beginning of this Section does demonstrate that even publicly accessible NDA information could, under the system proposed in this Part, prove useful to patentability assessments. But as long as patent prosecution and FDA approval continue to operate in inherently incompatible public and private spheres, USPTO Examiners can only ever be required to access publicly available information in FDA databases.

A second issue is that pharmaceutical companies typically file patent applications long before they submit corresponding NDAs.⁹⁷ To compound the problem, the FDA adds information about drug products to its publicly accessible databases only after the drug has been approved, which typically occurs six to ten months after NDA submission.⁹⁸ As a consequence, it is entirely possible that one or more patents will have already issued in a patent family covering the product for which the pharmaceutical entity later obtains regulatory approval.⁹⁹ For these patents, the new system of USPTO-FDA interaction and the existing FDA Orange Book would, in essence, become mirror images of each other: The USPTO’s file wrapper would contain NDA

95. *Id.*

96. See 37 C.F.R. § 1.11 (describing the public components of patent prosecution).

97. Consider the following sample timelines. According to data from 2000 to 2010, the time interval between filing a provisional patent application with the USPTO and receiving an IND effective date (which typically occurs up to 30 days after filing the initial IND application) for a New Chemical Entity at the FDA can be as long as 4.7 years without sacrificing market exclusivity. Michael K. Dunn, *Timing of Patent Filing and Market Exclusivity*, 10 NATURE REVS. DRUG DISCOVERY 487, 488 (2011). In contrast, in January 2023, the mean time between filing a nonprovisional application and receiving a final disposition (patent issuance or abandonment of the application) in the 1600 Technology Center was 2.3 years. *Patents Pendency Data October 2023*, USPTO, <https://www.uspto.gov/dashboard/patents/total-pendency-by-tc.html> (last visited Dec. 13, 2023). Thus, as a very rough estimate (assuming consistency over time, etc.), a typical pharmaceutical patentee will obtain a patent for their drug product more than one year before they file an IND application at the FDA. Note that pharmaceutical sponsors tend to submit an IND approximately 5–7 years before filing the corresponding NDA. See Martin S. Lipsky & Lisa K. Sharp, *From Idea to Market: The Drug Approval Process*, 14 J. AM. BD. FAMILY PRAC. 362, 365 (2001).

98. See FDA, STEP 4: FDA DRUG REVIEW (Jan. 4, 2018), <https://www.fda.gov/patients/drug-development-process/step-4-fda-drug-review#>.

99. See *Eli Lilly & Co. v. Actavis Elizabeth LLC*, 731 F. Supp. 2d 348, 376 (D.N.J. 2010) (“Indeed, most drugs are patented long before their commercial use is approved by the FDA.”).

information for approved drugs and the FDA Orange Book would provide relevant patent information for those same approved drugs.

Arguably, then, the real value of this new USPTO-FDA system lies in its potential for circumventing patent “evergreening.”¹⁰⁰ Because most drugs are covered by multiple patents, it is likely that a significant number of continuing applications¹⁰¹ will remain pending after initial FDA approval.¹⁰² For example, many pharmaceutical companies, when patenting a small-molecule drug product, will first patent the chemical entity, and then subsequently patent specific formulations, methods of treatment, and dosing.¹⁰³ NDA submissions often contain therapeutically specific safety and efficacy information that is more relevant to these later-issued patents.¹⁰⁴ It is therefore possible that later-filed patent applications covering these aspects of the invention will still be undergoing prosecution when the corresponding NDAs are submitted. Should that be the case, any information that is material to patentability—assuming it

100. See Rebecca S. Eisenberg, *The Role of the FDA in Innovation Policy*, 13 MICH. TELECOMM. & TECH. L. REV. 345, 354 (2007) (describing “evergreening” as the practice by which patentees seek “to prolong their effective periods of patent protection through . . . strategies that add new patents to their quivers as old ones expire”); see also WARD, HICKEY & RICHARDS, *supra* note 16, at 41–45 (providing more information about common “evergreening” practices).

101. Examples of “continuing applications” include continuation, divisional, and continuation-in-part applications. See MPEP § 201.02 (9th ed. Rev. 10.2019, June 2020); see generally Chen Chen, *Using Continuation Applications Strategically*, COOLEYGO, <https://www.cooleygo.com/using-continuation-applications-strategically/> (last visited Dec. 13, 2023) (describing continuation applications in the context of patent portfolio development).

102. See Robin Feldman, *May Your Drug Price Be Evergreen*, 5 J.L. & BIOSCIENCES 590, 601–02 (2018) (“Simple techniques can involve obtaining new protections on existing drugs by filing additional patents, sometimes on methods of producing or manufacturing the drugs More complex evergreening strategies involve developing new formulations, dosage schedules, or combinations that can be used to obtain new patents”); Uri Y. Hachoen, *Evergreening at Risk*, 33 HARV. J.L. & TECH. 479, 486 (2020) (“In [the pharmaceutical] industry, patents of negligible market value are sometimes disproportionately rewarded by allowing brand-name manufacturers to artificially extend their monopolies over existing drugs when their current legal protections are about to expire.”).

103. See M. David Weingarten & Shana K. Cyr, *Securing and Maintaining a Strong Patent Portfolio for Pharmaceuticals*, 10 ACS MEDICINAL CHEMISTRY LETTERS 838, 839 (2019) (“Once researchers begin to generate novel compounds that show relevant biological activity, patent applications may be filed on potential drug candidates, both specifically and generically, and their methods of use. As these potential drug candidates advance through preclinical and then clinical development, applications may be filed on further scientific advances such as new dosage forms, potential new uses, methods of administration, and possible novel drug combinations with other known drugs.”).

104. See FDA, BIOAVAILABILITY AND BIOEQUIVALENCE STUDIES SUBMITTED IN NDAS OR INDS—GENERAL CONSIDERATIONS (Mar. 2014), <https://www.fda.gov/media/88254/download> (describing the types of safety and efficacy information often included in FDA submissions).

is publicly available—can be assessed during the prosecution of those applications and, if the timing is right, prevent patents from issuing.¹⁰⁵

Lastly, this new system of USPTO-FDA interagency disclosure could suffer from noncompliance. To be sure, it would likely be more difficult for a patent applicant to offer a good-faith explanation of their decision to forgo a simple and explicit mandate—timely disclosure of FDA submission details to the USPTO—than it might be for them to excuse a failure to comply with a more amorphous duty to disclose.¹⁰⁶ But, as in any administrative system, at least some participants will default on their duties (intentionally or otherwise) and fail to provide relevant NDA information to the USPTO.

Taken together, issues of confidentiality, timing, and noncompliance would likely undermine, at least to some extent, the overall usefulness of the proposed system of USPTO-FDA interaction. This Note contends that such a system could nonetheless play an important role in undercutting inconsistent representation—especially in large, multi-generational patent families with drawn-out prosecution timelines. Further, all patent applicants would arguably be discouraged from making inconsistent representations in the first place: Because Examiners would have notice of and access to FDA records (later, if not sooner), there would be less incentive to try and game the system from the outset. But the proposed new system of USPTO-FDA interaction would certainly not be foolproof. What is needed, then, is a safety net. Accused patent infringers must have an effective means by which they can challenge the validity of a pharmaceutical patent obtained in spite of (if not because of) inconsistent representation at the USPTO and FDA.¹⁰⁷ And that, this Note proposes, is where the courts come in.

IV. TACKLING INCONSISTENT REPRESENTATIONS AFTER PATENT ISSUANCE

Under current law, if a court finds that a patentee engaged in inequitable conduct during patent prosecution, the whole patent is rendered

105. Again, that is not to say that patentees should not also be required to submit NDA information for *issued* patents. See *supra* note 70.

106. It is possible (though purely speculative) that similar reasoning underscored Director Vidal's decision to recently clarify that the duty to disclose already encompasses the need to make consistent representations to government agencies. See July 2022 Notice, *supra* note 56, at 45,764–67.

107. Issues of patentee noncompliance aside, an invalid patent might also be granted if the USPTO Examiner failed to recognize the materiality of information contained in an FDA submission.

unenforceable.¹⁰⁸ In theory, then, inequitable conduct doctrine *should* provide a useful mechanism for accused patent infringers to challenge the enforceability of a patent obtained through deception—including where that deception is evidenced by inconsistent representation at the USPTO and FDA.¹⁰⁹ But in practice, it is all but impossible for defendants in patent infringement lawsuits to raise a successful inequitable conduct claim because of the exceptionally high legal standards outlined by the Federal Circuit in *Therasense*.¹¹⁰

This Part argues that the Federal Circuit should revise its inequitable conduct doctrine to create a “pharma exception” to *Therasense*. The court should hold that, when an accused infringer shows that a patentee (1) failed to disclose to the USPTO references it shared with the FDA to support its case for regulatory approval, or (2) made inconsistent or contradictory statements to the USPTO and the FDA, there should be a rebuttable presumption that *both* the materiality *and* the intent prongs of the *Therasense* inequitable conduct test are satisfied.¹¹¹ By adopting this change, the court would revitalize an important post-patent-issuance mechanism for tackling the problem of inconsistent representation.

A. INEQUITABLE CONDUCT DOCTRINE UNDER *THERASENSE*

The remedy for a finding of inequitable conduct—whole-patent unenforceability—is the “atomic bomb” of patent law.¹¹² Patentees—even ones who are ultimately successful—must defend their good names against accusations of bad faith.¹¹³ The attorney who prosecuted the application will undoubtedly face devastating consequences to their professional reputation.¹¹⁴ Perhaps unsurprisingly, then, courts have struggled to strike the right balance between “ensur[ing] . . . candor and truthfulness” on the part of patent

108. *Aventis Pharma S.A. v. Amphastar Pharms., Inc.*, 525 F.3d at 1334, 1349 (Fed. Cir. 2008) (Rader, J., dissenting).

109. *See Therasense, Inc. v. Becton, Dickinson & Co.*, 649 F.3d 1276, 1287 (Fed. Cir. 2011) (en banc) (“As the inequitable conduct doctrine evolved . . . it came to embrace a broader scope of misconduct, including not only egregious affirmative acts of misconduct intended to deceive both the PTO and the courts but also the mere nondisclosure of information to the PTO.”).

110. *See id.* at 1290–95; Frederick Frei & Sean Wooden, *Inequitable Conduct Claims One Year After Therasense*, 221 *MANAGING INTELL. PROP.* 66, 66 (2012) (“After the holding in *Therasense*, it was widely believed that the court had sounded the death knell to the inequitable conduct defense by imposing evidentiary requirements that could rarely be met.”).

111. These two criteria mirror Director Vidal’s framing of the inconsistent representation problem. *See* July 2022 Notice, *supra* note 56, at 45,766.

112. *Aventis*, 525 F.3d at 1349 (Rader, J., dissenting).

113. *Therasense*, 649 F.3d at 1288.

114. *Id.*

applicants and nurturing the incentive to seek patent protection in the first place.¹¹⁵

The origins of inequitable conduct doctrine in patent law can be traced to the unclean hands doctrine.¹¹⁶ The Supreme Court laid the foundations of modern inequitable conduct doctrine in three germinal cases: *Keystone Driller Co. v. General Excavator Co.*,¹¹⁷ *Hazel-Atlas Glass Co. v. Hartford-Empire Co.*,¹¹⁸ and *Precision Instrument Manufacturing Co. v. Automotive Maintenance Machinery Co.*¹¹⁹ Shortly thereafter, the Patent Act of 1952 triggered the advent of whole-patent unenforceability as the remedy for inequitable conduct.¹²⁰ The creation of the Federal Circuit in 1982 brought much-needed uniformity to inequitable conduct doctrine.¹²¹ Because earlier cases had involved such flagrant misconduct, and the claims had arisen in equity, not law, the Supreme Court had been unable to articulate clear legal standards to guide lower courts.¹²² Fortunately, the Federal Court, since its inception, has been consistent in requiring that two elements be satisfied for a showing of inequitable conduct: materiality and intent.¹²³ Unfortunately, the consistency ends there.¹²⁴

The pleading and legal standards that govern inequitable conduct defenses have changed considerably and frequently over the past four decades.¹²⁵

115. See *Aventis*, 525 F.3d at 1349 (Rader, J., dissenting).

116. *Therasense*, 649 F.3d at 1285 (“Inequitable conduct is an equitable defense to patent infringement that, if proved, bars enforcement of a patent. This judge-made doctrine evolved from a trio of Supreme Court cases that applied the doctrine of unclean hands to dismiss patent cases involving egregious misconduct.”).

117. 290 U.S. 240 (1933).

118. 322 U.S. 238 (1944).

119. 324 U.S. 806 (1945).

120. 35 U.S.C. § 282(b)(1).

121. For a more detailed description of the evolution of inequitable conduct doctrine in lower courts prior to the establishment of the Federal Circuit, see Robert J. Goldman, *Evolution of the Inequitable Conduct Defense in Patent Litigation*, 7 HARV. J.L. & TECH. 37, 52–67 (1993).

122. Robert Swanson, *The Exergen and Therasense Effects*, 66 STAN. L. REV. 695, 700 (2014).

123. See *id.* at 701 (“For the entire duration of the Federal Circuit’s existence, it has been clear that inequitable conduct has two elements: materiality and intent.”).

124. See *id.* (explaining that, even after the creation of the Federal Circuit, “the elements needed to prove inequitable conduct were often vague and shifting”).

125. *Therasense, Inc. v. Becton, Dickinson & Co.*, 649 F.3d 1276, 1287–88 (Fed. Cir. 2011) (en banc) (“[T]he standards for intent to deceive and materiality have fluctuated over time.”); see, e.g., *Am. Hoist & Derrick Co. v. Sowa & Sons, Inc.*, 725 F.2d 1350, 1363 (Fed. Cir. 1984) (finding that materiality and intent exist on a sliding scale); *Hoffman-LaRoche Inc. v. Lemmon Co.*, 906 F.2d 684, 688 (Fed. Cir. 1990) (holding that gross negligence is insufficient for a finding of intent); *Ferring B.V. v. Barr Lab’s, Inc.*, 437 F.3d 1181, 1191 (Fed. Cir. 2006) (finding that intent may be presumed in the absence of a credible explanation for gross negligence on the part of the patentee); *Star Sci., Inc. v. R.J. Reynolds Tobacco Co.*, 537 F.3d

Exergen Corp. v. Wal-Mart Stores, Inc. marked a particularly notable shift.¹²⁶ In that case, the Federal Circuit held that, because inequitable conduct is a type of fraud, it demands a heightened pleading standard—specifically, it must be pleaded with particularity per Rule 9(b) of the Federal Rules of Civil Procedure.¹²⁷ Under materiality, the *Exergen* court held, the accused infringer’s plea must “identify the specific who, what, when, where, and how of the material misrepresentation or omission.”¹²⁸ Under intent, the party raising the inequitable conduct defense must include sufficient factual detail for a court to “infer that a specific individual (1) knew of the withheld information or of the falsity of the material misrepresentation, and (2) withheld or misrepresented this information with a specific intent to deceive the PTO.”¹²⁹

A second major change occurred after *Therasense*. In response to concerns over the perceived leniency of its inequitable conduct doctrine, the Federal Circuit heightened the legal standards for both the materiality and the intent prongs of its two-part test.¹³⁰ Now, under materiality, the defendant must show, by a preponderance of the evidence, that the information in question is but-for material to patentability such that the USPTO would not have allowed a claim if it had been aware of the information.¹³¹ Importantly, this but-for

1357, 1366–67 (Fed. Cir. 2008) (holding that materiality and intent must be established separately by clear and convincing evidence before the court can engage in balancing the facts and equities of unenforceability).

126. See 575 F.3d 1312 (Fed. Cir. 2009).

127. *Id.* at 1327.

128. *Id.* at 1328.

129. *Id.* at 1328–29.

130. *Therasense*, 649 F.3d at 1290 (“While honesty at the PTO is essential, low standards for intent and materiality have inadvertently led to many unintended consequences, among them, increased adjudication cost and complexity, reduced likelihood of settlement, burdened courts, strained PTO resources, increased PTO backlog, and impaired patent quality. This court now tightens the standards for finding both intent and materiality in order to redirect a doctrine that has been overused to the detriment of the public.”).

131. *Id.* at 1291–92. Prior to *Therasense*, the materiality standard had been pegged to the USPTO’s own (fluctuating) materiality standard under Rule 56. But now, the standard under *Therasense* is stricter than the USPTO’s Rule 56, which currently specifies that information is only material if it (1) is non-cumulative over the information already disclosed, and (2) either establishes a prima facie case of unpatentability or is inconsistent with a position patentee adopted during prosecution. In fact, the *Therasense* “but-for material” requirement for inequitable conduct is now virtually coextensive with the materiality standard needed to invalidate a claim. The only difference in the two tests is the standard of proof: but-for materiality must be proven by a preponderance of the evidence for a finding of inequitable conduct, but by clear and convincing evidence for a finding of invalidity. *Id.* Thus, a finding of invalidity in district court based on a withheld reference implies that the reference is necessarily but-for material under *Therasense*’s materiality standard. However, a withheld reference may still be but-for material in terms of inequitable conduct doctrine even if it is not sufficient to invalidate a claim in district court.

materiality is purely objective: Under this first prong of the two-part test, it matters not whether the patentee had knowledge of the information, let alone its materiality.¹³² The question to be answered is simply whether the information would have precluded patentability had the USPTO been aware of it.¹³³

Under the second prong of the *Therasense* test, the defendant must show, by clear and convincing evidence, that the specific intent to deceive or mislead the USPTO is the “single most reasonable inference” to be drawn.¹³⁴ In fact, when there are “multiple reasonable inferences . . . intent to deceive cannot be found.”¹³⁵ Writing for the majority in *Therasense*, Chief Judge Rader explained that, to satisfy the intent prong of an inequitable conduct defense, an accused infringer must show three things, each by clear and convincing evidence: (1) the patentee knew of the information, (2) the patentee knew that the information was material, and (3) the patentee made a deliberate decision to withhold the information from the USPTO.¹³⁶

Note that the first and second requirements under the intent prong of the inequitable conduct test create a separate and distinct materiality component.¹³⁷ As discussed above, the information withheld from the USPTO

132. See *Therasense*, 649 F.3d at 1291 (“[I]n assessing the materiality of a withheld reference, the court must determine whether the PTO would have allowed the claim if it had been aware of the undisclosed reference.”).

133. See *id.*

134. *Id.* at 1290; see *1st Media, LLC v. Elec. Arts, Inc.*, 694 F.3d 1367, 1374–75 (Fed. Cir. 2012) (“Knowledge of the reference and knowledge of materiality alone are insufficient after *Therasense* to show an intent to deceive [I]t is not enough to argue carelessness, lack of attention, poor docketing or cross-referencing, or anything else that might be considered negligent or even grossly negligent.”); *Western Plastics, Inc. v. DuBose Strapping, Inc.*, No. 2021-1371, 2022 WL 576218, at *1 (Fed. Cir. Feb. 25, 2022) (“[W]e agree with the district court that [the defendant] did not set forth evidence to meet the high [post-*Therasense*] standard of establishing that the patent applicant intended to deceive the Patent Office, as required to sustain an inequitable conduct defense.”).

135. *Therasense*, 649 F.3d at 1290–91.

136. *Id.* at 1290. Note that, though *Therasense* dealt with a patentee that withheld references from the PTO, the Federal Circuit has clarified that the *Therasense* standard also applies to factual misrepresentations (including representations about references that were actually submitted to the USPTO). See, e.g., *Ohio Willow Wood Co. v. Alps S., LLC*, 813 F.3d 1350, 1357 (Fed. Cir. 2016) (“A party seeking to prove inequitable conduct must show by clear and convincing evidence that the patent applicant made *misrepresentations or omissions* material to patentability, that he did so with the specific intent to mislead or deceive the PTO, and that deceptive intent was the single most reasonable inference to be drawn from the evidence.”) (emphasis added).

137. See, e.g., *Baxter Int’l, Inc. v. CareFusion Corp.*, No. 15 C 9986, 2022 WL 981115, at *8 (N.D. Ill. Mar. 31, 2022) (“The next requirement is that the inventors must have known

must first be *objectively* but-for material to patentability.¹³⁸ But a showing of intent requires the accused infringer to prove that the patentee had a *subjective* knowledge both of the information's existence and of its materiality.¹³⁹ Note also that the third requirement of the intent prong—a showing that the patentee made a “deliberate decision to withhold” the information from the USPTO—is one of purpose, not knowledge.¹⁴⁰ In his closing remarks in *Therasense*, Chief Judge Rader ordered that, on remand, the lower court should determine whether the patentee “made the conscious decision not to disclose [the relevant information] *in order to deceive* the PTO.”¹⁴¹ In other words, a showing that the patentee understood that their conduct would deceive the USPTO is not enough—the defendant must prove that the patentee had the express purpose of deception.

Thus, for post-*Therasense* defendants, the bar to raising an inequitable conduct defense is exceedingly high. Only a showing (under demanding evidentiary standards) that the patentee intentionally withheld or misrepresented information that would have precluded issuance of a patented claim will suffice. Indeed, it bears repeating: Not only must the accused infringer show that the information withheld from or misrepresented to the USPTO was objectively but-for material to patentability, but they must also show that the patentee had a subjective appreciation of the information's materiality and acted with the purpose of deceiving the USPTO when it withheld or misrepresented the material information. Importantly, the *Therasense* court also explicitly disavowed the sliding scale approach it had favored in the past, “where a weak showing of intent [could] be found sufficient based on a strong showing of materiality, and vice versa.”¹⁴² Instead, after *Therasense*, the “court must weigh the evidence of intent to deceive independent of its analysis of materiality.”¹⁴³

Some critics have suggested that the *Therasense* court was overzealous in its efforts to address the “plague” of inequitable conduct defenses that were

the [withheld information was] material. This is a requirement of knowledge, not of the separate inquiry of but-for materiality required for an inequitable conduct showing.”)

138. See *Therasense*, 649 F.3d at 1291 (“[I]n assessing the materiality of a withheld reference, the court must determine whether the PTO would have allowed the claim if it had been aware of the undisclosed reference.”).

139. See *id.* at 1290.

140. See *id.*

141. *Id.* at 1296 (emphasis added).

142. See *id.* at 1290.

143. *Id.*

common in patent litigation.¹⁴⁴ This view was seemingly shared by the dissent in *Therasense*, which, believing the majority's approach to be too restrictive, advocated for a more modest materiality test.¹⁴⁵ However, the data show that courts reject post-*Therasense* inequitable conduct defenses for lack of intent (87% of failed defenses) much more frequently than for lack of materiality (57% of failed defenses).¹⁴⁶ This trend may be attributable to the fact that "direct evidence of specific intent to deceive is difficult to find, so it is relatively simple for a judge to conclude that an accused infringer failed to prove intent."¹⁴⁷ As a result, even if the Federal Circuit were to now soften its materiality requirement, it is not clear that such a change, in the absence of a sliding scale, could revive the effectiveness of the inequitable conduct defense.¹⁴⁸

B. THREE CASE STUDIES IN INCONSISTENT REPRESENTATION: *BRUNO*, *BELCHER*, AND *BAXTER*

To understand the impact of *Therasense*, and the changes to the legal standards of the intent prong in particular, this Section considers three case studies: (1) *Bruno Independent Living Aids, Inc. v. Acorn Mobility Services, Ltd.*;¹⁴⁹ (2) *Belcher Pharmaceuticals, LLC v. Hospira, Inc.*;¹⁵⁰ and (3) *Baxter International, Inc. v. CareFusion Corp.*,¹⁵¹ each of which is discussed in turn below. All three cases (the first two from the Federal Circuit and the third from the Northern District of Illinois) turned on issues of inconsistent representation at the USPTO and FDA. Notably, *Bruno* and *Baxter* did not involve pharmaceutical patents. But the analysis and holdings in each case are nonetheless helpful for understanding the application of pre- and post-*Therasense* inequitable conduct

144. *Burlington Indus., Inc. v. Dayco Corp.*, 849 F.2d 1418, 1422 (Fed. Cir. 1988); see Swanson, *supra* note 122, at 720–24 (outlining criticism of the Federal Circuit's rationale for *Therasense*).

145. See *Therasense*, 649 F.3d at 1304 (Bryson, J., dissenting) (arguing that materiality should be measured by the PTO's Rule 56 standard).

146. Swanson, *supra* note 122, at 708. These findings are consistent with pre-*Therasense* data generated by Petherbridge and co-workers in 2011. See Lee Petherbridge, Jason Rantanen & Ali Mojibi, *The Federal Circuit and Inequitable Conduct: An Empirical Assessment*, 84 S. CAL. L. REV. 1293, 1319–21 (2011) ("[W]hen the Federal Circuit gives a single reason for patentee success, the reason is nearly two and a half times more likely to be lack of intent to deceive than it is to be lack of materiality.").

147. Swanson, *supra* note 122, at 709.

148. See *Therasense*, 649 F.3d at 1290 ("[A] court must weigh the evidence of intent to deceive independent of its analysis of materiality.").

149. 394 F.3d 1348 (Fed. Cir. 2005).

150. 11 F.4th 1345 (Fed. Cir. 2021).

151. No. 15 C 9986, 2022 WL 981115 (N.D. Ill. Mar. 31, 2022).

doctrine by the courts to claims of inconsistent representation at the USPTO and FDA.

1. Bruno Independent Living Aids, Inc. v. Acorn Mobility Services, Ltd.

In *Bruno*, a pre-*Therasense* case, the Federal Circuit held that a stairlift manufacturer had engaged in inequitable conduct¹⁵² by withholding from the USPTO material prior art they had previously disclosed to the FDA as part of a § 510(k) submission.¹⁵³ The analysis turned on disclosure of a competitor product, the “Wecolator.”¹⁵⁴ In seeking FDA approval to sell its stairlift, Bruno Independent Living Aids, Inc. (“Bruno”) made a claim of “substantial equivalence” between its product and the Wecolator.¹⁵⁵ However, the same information about the competitor product was never shared with the USPTO.¹⁵⁶

Adopting a pre-*Therasense* Rule 56 materiality standard,¹⁵⁷ the Federal Circuit held that, “[h]ad the Examiner known about the Wecolator . . . Bruno could not have touted the front offset swivel as a point of novelty.”¹⁵⁸ The Wecolator disclosure was thus material to patentability under Rule 56.¹⁵⁹ In fact, the Wecolator’s materiality was also crucial to the Federal Circuit’s analysis of intent. The court acknowledged that the district court had “provided little explicit support for its finding of intent.”¹⁶⁰ However, the court relied on the pre-*Therasense* materiality-intent sliding scale, finding that “the high materiality of the Wecolator” meant there was “sufficient evidence based upon which a fair inference of deceptive intent may be drawn.”¹⁶¹ Such a reliance would, as noted above, be impossible post-*Therasense*.¹⁶²

152. *Bruno*, 394 F.3d at 1355.

153. 510(k) NO. K921648, 510(k) PREMARKET NOTIFICATION (May 4, 1992), <https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfpmn/pmn.cfm?ID=K921648>. Note that § 510(k) submissions are used by medical device manufacturers to notify the FDA of their intent to market a medical device in the United States. See *510(k) Clearances*, FDA (Nov. 6, 2023), <https://www.fda.gov/medical-devices/device-approvals-denials-and-clearances/510k-clearances>.

154. *Bruno*, 394 F.3d at 1351–52.

155. *Id.* at 1352.

156. *Id.*

157. 37 C.F.R. § 1.56(b)(2)(ii) (2012).

158. *Bruno*, 395 F.3d at 1353.

159. *Id.* at 1354.

160. *Id.*

161. *Id.*

162. See *Therasense, Inc. v. Becton, Dickinson & Co.*, 649 F.3d 1276, 1290 (Fed. Cir. 2011) (en banc) (“[A] court must weigh the evidence of intent to deceive independent of its analysis of materiality.”).

The Federal Circuit likewise held that, because Bruno had “not proffered a credible explanation for the nondisclosure,” it was fair to make an inference of deceptive intent.¹⁶³ After *Therasense*, that type of inference cannot be made: Now, a “patentee need not offer any good faith explanation unless the accused infringer first . . . prove[s] a threshold level of intent to deceive by clear and convincing evidence.”¹⁶⁴ In other words, a patentee’s silence in the face of an inequitable conduct accusation leads to very different consequences pre- and post-*Therasense*. Before *Therasense*, the court was free to infer malintent from a lack of good-faith explanation; after *Therasense*, a patentee can hide behind their silence so long as the accused infringer fails to provide clear and convincing evidence of the intent to deceive the USPTO.

2. Belcher Pharmaceuticals, LLC v. Hospira, Inc.

In the *Belcher II* case, a post-*Therasense* Federal Circuit held a pharmaceutical patent unenforceable for inequitable conduct.¹⁶⁵ Recall that Belcher submitted a § 505(b)(2) NDA with the FDA for an injectable epinephrine formulation.¹⁶⁶ Belcher, in supporting its case for § 505(b)(2) approval, had disclosed to the FDA information about similar third-party products that it later withheld from the USPTO.¹⁶⁷ Likewise, Belcher, when corresponding with the FDA, referred to the pH range of one such competitor product as “old,” later asserting that that same pH range was “unexpectedly found to be critical” to its own invention when contesting an obviousness rejection at the USPTO.¹⁶⁸ The district court found—and the Federal Circuit agreed—that Belcher “did not merely withhold . . . information but also used emphatic language” to make inconsistent statements to the USPTO and FDA.¹⁶⁹

Applying *Therasense*, the Federal Circuit found that Belcher had withheld multiple pieces of information, including knowledge of third-party products, that were but-for material to patentability.¹⁷⁰ The Federal Circuit rejected Belcher’s argument that it only withheld information that it believed to be cumulative over the art already on record.¹⁷¹ In the court’s view, Belcher’s argument was unpersuasive because it was “directly at odds” with Belcher’s

163. *Bruno*, 395 F.3d at 1354.

164. *Therasense*, 649 F.3d at 1291 (citing *Star Sci., Inc. v. R.J. Reynolds Tobacco Co.*, 537 F.3d 1357, 1368 (Fed. Cir. 2008)).

165. *Belcher Pharms., LLC v. Hospira, Inc. (Belcher II)*, 11 F.4th 1345, 1354 (Fed. Cir. 2021).

166. *Id.*

167. *Id.*

168. *Id.* at 1350–51.

169. *Id.* at 1352.

170. *Id.* at 1353.

171. *Id.*

assertion during patent prosecution that the claimed pH range was critical.¹⁷² Likewise, the Federal Circuit affirmed the district court's finding of intent.¹⁷³ Recognizing that it is often difficult to find direct proof of intent, the court pointed to evidence in the record (for example, Belcher's knowledge of third-party products and its assertions relating to the criticality of the pH range) that supported a finding that "the single most reasonable inference is that [Belcher] possessed the specific intent to deceive the PTO."¹⁷⁴

3. Baxter International, Inc. v. CareFusion Corp.

Most recently, in *Baxter*, a district court in Illinois found that an infusion-pump manufacturer did not engage in inequitable conduct when it failed to disclose to the USPTO information about competitor products that it had described as "substantially equivalent" to its own device as part of a § 510(k) submission to the FDA.¹⁷⁵

The court analyzed each of the three intent components (subjective knowledge of the existence of the information, subjective knowledge of the materiality of the information, and specific intent to deceive the USPTO) for each of the three Baxter International ("Baxter") inventors in turn.¹⁷⁶ Judge Kendall acknowledged that there were genuine issues of fact as to whether certain inventors knew of the existence of the withheld information, its materiality, or both.¹⁷⁷ However, the court ultimately concluded that such factual disputes were not dispositive because CareFusion Corporation ("CareFusion") could not "set forth evidence that any Baxter Inventor made a deliberate decision to deceive the USPTO."¹⁷⁸

In reality, the court found, "even if a factfinder were to disbelieve the Baxter Inventors, the 'single most reasonable inference' would still not be" one of deliberate deception.¹⁷⁹ Because a reasonable factfinder could equally conclude that the nondisclosure was due to, for example, gross negligence or incompetence, it would not be possible for that same factfinder to conclude that Baxter had engaged in inequitable conduct under the standards set forth in *Therasense*.¹⁸⁰ Accordingly, Judge Kendall granted Baxter's Motion for Partial

172. *Id.*

173. *Id.* at 1354.

174. *Id.*

175. *Baxter Int'l, Inc. v. CareFusion Corp.*, No. 15 C 9986, 2022 WL 981115, at *6 (N.D. Ill. Mar. 31, 2022).

176. *Id.* at *6–7.

177. *Id.* at *7.

178. *Id.*

179. *Id.*

180. *Id.*

Summary Judgment of No Inequitable Conduct, noting that, even if it were true that summary judgment motions for no inequitable conduct were rarely granted pre-*Therasense*, that “is no longer the case.”¹⁸¹

Judge Kendall specifically disparaged CareFusion’s attempt to analogize the factual and legal issues in *Baxter* to the Federal Circuit’s analysis in *Bruno*.¹⁸² Kendall distinguished *Baxter* from *Bruno* in two key ways. First, she noted that, post-*Therasense*, use of a materiality-intent sliding scale was “improper.”¹⁸³ Second, she explained that *Therasense* voided the possibility of inferring deceptive intent from the absence of a good-faith explanation from the patentee for their nondisclosure.¹⁸⁴

Taken together, *Bruno*, *Belcher II*, and *Baxter* provide a number of important insights into the development of inequitable conduct doctrine over time, especially with respect to the courts’ understanding of the intent requirement. For one thing, it is clear that, based on the evidence presented at trial, the pre-*Therasense Bruno* court would almost certainly have been unable to find that the patentee had the intent to deceive if the court had instead been operating under a *Therasense* standard that did not permit the use of a materiality-intent sliding scale.¹⁸⁵ But even in light of *Belcher II*, it is not so clear exactly how a court *can* find intent in cases of inconsistent representation at the USPTO and FDA in a post-*Therasense* world.¹⁸⁶ In fact, *Belcher II* may be most notable because it is

181. *Id.* at *8.

182. *See supra* Section IV.B.1.

183. *Baxter*, No. 15 C 9986, 2022 WL 981115, at *7.

184. *Id.* at *8.

185. *See Therasense, Inc. v. Becton, Dickinson & Co.*, 649 F.3d 1276, 1290 (Fed. Cir. 2011) (en banc) (“[A] court must weigh the evidence of intent to deceive independent of its analysis of materiality.”).

186. For examples of other cases in which federal courts have failed to find inequitable conduct post-*Therasense*, see *Exergen Corp. v. Kaz USA, Inc.*, 120 F. Supp. 3d 1, 7 (D. Mass. 2015) (“Because Kaz has not adduced competent evidence to establish the intent element of its inequitable conduct claim, the claim is not viable as a matter of law and must be dismissed.”); *Galderma Lab’sys, L.P. v. Tolmar, Inc.*, 891 F. Supp. 2d 588, 649–50 (D. Del. 2012) (“The alleged failure to disclose the Phase III clinical trial data [to the USPTO] was not but-for material . . . [T]he Court concludes that the [patentee’s conduct] does not rise to the level of an affirmative egregious act of misconduct . . . [I]nequitable conduct fails for the additional reason that the evidence does not persuade the Court that the inventors acted with an intent to deceive the PTO.”); *Sun Pharma Glob. Fze v. Lupid Ltd.*, No. CV 18-2213 (FLW), 2021 WL 4473411, at *34 (D.N.J. Sept. 30, 2021) (finding that, because PTO and FDA disclosures “serve very different purposes,” intent to deceive could not be inferred from a decision to withhold a reference from the PTO when the reasons for doing so were plausible); *ProStrakan, Inc. v. Actavis Lab’sys UT, Inc.*, No. 2:16-CV-00044-RWS, 2018 WL 11363829, at *72 (E.D. Tex. Sept. 28, 2018) (“[E]ven assuming that the data in the [patent] was material to patentability, Actavis has no evidence—either express or inferred—that anyone associated

rare. In preparing this Note, not a single other instance of a post-*Therasense* court finding inequitable conduct based on inconsistent representation at the USPTO and FDA was found at the trial or appellate level.¹⁸⁷ Thus, this Note argues, post-*Therasense* inequitable conduct doctrine—at least in its current form—is wholly inadequate for tackling the inconsistent-representation problem in the pharmaceutical industry. The Federal Circuit must revisit its inequitable conduct doctrine to uphold the integrity of the patent system and promote public access to innovation.

C. A NEW “PHARMA EXCEPTION” TO *THERASENSE*

The Federal Circuit should revise its inequitable conduct doctrine to create a “pharma exception” to the otherwise exceedingly high legal standards outlined in *Therasense*. Specifically, the court should hold that, when an accused infringer shows that a patentee (1) failed to disclose to the USPTO references it shared with the FDA to support its case for regulatory approval, or (2) made inconsistent or contradictory statements to the USPTO and the FDA, there should be a rebuttable presumption that *both* the materiality *and* the intent prongs of the *Therasense* inequitable conduct test are satisfied.

Adopting a “pharma exception” to *Therasense* will help address the problem of inconsistent representation in at least three ways. First, creating a presumption of inequitable conduct in cases of inconsistent representation will encourage accused infringers (including generics manufacturers) to raise inequitable conduct defenses when their products are unfairly blocked by invalid patents.¹⁸⁸ Second, once such defenses are raised, the accused infringer will have a greater chance of success.¹⁸⁹ Third, the combined effect of an accused infringer being both more likely to raise and to win on a claim of

with the prosecution of the [patent]—including the named inventors—intended to deceive the PTO.”).

187. The term (+ “inequitable conduct” + “patent” + “FDA”) was searched in Westlaw on October 5, 2022. The search was limited to “All Federal” cases that were decided on or after May 26, 2011 (*Therasense* was decided on May 25, 2011). The search returned 109 hits, each of which was reviewed individually for (1) factual issues of inconsistent representation, (2) a finding of inequitable conduct, (3) a finding of materiality, (4) reliance on the “egregious misconduct” exception to but-for materiality, and (5) a finding of intent. Of the 109 hits, seven cases were found to turn on issues of inconsistent representation and were ultimately decided on the merits (which, for the sake of this review, included summary judgment for no inequitable conduct but did not include, for example, Rule 12(b)(6) motions).

188. *See* Frei & Wooden, *supra* note 110, at 66 (“After the holding in *Therasense*, it was widely believed that the court had sounded the death knell to the inequitable conduct defense by imposing evidentiary requirements that could rarely be met.”).

189. *See id.*

inequitable conduct will deter pharmaceutical patentees from making inconsistent representations in the first place.¹⁹⁰

The Federal Circuit has already shown its willingness to create exceptions to *Therasense*. Despite the court's determination to heighten the legal standards for inequitable conduct, *Therasense* preserved an "egregious misconduct" exception which, when triggered, infers per se materiality.¹⁹¹ In the court's view, the exception "strikes a necessary balance between encouraging honesty before the PTO and preventing unfounded accusations of inequitable conduct."¹⁹² However, the egregious misconduct exception is wholly inadequate for tackling *Belcher*-style inconsistent representation in the pharmaceutical industry for at least two reasons.

First, the *Therasense* court was clear that the "mere nondisclosure of prior art references" to the USPTO is not egregious misconduct.¹⁹³ Second, because the *Therasense* court specifically disavowed the "sliding scale" approach that had existed in the past, per se materiality in light of egregious misconduct infers nothing about intent: the "single most reasonable inference" standard remains unaltered.¹⁹⁴ Thus, even if a court were to find that making inconsistent representations to the USPTO and FDA denoted misconduct that was sufficiently egregious to infer materiality, the court could still conclude that the patentee lacked the intent to render such misconduct inequitable.¹⁹⁵ It follows,

190. See Louis Kaplow, *On the Optimal Burden of Proof*, 119 J. POL. ECON. 1104, 1104 (2011) ("The optimal strength of the burden of proof . . . involves trading off deterrence and the chilling of desirable behavior.").

191. *Therasense, Inc. v. Becton, Dickinson & Co.*, 649 F.3d 1276, 1292 (Fed. Cir. 2011) (en banc) ("Although but-for materiality generally must be proved to satisfy the materiality prong of inequitable conduct, this court recognizes an exception in cases of affirmative egregious misconduct When the patentee has engaged in affirmative acts of egregious misconduct . . . the misconduct is material."); see also *Regeneron Pharms., Inc. v. Merus B.V.*, 144 F. Supp. 3d 530, 585 (S.D.N.Y. 2015) ("The Court finds by clear and convincing evidence . . . that Regeneron made false and misleading statements. The Court finds by clear and convincing evidence that this constitutes egregious affirmative misconduct."); *Apotex, Inc. v. UCB, Inc.*, 970 F. Supp. 2d 1297, 1328 (S.D. Fla. 2013) ("I find that this case is one of those exceptional cases where, as discussed above, a finding of materiality is not necessary. Specifically, I find that [patentee] engaged in affirmative and egregious misconduct.").

192. *Therasense*, 649 F.3d at 1293.

193. See *id.* at 1292.

194. See *id.* at 1290 ("[A] court must weigh the evidence of intent to deceive independent of its analysis of materiality.").

195. See, e.g., *Outside the Box Innovations, LLC v. Travel Caddy, Inc.*, 695 F.3d 1285, 1294 (Fed. Cir. 2012) ("Although on its face, it appears that a false declaration of small entity status would fall within the definition of an 'unmistakably false affidavit,' . . . we need not decide that question. Even if a false assertion of small entity status were per se material, the requirements of *Therasense* are not met here because there was no clear and convincing evidence of intent to deceive the PTO.").

then, that, if a revamped inequitable conduct doctrine is to be a truly useful tool for tackling inconsistent representation at the USPTO and FDA, any worthwhile proposal must alter both components of *Therasense*'s two-prong test.¹⁹⁶

It is reasonable to presume both materiality and intent in cases of inconsistent representation at the USPTO and FDA. As outlined in Part II, patentability and regulatory approval often turn on similar issues—including novelty and nonobviousness—but to opposite ends. A patent applicant needs to convince the USPTO that their drug is both novel and nonobvious over the prior art.¹⁹⁷ But that same patent applicant may want to point to similarities between their product and existing alternatives when they seek FDA approval, e.g., through the § 505(b)(2) pathway.¹⁹⁸ Thus, it is fair to presume that information is material to patentability when it has been (1) disclosed to the FDA but not to the USPTO, or (2) characterized inconsistently (or, in some cases, contradictorily) before each entity.¹⁹⁹

Likewise, it is reasonable to presume intent. In recognizing the need for a finding of per se materiality in cases of egregious misconduct, the Federal Circuit explained that “a patentee is unlikely to go to great lengths to deceive the PTO with a falsehood unless it believes that the falsehood will affect issuance of the patent.”²⁰⁰ In other words, it is reasonable to presume that patentees engage in risky and deceptive tactics only with respect to information that is material to patentability. But the inverse is also true. A patentee is unlikely to withhold or misrepresent material information at the USPTO unless they wish to “deceive the PTO with a falsehood [that] will affect issuance of the patent.”²⁰¹ In that sense, it is fair to presume both materiality and intent when there is evidence of inconsistent representation.

Creating a new “pharma exception” in cases of inconsistent representation would also help “strike a necessary balance” between encouraging honesty before the USPTO and triggering a new “plague” of inequitable conduct defenses.²⁰² First, the exception would help nurture a culture of honesty among

196. Recall that 87% of unsuccessful post-*Therasense* inequitable defense claims fail the intent prong. Swanson, *supra* note 122, at 708.

197. See *supra* Section II.A.

198. See *id.*

199. See July 2022 Notice, *supra* note 56, at 45,766 (discussing these two scenarios in the context of the duty to disclose).

200. *Therasense, Inc. v. Becton, Dickinson & Co.*, 649 F.3d 1276, 1292 (Fed. Cir. 2011) (en banc).

201. See *id.*

202. See *id.* at 1293.

pharmaceutical patentees—somewhere it is currently known to be lacking.²⁰³ There would be little point in making inconsistent representations at the USPTO and FDA if an accused infringer could then rely on such representations as the basis of an inequitable conduct defense that would now be more likely to succeed.

Second, creating a “pharma exception” is unlikely to produce an overwhelming increase in the number of inequitable conduct defenses being raised. For one thing, the exception would be limited to instances of inconsistent representation by pharmaceutical patentees at the FDA and USPTO.²⁰⁴ That is, of course, not to say that variations of this exception could not apply elsewhere. There may well be other instances of interagency inconsistent representation for which it would be reasonable to presume inequitable conduct—for example, when a patentee withholds material information about the novelty of a medical device from the USPTO while disclosing that same information to the FDA for regulatory approval.²⁰⁵ However, this Note focuses on tackling inconsistent representation by pharmaceutical patentees at the USPTO and FDA because (1) this particular problem is causing such notable concern across the full spectrum of parties involved in pharmaceutical patenting and regulation, including the agencies themselves; and (2) the courts may be more receptive to a narrowly tailored solution that addresses a highly specific grievance.²⁰⁶

Likewise, *Exergen’s* heightened pleading standards could continue to provide a gatekeeping mechanism that discourages frivolous claims of inequitable conduct.²⁰⁷ An accused infringer invoking the “pharma exception”

203. *See supra* Part II.

204. For a discussion of the benefits and drawbacks of sector-specific (as opposed to uniform) intellectual property regimes, *compare* Dan L. Burk & Mark A. Lemley, *Is Patent Law Technology-Specific?*, 17 BERKELEY TECH. L.J. 1155, 1159–60 (2002) (discussing the downsides of uniform intellectual property regimes), *and* Michael W. Carroll, *One for All: The Problem of Uniformity Cost in Intellectual Property Law*, 55 AM. U.L. REV. 845, 849–50 (2006) (explaining that uniform intellectual property rights necessarily result in deadweight loss), *with* ADAM B. JAFFE & JOSH LERNER, INNOVATION AND ITS DISCONTENTS 203–05 (2004) (arguing for the need for uniform treatment of intellectual property), *and* R. Polk Wagner, *(Mostly) Against Exceptionalism* 8–17 (U. Penn. L. Sch. Inst. for L. & Econ., Rsch. Paper No. 02-18, 2002), <https://ssrn.com/abstract=321981> (advocating against “Type II Exceptionalism,” which “shifts consideration of the patent law from a general background principle of property rights to a vehicle for a particularistic, technology-specific innovation policy choices”).

205. For more information on the approval of medical devices by the FDA, see 510(k) CLEARANCES, FDA (Nov. 6, 2023), <https://www.fda.gov/medical-devices/device-approvals-denials-and-clearances/510k-clearances>.

206. *See supra* Section II.B.

207. *See Exergen Corp. v. Wal-Mart Stores, Inc.*, 575 F.3d 1312, 1328–29 (Fed. Cir. 2009) (explaining that inequitable conduct must be pleaded with particularity under Federal Rule of

would now be required to “identify the specific who, what, when, where, and how” of the inconsistent representation at the USPTO and FDA.²⁰⁸ As a result, only those accused infringers who are able to point to the particularities of a meaningful instance of interagency misrepresentation will be able to successfully plead a “pharma exception” to *Therasense*.

Third, rebuttable presumptions of intent and materiality are, as the name would suggest, rebuttable. If a patentee is able to show, by a preponderance of the evidence, that the information that they withheld or characterized inconsistently before the USPTO and FDA is not objectively but-for material to patentability, there will be no finding of inequitable conduct.²⁰⁹ Likewise, if a patentee can show, by a preponderance of the evidence, that the intent to deceive the PTO is not the “single most reasonable inference,” their patent will not be unenforceable.²¹⁰ Consequently, a patentee acting in good faith has nothing to fear. Likewise, an accused infringer has no incentive to raise an inequitable conduct defense over matters of inconsistent representation if their claim is merely frivolous.

Note that, when the “pharma exception” is invoked, the evidentiary standard under the materiality prong stays the same, whereas the evidentiary standard under the intent prong changes. Recall that, under *Therasense*, the defendant bears the burden of proving (1) materiality by a preponderance of the evidence, and (2) intent to deceive by clear and convincing evidence.²¹¹ This Note suggests that, when, under the “pharma exception,” the burden of proof flips from the infringer-defendant to the patentee-plaintiff, the evidentiary standard under intent should shift, too—specifically, to require a showing of intent by a preponderance of the evidence.²¹² In that way, the

Civil Procedure 9(b)). But Swanson suggests that courts should exercise caution when applying post-*Exergen* pleading standards as a means of safeguarding meritorious claims. Swanson, *supra* note 122, at 723 (“[C]ourts must permissively grant leave to amend pleadings to add inequitable conduct defenses.”).

208. *See Exergen*, 575 F.3d at 1328.

209. *See Therasense, Inc. v. Becton, Dickinson & Co.*, 649 F.3d 1276, 1291 (Fed. Cir. 2011) (en banc) (“[I]n assessing the materiality of a withheld reference, the court must determine whether the PTO would have allowed the claim if it had been aware of the undisclosed reference.”).

210. *See id.* at 1290.

211. *See id.* at 1290–92.

212. This Note declines to recommend a change in the evidentiary standard for the materiality prong because, as noted by the *Therasense* court, the evidentiary standard under materiality mirrors that used by the USPTO. *See id.* at 1291–92 (“[I]n assessing the materiality of a withheld reference, the court must determine whether the PTO would have allowed the claim if it had been aware of the undisclosed reference. In making this patentability determination, the court should apply the preponderance of the evidence standard and give claims their broadest reasonable construction.”).

“pharma exception” maintains some of *Therasense*’s pro-patentee skew: Under the “pharma exception,” it is easier for the patentee-plaintiff to rebut a presumption of intent than it is for a traditional infringer-defendant to raise a viable inequitable conduct claim in the first place.

What would a successful rebuttal look like in practice? A patentee could successfully rebut a presumption of materiality by showing, by a preponderance of the evidence, that the inconsistent representation relates to information that is not objectively but-for material to patentability.²¹³ The patentee could demonstrate, for instance, that the information in question neither anticipates the claimed invention nor renders it obvious.²¹⁴ Likewise, a patentee could successfully rebut a presumption of intent by demonstrating, by a preponderance of the evidence, that the intent to deceive is not the “single most reasonable inference” to be drawn.²¹⁵ To do this, the patentee would simply show that at least one other inference is as equally reasonable as the intent to deceive.²¹⁶ For example, the patentee could point to internal communications between the inventors that indicate a failure to subjectively appreciate the materiality of the reference to patentability.²¹⁷ In that case, a court may well find it at least equally reasonable to attribute the patentee’s actions to incompetence or ignorance as opposed to purposeful deception, meaning the presumption of intent can be rebutted.²¹⁸

Grounding the new “pharma exception” in rebuttable presumptions of materiality and intent (rather than, say, strict liability) and maintaining some of *Therasense*’s pro-patentee skew is likely to make the “pharma exception” more doctrinally palatable for the courts, too. As noted above, under *Therasense*, the court is already willing to infer strict liability with respect to materiality in cases of egregious misconduct—but, even in those instances, the burden of proving the intent to deceive as the “single most reasonable inference” remains with

213. *See id.* at 1291 (“[I]n assessing the materiality of a withheld reference, the court must determine whether the PTO would have allowed the claim if it had been aware of the undisclosed reference.”).

214. *See supra* Section IV.A.

215. *See Therasense*, 649 F.3d at 1290.

216. *See id.* at 1290–91 (noting that, when there are “multiple reasonable inferences . . . intent to deceive cannot be found.”); *see also* Sun Pharma Glob. Fze v. Lupid Ltd., No. CV 18-2213 (FLW), 2021 WL 4473411, at *34 (D.N.J. Sept. 30, 2021) (“FDA and PTO disclosures serve very different purposes, and Defendants have not presented evidence to suggest that they must or should overlap in this case.”).

217. *See Therasense*, 649 F.3d at 1290 (“[T]he accused infringer must prove by clear and convincing evidence that the applicant knew of the reference, *knew that it was material*, and made a deliberate decision to withhold it.”) (emphasis added).

218. *See id.* at 1290–91 (noting that when there are “multiple reasonable inferences . . . intent to deceive cannot be found”).

the accused infringer.²¹⁹ The Federal Circuit thus balanced strict liability in the materiality prong against a much more patentee-friendly intent standard. Meanwhile, under the new “pharma exception,” materiality and intent are both presumed. In other words, inconsistent representation does not go so far as to trigger strict liability in either prong of the two-part test—but it does create presumptions of both materiality and intent that are relatively easy for a good-faith patentee to rebut. In that way, the new “pharma exception” maintains the balance advocated by the *Therasense* court: It encourages honesty before the USPTO without risking a deluge of inadequately robust inequitable conduct claims.

V. CONCLUSION

This Note opened with a question: How can we address the problem of pharmaceutical companies making inconsistent representations to the USPTO and FDA? This Note, in response, offered two solutions. Part III outlined a new system of USPTO-FDA interaction that, to the extent possible, undercuts inconsistent representation before a patent issues. But because the system in Part III would likely suffer from issues of confidentiality, timing, and noncompliance, Part IV offered a post-patent-issuance safety net. Specifically, Part IV proposed that the Federal Circuit revise its inequitable conduct doctrine—by creating a new “pharma exception” to *Therasense*’s strict materiality-plus-intent test—to make it easier for accused infringers to raise claims of inequitable conduct and undermine the enforceability of pharmaceutical patents obtained through deception.

Importantly, the solutions outlined in Part III and Part IV need not be mutually exclusive. In fact, they could be synergistic. First, if a patentee failed to comply with their duty of disclosure in the new USPTO-FDA system of Part III, the patentee’s noncompliance could weigh in favor of triggering the dual presumptions under the new *Therasense* exception outlined in Part IV. Second, adopting the proposed system of Part III could help maintain the balance between encouraging honesty before the USPTO and preserving judicial resources for reviewing inequitable conduct claims. Because the new system would undercut at least some acts of inconsistent disclosure before patent issuance, fewer unenforceable patents will issue in the first place, thereby lessening the need for post-issuance judicial remedies.

Both solutions described herein can play a meaningful role in tackling the problem of inconsistent representation by pharmaceutical patentees. But whichever solutions the USPTO, the FDA, and the courts adopt, one thing is

219. *Id.*

certain: Meaningful change is needed to uphold the integrity of the patent system and promote public access to generic drug products that are unfairly blocked by invalid patents. Reform can come from the federal agencies, or the courts, or both. But it must indeed come—and soon.

TAKE A PICTURE: COPYRIGHT AND STATE SOVEREIGN IMMUNITY

Sarah Davidson[†]

TABLE OF CONTENTS

I.	INTRODUCTION	1170
II.	STATE SOVEREIGN IMMUNITY AND COPYRIGHT.....	1172
A.	THE STATE OF STATE SOVEREIGN IMMUNITY.....	1173
1.	<i>History and Purpose of State Sovereign Immunity.....</i>	1173
2.	<i>The Application of Sovereign Immunity to the State’s Instrumentalities.....</i>	1175
3.	<i>The Scope of and Exceptions to State Sovereign Immunity.....</i>	1175
a)	Congressional Abrogation	1176
B.	HISTORY AND PURPOSES OF COPYRIGHT	1177
C.	COPYRIGHT AND STATE SOVEREIGN IMMUNITY.....	1179
1.	<i>Congress’ Attempt to Abrogate and the Rise of the CRCA.....</i>	1180
2.	<i>Seminole Tribe, Florida Prepaid Cases, and the Fall of the CRCA</i>	1181
a)	<i>Allen v. Cooper: The Death Knell of the CRCA.....</i>	1182
3.	<i>Availability of Suit Against State Infringers After the CRCA.....</i>	1184
III.	TAKINGS CLAUSE AND COPYRIGHT	1185
A.	TAKINGS CLAUSE JURISPRUDENCE	1186
1.	<i>Physical Takings and Per Se Takings Claims.....</i>	1187
2.	<i>Regulatory Takings Claims.....</i>	1188
3.	<i>Expansion of “Physical” Takings: Cedar Point Nursery</i>	1189
B.	COPYRIGHT & TAKINGS CLAIMS	1190
1.	<i>Copyright is Likely Protectable Under the Takings Clause.....</i>	1191
2.	<i>But There is Uncertainty in the Scope of Protection.....</i>	1191
C.	THE VIABILITY OF THE COPYRIGHT TAKINGS CLAIMS IN <i>JIM OLIVE</i> <i>PHOTOGRAPHY</i> BEFORE AND AFTER <i>CEDAR POINT NURSERY</i>	1193
1.	<i>Before Cedar Point Nursery.....</i>	1194
2.	<i>After Cedar Point Nursery.....</i>	1196

DOI: <https://doi.org/10.15779/Z385717P8R>

© 2023 Sarah Davidson.

† J.D. 2024, University of California, Berkeley, School of Law. Sincere thanks to Professor Talha Syed, Shih-Wei Chao, my fellow students in the 2022 Law & Technology Writing Workshop, and the *Berkeley Technology Law Journal* editors.

IV. SOVEREIGN IMMUNITY AND COPYRIGHT TAKINGS.....	1198
V. ADEQUATE ALTERNATIVES AND PATHWAY TO CONGRESSIONAL ABROGATION	1200
A. ADEQUACY OF ALTERNATIVES.....	1200
1. Breach of Contract Claims.....	1200
2. Ex parte Young Claims.....	1201
3. Personal-Capacity Suits Against State Officials.....	1202
4. Waiver.....	1202
B. PATHWAY TO CONGRESSIONAL ABROGATION: RECOMMENDATIONS FROM THE COPYRIGHT OFFICE REPORT	1203
VI. CONCLUSION.....	1205

I. INTRODUCTION

Should individual states be allowed to use the intellectual property of their citizens without payment? Suppose that to boost its public treasury, the state of Texas downloaded copies of Billboard’s Top 100 and used these hit songs to create its own subscription-based music-sharing platform to rival Spotify. Do the songs become a “public good” if Texas’ use leads to more funding for local services? What if the use simply bolsters the marketability and reputation of local universities?

In *Jim Olive Photography v. University of Houston*, this was the question before the Texas Supreme Court.¹ Jim Olive is a Houston-based photographer, specializing in aerial photography.² Several of the photographs that Olive produces and licenses are of the downtown Houston skyline.³ These images require Olive to rent a helicopter at \$2,500 per hour, invest in specialized photography equipment, strap into a harness, and dangle below the helicopter while he is taking the photographs.⁴ After he is grounded, Olive painstakingly edits his images, including adding his watermark, and files the images with the Copyright Office. Protective over the work, Olive hired a copyright infringement scanner called Image Rights that alerted him to the unauthorized use of one of his skyline photographs, “The Cityscape,” by the University of

1. 624 S.W.3d 764, 768 (Tex. 2021), *cert. denied*, 142 S. Ct. 1361 (2022).

2. *Id.*

3. *Id.*

4. Ryan Hughes, *University Sued for Image of Houston Skyline*, BYU COPYRIGHT LICENSING OFF. (Feb. 1, 2018), <https://copyright.byu.edu/university-sued-for-image-of-houston-skyline>, (listing some of Jim Olive’s expenses in creating and protecting his works).

Houston (UH)'s C.T. Bauer College of Business.⁵ UH replaced Olive's watermark from the image with its own logo and stripped its metadata, making it untraceable back to Olive, and provided him neither compensation nor attribution.⁶ UH used the images repeatedly on its website and social media pages, and even provided the photo to *Forbes* magazine who published the photograph and credited the work to the Bauer College of Business.⁷ When Olive's attorney reached out to the university to negotiate a license fee for the use, UH responded, "[y]ou can't sue us; we have sovereign immunity."⁸ The University's dismissal of a federal copyright infringement suit inspired Olive to pursue Takings Clause claims at the state-level instead, testing the Court's likelihood of extending the Takings⁹ doctrine to cover instances of state infringement.¹⁰

Under the doctrine of sovereign immunity, states may not be held liable for torts like copyright infringement.¹¹ In an attempt to hold states liable for the unauthorized taking of a citizen's copyrighted material, Congress passed the Copyright Reform Clarification Act (CRCA) in 1990. Thirty years later, the Supreme Court declared the statute an unconstitutional abrogation of state sovereign immunity that encroached on states' rights under federalism in *Allen v. Cooper*.¹² After the fall of the CRCA in 2020, copyright owners are left with little recourse against even repeated and intentional state infringers.

5. Alex Meyer, *Photographer's Copyright Lawsuit Calls Bauer Ethics into Question*, DAILY COUGAR (Mar. 8, 2017), <http://thedailycougar.com/2017/03/08/photographers-copyright-lawsuit-calls-bauer-ethics-into-question/> (interviewing Jim Olive about his investment in monitoring techniques to prevent copyright infringement).

6. *Id.* (explaining that Olive discovered that UH infringed "The Cityscape" after hiring professional copyright monitoring services).

7. *Id.* (documenting Jim Olive's discovery of altered versions of his image appearing on UH websites and on *Forbes*).

8. *Id.* (quoting Dana Andrew LeJune, attorney for Jim Olive on the response received from UH).

9. This Note uses "Takings" (with a capital T) to refer specifically to the legal process outlined by the Fifth Amendment and similar state provisions, involving the government's formal acquisition of private property for public use with just compensation, and "takings" (lowercase) more broadly refers to any instance where a property right is deprived from an owner, regardless of the legal mechanism or justification.

10. *Jim Olive Photography*, 624 S.W.3d at 768.

11. See Michael Shaunessy & Ian M. Davis, *The Lion's Share: Federal Law, State Sovereign Immunity, and Intellectual Property*, in *ADVANCED INTELLECTUAL PROPERTY LITIGATION 2021* 1, 6 (State Bar of Texas Course Book, 2021), https://www.mcginislaw.com/media/publication/15313_15312_14_Shaunessy.pdf [hereinafter *Lion's Share*] (explaining that direct relief against state entities is barred by sovereign immunity following *Allen v. Cooper*).

12. See Tom James, *Digital Blackbeards: Copyright Infringement by States and the "Congruence and Proportionality" Test in Allen v. Cooper*, 36 *BERKELEY TECH. L.J.* 1375, 1377–91 (2021) (explaining the Court's decision to strike down the CRCA as an unconstitutional abrogation

One pathway for direct state liability and compensation for appropriation of the right to exclude under the Takings Clause may remain at the federal level.¹³ The Court's recent decision in *Cedar Point Nursery v. Hassid* extends the state's liability under the federal Takings Clause doctrine to encompass cases where the state has provided third parties access to a plaintiff's property; however, it is unclear if this case will apply to intellectual property, because intellectual property is nonrivalrous unlike real property and cannot be physically occupied.¹⁴ As such, this Note argues a more narrowly-tailored version of the CRCA is necessary to balance the interests of copyright owners with the important research, archival, and educational work of state universities and libraries in response to the growing record of copyright infringement by state actors.

This Note examines the interplay between copyright, state sovereign immunity, and the federal Takings Clause. In Part II, the Note surveys the history and purposes of state sovereign immunity and copyright in the United States. Part III questions the potential viability of Takings Clause claims as an alternative pathway to compensation for state appropriations of exclusive rights under the Copyright Act. Part IV addresses the threat that even if copyright Takings Clause claims may be viable, states may still claim sovereign immunity as a defense. Part V explores potentially viable alternative pathways to compensation that are less desirable than congressional abrogation. Finally, Part VI concludes by proposing a constitutional pathway for congressional abrogation in revising the Copyright Remedy Clarification Act.

II. STATE SOVEREIGN IMMUNITY AND COPYRIGHT

State sovereign immunity doctrines prevent copyright holders from pursuing copyright infringement cases against states.¹⁵ States are independent sovereigns within the United States and therefore generally cannot be sued in

both under Article I's Intellectual Property Clause powers and under § 5 of the Fourteenth Amendment which vests Congress with enforcement powers to ensure that no state "deprives any person of life, liberty, or *property without due process of law*" because the CRCA is too broad to meet the congruence and proportionality test established in *City of Boerne v. Flores* (emphasis added).

13. U.S. COPYRIGHT OFF., COPYRIGHT AND SOVEREIGN IMMUNITY 65–67 (2021) [hereinafter COPYRIGHT AND SOVEREIGN IMMUNITY] (explaining that copyright owners may still pursue federal and state Takings Clause claims and these claims are still considered viable, but that the chances of success are questionable).

14. 141 S. Ct. 2063, 2073–74 (2021); *see also* discussion *infra* Section III.E.

15. *See Lion's Share*, *supra* note 11 (surveying the origins and purpose of state sovereign immunity).

state or federal court without their consent.¹⁶ Congress possesses limited powers to pierce this immunity against the states' will.¹⁷

A. THE STATE OF STATE SOVEREIGN IMMUNITY

Under the doctrine of state sovereign immunity, a state cannot be sued in federal or state court without the state's consent.¹⁸ This doctrine emerges both from the text of the Eleventh Amendment and from the structure of the original Constitution itself, which the Supreme Court has interpreted to broaden the scope of the immunity to prevent suit from citizens within the state.¹⁹ States may choose to waive their immunity or Congress may abrogate the immunity under limited circumstances prescribed by the Court.²⁰ As a result, states enjoy immunity from liability under otherwise valid federal laws.²¹

1. *History and Purpose of State Sovereign Immunity*

Sovereign immunity descends from English common law and from the adage that “the King can do no wrong.”²² In the early United States, Anti-Federalists argued that the proposed constitution failed to adequately protect the doctrine. In particular, they feared that Article III of the proposed Constitution, which states that the power of the judiciary extends to controversies “between a state and citizens of another state” would allow private lawsuits against states.²³ On the other hand, Federalists claimed that Article III would not be interpreted to allow citizens to sue states without their consent.²⁴ However, other Federalists conceded that Article III allowed suit against states and reasoned instead that justice requires states to be held

16. KEVIN J. HICKEY, CONG. RES. SERV., COPYRIGHT AND STATE SOVEREIGN IMMUNITY: THE *ALLEN V. COOPER* DECISION 1 (2020), <https://crsreports.congress.gov/product/pdf/LSB/LSB10465>.

17. *See id.* at 1–4 (explaining that Congress' ability to abrogate is limited only to cases where Congress can show “congruence and proportionality” between the constitutional harm and the proposed remedy).

18. *See* Bradford R. Clark & Vicki C. Jackson, *The Eleventh Amendment*, NAT'L CONST. CTR., <https://constitutioncenter.org/the-constitution/amendments/amendment-xi/interpretations/133> (last visited Apr. 30, 2023) (explaining that generally the states enjoy broad immunity from suit, although some circumstances allow suit against state officers).

19. *Alden v. Maine*, 527 U.S. 706, 728 (1999) (holding that the Eleventh Amendment bars suits from individuals and corporations within the state and from other states when looking beyond the text of the Amendment to the overall structure of the constitution).

20. *See* HICKEY, *supra* note 16, at 1–2.

21. *See id.* at 1.

22. Shaunessy & Davis, *supra* note 15.

23. U.S. CONST. art. III, § 2; *see also* Clark & Jackson, *supra* note 18 (recounting the debate between Anti-Federalists and Federalists over the scope of sovereign immunity afforded by the text of Article III).

24. *Id.*

accountable when they violate valid laws.²⁵ Shortly after the Constitution's ratification, the Supreme Court addressed the power of the doctrine under Article III in the landmark case *Chisholm v. Georgia*.²⁶

In that case, a citizen from North Carolina sought damages from Georgia for unpaid goods sold to the state during the Revolutionary War.²⁷ The Court held that under Article III, a state can be held liable for harms caused to private citizens.²⁸ In its analysis, the Court explained that the sovereign immunity guaranteed by the Constitution differs from that in Great Britain, because citizens of the United States are not subjects. Rather, all citizens are equal, and thus state governments, as representatives of those citizens are not immune from private suit brought by citizens from another state.²⁹

Concerned by the precedent set by *Chisholm v. Georgia*, Congress rapidly ratified the Eleventh Amendment to protect sovereign immunity.³⁰ The text of the Eleventh Amendment provides: "The Judicial power of the United States shall not be construed to extend to any suit in law or equity, commenced or prosecuted against one of the United States by Citizens of another State, or by Citizens or Subjects of any Foreign State."³¹ Following its ratification, pending suits against states were "generally dropped"³² and private citizens were left without recourse for past wrongs committed by the states.³³

While the text of the Amendment only protects states from litigation against out-of-state or foreign plaintiffs, the Supreme Court has expanded the doctrine to include suits originating from a state's own citizens.³⁴ The Court continued this expansion in *Alden v. Maine*, enshrining state sovereign immunity even for cases brought within the state's own court system.³⁵

25. *Id.*

26. *See generally* *Chisholm v. Georgia*, 2 U.S. 419, 430–38 (1793).

27. *See id.* at 419.

28. *See id.* at 452.

29. *See id.* at 472.

30. *See id.* (explaining that Congress passed the Eleventh Amendment to overturn *Chisholm v. Georgia* and enshrine state sovereign immunity).

31. U.S. CONST. amend. XI.

32. Clark & Jackson, *supra* note 18.

33. *See* CONG. RES. SERV., INFRINGEMENT OF INTELLECTUAL PROPERTY RIGHTS AND STATE SOVEREIGN IMMUNITY 1, 4 (2010), https://www.everycrsreport.com/files/20100730_RL34593_d9260ec5494a86ce9adcd94ce633330d05b268f.pdf (explaining that IP holders generally have no recourse against state infringers due to the Supreme Court's precedent regarding the Eleventh Amendment).

34. *See id.* at 4; *see generally* *Hans v. Louisiana*, 134 U.S. 1, 10 (1890) (holding that a state cannot be sued in federal court by one of its own citizens, even if the conflict arose under federal law).

35. 527 U.S. 706, 730–32 (1999) (holding that the Constitution, "by delegating to Congress the power to establish the supreme law of the land when acting within its enumerated

Generally, courts recognize two purposes behind sovereign immunity.³⁶ The first reason is to promote faith in governmental decisions by barring lawsuits challenging their discretionary decisions and actions, such as allocating resources or specific political decisions.³⁷ The second reason is to protect the public treasury from the costs and consequences of governmental infringement, as allowing citizens to sue state actors could cause states to allocate money from public goods to defend the lawsuit and pay damages.³⁸

Today, states are immune from suits originating from out-of-state or foreign citizens, or its own citizens, in federal or state court, and may define their own state laws surrounding sovereign immunity.³⁹ In effect, state sovereign immunity doctrine generally protects states from suit by any private citizen, unless the state specifically consents to waive this immunity or unless Congress has specifically abrogated the immunity.⁴⁰

2. *The Application of Sovereign Immunity to the State's Instrumentalities*

Under the Eleventh Amendment, state sovereign immunity also extends to a state's "lesser entities" serving as "arms" or "instrumentalities" as defined by that state.⁴¹ While the Supreme Court did not offer a universal test, whether an entity is immune as "arms" or "instrumentalities" depends on the "relationship between the sovereign and the entity in question" and the "essential nature and effect of the proceeding."⁴² For example, the Fifth Circuit utilizes a six-part test originating in *Clark v. Tarrant County* to determine whether or not a state entity also has immunity where the most important factor is the extent that the entity is funded by the state.⁴³

3. *The Scope of and Exceptions to State Sovereign Immunity*

In general, a state's sovereign immunity protection supersedes an individual's copyright protection unless the state waives their immunity. The

powers, does not foreclose a State from asserting immunity to claims arising under federal law merely because that law derives not from the State itself but from the national power").

36. Miles McCann, *State Sovereign Immunity*, NAT'L ASS'N ATT'YS GEN., (Nov. 11, 2017) <https://www.naag.org/attorney-general-journal/state-sovereign-immunity/> (explaining the history, purposes, and exceptions to the state sovereign immunity doctrine).

37. *See id.* (explaining that the Eleventh Amendment serves, in part, to shield states from legal challenges to their political decisions promoting faith in state governments).

38. *See id.* (positing that abrogating or further diminishing state sovereign immunity may threaten a state's treasury).

39. *See id.* (concluding that "only in limited instances can the state itself be sued against its will and even the doctrine's many wrinkles tend to favor of the state as sovereign").

40. *See Hans*, 134 U.S. at 13; *Alden*, 527 U.S. at 728.

41. *Alden*, 527 U.S. at 756.

42. *Regents of the Univ. of Cal. v. Doe*, 519 U.S. 425, 429 (1997).

43. 798 F.2d 736 (5th Cir. 1986).

Copyright Act vests exclusive jurisdiction over copyright infringement cases in federal court.⁴⁴ However, federal courts are precluded by the Eleventh Amendment from exercising jurisdiction over state defendants.⁴⁵ In the absence of state consent or Congressional abrogation, states are effectively “immune from any monetary liability for infringing federal copyrights.”⁴⁶

There are limited exceptions to state sovereign immunity, including: the *Ex parte Young* exception, waiver by the state, and congressional abrogation. Under the *Ex parte Young* exception, private litigants may sue a state actor for prospective injunctive relief to end a “continuing violation of federal law” such as on-going copyright infringement.⁴⁷ Additionally, states may voluntarily waive their sovereign immunity; however, they are unlikely to do so without an incentive.⁴⁸ A state may waive its sovereign immunity by issuing a state statute, proclaiming waiver in its state Constitution, or by accepting federal funds contingent on waiving state sovereign immunity as part of a federal program.⁴⁹ Finally, Congress may abrogate sovereign immunity under strict guidelines.

a) Congressional Abrogation

The Supreme Court places two stringent requirements on Congress’ ability to abrogate state sovereign immunity. Congress must first have “unequivocally expresse[d] its intent to abrogate [sovereign] immunity” within the specific legislation.⁵⁰ Second, Congress must have acted “pursuant to a valid exercise of power.”⁵¹ Determining whether Congress has genuinely acted within its powers has proven complicated, even more so after the rise of the Congruence and Proportionality test in *City of Boerne v. Flores*. In that case, the Court held that for sovereign laws passed under the Fourteenth Amendment, abrogating legislation is valid only if “there is congruence and proportionality between the injury to be prevented or remedied and the means adopted to that end.”⁵²

The Copyright Office discussed the ambiguity of the Congruence and Proportionality test in its 2020 Copyright and State Sovereign Immunity

44. See 17 U.S.C. § 301(a) (declaring that all claims arising from the violation of an exclusive right under the Copyright Act are governed exclusively by the Act).

45. *Hans*, 134 U.S. at 10–11.

46. John T. Cross, *Suing the States for Copyright Infringement*, 39 BRANDEIS L.J. 337, 339 (2000).

47. McCann, *supra* note 36.

48. See *id.* (listing reasons why a state may waive sovereign immunity, such as by accepting funds through a federal program).

49. *Id.*

50. *Seminole Tribe v. Florida*, 517 U.S. 44, 55 (1996).

51. *Id.*

52. *City of Boerne v. Flores*, 521 U.S. 507, 521 (1997).

Report.⁵³ In more than two decades since *City of Boerne*, the Court had only upheld congressional abrogation on two occasions: abrogation involving either disability- or sex-based discrimination.⁵⁴ The Copyright Office noted that while the Court had encouraged Congress to consider revising the CRCA to align with the *City of Boerne* test, the Court had “provided less guidance as to the nature and volume of evidence” that would support a finding that state infringement is a widespread and persistent problem, that states are acting intentionally or at least recklessly, and that the bill is narrowly tailored to address those injustices as necessary for Congressional abrogation.⁵⁵

B. HISTORY AND PURPOSES OF COPYRIGHT

The copyright system secures the legal rights of a creator’s investment into their innovative practice.⁵⁶ This system serves as a vehicle of free expression and provides economic incentive to create and spread new and innovative ideas.⁵⁷ This system grants copyright owners a bundle of exclusive rights, similar to the rights granted to other property owners.⁵⁸ Artists, authors, musicians, photographers, actors, programmers, and other creatives rely on these rights to protect and monetize their otherwise nonexcludable works.⁵⁹

These exclusive rights are almost as old as the nation itself. The “IP Clause” of the United States Constitution empowers Congress to “promote the Progress of Science and useful Arts, by securing for limited Times to Authors and Inventors the exclusive Right to their respective Writings and Discoveries.”⁶⁰ Prior to the ratification of the Bill of Rights in 1791, the only direct mention of the word “right” in the Constitution laid in the IP Clause.⁶¹

53. COPYRIGHT AND SOVEREIGN IMMUNITY, *supra* note 13, at 72.

54. *Id.*

55. *Id.* at 72–73.

56. ALDEN ABBOTT, KEVIN MADIGAN, ADAM MOSSOFF, KRISTEN OSENGA & ZVI ROSEN, FEDERALIST SOC., HOLDING STATES ACCOUNTABLE FOR COPYRIGHT PIRACY 1 (2021), <https://regproject.org/wp-content/uploads/Paper-Holding-States-Accountable-for-Copyright-Piracy.pdf> [hereinafter *Holding States Accountable*].

57. *See* Harper & Row Publishers, Inc. v. Nation Enters., 471 U.S. 539, 558 (1985) (finding “copyright itself to be the engine of free expression” and that “copyright supplies the economic incentive to create and disseminate ideas”).

58. *Jim Olive Photography*, 580 S.W.3d at 376 n.26 (Tex. 2021) (citing *Harper & Row*, 471 U.S. at 546) (“Section 106 of the Copyright Act confers a bundle of exclusive rights to the owner of the copyright.”).

59. *Id.*

60. U.S. CONST. art. I, § 8, cl. 8; *see* ADAM MOSSOFF, HERITAGE FOUND., THE CONSTITUTIONAL PROTECTION OF INTELLECTUAL PROPERTY 2–3 (2022), <https://www.heritage.org/sites/default/files/2021-03/LM282.pdf> (explaining Congress’ authority under the Constitution to pass the Copyright Act).

61. *See* MOSSOFF, *supra* note 60, at 3.

Heeding the call from the Constitution, the First Congress immediately enacted the original Copyright Act in 1790.⁶² Early courts interpreted and enforced the evolving copyright statutes to ensure creatives “were provided reliable and effective property rights in the fruits of their productive labors.”⁶³ Having secured these rights, creatives and inventors invested their time and wealth into their arts, sparking many of the technological and cultural advances of the 19th century.⁶⁴ Inspired by the industry of this period, the American system of awarding property rights to creators became the “gold standard” across the globe.⁶⁵

A copyright is a bundle of federal statutory rights over an expressive work or “original work of authorship.”⁶⁶ Copyright protects the abstracted creative work, not simply the individual mediums displaying the work.⁶⁷ The principal rights granted by the Copyright Act include the exclusive right to reproduce and display the work.⁶⁸ These rights initially vest in the creator,⁶⁹ but like property rights, they may be sold, licensed, left in a will,⁷⁰ or used as collateral.⁷¹ Copyright is “defined by the right to exclude,” which the Supreme Court has recognized as one of the most important “sticks” in the “property bundle.”⁷²

There is a tension between the public’s desire to access the work to spread ideas and culture, and the artist’s desire for economic compensation and recognition to create future works.⁷³ To ease these competing desires, the copyright system provides for a time-limited right for the artist to exclude the public from accessing and altering the work that extends beyond the duration of the creator’s lifetime.⁷⁴ After the copyright protection lifts, the work enters

62. *See id.* at 1–2.

63. *Id.*

64. *See id.* at 2 (explaining how “[t]hese intellectual property rights spurred the explosive growth in the U.S. innovation economy from the 19th century through today”).

65. *See id.*; Adam Mossoff, *Institutional Design in Patent Law: Private Property Rights or Regulatory Entitlements*, 92 SO. CAL. L. REV. 921, 936–37 (2019).

66. 17 U.S.C. § 102.

67. *See id.* § 202 (“Ownership of a copyright, or of any of the exclusive rights under a copyright, is distinct from ownership of any material object in which the work is embodied.”).

68. *See id.* § 106 (listing the exclusive rights granted by the Copyright Act).

69. *Id.* § 201(a).

70. *Id.* § 201(d)(1).

71. *See Copyright Reform and the Takings Clause*, 128 HARV. L. REV. 973, 978 (2015).

72. *Id.* at 980.

73. *See* Alexander Cuntz, *Copyright and the Currency of Creativity: Beyond Income*, WIPO MAG. (June 2019), https://www.wipo.int/wipo_magazine/en/2019/03/article_0003.html (criticizing solely income-based rationales for copyright that fail to consider the artist’s need for attribution and that some artists are intrinsically motivated).

74. DILAN J. THAMPAPILLAI, CORNELL LAW SCHOOL GRADUATE STUDENT PAPERS, THE BALANCING ACT OF COPYRIGHT: THE COPYRIGHT LAWS OF AUSTRALIA AND THE

into the public domain, where the public can access it generally and incorporate it into their own expressive works or use it to promote their business.

C. COPYRIGHT AND STATE SOVEREIGN IMMUNITY

For the majority of American history, copyright law did not directly address the liability of state actors.⁷⁵ The law instead stated that “anyone” or “any person” infringing a copyright was subject to liability.⁷⁶ In the absence of clear guidelines, courts turned to the Supreme Court’s abrogation jurisprudence to determine whether states should be held liable for infringement and came to “inconsistent conclusions.”⁷⁷

For example, in 1962, the Eighth Circuit in *Wibtol v. Crow* dismissed a copyright claim brought against a school district, because the district was found to be an instrumentality of a state’s educational system, and because the Eleventh Amendment provided that “a state cannot be sued without its consent.”⁷⁸ The court did not determine whether Congress had “made its intent to abrogate sufficiently clear” or on what basis Congress could abrogate this immunity.⁷⁹

However, in 1979, the Ninth Circuit in *Mills Music, Inc. v. Arizona* affirmed an award for copyright damages against a state infringer.⁸⁰ In its analysis, the court found that the Copyright Act of 1909’s language was “sweeping and without apparent limitation,” meaning that the words “any person” in “any person [who] shall infringe” should be interpreted broadly and encompass state infringers.⁸¹ The court further held that Congress had “inherent” authority over legislating copyrights and abrogating sovereign immunity under the IP Clause, and concluded that a state may not “in any way diminish the federally granted and protected rights of a copyright holder.”⁸²

In an attempt to directly abrogate sovereign immunity and clarify the ambiguity throughout the circuits, Congress asked the Copyright Office to study the extent of the problem in the 1980s, which resulted in the Oman

UNITED STATES IN THE DIGITAL ERA 1, 3–4 (2003), http://scholarship.law.cornell.edu/lps_papers/7.

75. COPYRIGHT AND SOVEREIGN IMMUNITY, *supra* note 13, at 14.

76. *Id.*

77. *Id.* at 15.

78. *Id.* at 15 (citing *Wibtol*, 309 F.2d 777, 781 (8th Cir. 1962)).

79. *Id.*

80. 591 F.2d 1278, 1284–85 (9th Cir. 1979).

81. *Id.*; see also COPYRIGHT AND SOVEREIGN IMMUNITY, *supra* note 13, at 15.

82. COPYRIGHT AND SOVEREIGN IMMUNITY, *supra* note 13, at 15 (citing *Mills Music*, 591 F.2d at 1285).

Report,⁸³ after which Congress passed the Copyright Remedy Clarification Act (CRCA).⁸⁴ Between 1990 and 2020, this Act specifically held states liable for copyright infringement; however, the CRCA was consistently weakened by court rulings on a patent act.⁸⁵ Ultimately, in 2020, the Supreme Court found that Congress' attempt at abrogation through the CRCA was unconstitutional, restoring states' immunity from copyright infringement suits in federal court.⁸⁶

1. *Congress' Attempt to Abrogate and the Rise of the CRCA*

In 1976, Congress reconstructed the nation's copyright laws but did not expressly address the issue of state sovereign immunity.⁸⁷ The 1976 Copyright Act instead held "anyone" violating the exclusive rights of the copyright owner liable for infringement. The statute's ambiguity led to inconsistent application of sovereign immunity across the lower courts.⁸⁸ Further, the ambiguous language of the statute failed to provide a "clear statement" of Congress' intent to abrogate and could not be used to hold states liable for infringement.⁸⁹

By 1987, Congress grew wary of states' unchecked ability to freely infringe copyrighted material and the House Judiciary Committee asked the Copyright Office to conduct a report on the "interplay between copyright infringement and the Eleventh Amendment."⁹⁰ In its final report (the "Oman Report"), the Copyright Office published the results of its study after conducting a thorough review of public comments, analyzing the Eleventh Amendment case law, and examining state waiver of sovereign immunity.⁹¹ Its investigation yielded only five instances of copyright infringement by a state actor where the copyright

83. U.S. COPYRIGHT OFF., COPYRIGHT LIABILITY OF STATES AND THE ELEVENTH AMENDMENT 1 (1988), <https://www.copyright.gov/reports/copyright-liability-of-states-1988.pdf> [hereinafter Oman Report].

84. *See id.* at 14.

85. *See, e.g., Fla. Prepaid Postsecondary Educ. Expense Bd. v. Coll. Sav. Bank*, 527 U.S. 627, 636 (1999) (finding that the Court's precedent in *Seminole Tribe* determines that Congress may not abrogate sovereign immunity under its Article I powers, including the Commerce Clause and the IP Clause).

86. *Allen v. Cooper*, 140 S. Ct. 994, 1007 (2020) (concluding that *Florida Prepaid* "all but prewrote" the decision to overturn the CRCA because that case held that "Article I's Intellectual Property Clause could not provide the basis for an abrogation of sovereign immunity" and that the current record of state infringement under the Oman report does not provide evidence of a widespread deprivation of due process of property rights to support abrogation under § 5 of the Fourteenth Amendment).

87. 17 U.S.C. §§ 101–801 (1976).

88. *See* COPYRIGHT AND SOVEREIGN IMMUNITY, *supra* note 13, at 15.

89. *See id.* at 16 (expressing the House Judiciary Committee's concern that the 1976 Act's "anyone" language was insufficiently clear).

90. Oman Report, *supra* note 83, at 7.

91. *Id.*

owner documented actual problems enforcing their claims.⁹² The report further found in its analysis of the relevant Eleventh Amendment case law that the text of the Copyright Act was not “sufficiently clear” in expressing its intent to abrogate state sovereign immunity.⁹³ In its conclusion, the Oman Report recommended that Congress amend the Copyright Act to clearly state its intention to subject states to liability for copyright damages.⁹⁴

Congress followed the recommendation and proposed the Copyright Remedy Clarification Act (CRCA) under its Article I abrogation power in 1989.⁹⁵ It modeled the language of the final bill on a previous law that the Supreme Court had “twice cited as an example of Congress’ ability to abrogate the Eleventh Amendment when it wanted to do so.”⁹⁶ The final bill had two provisions. The first amended § 501(a) of the Copyright Act to clarify that the term “anyone” included both states and their instrumentalities.⁹⁷ The second amended § 511 to specifically mandate that states “shall not be immune under the Eleventh Amendment” and would be liable “to the same extent” as private actors.⁹⁸ The CRCA was signed into law on November 15, 1990.⁹⁹

2. Seminole Tribe, Florida Prepaid Cases, and the Fall of the CRCA

Over the course of three decades, the Supreme Court repeatedly challenged Congress’ authority to abrogate sovereign immunity in intellectual property infringement cases under Article I. In 2019, the Court ultimately declared the CRCA to be unconstitutional in *Allen v. Cooper*.¹⁰⁰

The Court first cast doubt on the CRCA with its ruling in *Seminole Tribe of Florida v. Florida*. In that case, the Court explained that Article I of the Constitution could not be interpreted to abrogate sovereign immunity and that intended abrogation would require Congress to use its § 5 powers under the Fourteenth Amendment instead.¹⁰¹ Following this revelation, the Court further considered whether Congress had appropriately abrogated state

92. *Id.* at 104.

93. *Id.*

94. *Id.*

95. *Id.* at 19 (citing Copyright Remedies Clarification Act, H.R. 3045, 101st Cong. § 2(a)(2), 101 Stat. 2749, 2749 (1990)).

96. *Id.*

97. Copyright Remedies Clarification Act, H.R. 3045, 101st Cong. § 2(a)(1), 101 Stat. 2749, 2749 (1990).

98. *Id.* § 2(a)(1).

99. *Id.* §§ 1–3.

100. 140 S. Ct. at 1000.

101. *Florida Prepaid*, 527 U.S. at 635–36 (“Congress may not abrogate state sovereign immunity pursuant to its Article I powers; hence the Patent Remedy Act cannot be sustained under either the Commerce Clause or the [Intellectual Property] Clause.”).

sovereign immunity under its § 5 powers in a pair of cases referred to as the *Florida Prepaid* cases. In those cases, the Court concluded that the language of the Patent and Plant Variety Protection Remedy Clarification Act (which functioned like the CRCA) was sufficiently clear as to Congress' intent to abrogate.¹⁰² However, the Court found that the statute failed the “congruence and proportionality” test from *City of Boerne*, as the few constitutional violations sought to be remedied were incongruent with the requirement of abrogation.¹⁰³ In particular, the Court held that Congress had not identified a pattern of unconstitutional infringement and narrowly tailored its abrogation to that finding.

The majority opinion in *Florida Prepaid* detailed a pathway for congressional abrogation. First, Congress had to establish sufficient evidence of the state actors' infringement.¹⁰⁴ Second, Congress had to consider the adequacy of state-law remedies for state infringement.¹⁰⁵ Third, Congress must find more than “a handful of instances” of infringement and must establish a record of intentional or reckless infringement.¹⁰⁶ Fourth, legislation must “not be limited to ‘cases involving arguable constitutional violations, such as where a State refuses to offer any state-court remedy.’”¹⁰⁷ Under this framework, Congress' efforts to abrogate under the CRCA were insufficiently congruent and proportional to amount to a constitutional harm sufficient for Congress' abrogation of sovereign immunity under the Fourteenth Amendment.¹⁰⁸

a) *Allen v. Cooper*: The Death Knell of the CRCA

The CRCA was officially declared unconstitutional in *Allen v. Cooper*.¹⁰⁹ The petitioners, videographer Frederick Allen and his production company Nautilus Productions, LLC, entered into an exclusive contract with the state of North Carolina to film the historic restoration of the *Queen Anne's Revenge*, the ship captained by the pirate Blackbeard.¹¹⁰ Allen retained the copyright in these videos.¹¹¹ Despite previously agreeing to a settlement compensating Allen for the use of the work, the state further posted the copyrighted material

102. *Id.* at 635.

103. *Id.* at 639.

104. *Id.* at 640.

105. *Id.* at 643–44.

106. *Id.* at 646–47.

107. *Id.*

108. COPYRIGHT AND SOVEREIGN IMMUNITY, *supra* note 13, at 21.

109. 140 S. Ct. at 1006–07.

110. *Id.* at 999.

111. *Id.*

to its website without Allen's knowledge and in violation of the parties' agreement.¹¹²

After discovering the infringement, Allen sued the state for copyright infringement in 2013.¹¹³ As the CRCA amended the Copyright Act to eliminate states' sovereign immunity, Allen was able to bring the case directly in federal district court.¹¹⁴ The case settled, and North Carolina agreed not to further infringe the copyrighted material.¹¹⁵ Yet, to circumvent this settlement, the state passed Blackbeard's Law, which declared all video footage of the historic restoration part of the public domain, and impacted only the petitioners due to their exclusive contract.¹¹⁶

Seeking relief from the state's intentional infringement, Allen again sued the state in federal district court in 2015. The state argued that the suit was barred due to the Eleventh Amendment's promise of state sovereign immunity and that the CRCA was beyond the scope of Congress' constitutional authority to abrogate and render the state liable for copyright infringement.¹¹⁷ The district court rejected this argument and denied the state's motion to dismiss.¹¹⁸ The Fourth Circuit reversed this decision and held Congress' attempt to abrogate state sovereign immunity with the CRCA unconstitutional.¹¹⁹ With a constitutional question raised, the case headed to the Supreme Court.

Even though Justice Kagan's majority opinion recognizes that North Carolina has committed a "modern form of piracy,"¹²⁰ the Court offered Allen no relief. The Court found that Congress, in passing the CRCA, failed to establish a record of intentional and unconstitutional state infringement required by the congruence and proportionality test established in *Flores*. Without satisfying the *Flores* test, the CRCA fell outside of the scope of congressional authority and was struck down by the Court in 2020.¹²¹ The Court concluded by suggesting that Congress pass more tailored legislation in

112. *Id.*

113. *Id.*

114. *Allen v. Cooper*, CONST. ACCOUNTABILITY CTR., <https://www.theusconstitution.org/litigation/allen-v-cooper/> (last visited May 9, 2023).

115. *Id.*

116. *Id.*; see also Hudson Institute, *IP Infringement and State Sovereign Immunity*, YOUTUBE (Oct. 25, 2021), <https://www.youtube.com/watch?v=NyVJW0T6ejM&t=3060s> (starting at 23:30) (Frederick Allen discussing the sole impact of "Blackbeard's Law" (N.C. § 121-25(b)) which eliminated Allen's ability to enforce his copyright).

117. COPYRIGHT AND SOVEREIGN IMMUNITY, *supra* note 13, at 22.

118. *Id.*

119. *Id.*

120. *Allen*, 140 S. Ct. at 999.

121. *Id.* at 1007.

the future to accomplish its goal of properly abrogating state sovereign immunity.¹²²

3. *Availability of Suit Against State Infringers After the CRCA*

The Supreme Court's decision leaves Allen and other creators with little to no recourse for "blatant and intentional theft" of their copyrighted works.¹²³ After the decision, Allen reported to the Copyright Office that "because of current law and Supreme Court precedent, [he was] powerless to enforce [his] constitutionally granted intellectual property rights against infringement by States."¹²⁴ Beyond the lost licensing opportunities Allen had suffered from the state's piracy, Allen had incurred "hundreds of thousands of dollars in legal expenses" attempting to secure his investment, and had spent countless hours engaging in monitoring and enforcement rather than creating new works.¹²⁵

Congress appeared interested in remedying the issue. Just weeks after the CRCA was struck down, Senators Thom Tillis (R-NC) and Patrick Leahy (D-VT) wrote to both the U.S. Copyright Office and the U.S. Patent and Trademark Office, requesting a study on the impact of infringement by states.¹²⁶ The pair of Senators formed a bipartisan team committed to combatting intellectual property theft.¹²⁷ The report sought to lay out a path for congressional abrogation consistent with the decision in *Allen v. Cooper*.¹²⁸

In August 2021, the Copyright Office released a report entitled *Copyright and State Sovereign Immunity* exploring the interplay between the doctrines and the extent of state actors' copyright infringement.¹²⁹ In large part, the report focused on the lack of adequate remedies from states' infringement. Overall, the report found that the number of infringement allegations substantially

122. *Id.*

123. *Holding States Accountable*, *supra* note 56, at 4.

124. *Id.*

125. *Id.*

126. Thom Tillis & Patrick Leahy, Letter to Maria Strong, Acting Register of Copyrights and Director, U.S. Copyright Office (Apr. 28, 2020), <https://www.copyright.gov/policy/state-sovereign-immunity/letter.pdf>.

127. *See, e.g.*, Tillis and Leahy Introduce Bipartisan Legislation to Combat Copyright Theft, Enhance Content Sharing, and Hold Tech Accountable, THOM TILLIS: U.S. SEN. N.C. (Mar. 18, 2022), <https://www.tillis.senate.gov/2022/3/tillis-and-leahy-introduce-bipartisan-legislation-to-combat-copyright-piracy-enhance-content-sharing-and-hold-tech-accountable>; Tillis and Leahy Introduce Bipartisan Legislation to Improve Patent Quality, THOM TILLIS: U.S. SEN. N.C. (Aug. 2, 2022), <https://www.tillis.senate.gov/2022/8/tillis-and-leahy-introduce-bipartisan-legislation-to-improve-patent-quality>.

128. *See* COPYRIGHT AND SOVEREIGN IMMUNITY, *supra* note 13, at 1 (responding to Congressional requests to undertake a study to determine whether Congress could legislatively abrogate sovereign immunity consistent with the Court's analysis in *Allen v. Cooper*).

129. *See id.* at 1–3.

grew since Congress passed the CRCA and that “evidence indicates that state infringement represents a legitimate concern for copyright owners.”¹³⁰

However, because of the high standard set by *Allen v. Cooper* and because of the ambiguity of the doctrine, the report admitted that it cannot conclude that even the more robust record of states’ piracy it established would meet constitutional muster.¹³¹ Nonetheless, the Copyright Office still “believes that infringement by state entities is an issue worthy of congressional action” and asked Congress to consider renewed abrogation attempts or alternative approaches to establishing liability for state infringement.¹³² One approach is through takings claims.

III. TAKINGS CLAUSE AND COPYRIGHT

The government has the authority to dispossesses private citizens of personal property to allocate it for the “public good” through its power of eminent domain.¹³³ In granting the state this power, the Takings Clause signals that “individual property rights are subordinate to the good of the polity.”¹³⁴ However, the state must provide just compensation for property that it takes, and it can only take property when it is justified by an expected social benefit.¹³⁵ Similar frameworks for compensation have been in place since Ancient Roman times, such as the *Magna Carta* which provided that the King must compensate private citizens before taking any property for the common good.¹³⁶ However, this doctrine originally applied only to tangible or physical property, not intellectual property such as copyright.¹³⁷

After the fall of the CRCA, copyright owners began pursuing copyright-takings claims to seek damages resulting from state infringement. So far, these claims have been unsuccessful in both state and federal circuit courts because no court has been willing to both acknowledge (1) copyright as a protectable property interest under the Fifth Amendment and (2) that state infringement impermissibly interferes with a property owner’s ability to concurrently

130. *Id.* at 2.

131. *Id.*

132. *Id.*

133. Micah Elazar, “Public Use” and the Justification of Takings, 7 U. PA. J. CONST. L. 249, 249 (2004).

134. *Id.*

135. *Id.*

136. MAGNA CARTA cl. 28 (“No constable or other bailiff of ours is to take the corn or other chattels of anyone, unless he immediately gives money for this, or is able to have a delay with the consent of the seller.”).

137. Leroy J. Ellis V, *Copyright and Federalism: Why State Waiver of Sovereign Immunity is the Best Remedy for State Copyright Infringement*, 20 NW. J. TECH. & INTELL. PROP. 1 (2022).

exercise a fundamental property right. However, the Supreme Court's recent decision in *Cedar Point Nursery v. Hassid* may enable copyright infringement to be viewed as a temporary appropriation of one's right to exclude third parties, which would be compensable under the Takings Clause. Nonetheless, after *Jim Olive Photography*, the viability of the copyright-takings claim is unreliable, as the Supreme Court denied Jim Olive's Petition for Writ of Certiorari that asked the Court to consider whether the *Cedar Point Nursery* holding extends to nonrivalrous property like copyright.¹³⁸

This Part proceeds in three Sections. The first Section explains the state of Takings Clause jurisprudence in three contexts: physical or "per se" takings, regulatory or constructive takings, and appropriation of access rights under *Cedar Point Nursery*. The second Section argues that intangible property such as copyright are protected by the Takings Clause. The final Section explores the Court's decision in *Cedar Point Nursery* which interprets the Takings Clause to include appropriations of the right to exclude third parties from access under a "per se" rule and its application to *Jim Olive Photography*.

A. TAKINGS CLAUSE JURISPRUDENCE

The Fifth Amendment provides that "private property [shall not] be taken for public use, without just compensation."¹³⁹ This amendment is applied to the states through the Fourteenth Amendment.¹⁴⁰ The Constitution does not define the term "takes" in the context of the Takings Clause, but courts have interpreted it to include government actions that seize or occupy private property for a public purpose or that otherwise substantially diminish its value or use.¹⁴¹ The Supreme Court has gradually expanded the definition of "takings" under the Fifth Amendment to include not only physical takings, where the government physically seizes or occupies a property, but also regulatory takings, where the government's actions restrict a property owner's use of their property to a similar degree as a physical seizure. Most recently in *Cedar Point Nursery*, the Supreme Court further expanded the definition of "physical takings" to include state actions that temporarily take away or appropriate a property owner's right to exclude third parties from accessing their property.¹⁴²

138. 624 S.W.3d 764 (Tex. 2021), *cert. denied*, 142 S. Ct. 1361 (2022).

139. U.S. CONST. amend. V.

140. U.S. CONST. amend. IV.

141. Douglas T. Kendall & Charles P. Lord, *The Takings Project: A Critical Analysis and Assessment of the Progress So Far*, 25 B.C. ENVTL. AFF. L. REV. 509, 515 (1998).

142. *Allen*, 140 S. Ct. at 999.

1. *Physical Takings and Per Se Takings Claims*

Generally, the Court has recognized two varieties of physical takings: occupations and appropriations.¹⁴³ An occupation is a physical invasion of a private property (usually land) by either a state actor or a third-party acting with governmental authority, or a physical placement of objects on private personal property by those actors. Whereas, an appropriation refers to “a government order or other action that either explicitly or effectively divests an owner of her interest in property and transfers ownership to the government” or a third-party.¹⁴⁴ Sometimes courts expressly distinguish between the two types of physical takings within a particular case; however, courts also often conflate these categories and refer to them by either term, or generically as a “physical” taking.¹⁴⁵ The Supreme Court has repeatedly stated that a “per se” (or “categorical”) rule governs most physical takings.¹⁴⁶ This means that in most cases, the government must compensate a private citizen for the use of their property if it physically takes or occupies that property, or if it appropriates property rights onto itself or a third party.¹⁴⁷

In *Loretto v. Teleprompter Manhattan CATV Corp.*,¹⁴⁸ the Court held that any permanent physical occupation authorized by the government would be a per se taking that requires just compensation regardless of the public interest it furthers or the economic interest of the property owner.¹⁴⁹ The Court also noted it is immaterial that the appropriation takes over only a small area, so long as the government action caused the permanent occupation of a space.¹⁵⁰

In 2015, the Court in *Horne v. Department of Agriculture* extended the per se rule from land to personal property. The Court considered whether a mandate requiring property owners to set aside personal property, in this case raisins, for governmental use could be compensable under the per se analysis.¹⁵¹

143. John D. Echeverria, *What is a Physical Taking?*, 54 U.C. DAVIS L. REV. 731, 747 (2020).

144. *Id.* at 747–48.

145. *Id.* at 748.

146. *Id.* at 745.

147. *Id.* at 745.

148. 458 U.S. 419, 441 (1982). In that case, New York law required landlord companies to allow the installation of cable equipment in rented properties. Seeing this permanent occupation as an appropriation of their right to do as they please with their property, a landlord brought a per se Takings Clause claim. The court agreed with the landlord that a permanent physical occupation appropriates three of their basic property rights: (1) the right to fully possess the property or exclude others from possessing it; (2) the ability to exclude others from using the property and an inability to make personal use of the property; and (3) the ability to sell or dispose of the property due to its decreased economic value.

149. *Id.* at 441.

150. *Id.* at 421.

151. *Horne*, 576 U.S. at 354–55.

Finding the reserve requirement to be akin to a physical taking, the Court interpreted the Fifth Amendment to prevent the government from appropriating any part of a person's private property without compensation.¹⁵² The case further clarified that "nothing in text or history of the Takings Clause" prevents considering appropriations of personal property under the per se analysis, and that personal property is no less protectable than real property such as land.¹⁵³

2. *Regulatory Takings Claims*

Courts may also find that the government must compensate a property owner for a regulatory taking (also known as a constructive taking) if the government's action goes "too far" and restricts the property owner's rights to the point where they are functionally equivalent to a physical seizure or occupation.¹⁵⁴ Although it is more inherently difficult to determine whether a government action has exceeded its ability to regulate public property than whether a property has been physically occupied or seized, the Court has developed several tests to guide this inquiry over the past century.

The Court began developing the regulatory takings doctrine in 1922, in *Pennsylvania Coal v. Mahon*, where Court first interpreted the Takings Clause to encompass particularly oppressive regulatory takings of private property generally.¹⁵⁵ In that case, the Court found that regulation had "very nearly the same effect for constitutional purposes as appropriating or destroying [the estate]" and that a regulation should be considered a taking when it "goes too far" in depriving the property owner of the enjoyment of any of the property rights.¹⁵⁶

Modern jurisprudence utilizes the balancing test expressed in *Penn Central Transportation Co. v. New York City* to determine whether a regulatory taking has occurred and is compensable.¹⁵⁷ In that case, Justice Holmes clarified the holding in *Mahon* by first reminding of its warning that governing could hardly "go on if to some extent values incident to property could not be diminished without paying for every such change in the general law."¹⁵⁸ Rather, to prevail on regulatory takings claims, plaintiffs must show that "government regulations work a significant deprivation of property right."¹⁵⁹ The decision

152. *Id.*

153. *Id.* at 358.

154. Echeverria, *supra* note 143, at 747.

155. *Pennsylvania Coal Co. v. Mahon*, 260 U.S. 393, 416 (1922).

156. *Id.* at 414–15.

157. 438 U.S. 104, 124–25 (1978).

158. *Id.* at 124.

159. *Id.*

recognizes that states are not financially responsible for every economic effect of their legislative actions, but that citizens should be compensated when a regulation goes “too far” so as to prevent governmental abuse of its ability to regulate private property.

Ultimately, the Court articulated and applied a three-factor ad-hoc balancing approach to determine whether the government action required compensation.¹⁶⁰ These factors include: (1) the economic impact of the regulation, (2) the extent the regulation interferes with “distinct investment-backed expectations,” and (3) “the character of the governmental action.”¹⁶¹ Under this approach, no factor is definitive, and significant diminutions in property value are generally permissible without compensation.¹⁶² In the absence of a physical conversion of the property, takings are compensable when they go “too far” and amount to a “significant deprivation of a property right.”¹⁶³

Plaintiffs often struggle to meet the high standard for relief demanded by the *Penn Central* test, which has been described as “maddeningly unpredictable” and “favoring the government in most cases.”¹⁶⁴ In 2017, the Court made the *Penn Central* test even more difficult to satisfy with its decision in *Murr v. Wisconsin* that established a new threshold for property owners to meet before reaching the *Penn Central* test.¹⁶⁵ As such, regulatory takings claims are more difficult for plaintiffs to succeed on than per se takings claims.

3. Expansion of “Physical” Takings: Cedar Point Nursery

Recently, the Court issued a unanimous opinion in *Cedar Point Nursery v. Hassid*, further expanding the scope of physical takings and potentially providing an alternative pathway for compensation for copyright holders under the Takings Clause.¹⁶⁶ The case surrounded an access regulation in California that required agricultural employers to allow labor organizers temporary access to and use of the employers’ land to meet with employees

160. *Id.* at 123–24 (“[T]his Court, quite simply, has been unable to develop any ‘set formula’ for determining when ‘justice and fairness’ require that economic injuries caused by public action be compensated by the government.”).

161. *Id.* at 124.

162. *Id.* at 124–25.

163. *Id.* at 125.

164. Timothy M. Harris, *No Murr Tests: Penn Central is Enough Already!*, 48 GEO. ENVTL. L. REV. 605, 609 (2018).

165. *Id.* at 607 (citing *Murr v. Wisconsin*, 137 S. Ct. 1933, 1945–46 (2017) which requires courts to also consider (1) the property’s treatment under state and federal law, (2) the property’s physical properties, and (3) the property’s value).

166. *Cedar Point Nursery*, 141 S. Ct. 2063, 2066.

and promote their union.¹⁶⁷ Cedar Point Nursery, a strawberry farm, argued that California's access regulation violated the Takings Clause of the Fifth Amendment, which provides that private property shall not "be taken for public use, without just compensation."¹⁶⁸

The question before the Court was whether the temporary access regulation amounted to a per se physical taking.¹⁶⁹ Answering affirmatively, the Court expanded the definition of per se takings to include actions by governmental actors that appropriates a right to exclude third parties from accessing physical, private property.¹⁷⁰ The Court explained that rather than restricting the owners' use, the government instead appropriated the property for the enjoyment of a third party, which deprived the owners of their right to exclude—a fundamental property right.¹⁷¹ Even though the access was temporary, the Court found this to be inconsequential because of its ruling in *United States v. Dow*, which found that "physical appropriation is a taking whether it is permanent or temporary; the duration of the appropriation bears only on the amount of compensation due."¹⁷² Further, the Court refused to adopt the theory that the access regulation merely regulates without appropriating the growers' right to exclude and that "the right to exclude is not an empty formality that can be modified at the government's pleasure."¹⁷³

Further, the case answered the question raised by the Supreme Court of Texas in *Jim Olive Photography* as to whether governmental appropriations of property rights, such as the right to exclude third parties, are best analyzed under a per se rule or the regulatory framework proposed in *Penn Central*.¹⁷⁴ Definitively ruling that "when the government physically appropriates property, *Penn Central* has no place—regardless whether the government action takes the form of a regulation, statute, ordinance, or decree."¹⁷⁵ Less definitely is whether or not courts will apply a per se rule to protect non-physical properties such as those protected by the Copyright Act.

B. COPYRIGHT & TAKINGS CLAIMS

Courts award compensation to owners of intangible and nonrivalrous property under the Takings Clause, just like they do for owners of traditional,

167. *Id.*

168. U.S. CONST. amend. V.

169. *Cedar Point Nursery*, 141 S. Ct. at 2066.

170. *Id.* at 2072.

171. *Id.* at 2076.

172. *Id.* (citing *United States v. Dow*, 357 U.S. 17, 26 (1956)).

173. *Id.* at 2077.

174. *Id.* at 2072.

175. *Id.*

physical property. However, courts might award a narrower scope of protection for copyright than for physical property. This is because the Court grants greater deference to physical property rights, and because copyright rights can be segmented. In other words, a state's violation of one of the exclusive rights granted by the Copyright Act does not destroy the copyright owner's enjoyment of other rights.

1. *Copyright is Likely Protectable Under the Takings Clause*

Courts have found that both intangible and nonrivalrous properties are compensable in the case of state occupation.¹⁷⁶ For example, the Supreme Court in *Ruckelshaus v. Monsanto Co.* recognized trade secrets as a property interest protected by the federal Takings Clause.¹⁷⁷ Despite trade secrets' intangible and nonrivalrous nature, which allows them to be used by another party without simultaneously depriving anyone else of their use,¹⁷⁸ the Court found that the intellectual property taking was compensable because the property holder had a reasonable investment-backed expectation in the exclusive use of the property.¹⁷⁹ After the Court determined that trade secrets were a form of property protected under Takings Clause jurisprudence, it extended property status to copyright in dicta, as it is a more "durable" form of intellectual property protection than trade secrets.¹⁸⁰ However, not all courts have accepted this dicta as law, refusing to protect copyright under the federal Takings Clause.¹⁸¹ Ultimately, the Supreme Court after *Monsanto* has accepted that patents, which are legislatively similarly to copyrights, may be considered property in the context of the Fifth Amendment.¹⁸² As such, courts after *Monsanto* usually assume copyright to be a protectable form of property under the federal Takings Clause.¹⁸³

2. *But There is Uncertainty in the Scope of Protection*

Not all property rights are held equal in the eyes of courts. As physical property historically forms the core of Takings Clause jurisprudence, physical

176. *Ruckelshaus v. Monsanto Co.*, 467 U.S. 986, 1001–02 (1984).

177. *Id.*

178. *Id.*

179. *Id.*

180. *Id.* at 1003.

181. *See, e.g.*, *Univ. of Hous. Sys. v. Jim Olive Photography*, 580 S.W.3d 360, 363 (Tex. App. 2019).

182. *Fla. Prepaid Postsecondary Educ. Expense Bd. v. Coll. Sav. Bank*, 527 U.S. 627, 642 (1999).

183. *See, e.g.*, *Jim Olive Photography v. Univ. of Houston Sys.*, 624 S.W.3d 764, 770 (Tex. 2021).

property is held in higher regard than personal and intangible properties.¹⁸⁴ However, physical and intangible property share many of the same fundamental property rights.¹⁸⁵ For example, the Supreme Court in *Monsanto* explained that trade secrets share many characteristics with real property such as their transferability and ability to be used as collateral.¹⁸⁶ Like any other form of property, copyright is a bundle of rights, including the fundamental right to exclude.¹⁸⁷ However, the Court has recognized that intellectual property enjoys weaker Takings Clause protection than physical property because it is intangible and nonrivalrous, which means that it can be easily appropriated and used by a government actor without wholly depriving the owner of its value.¹⁸⁸

All property can be best thought of as a bundle of rights protected by the Courts, although this makes it susceptible to segmentation.¹⁸⁹ However, the Court requires that Takings claims consider the proportion of the size of the damages over the total property's value.¹⁹⁰ While this analysis requires consideration of the whole property, the Court has sometimes divided property for the Takings Clause purposes.¹⁹¹ One of the factors it uses to divide property rights is by the property interest impacted.¹⁹² In regulatory claims, the Court has found that interference with a single key right, such as the right to exclude, amounts to a compensable taking.¹⁹³ However, in other cases the Court has insisted in viewing property as a bundle of rights, in which the loss of a single stick does not amount to a taking.

A copyright itself is likewise a “bundle of exclusive rights” established and governed by the Copyright Act.¹⁹⁴ Under the Act, the government's violation of those rights does not destroy them. Copyright in a digital good differs from rights in real property in that digital goods are intangible and nonrivalrous. Despite these differences, the Supreme Court of Texas in *Jim Olive Photography* found that because of its transferability and other property attributes, the “copyright owner thus retains the key legal rights that constitute property for

184. *Copyright Reform and the Takings Clause*, *supra* note 71, at 975.

185. *Id.*

186. *Id.*

187. *Id.*

188. See generally Davida H. Isaacs, *Not All Property is Created Equal: Why Modern Courts Resist Applying the Takings Clause to Patents and Why They Are Right to Do So*, 15 GEO. MASON L. REV. 1, 16–28 (2007) (explaining how intangible and nonrivalrous property such as patent receives limited protection compared to traditional forms of property).

189. *Copyright Reform and the Takings Clause*, *supra* note 71, at 978.

190. *Id.* at 978–79.

191. *Id.*

192. *Id.*

193. *Kaiser Aetna v. United States*, 444 U.S. 164, 179–80 (1979).

194. *Harper & Row Publishers, Inc. v. Nation Enters.*, 471 U.S. 539, 546–47 (1985).

purposes of a per se takings analysis.”¹⁹⁵ Although a copyright is a bundle of rights and not actually the thing itself,¹⁹⁶ the distinction between things and property is often of little consequence in the typical takings case. If the State seizes one’s automobile, it has also by definition interfered with one’s “property”: one’s right to possess, use that automobile, and exclude other from using the automobile. In Takings Clause claims involving intellectual property, however, the distinction between things and property rights becomes more important. Because the “thing” is intangible, an unauthorized user of that thing need not physically seize or take the property to appropriate one or more of the owner’s rights to the “thing” under the Copyright Act.¹⁹⁷

Prior to *Cedar Point Nursery*, Courts applied both regulatory and physical tests to determine whether copyrighted material has been “taken” by a state’s action because of the difficulty of distinguishing between the seizure of an intangible thing like a digital photography or a rather an interference of the exclusive rights to the underlying work protected by the copyright. After *Cedar Point Nursery*, plaintiffs in cases such as *Allen v. Cooper* and *Jim Olive Photography* have emulated the Supreme Court’s reasoning in *Cedar Point Nursery* and relied on its holding to bring copyright-takings claims where a state agent has appropriated any of their exclusive rights under the Copyright Act and enabled unauthorized third-party access to the copyrighted material.

C. THE VIABILITY OF THE COPYRIGHT TAKINGS CLAIMS IN *JIM OLIVE PHOTOGRAPHY* BEFORE AND AFTER *CEDAR POINT NURSERY*

A copyright owner whose work has been infringed by a state actor may potentially bring a suit alleging that “the infringement constitutes a taking of property without just compensation, in violation of the Fifth Amendment of the U.S. Constitution or a state constitution.”¹⁹⁸ The Court in *Florida Prepaid* first proposed a Takings Clause claim as a possible remedy for state intellectual property infringement. The Supreme Court has only ruled once on whether intellectual property, specifically trade secrets, can be “taken” by a governmental actor and are thus eligible for just compensation under the Takings Clause in *Ruckelshaus v. Monsanto Co.*¹⁹⁹

This theory has rarely been tested.²⁰⁰ When it has been, federal and state courts have ruled inconsistently, and scholars have debated whether copyrights

195. *Jim Olive Photography v. Univ. of Houston Sys.*, 624 S.W.3d 764, 771 (Tex. 2021).

196. *See Lightning Oil Co. v. Anadarko E&P Onshore, L.L.C.*, 520 S.W.3d 39, 48 (Tex. 2017) (describing property “as a bundle of rights, or a bundle of sticks”).

197. *Id.*

198. COPYRIGHT AND SOVEREIGN IMMUNITY, *supra* note 13, at 65.

199. *Monsanto*, 467 U.S. at 986.

200. COPYRIGHT AND SOVEREIGN IMMUNITY, *supra* note 13, at 66.

are property protectable under the Fifth and Fourteenth Amendments and whether or not temporary appropriation of non-physical property (such as a copyright) amounts to a taking by that state actor.²⁰¹ Neither of these questions are explicitly addressed in *Cedar Point Nursery*, which leaves them open for litigation and on-going debate. In the meantime, “the viability of such a claim remains uncertain.”²⁰²

For example, attorneys for Jim Olive Photography tested two variations of this legal theory under both federal and state Takings Clauses twice against UH. Olive’s original takings claims made their way to the Supreme Court of Texas, where the court ultimately found that Olive was not entitled to compensation because he retained some of his property rights despite the infringement—mainly the right to exclude further private parties from using the photographs, and because the infringement was temporary in that UH took down the infringing copies. Just days after the Texas court issued its opinion, the Supreme Court decided *Cedar Point Nursery v. Hassid*, which found that a property owner may bring a per se Takings Clause claim when the state appropriates only some of the owner’s right and for only a limited amount of time. Despite the ambiguity as to whether this holding is only meant to apply to physical takings, the Supreme Court denied Jim Olive Photography’s petition for writ of certiorari.²⁰³ As such, it remains unclear if the holding in *Cedar Point Nursery* would change the outcome in *Jim Olive Photography*.

1. Before Cedar Point Nursery

Olive’s case was much weaker before *Cedar Point Nursery*. On first appeal from district court, the Court of Appeals of Texas acknowledged that as of 2008, the question whether “copyright is property under the takings clause is ‘as of yet unlitigated.’”²⁰⁴ The court further relied on a student law review article arguing that the Supreme Court’s “definition of ‘property’ does not appear to shelter copyright” and that “copyrights exist only by the grace of the Constitution.”²⁰⁵ While Olive proposed that *Horne* extends protection to other

201. See, e.g., *Jim Olive Photography v. Univ. of Houston Sys.*, 624 S.W.3d 764, 774–77 (Tex. 2021) (assuming without deciding that copyright qualifies as property for the purposes of takings law); but see Tom W. Bell, *Copyright as Intellectual Property Privilege*, 58 SYRACUSE L. REV. 523, 539–40 (2008) (arguing copyrights exist only by the grace of the Constitution and are not property for Takings Clause purposes).

202. COPYRIGHT AND SOVEREIGN IMMUNITY, *supra* note 13, at 66.

203. *Jim Olive Photography*, 624 S.W.3d at 764.

204. *Univ. of Hous. Sys. v. Jim Olive Photography*, 580 S.W.3d 360, 361 (Tex. App. 2019).

205. *Id.* at 368 (citing *Copyright Reform and the Takings Clause*, *supra* note 71, despite acknowledging that the author ultimately argued in favor of Takings Clause protection for copyrighted material).

forms of property including personal property, the court refused to apply the Takings Clause to intangible intellectual property.²⁰⁶

Principally, Olive originally relied on reasoning from *Pruneyard Shopping Center v. Robins*, which defined “property” under the Takings Clause as “the group of rights inhering in the citizen’s relation to the physical thing, as the right to possess, use and dispose of it” and clarified that “[t]he constitutional provision is addressed to every sort of interest the citizen may possess.”²⁰⁷ This language was also relied upon in *Ruckelshaus* to equate intellectual property rights with the property rights protected by the Takings Clause.²⁰⁸ The district court found this citation unpersuasive because neither *Pruneyard* nor *Ruckelshaus* addressed the question of copyright infringement by a state actor.²⁰⁹

Olive also cited to *James v. Campbell*, which provides Supreme Court precedent for an IP-takings claim because it “purports to protect patents from a government-taking without just compensation.”²¹⁰ The Texas appellate court rejected this argument because it found that the Supreme Court has “never definitively held that a patent holder’s recourse against the government for infringement is a *constitutional* takings claim.”²¹¹

Finally, the Texas Appeals Court found that even if copyright were a form of protectable property, a Takings claim would still not succeed because what the university committed was copyright infringement, not a taking. As Olive did not lose his right to use or otherwise license his photograph, the court found that Olive had only lost a licensing fee, rather than the use or possession of his property.²¹² The court ultimately found this to be a case of “transitory common law trespass—a government interference with real property that may not amount to a taking at all.”²¹³

Appealing the case to the Texas Supreme Court yielded no better result for Olive. The Supreme Court of Texas acknowledged the ambiguity of copyright as property, and decided to assume for its analysis that a copyright is property and entitled to protection under the Fifth Amendment.²¹⁴ UH argued that even

206. *Id.* at 369.

207. *Id.* at 371 (citing *Pruneyard Shopping Ctr. v. Robins*, 447 U.S. 74, 82–84 (1980)).

208. *Id.* at 369–70.

209. *Id.*

210. *Id.* at 372–73.

211. *Id.* at 373.

212. *Id.* at 376.

213. *Id.*

214. *Jim Olive Photography v. Univ. of Houston Sys.*, 624 S.W.3d 764, 770 (Tex. 2021).

if copyright is property and is entitled to protection, the act of copyright infringement does not rise to a taking.²¹⁵

UH distinguished its infringement from a taking because the university did not “confiscate or appropriate those rights” in the photograph as necessary for a physical taking under *Loretto* and *Horne*.²¹⁶ Olive contended that UH’s infringement deprived Olive of the exclusive right to control his work, not just use of the photo itself.²¹⁷ Further, Olive argued that just as each raisin in *Horne* was Horne’s personal property, each reproduction and display of The Cityscape photograph is Olive’s.²¹⁸ The Supreme Court of Texas affirmed the lower court’s analysis that *Horne* does not apply because the case is silent on the question as to whether “state action may be asserted as a per se taking of an intellectual property right.”

Further, the court assumed that the reasoning in *Horne* only protects forms of tangible property, such as grapes, from state appropriation; however, this conflicts with the Supreme Court’s decision in *Monsanto*, which extends Takings Clause protection to intangible property such as trade secrets.²¹⁹ Finally, the court rejected Olive’s argument that UH’s infringement appropriated any strands within his property rights bundle. While Olive argued that “exclusivity” is the core component of each specific right granted under the Copyright Act, the court found that UH’s infringement does not “indicate the existence of a per se taking” because it does not necessarily destroy any of the rights to possess and use the work. Because Olive may still exclude third parties from using or displaying The Cityscape, the court found that the State had not appropriated these rights in Olive’s bundle of exclusive rights.²²⁰ Ultimately, Olive found no remedy for UH’s infringement within the Texas court system. However, the Supreme Court’s later decision in *Cedar Point Nursery* would extend the scope of the per se takings analysis to include state actions, like the University’s, which appropriate a property owner’s right to exclude third parties.

2. *After Cedar Point Nursery*

The majority opinion in *Cedar Point Nursery* opened up the potential that a state’s temporary use of copyrighted material for a “public good” may constitute a per se taking because it temporarily encroaches on the creator’s right to exclude third parties from accessing the work. Although courts may

215. *Id.* at 771.

216. *Id.*

217. *Id.*

218. *Id.*

219. *Id.* at 773.

220. *Id.*

find this to be a transitory appropriation and award just compensation, some scholars including Professor Michael McConnell interpret the case to dramatically reduce or permanently curtail the regulatory takings framework.²²¹ Dean Erwin Chemerinsky believes that this case will lead to much more litigation because it leaves many of these questions unanswered, but that it ultimately reflects a Supreme Court that is more protective of copyrights.²²²

Just days after the Texas court system failed to remedy the harm caused by the University's infringement, Olive revised his taking claims in light of the Supreme Court's holding in *Cedar Point Nursery* and petitioned for writ of certiorari. Because the Supreme Court did not grant the petition, it is unclear how it would have ruled. The Court's reluctance to accept the case may suggest unwillingness to overturn the Supreme Court of Texas' opinion.²²³

In the petition and accompanying support briefs, Olive argues that the "right to exclude" is the core property interest created by the Copyright Act and violated by acts of copyright infringement.²²⁴ Olive points to language from *Cedar Point* explaining that the right to exclude is one of the most treasured rights and "that sole and despotic dominion which one man claims and exercises over the external things of the world, in *total exclusion of the right of any other individual in the universe.*"²²⁵ Specifically, Olive claims that the Supreme Court of Texas ignored his right to exclude in light of *Cedar Point Nursery* because a taking under *Cedar Point* does not require the University to acquire legal title, and also does not require the complete destruction of the bundle of property rights, only the appropriation of the right to exclude.²²⁶ In support of this argument, Olive again cites *Cedar Point* for the rule that "even if the Government physically invades only an easement in property, it must

221. United States Courts, *Term Talk (2020-2021): Cedar Point Nursery v. Hassid*, YOUTUBE (Oct. 28, 2021), <https://www.youtube.com/watch?v=L6uIHgrKorw&t=577s>. (discussing at timestamp 9:25 how, after *Cedar Point Nursery*, "regulation-goes-too-far-type claims are . . . on the road to oblivion").

222. *Id.* (explaining at timestamp 9:47 that *Cedar Point Nursery* reflects that a "conservative majority [in the Supreme Court] wants much more protection of intellectual property" and leaves open the question of what comprises just compensation for temporary appropriations, which will lead to increased litigation).

223. See Mark Brodt, *What the Constitution Giveth, Texas May Taketh Away*, FORDHAM INTELL. PROP. MEDIA & ENT. L.J. (May 10, 2022), <http://www.fordhamiplj.org/2022/05/10/what-the-constitution-giveth-texas-may-taketh-away/> (arguing while it is unclear whether Jim Olive would prevail on a per se taking claim, "with the Supreme Court denying certiorari at what, presumably, was the most opportune time for such a case to be heard, states remain free and protected for the foreseeable future").

224. Petition for Writ of Certiorari at 18–21, *Jim Olive Photography v. Univ. of Hous. Sys.*, 142 S. Ct. 1361 (2022) (No. 21-735).

225. *Id.* at 20.

226. *Id.*

nonetheless pay just compensation.”²²⁷ In sum, Olive contends that copyright infringement constitutes a compensable easement onto its property because the state has interfered with his right to exclude.

The University in response contends that infringement is more like a trespass which is a tort, not a taking. The University alleges that *Cedar Point Nursery* actually bolsters the Supreme Court of Texas’ analysis because it reaffirms the trespass-takings distinction.²²⁸ Further, it argues that because Olive retains the right to exclude third parties, he retains all of his rights under the Copyright Act as they apply to third parties.

As the Supreme Court chose not to take the appeal, it is uncertain whether it would have ruled in favor of creators like Jim Olive, or state actors like the University of Houston. In part, the Court’s decision would likely turn on whether *Cedar Point* extends to digital appropriations of nonrivalrous goods. Even if its holding extends to intellectual property, creators will still need to overcome state sovereign immunity before they can recover on Takings Clause claims. Ultimately, although the Supreme Court will not shed light on the murky questions raised by *Jim Olive Photography*, Olive’s plight at least adds to the growing record of state actors’ intentional infringement that could allow Congress to abrogate sovereign immunity for copyright infringement directly rather than shape Takings Clause jurisprudence to address what it is still, at its heart, the tort of copyright infringement.

IV. SOVEREIGN IMMUNITY AND COPYRIGHT TAKINGS

If state sovereign immunity can preclude a Takings Clause claim, then how can any Takings Claim be brought against a state actor? So far, the Supreme Court has never granted sovereign immunity in a Takings Clause case against a state actor.²²⁹ Further, the Court has questioned whether “sovereign immunity retains its vitality” against the basic policy rationale of the Takings Clause.²³⁰ However, private citizens have historically been limited in their ability to sue their own state governments for Takings Clause claims because states rarely consent to suit arising from their citizens, and Congress has not abrogated state sovereign immunity for state Takings Clause claims.²³¹ Some argue that the federal Takings Clause, in absence of other adequate remedies,

227. *Id.* at 26.

228. *Id.* at 15.

229. J.P. Burleigh, *Can State Governments Claim Sovereign Immunity in Takings Cases?*, U. CINCINNATI L. REV. (Jan. 15, 2020), <https://uclawreview.org/2020/01/15/can-state-governments-claim-sovereign-immunity-in-takings-cases/>.

230. *City of Monterey v. Del Monte Dunes at Monterey, Ltd.*, 526 U.S. 687, 714 (1999).

231. *Hutto v. S.C. Ret. Sys.*, 773 F.3d 536, 551–52 (4th Cir. 2014).

exists to force non-consenting states into their state courts to hold them accountable for violations of private property rights.²³² However, according to the Court's interpretation in *Alden v. Maine*, the Eleventh Amendment protects states from suit by private parties in its own courts without its consent even when the suit derives from federal law—such as an infringement of the Federal Copyright Act or a violation of the federal Takings Clause.²³³

However, there is ample academic support for holding the states accountable. Eric Berger explains that the Takings Clause should be interpreted as self-executing because of its unique nature and intent to hold states accountable which “naturally supersedes” the Eleventh Amendment removing the state's grant of immunity.²³⁴ However, Berger speculates that in temporary takings claims the Court may allow the state to retain its immunity to retain the health of federalism.²³⁵ Professor Richard H. Seamon further clarifies that although states can claim sovereign immunity against copyright infringement cases brought in federal court, due process concerns may force non-consenting states into just compensation suits in state courts.²³⁶ This means that if a state fails to create an adequate remedy, the Due Process Clause of the Fourteenth Amendment requires that states hear just compensation cases within their courts regardless of any claims of sovereign immunity.²³⁷

In another copyright-takings case involving a Texas university, the Fifth Circuit Court of Appeals held that states are entitled to sovereign immunity from federal takings claims in federal court, and from state takings claims in both federal and state courts if not waived.²³⁸ In *Canada Hockey L.L.C. v. Texas A&M University Athletic Department*, Michael Bynum, a sportswriter and editor, sued Texas A&M University (TAMU) after it published a key part of Bynum's forthcoming book without permission and retyping the byline of the portion to indicate university sponsorship.²³⁹ The Fifth Circuit upheld the lower court's ruling that TAMU's federal sovereign immunity cannot be abrogated from either the federal or state takings claims, which were pleaded in the alternative to the copyright infringement claims.²⁴⁰

232. Richard H. Seamon, *The Asymmetry of State Sovereign Immunity*, 76 WASH. L. REV. 1067–69 (2001).

233. *Alden v. Maine*, 527 U.S. 706, 712 (1999).

234. *Id.*

235. *Id.* at 601.

236. *See* Seamon, *supra* note 232, at 1069.

237. *Id.*

238. *Can. Hockey, L.L.C. v. Tex. A&M Univ. Ath. Dep't*, U.S. App. LEXIS 3976, at *25–*26 (5th Cir. Feb. 14, 2022).

239. *Id.* at *2–*6.

240. *Id.* at *25–*26.

From a policy perspective, the juxtaposition of these two competing doctrines is about balancing the desire to protect the public treasury from the cost of countless lawsuits and violations, and the desire to protect private citizens from their own states' theft. If the Supreme Court continues to not provide guidance on how to balance these interests, then courts may adopt the Fifth Circuit's rationale and prevent copyright owners from recovering on any copyright-takings claims. As such, the Takings Clause does not currently provide copyright owners with just compensation when states violate private citizen's rights under the Copyright Act.

V. ADEQUATE ALTERNATIVES AND PATHWAY TO CONGRESSIONAL ABROGATION

After *Jim Olive Photography*, few pathways remain for just compensation for creators whose copyright have been infringed by states, even intentionally and repeatedly. While attorneys for Frederick Allen, Jim Olive, and Michael Bynum have argued that the Takings Clause may provide a pathway for just compensation, this is unlikely even after the expansion of the doctrine in *Cedar Point Nursery* as the Supreme Court declined to overturn both the Texas Supreme Court's decision in *Jim Olive Photography* and the Fifth Circuit's decision in *Canada Hockey L.L.C.* Other alternative pathways for compensation are largely inadequate. If all alternative pathways for compensation including takings claims are found to be inadequate to remedy creators for widespread and persistent, intentional, or at least reckless infringement by the states, the Court should uphold the constitutionality of a revised Copyright Remedy Clarification Act that is narrowly tailored to curb the worsening pattern of state infringement sparked after the *Florida Prepaid* cases.

A. ADEQUACY OF ALTERNATIVES

This Section lists alternative pathways for just compensation and assesses their adequacy in remedying the constitutional harms state infringement has inflicted. Beyond takings claims, which Part III has discussed, alternative pathways include: breach of contract, injunctive relief under *Ex parte Young*, personal-capacity suits against state officials, and waiver.

1. Breach of Contract Claims

Breach of contract claims may arise where the creator and infringer have an existing contract or an implied-in-fact contract.²⁴¹ However, some courts “will reject contract claims that are at their core about copyright violation[s]”

241. COPYRIGHT AND SOVEREIGN IMMUNITY, *supra* note 13, at 64.

and find instead that they should be preempted by the Copyright Act's express preemption provision.²⁴² Because breach of contract claims add an extra element, the existence of a contract, to the copyright infringement analysis, these claims are less likely to be preempted.²⁴³ The state cannot claim its sovereign immunity when entering into contracts because it "bind[s] itself like any other party to the terms of the agreement."²⁴⁴

Breach of contract claims are largely inadequate because they are not available for the creators who have not entered into a contract with the state infringer. This is particularly pertinent as the "vast majority" of infringements do not involve a contractual relationship as the state entity "may obtain copies from sources such as Google [I]mages, social media, websites," and exploit those copies "without the creators' knowledge."²⁴⁵ Additionally, these claims are disfavored because they add an additional element that a plaintiff must prove in order to properly allege copyright infringement.²⁴⁶ Finally, even if a copyright owner prevails on a breach of contract claim, they are not eligible for the statutory remedies under the Copyright Act.²⁴⁷

2. *Ex parte Young Claims*

Under the doctrine of *Ex parte Young*, creators may seek injunctive relief that prevents the state's further use of the copyrighted work.²⁴⁸ States have long argued that this protection is enough to justify state sovereign immunity. In support of their argument, states cite to a survey conducted by the Copyright Alliance which indicates that fifty percent of respondents would be willing to accept injunctive relief alone.²⁴⁹ Further, they argue that injunctive relief proceedings allow creators to prove infringement on the merits and obtain an injunction.²⁵⁰ Lastly, they allege that the high monetary cost of defending against injunctive claims and the negative publicity that such suits may spark is a powerful enough deterrent to keep states from committing particularly egregious forms of infringement.²⁵¹

242. *Id.*

243. *See id.* (noting that the Ninth Circuit, in *Ryan v. Editions Ltd.*, 786 F.3d 754, 761 (9th Cir. 2015), recognized that a contractual-based claim contains the extra element necessary to distance the claim from the express preemption provision of the Copyright Act).

244. *Tooke v. City of Mexia*, 197 S.W.3d 325, 332 (Tex. 2006).

245. *Id.*

246. *Id.*

247. *Id.* at 81.

248. *See Ex parte Young*, 209 U.S. 123, 166 (1908).

249. COPYRIGHT AND SOVEREIGN IMMUNITY, *supra* note 13, at 67.

250. *Id.*

251. *See id.* (finding many commenters citing the monetary cost of defending against these suits and the adverse publicity that could result from these claims to be "powerful deterrents").

Nonetheless, injunctive relief is inadequate because it only prevents future infringement and does nothing to remedy damage done from past infringement.²⁵² This is particularly inadequate when the infringement depletes the expressive work of all its economic value, like in the case of *Allen v. Cooper*.

3. *Personal-Capacity Suits Against State Officials*

Another alternative remedy is the ability to bring a copyright infringement suit against a state official in their personal capacity in federal court. The Ohio Attorney General's Office described these suits as a "pretty easy workaround" that enables recovery against the state, and that the state can indemnify the named defendant for any resulting damages if it chooses to.²⁵³

This pathway is also inadequate for three reasons: (1) if the state does not indemnify, "the copyright owner's ability to recover damages may be limited by the individual official's ability to satisfy a judgment";²⁵⁴ (2) often it will be difficult or impossible for the copyright owner to uncover the identity of the individual that committed the infringement as required by the suit;²⁵⁵ (3) the individual state official, if identified, may be protected from liability by qualified immunity which shields state actors if "their conduct does not violate clearly established statutory or constitutional rights of which a reasonable person would have known."²⁵⁶ In many cases, copyright owners have been unable to prove that the infringing individuals should have known they were committing copyright infringement due to the fair use defense.²⁵⁷

4. *Waiver*

Finally, states may choose to waive their sovereign immunity for copyright infringement suits.²⁵⁸ Most states have waived their sovereign immunity to state-law claims such as tort actions or contractual violations; however, the procedural and substantive requirements for bringing these actions vary. For example, three states have state constitutional protection from becoming a defendant in its courts.²⁵⁹ Currently, thirty-seven states waive sovereign immunity for torts like copyright infringement, and fifteen allow plaintiffs to

252. *Id.* at 67.

253. *See id.* at 69.

254. *Id.*

255. *Id.*

256. *Id.*

257. *See id.* at 69–70 (determining that several copyright owners have been unable to approve that a state actor should have known that their actions infringed a valid copyright because of the "unsettled nature of the legal issue involved").

258. *Id.* at 60.

259. *See id.* (referencing the Constitutions of Alaska, Arkansas, and West Virginia).

bring contract claims.²⁶⁰ When states choose to do so, waiver is adequate because it allows the consensual abrogation of sovereign immunity and provides a pathway for creators to be made whole following a copyright infringement suit. However, this pathway remains inaccessible for many copyright holders across the country living and creating in states that have not waived immunity.

B. PATHWAY TO CONGRESSIONAL ABROGATION: RECOMMENDATIONS FROM THE COPYRIGHT OFFICE REPORT

After *Allen v. Cooper*, congressional leaders requested that the Copyright Office conduct a study to “determine whether there is sufficient basis for federal legislation abrogating State sovereign immunity when States infringe copyrights.”²⁶¹ Specifically, the Office studied (1) “the extent to which copyright owners are experiencing infringements by state entities without adequate remedies under state law” and (2) “the extent to which such infringements appear to be based on intentional or reckless conduct.”²⁶²

In response to these inquiries, the Office reported that the number of allegations of state infringement has substantially increased since the release of the Oman Report, which the CRCA was passed on.²⁶³ Infringement has increased since the *Florida Prepaid* cases began casting doubt on the validity of the CRCA.²⁶⁴ Additionally, the majority of respondents indicated that their works have been infringed repeatedly by state actors, which establishes a pattern of infringement.²⁶⁵

The Office contended that “state infringement constitutes a legitimate concern for copyright owners.”²⁶⁶ Largely, copyright owners are left with no adequate remedies. The Office illustrated how recent cases “cast doubt on the viability of claims seeking to recover under a takings theory.”²⁶⁷ While some creators may be able to bring a breach of contract claim, these claims may be preempted by the Federal Copyright Act, and even if successful, they do not

260. *See id.* (finding that thirty-seven states waive immunity for tort actions against state officials and that fifteen permit contract claims).

261. *Id.* at 70.

262. *Id.*

263. *See id.* at 71–72 (determining that over 130 copyright infringement suits against state entities were brought between 2000 and 2020, compared with the “no more than ‘half a dozen’ examples” that Congress had gathered to support passing the CRCA).

264. *See id.* (finding the instances of documented infringement increased “substantially” between 2000 when the Court decided the *Florida Prepaid* cases and 2020 when the Report was released).

265. *Id.* at 71.

266. *Id.*

267. *Id.*

provide the same remedies as copyright infringement.²⁶⁸ Injunctive relief only prevents future harm, and does not remedy previous harm even if state action depletes the copyrighted work of all economic value.²⁶⁹ Lastly, suits against individuals are inadequate where qualified immunity or a lack of funding prevents economic relief.²⁷⁰

Additionally, the Office contemplated that evidence of intentional or reckless state infringement may be “elusive.”²⁷¹ The state sovereign immunity defense may deter copyright owners from suing, which makes it difficult to establish a conclusive record of unconstitutional conduct.²⁷²

The Office explained that the evidentiary standard that the Court is seeking to allow abrogation is unclear.²⁷³ Since the Court articulated its “congruence and proportionality test” in *City of Boerne* over two decades ago, only two cases based on discrimination claims have upheld Congress’ ability to abrogate immunity.²⁷⁴ While the Court had articulated that the record established in the Oman Report was insufficient, it has not issued guidance on the “nature and volume of evidence that would be sufficient in this area”—although it appears that the standard is quite high.²⁷⁵ As such, the Office was ultimately “unable to conclude with certainty that the evidence provided in this study would be held sufficient to establish a pattern of unconstitutional conduct.”²⁷⁶ Nonetheless, the Office believed that, equipped with the heightened record and the growing number of plaintiffs pursuing copyright-takings claims, Congress may still proceed with proposing new abrogation legislation.²⁷⁷ The Copyright Office and numerous amicus organizations continue to “believe that infringement by state entities is an issue worthy of congressional action.”²⁷⁸

State universities and libraries, which currently benefit from immunity, contend that they adopt rigorous policies and educational programs to alert

268. *Id.* at 64.

269. *Id.* at 67–68.

270. *See id.*

271. *Id.*

272. *See id.* at 71–72 (discussing the difficulty of compiling conclusive evidence of intentional or reckless infringement because the presence of sovereign immunity claims either dissuades potential plaintiffs from bringing suit or may cause the suit to not be adjudicated).

273. *See id.* at 72 (explaining that the standard to which the evidence is to be weighed after *City of Boerne* is unclear but appears to be “set quite high.”).

274. *See id.* (discussing that the only two cases to uphold congressional abrogation in response to either disability- or sex-based discrimination).

275. *See id.*

276. *See id.* at 73.

277. *See id.*

278. *Id.*

their staff about potential copyright infringements.²⁷⁹ Predominantly, these states entities are concerned about the impact on their finances if they must defend themselves in infringement suits.²⁸⁰ However, these entities may generally be about to invoke protection under existing protections and limitations of the Copyright Act, such as “fair use, exceptions for reproduction by libraries and archives, and limitations on remedies” and do not require sovereign immunity to protect against non-meritorious copyright infringement claims.²⁸¹

Congress should choose to pass legislation to ensure that creators have adequate remedies to protect their work against other state entities such as athletic departments and radio stations which may currently make use of copyrighted works to commercial purposes that affect the works’ marketability. If Congress chooses to abrogate, it must narrowly tailor its legislation to the extent that copyright owners are experiencing widespread, persistent, and intentional—or at least reckless—infringement by state actors that would subject private parties to liability. However, even if Congress passes legislation that seeks to abrogate state sovereign immunity, the Court may still find that the record established by the Copyright Office is insufficient to support abrogation.²⁸²

If Congress chooses not to pursue full abrogation because of the lack of clarity from the Court, it should consider a waiver-based framework proposed by previous Congresses.²⁸³ Under this framework, “a state’s ability to recover damages for infringement of its own intellectual property rights would be conditioned on its waiving sovereign immunity from infringement suits.”²⁸⁴

VI. CONCLUSION

Because the Copyright Office’s report on Copyright and State Sovereign Immunity greatly expanded the record of state infringement and established a pattern of increased infringement following the fall of the CRCA, Congress should pass a revised, narrowly tailored version of the CRCA that holds state actors liable from intentionally or recklessly engaging in the same conduct that

279. *See id.* at 2 (recognizing that many state entities have taken “significant steps to ensure respect for copyright including implementing policies, procedures, and social norms).

280. *Id.* at 67.

281. *Id.* at 74.

282. *See id.* at 73 (concluding that the Copyright Office is unable to conclude with any certainty that the Court would uphold any legislation abrogating sovereign immunity under the current record).

283. *See id.* at 74.

284. *Id.*

would subject private parties to liability under the Copyright Act. This bill should create an explicit safe harbor for public schools, libraries, and museums that implement strenuous copyright trainings and policies to prevent infringement. This approach is narrowly tailored to prevent further harm to creators that generate art and culture without placing an undue burden on state institutions that curate and disseminate culture and knowledge. Although it is untested whether the present record would be sufficient enough to support abrogation, at the very least, passing this legislation would likely prompt the Supreme Court to provide clearer guidance on Copyright Takings Claims following *Cedar Point Nursery* and *Jim Olive Photography*.

RESOLVING COPYRIGHT'S DISTORTIONARY EFFECTS

Akshat Agrawal[†]

TABLE OF CONTENTS

I.	INTRODUCTION	1208
II.	IN QUEST FOR CULTURAL DIVERSITY.....	1212
	A. GOALS OF CULTURAL POLICY	1212
	B. ROLE OF COPYRIGHT IN CULTURAL POLICY.....	1214
	C. COPYRIGHT'S TOOLS.....	1217
	1. <i>Rights to Exclude</i>	1217
	2. <i>Pitfalls of Expanding the Right to Exclude</i>	1218
III.	SCOPE OF RIGHTS.....	1219
	A. RIGHT TO EXCLUDE MAKING OF A DERIVATIVE WORK.....	1219
	1. <i>Contours</i>	1219
	2. <i>Get a License or Do Not Sample!</i>	1220
	3. <i>Is De Minimis Helpful? Dissonance with Compositional Logic</i>	1223
	B. THE RIGHT TO EXCLUDE REPRODUCTIONS	1224
	1. <i>Contours</i>	1224
	2. <i>Protection of Fragments: Joyful Noise?</i>	1226
IV.	INDIAN CLASSICAL MUSIC—EXPRESSION OR THIEVERY? 1228	
	A. THE RAGA SYSTEM: RULES OF COMPOSITION.....	1228
	B. IMPLICATIONS OF THESE RULES ON PHRASING.....	1231
	C. DISSONANCE WITH THE SCOPE OF RIGHTS.....	1233

DOI: <https://doi.org/10.15779/Z38DR2P978>

© 2023 Akshat Agrawal.

† LL.M., Class of 2023, University of California, Berkeley, School of Law. I am enormously grateful to Profs. Talha Syed and Oren Bracha for very deep engagement with this piece in all its phases. Their inputs have substantially shaped my thoughts. I am also thankful to Prof. Molly V. Houweling, Mark Lemley, Madhavi Sunder, Carys Craig, Glynn Lunney, Zahr Said, O. Arewa, Anjali Vats and the participants of the Works in Progress Intellectual Property Scholars Conference 2023 as well as Race+IP '23 conference where I had the opportunity to present drafts of this Note. I would also like to thank Swaraj Barooah and Lokesh Vyas for playing a substantial role in all thoughts that led to this piece. Finally, without the enormous support of Shih-Wei Chao and members of the Law and Technology Writing Workshop, this Note would not have been possible. This Note won Google's Global Inclusive Copyright Thesis Award, 2023.

V.	DISSONANCE LEADS TO DISTORTION	1235
A.	DISTORTIONARY EFFECTS	1235
B.	RELEVANCE IN THE GLOBAL POLITICAL ECONOMY	1239
VI.	POLICY SIGNIFICANCE AND PRESCRIPTIONS.....	1241
A.	RESIST EXPANSIONARY IMPULSE.....	1241
B.	STRUCTURALLY SCALE BACK EXCLUSIONARY RIGHTS.....	1241
1.	<i>Limiting Derivative Markets to Alternate “Forms” of Representation</i>	1243
2.	<i>De-Fragmenting the Work and Its Primary Market</i>	1245
VII.	CONCLUSION.....	1248

I. INTRODUCTION

“Culture is a source of the self.”¹ But does copyright policy embrace cultural difference? This Note argues that copyright law entrenches investment bias in favour of genres of cultural expression that are easier and less costly to legally exclude and appropriate returns from, and freezes out genres that are not.

Legal excludability (“excludability”)² refers to the ability to exclude others from using and accessing copyrighted works without authorization. It is similar to saying—*Do not enter!* Appropriability refers to the ability to use or leverage legal excludability to seek payment or license fees for access and use, and attain economic returns preventing free consumption, reproduction, and dissemination. It is similar to saying—*You can only enter if you pay up!* Cultural expressions stemming from compositional practice that is ontologically derivative or dialogic, or that inherently involves perceivable similarity with previous works because of cultural norms, are (1) *either* costlier to produce due to licensing costs, *or* (2) have relatively lower potential of appropriability. Thus, such expressions potentially lose out on effective market circulation.

In this piece, I analyze compositional norms prevalent in Indian Classical Musical practice, and its *Raga* system, to show how multiple compositions in this genre inherently: (1) involve desirable similarity that is easily perceivable for listeners and the performer; (2) follow strictly defined rules of phrasing, sequencing, and performing compositions in a particular *Raga*; and (3)

1. Jack M. Balkin, *Digital Speech and Democratic Culture: A Theory of Freedom of Expression for the Information Society*, 79 N.Y.U. L. REV. 1, 33 (2004).

2. Throughout this Note, the use of “excludability” refers only to the ability to exclude using exclusionary/property rights that are statutorily granted. This is not to be confused with the ability to exclude others using technology.

voluminously incorporate pre-existing expression as critical to any composition in a particular *Raga*. These characteristic traits do not curtail the ability of producing limitless compositions within the genre, or even within the same *Raga*. However, they do make substantial similarity in multiple expressions, as is currently understood by courts, inevitable. Thus, at the level of copyright protection that currently persists, these expressions, due to their cultural traits, either involve high licensing costs (in case copyright law continues to protect all elements of the works in spite of the *inherence* of similarity) or high amounts of *scenes a faire* elements (in case copyright law renders motivic phrases, arpeggios, sequences and intonations, that are essential to be followed in a *Raga*, as *scenes a faire*) that are important to any and every composition in a genre. They are thus *either* costlier to produce, or relatively less excludable and appropriable than expressions from other genres of music. It is a lose-lose.

Current copyright policy, therefore, asymmetrically shapes the extent to which one can internalize market demand—in favour of highly excludable expressions,³ which estranges cultural expressions from genres like Indian classical music that are relatively dissonant and less legally excludable. Due to its sole focus on producing efficiency based on internalizing market value, copyright, in its present scope, is biased against works that have a lower ability to internalize market demand due to their potential of excludability and appropriability but may have high normative and cultural value in fostering self-determination.

These are copyright's *distortionary effects* that arise out of copyright's specific scope of exclusionary rights.⁴ They draw hierarchies in cultural aspirations and experience, while contributing to conformity in cultural practice.

As a solution, this Note advocates to structurally limit the scope of copyright's exclusionary rights across genres of cultural composition, by (1) limiting the overall scope of the derivative right to only cover adaptations in different mediums of representation,⁵ and (2) curtailing the scope of the reproduction right (a) to only protect the work *as a whole*, as against its

3. See generally Amy Kapczynski & Talha Syed, *The Continuum of Excludability and the Limits of Patents*, 122 YALE L.J. 1900 (2013) (making a similar case in context of checklist interventions and natural medicines that might have higher net social benefits, but are estranged from circulation as they are relatively less excludable and appropriable using patent rights).

4. Glyn S. Lunney Jr., *Reexamining Copyright's Incentives-Access Paradigm*, 49 VAND. L. REV. 483, 582 (1996).

5. See generally Talha Syed & Oren Bracha, *Copyright Rebooted*, Presentation at the 2022 Stanford University Law School Intellectual Property Scholars Conference (Aug. 12, 2022) (unpublished manuscript on file with author) (proposing a similar prescription, although justified by other reasons).

fragments or elements, and (b) to limit infringement only to when there is potential of market substitutability of the *overall* aesthetic experience of the primary work in the minds of perceivers.⁶ I argue that lack of an overly broad derivative right—that includes expressions in the same form, medium of representation, and market as the primary work—would eradicate the overt licensing cost involved in cultural compositions that inherently require re-mixing or use of pre-existing expression. Further, cutting down the scope of an overly broad reproduction right, which currently even protects fragments of expressions, would eradicate any relative lack of appropriability that exists due to a high volume of *scenes a faire* elements, and will protect the primary market of the work equitably if there is potential of aesthetic substitution of the work.

These structural changes, applicable across all genres of cultural expression, may reduce the potential to appropriate the highest possible economic value from a single highly excludable expression. However, in parallel, they enlarge the cultural breadth or diversity of expressions that would potentially be invested upon. These changes push towards an egalitarian position or starting point for a cultural speaker irrespective of the kind of cultural composition that they practice.

Part II of this Note explores the role of copyright policy and its tools in *enabling* production of diverse cultural expressions. Section II.A expands on the goals of cultural policy, namely, to allow diverse participation and exposure to expressions for autonomous, yet social, self-determination. Section II.B sketches the role of copyright law as a policy tool in fulfilling these goals. Section II.C, first, highlights the tools that copyright policy employs towards its instrumental purpose—the rights to exclude. Second, it critiques the myopic focus of policy on broad exclusionary rights and social value appropriation, highlighting its pitfalls: the *price tag* effect, *privilege expanding* effect, and finally, the focus of this Note, its *distortionary* effect.

Part III illuminates the doctrine and scope of exclusionary rights that produce these pitfalls—(1) the right to exclude the making of a derivative work and (2) the right to exclude reproductions. Section III.A first explores the

6. *Id.*; see also Talha Syed & Oren Bracha, *Copyright's Atom: The Expressive Work as the Basic Unit of Analysis*, Presentation at Philosophical Methods in IP Colloquium (June 2021) (unpublished manuscript on file with author) (providing a similar prescription to structurally scale down the scope of work in copyright infringement analysis, although justified by other reasons); Carys J. Craig, *Transforming "Total Concept and Feel": Dialogic Creativity and Copyright's Substantial Similarity Doctrine*, 38 CARDOZO ARTS & ENT. L.J. (forthcoming), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3691280 (laying out a novel three-step test focusing on a holistic analysis to identify substantial differences, dissective analysis, and comparison).

contours of the derivative right, and then analyses its application to music sampling cases, where courts have effectively deemed recognisable digital sampling to be infringement unless one licenses the samples used—which adds to costs of production. Section III.B first traces the contours of the reproduction right and its expansion over time. It further analyses the case of *Gray v. Hudson*, where the Ninth Circuit Court of Appeals extended protection of a work to its fragments and held that even similarity of elements of musical works are infringing, unless these elements are *scenes a faire*.⁷ It thus created a distinction in the appropriability of cultural expressions that include lesser or no *scenes a faire* elements as against those which include more—by extending protection to not just the whole work but also its fragments or elements.

Part IV shows how, in the current state of copyright law, every expression in Indian classical music will *either* be an infringement of the derivative right and the reproduction right, *or* would have an enormous volume of *scenes a faire* elements which would render a composition inherently less excludable or appropriable. Section IV.A expands on the *Raga* system followed by Indian classical music, showing its emphasis on strict rules of composition. Section IV.B shows implications of these rules and the inherence of perceivable similarity in expressions in the same *Raga*. Section IV.C lays out how these cultural norms run dissonant with current copyright policy.

Part V shows how this dissonance of law with cultural practice leads to distortion of investment decisions away from these works and is complicit to freezing their visibility in global cultural markets. Section V.A argues that such a system of ordering creation and dissemination of cultural expressions distorts investments away from expressions that, due to copyright's scope, enable relatively lesser internalization of market value, are costlier to produce and are less excludable and appropriable. Section V.B shows the impact of these distortionary effects on global cultural practice and cultural dissemination considering global enforcement of minimum copyright standards. I contextualize the relevance of the case study on Indian classical music to argue in favour of revamping U.S. copyright law by taking lessons for inherently derivative art forms like music sampling that are predominantly practiced in contemporary American culture.

Finally, Part VI lays down preliminary prescriptions to resolve these distortionary effects by fine-tuning the scope of rights. Section VI.A addresses the need to resist expansion of rights as a solution, as it further weakens and ignores cultural norms and practices. Section VI.B elaborates on the structural limits that I propose regarding the overall scope of the derivative and

7. 28 F.4th 87, 97 (9th Cir. 2022).

reproduction rights. These structural limits resolve copyright's distortionary effects, and further copyright's goal of enabling *performers* of diverse cultural diversity.

II. IN QUEST FOR CULTURAL DIVERSITY

Culture plays a significant role in shaping us as political beings. It constitutes the interactive processes that facilitate “forg[ing], communicat[ing], enact[ing], interpret[ing], adapt[ing], challeng[ing], revis[ing] and recomb[ing]” meanings.⁸ Copyright policy is globally supposed to play an important role in enabling diverse cultural production. Towards this end, it provides exclusionary rights to producers. These exclusionary rights, however, come with significant costs that antithetically constrain cultural diversity. This Part expands on copyright's purposive end and drawbacks of its dependence on broad exclusionary rights.

A. GOALS OF CULTURAL POLICY

Cultural interactions are shaped by “expressions,” which allow participation and exposure to a variety of narratives and meanings.⁹ These expressions contribute to underwriting the meaning of creativity.¹⁰ They are agents of participating in social interactions. Such agency is not merely a medium of self-determination¹¹ and free choice but is its essential precondition.¹²

Law and Culture scholars postulated cultural interaction to be representative of *semiotic* democracy,¹³ something equally essential as, if not

8. Oren Bracha & Talha Syed, *Beyond Efficiency: Consequence-Sensitive Theories of Copyright*, 29 BERKELEY TECH. L.J. 229, 255 (2014).

9. HENRY JENKINS, KATIE CLINTON, RAVI PURUSHOTMA, ALICE J. ROBISON & MARGARET WEIGEL, CONFRONTING THE CHALLENGES OF PARTICIPATORY CULTURE: MEDIA EDUCATION FOR THE 21ST CENTURY 5–6 (2009), https://www.macfound.org/media/article_pdfs/jenkins_white_paper.pdf.

10. Yong Shao, Chenchen Zhang, Jing Zhou, Ting Gu & Yuan Yuan, *How Does Culture Shape Creativity? A Mini-Review*, 10 FRONTIERS PSYCHOL. 1219, 1221 (2019).

11. Bracha & Syed, *supra* note 8, at 251, 252 n.64.

12. *See id.* at 252 n.64.

13. *See* JOHN FISKE, TELEVISION CULTURE 236–39 (2d ed. 1987). The concept of “semiotic democracy” originates in the writings of John Fiske; in context of copyright scholarship, see Madhavi Sunder & Anupam Chander, *Everyone's a Superhero: A Cultural Theory of “Mary Sue” Fan Fiction as Fair Use*, 95 CALIF. L. REV. 597, 624 (2007); Madhavi Sunder, *IP*³, 59 STAN. L. REV. 257, 279–80 (2006); William Fisher, *Theories of Intellectual Property*, in NEW ESSAYS IN THE LEGAL AND POLITICAL THEORY OF PROPERTY 23 (Stephen R. Munzer ed., 2001); William Fisher, *When Should We Permit Differential Pricing of Information?*, 55 UCLA L. REV. 1, 34 (2007).

more essential than, political democracy.¹⁴ It reflects a commitment to decentralized meaning-making by ensuring an effective opportunity to participate in the shaping of social subjectivity and norms.¹⁵ People tend to internalize the narratives they are exposed to.¹⁶ Our preferences, habits, and thoughts are often shaped by the kind of cultural communication and exchange that we are exposed to.¹⁷ By participating in cultural exchange, we not only absorb or shape the culture around us but are also intrinsically shaped by our exposure.¹⁸

Participation is a source of voice and perspective. It directly and proportionally affects “choice” in forming collective will.¹⁹ For “choice” to be free, moreso than quantity of participants, meaningful diversity is necessary.²⁰ Being exposed to diverse lifestyles and cultural expressions significantly enables autonomous self-determination because it is informed by a more diverse and meaningful *range* of options.

The diversity I refer to here is twofold: (1) diverse participation and (2) diverse exposure through access. A diverse expressive environment free from control and manipulation offers opportunities for critical reflection and an arena of meaningful self-determination.²¹ However, systemic control over expression—that individuals can visibly access or use for downstream creation—tramples on autonomy of individuals and communities. Control, in the hands of privately coordinated entities, working in profit-enhancing bubbles, constrains and manipulates the process of forming cultural preferences.²² It could potentially curtail minority speech that private actors deem less or not profitable.²³

14. Sunder, *IP*³, *supra* note 13, at 325.

15. Bracha & Syed, *supra* note 8, at 256.

16. See Tania Zittoun & Alex Gillespie, *Internalization: How Culture Becomes the Mind*, 21 *CULTURE & PSYCHOL.* 477, 484 (2015).

17. See Balkin, *supra* note 1, at 36 (arguing that the “various processes of communication and cultural exchange are the sources of the self and its development over time,” and that we produce our ideas, habits, thoughts and selves through communication).

18. Bracha & Syed, *supra* note 8, at 254–55.

19. *Id.* at 253, 262.

20. Fisher, *Differential Pricing*, *supra* note 13, at 34 (citing JOHN STUART MILL, *PRINCIPLES OF POLITICAL ECONOMY* 252–57 (5th ed. 1909); see also William Fisher, *Reconstructing the Fair Use Doctrine*, 101 *HARV. L. REV.* 1659, 1751–52, 1772 (1988) (harping on the importance of cultural diversity and the role of meaningful choice in accessing cultural expressions).

21. Bracha & Syed, *supra* note 8, at 252.

22. ASTRA TAYLOR, *THE PEOPLE'S PLATFORM: TAKING BACK POWER AND CULTURE IN THE DIGITAL AGE* 186 (2014).

23. Balkin, *supra* note 1, at 28; see also Elizabeth Rosenblatt, *Copyright's One-Way Racial Appropriation Ratchet*, 53 *U.C. DAVIS L. REV.* 591, 618 (“Proprietary ownership of traditional creative processes can hardly be said to promote ‘progress.’”).

Here is where the law steps in. For a cultural dialogue to be fair, representative, diverse, and dialogic, it is imperative that the law substantively equalizes the position of cultural speakers, as well as the forms of expressions produced. The goal of cultural policy thus, as a legal matter, is *enablement* of those who wish to expressively produce for their own self-determination as well as the social meaning-making.

B. ROLE OF COPYRIGHT IN CULTURAL POLICY

Copyright policy is no bystander to cultural expressions and their engagement,²⁴ and significantly affects the shape and vision of culture.²⁵

Traditionally, copyright has been justified by the incentive-access paradigm.²⁶ As information is inexpensive to copy, to ensure creators can recoup costs of development, they receive exclusionary rights²⁷ that allow them to charge a price for using their works.²⁸ As information is arguably non-rival²⁹ (one person's use does not disable someone else's use), exclusion creates deadweight loss—often called copyright's static inefficiency on the consumption and use side, and dynamic inefficiency on the side of downstream creation.³⁰ However, this is often justified by claims of dynamic efficiency and a larger corpus of output that is supposedly produced through exclusion.³¹ Copyright thus is often a complex compromise that involves constant tradeoff between: (1) relative social costs, i.e., its static and dynamic inefficiencies; and (2) benefits, i.e., its dynamic efficiency.³² This traditional economic view, focused on recouping highest possible value as a means of inducing more creation, a view that spins out of methodological individualism, has pervaded copyright's purpose for more than five decades.³³

24. Sunder, *IP*³, *supra* note 13, at 322.

25. Fisher, *Reconstructing the Fair Use Doctrine*, *supra* note 20, at 1696.

26. See Talha Syed & Oren Bracha, *Beyond the Incentive-Access Paradigm? Product Differentiation and Copyright Revisited*, 92 TEX. L. REV. 1841, 1843 (2014); see also Amy Kapczynski, *The Cost of Price: Why and How to Get Beyond Intellectual Property Internalism*, 59 UCLA L. REV. 970, 975 (2012).

27. See Syed & Bracha, *supra* note 26, at 1843.

28. *Id.*

29. *Id.* at 1848–49; Kapczynski, *supra* note 26, at 974; Swaraj Paul Barooah, *Looking Beyond IP Internalism*, SPICY IP (Sept. 21, 2012), <https://spicyip.com/2012/09/looking-beyond-ip-internalism.html>.

30. Kapczynski, *supra* note 26, at 974.

31. *Id.*

32. Syed & Bracha, *supra* note 26, at 1844; see Fisher, *Reconstructing the Fair Use Doctrine*, *supra* note 20, at 1703.

33. Syed & Bracha, *supra* note 26, at 1844.

More recently, however, IP and social justice scholars have been arguing for a culturally conscious account of copyright.³⁴ They argue copyright policy's goal to be to foster cultural flourishing by inducing participation, production and dissemination of diverse cultural expressions,³⁵ especially ones that “talk back” to normative cultural conceptions.³⁶ Referring to it as copyright's new

34. See, e.g., Neil Netanel, *Copyright and a Democratic Civil Society*, 106 YALE L.J. 283 (1996) (arguing that copyright's fundamental purpose to underwrite speech competence to contribute to a democratic civil society, focusing on multiplicity of expressive outlets, is more conducive to market diversity than concentrated markets); Fisher, *Reconstructing the Fair Use Doctrine*, *supra* note 20 (arguing for a democratic and culturally conscious account of copyright's economic basis); see also Sunder, *IP*, *supra* note 13, at 269 (arguing for copyright law to be understood as a legal vehicle for facilitating recognition of diverse contributors to cultural discourse); Bracha & Syed, *supra* note 8 (arguing for copyright to be understood from a consequence-sensitive lens focusing on broadening its purpose to further autonomous self-determination and cultural democracy which comprise important determinants of efficiency, complementary to economic concerns); Julie Cohen, *Creativity and Culture in Copyright Law*, 40 U.C. DAVIS L. REV. 1151, 1197 (2007) (criticizing the standard economic account of copyright as being counterintuitive to advancing its goals of promoting a breathing room for autonomous creative practice); BJ Ard, *Taking Access Seriously*, 8 TEX. A&M L. REV. 225, 269 (2021) (arguing access of diverse perspectives to distribution markets as being instrumental to achieving copyright's goals of more democratic and participatory culture); James Boyle, *Cultural Environmentalism and Beyond*, 70 L. & CONTEMP. PROBS. 5, 10–14 (2007) (criticizing the solely economic account of copyright as being counter-intuitive to the intent of cultural policy, or what the author refers to as “cultural environmentalism”); LAWRENCE LESSIG, *FREE CULTURE* 8, 192–93 (2004) (arguing that copyright imposes a permission culture on the kinds of cultural expressions that are practiced); YOCHAI BENKLAR, *THE WEALTH OF NETWORKS: HOW SOCIAL PRODUCTION TRANSFORMS MARKETS AND FREEDOM* 285 (2006) (critiquing copyright exclusivities as denying opportunities to many for participating in cultural practice and instead arguing for a commons-based regime); Elizabeth Rosenblatt, *Social Justice and Copyright's Excess*, 6 TEX. A&M J. PROP. L. 5 (2020) (urging courts to consider impact on social justice while adjudicating copyright policy's scope and its promotion of progress); CARYS CRAIG, *COPYRIGHT COMMUNICATION AND CULTURE: TOWARDS A RELATIONAL THEORY OF COPYRIGHT LAW* 38–42 (2011) (discussing the purpose of copyright from the perspective of nature of creativity being inherently dialogic and cumulative); Rosemary Coombe & Susannah Chapman, *Ethnographic Explorations of Intellectual Property*, OXFORD RES. ENCYCLOPEDIA: ANTHROPOLOGY (2020), <https://doi.org/10.1093/acrefore/9780190854584.013.115> (suggesting that the purpose of copyright should be cognizant of the actual reality of how culture is practiced rather than molding cultures to fit in into a solely economic narrative); ANJALI VATS, *THE COLOR OF CREATORSHIP: INTELLECTUAL PROPERTY, RACE AND THE MAKING OF AMERICANS* 204–08 (2020) (criticizing the economic account of copyright from a critical race lens and arguing that it is important to consider citizenship implications of copyright and to be cognizant of the participants it deems to be cultural citizens and those it ignores).

35. Sean A. Pager, *Does Copyright Help or Harm Cultural Diversity in a Digital Age?*, 32 KRITIKA KULTARA 397, 400–02 (2019), <https://ajol.ateneo.edu/kk/articles/83/1007>.

36. Rosenblatt, *supra* note 23, at 619 (citing Keith Aoki, *Adrift in the Intertext: Authorship and Audience “Recoding” Rights—Comment on Robert H. Rotstein, “Beyond Metaphor: Copyright Infringement and the Fiction of the Work,”* 69 CHI-KENT L. REV. 805, 836 (1992)).

enlightenment in a participation age,³⁷ this view postulates copyright law as a means to provide optimal control over creative expressions to recoup *fair* economic returns, so that one is not dis-incentivized from producing, or is enabled to produce, creative expressions and is not forced to shift to marginal sources of revenue. This furthers the *end* goal³⁸ of providing a wide variety of cultural expressions for people to access.³⁹ This is the presumptive goal of copyright that I base this Note on.⁴⁰

The theory that I postulate here, distinct from a theory of copyright as rewards or the law and economics justification of incentives, is that copyright, as a matter of legal policy, is a *historically specific*⁴¹ tool of *enablement* to allow for *human flourishing*. It is a tool meant to ensure that those who wish to expressively produce are free (or have the agency) to do so without worrying about fulfilling their basic economic needs in a modern market society. It is a tool to *affirmatively protect* those who wish to produce expressions from *involuntary subjection* to the logic of a *historically specific* market society⁴² where realization of basic needs, that constitute *human flourishing*,⁴³ is dependent on market

37. Sunder, *IP*³, *supra* note 13, at 264.

38. See MADHAVI SUNDER, FROM GOODS TO A GOOD LIFE: INTELLECTUAL PROPERTY AND GLOBAL JUSTICE 100 (2012) (discussing whether intellectual property is an end or the means).

39. See *Sony Corp. of Am. v. Universal City Studios, Inc.*, 464 U.S. 417, 429 (1984).

40. Although I rely on this cultural justification of copyright policy, scholars continue to debate on the fundamental justifications of copyright policy. See, e.g., William Fisher, *Theories of Intellectual Property*, in NEW ESSAYS IN THE LEGAL AND POLITICAL THEORY OF PROPERTY (Stephen Munzer ed., 2000). Even if we change the premise of analysis to copyright's labour justification or its justification focused on protecting the personality of the author, the distortionary effects of copyright policy continue to sustain.

41. Oren Bracha, *The History of Intellectual Property as The History of Capitalism*, 71 CASE W. RESV. L. REV. 547, 574–75 (2020) (tracing the history of IP to the process of commodification which is an output of an ensemble of social relations that constitute capitalism and found specific phenomenological presence only during the 17th Century, while also trying to denaturalize IP law).

42. ELLEN MEIKSINS WOOD, THE ORIGIN OF CAPITALISM: A LONGER VIEW 106 (1999).

43. What components constitute Human Flourishing can be widely debated, but the ones I specifically mean to refer to here are a combination of the spirit of the components endorsed in two texts—both of which specifically reject a notion of methodological individualism and endorse fulfillment of these basic components through the instrument of the law—one in context of real property law, and the other in context of copyright law. See Gregory S. Alexander, *Ownership and Obligations: The Human Flourishing Theory of Property* 2, 5, CORNELL LAW FACULTY PUBLICATIONS, PAPER 653 (2013), <https://scholarship.law.cornell.edu/facpub/653/> (emphasizing life, freedom, practical reasoning, and sociability as four essential capabilities of human flourishing); see also Bracha & Syed, *supra* note 8, at 256–57 (distilling three elements of human flourishing, which are self-determination, meaningful activity, and sociability).

competition.⁴⁴ As a matter of legal policy, it specifically aims to *enable* (as against incentivize) those who perform expressions to sustain economically (i.e., at the least fulfill basic needs), be recognized, as well as flourish (as distinguished from theories of individual welfare) in a market society. It is in consonance with this theory that I distill the relevant scope of copyright that adequately, optimally, and equitably enables a diverse set of performers without creating distortions based on the ability to internalize market value out of these expressions.

Using this premise, the central thesis of this Note is that current copyright policy, which is pervaded by neoliberal concepts of efficiency instituted by Coase-ian and Hayekian economics, creates asymmetrical demand in favour of highly excludable expressions, and therefore estranges cultural expressions from genres like classical music and hip hop that are relatively dissonant and less legally excludable. These are copyright's distorting effects that arise out of copyright's specific scope of exclusionary rights. They draw hierarchies in cultural aspirations and experience and contribute to conformity in creative cultural practice.

C. COPYRIGHT'S TOOLS

To fulfill its purpose, copyright policy confers “rights to exclude” to producers of original expression.⁴⁵ However, such exclusionary rights have significant adverse effects on the *nature* of expressions produced, the kind of participants who produce, as well as the exposure and access to expressions that are produced.

1. *Rights to Exclude*

Rights to exclude, or exclusionary rights, allow those who produce expressions to commodify their output—namely restrict unauthorized use and access and reap economic returns by selling the output in the market.⁴⁶ Its reasoning flows from Harold Demsetz's influential statement that exclusionary rights have a fundamental advantage in dictating efficient production, as production is guided by market signals which drive investment towards content in demand.⁴⁷ It reflects the philosophy that exclusionary rights best

44. See Talha Syed, *Capital as a Social Relation* (unpublished manuscript) (on file with author); see also Talha Syed, *The Horizontal and Vertical in Capitalism* (unpublished manuscript) (on file with author).

45. Kapczynski & Syed, *supra* note 3, at 1908.

46. *Id.*

47. See Kapczynski, *Cost of Price*, *supra* note 26, at 982 (citing Harold Demsetz, *Information and Efficiency: Another Viewpoint*, 12 J.L. & ECON. 1, 9 (1969)); see also PAUL GOLDSTEIN, *COPYRIGHT'S HIGHWAY: FROM GUTENBERG TO THE CELESTIAL JUKEBOX* 146 (rev. ed.

foster investment directed to valuable expressions⁴⁸ and they help creators internalise a substantial part of the social value.⁴⁹

Over time, these rights to exclude have become broader, deeper, and more severe⁵⁰ in their scope, duration, breadth of entitlements, as well as the remedies offered—they may impose even criminal punishments for those who hinder potential of appropriating economic value. The law's shift in focus from *enablement* towards extraction of surplus value has expanded the scope of exclusionary rights to cover fragments,⁵¹ as well as potential variations to the original work that the producer may not have originally conceived. Such expansion comes with at least three pitfalls: (1) the *price-tag* effect; (2) the *privilege-expanding* effect; and (3) the *distortionary* effect.

2. Pitfalls of Expanding the Right to Exclude

The first adverse consequence of expanding exclusionary rights in knowledge and culture is the *price-tag effect*. This *price-tag* effect reflects two key social concerns: (1) many consumers of cultural expressions are denied access to works that could potentially define their cultural selves, due to being priced out; (2) many downstream creators unable to pay licensing costs to access and use pre-existing works are denied that creative opportunity.⁵² Thus, welfare-maximising effects of expanding exclusionary rights ignore distributive realities and impair the stimulus of creativity, especially for those who structurally lack the ability to pay.

The second adverse consequence is what I refer to as the *privilege-expanding* effect, drawing from the work of Amy Kapczynski.⁵³ Professor Kapczynski argues that the *kind* of expressions that are available often depend on the choices and preferences of those with the highest ability to pay.⁵⁴ This reinforces privilege and curates a homogenous bubble around culture.

2003) (presenting a commodity strategy of transacting with prices appropriate to drive investment in desirable directions, reflecting the social value of the information and expression).

48. Kapczynski, *Cost of Price*, *supra* note 26, at 983.

49. Syed & Bracha, *supra* note 26, at 1843.

50. Amy Kapczynski, *The Access to Knowledge Mobilization and the New Politics of Intellectual Property*, 117 YALE L.J. 804, 821 (2008).

51. See Syed & Bracha, *Copyright's Atom*, *supra* note 6; see Oren Bracha, *The Ideology of Authorship Revisited: Authors, Markets and Liberal Values in Early American Copyright*, 118 YALE L.J. 186, 238–46 (2008).

52. Syed & Bracha, *supra* note 26, at 1843–44.

53. Kapczynski, *Cost of Price*, *supra* note 26, at 978–79 (discussing how IP systems influence which goods are produced by prioritizing the needs of the wealthy).

54. See *id.*; see also Boyle, *supra* note 34, at 12 (explaining how a hair-loss drug could be more valuable than the cure for most tropical diseases, if value is defined as what current

The focus of this Note, however, is on the third adverse consequence—copyright's *distortionary* effect. Copyright systems overtly reliant on broad exclusionary rights systematically distort private investment decisions and resources towards expressions that are highly excludable and appropriable, allowing creators of such works to internalise a higher relative share of market value relative to other works. Thus, it portrays a perception of a relative lesser value of works that are relatively less excludable⁵⁵ due to their inherent cultural norms of practice, in spite of their normative significance. Consequently, this system fails to enable investment in expressions that potentially offer more or equal net social benefits because these investments are *either* costlier *or* have lower potential of extracting surplus value through the current scope of rights.⁵⁶

Before explaining these distortionary effects and contextualising them in the example of Indian classical music, I will analyse how courts in the United States have interpreted specific exclusionary rights granted under the Copyright Act, the scope of which significantly contributes to these distortions.

III. SCOPE OF RIGHTS

In this Part, I will explore contours of (1) the right to exclude making of a derivative work, meant to exclude works that impinge the original's secondary market, and (2) the right to exclude making of reproductions in the original's primary markets. Expansive doctrinal framings and misguided interpretations of both these rights produce copyright's *distortionary* effects.

A. RIGHT TO EXCLUDE MAKING OF A DERIVATIVE WORK

1. *Contours*

The concept of a derivative right⁵⁷ exists to extend exclusionary rights to the secondary market of the original work—where merely the “form” or medium of representing the primary work is altered.⁵⁸ However, courts have extended the scope of this right to allow the owner to exclude all utilizations

market participants would value the most, which depends on the existing distribution of wealth and of rights).

55. Kapczynski & Syed, *supra* note 3, at 1938 (recognizing similar distortion of investment decisions in case of those innovations that are highly excludable and appropriable using patent law, and those, due to social norms or the nature of the practice involved, are relatively less excludable and commodifiable).

56. *Id.* at 1942–48.

57. 17 U.S.C. §106 (2016).

58. *Id.* (providing definition of “derivative work”).

of their work that are even for the same medium, form, or purpose as the original.⁵⁹ Courts say that as long as there are “recognizable blocks of expression” from the primary work, there is no need for a change in medium of representation.⁶⁰ This expansion has adversely affected the ability of follow-on creators who use pre-existing works to reinterpret and provide alternate expressions through the same medium of representation.⁶¹ Consequently, the scopes of the reproduction right (supposed to cater to the primary market or the same medium of representation) and this derivative right often overlap, resulting in what Professors Talha Syed and Oren Bracha call a “freewheeling reproduction-derivative super right.”⁶²

The United States was the first country in the world to adopt an arguably open-ended derivative right, distinct from even the Berne Convention that restricts the right to exclude only to certain specific kinds of adaptations of works.⁶³ The logic is to expand exclusionary rights to all channels which expose even fragments of the primary work to the public.⁶⁴

2. *Get a License or Do Not Sample!*

Such broad reading of the derivative right has adversely impacted the practice of hip hop music producers who significantly rely upon use of pre-

59. See Neil Weinstock Netanel, *Making Sense of Fair Use*, 15 LEWIS & CLARK L. REV. 715, 747–48 (2011) (observing that for most courts, the rule is that “new expressive content, even a fundamental reworking of the original, is generally insufficient for the use to be transformative absent a different expressive purpose”); see also R. Anthony Reese, *Transformativeness and the Derivative Work Right*, 31 COLUM. J.L. & ARTS 467, 484–85 (2008) (explaining that courts focus on the transformativeness of the purpose in using the underlying work, instead of transformation of the content); Rebecca Tushnet, *Content, Purpose, or Both?*, 90 WASH. L. REV. 869, 876 (2015) (arguing that courts have mostly required the allegedly infringing work to have a different purpose to not infringe the derivative right, opposing the idea of “content transformativeness”); Amy Adler, *Fair Use and the Future of Art*, 91 N.Y.U. L. REV. 559, 578 n.83 (2016) (collecting cases where courts have asked for a different purpose for use to not be infringement).

60. See, e.g., *Litchfield v. Spielberg*, 736 F.2d 1352, 1357 (9th Cir. 1984); see also *Berkic v. Crichton*, 761 F.2d 1289, 1291 (9th Cir. 1985).

61. *Andy Warhol Found. for the Visual Arts, Inc. v. Goldsmith*, 11 F.4th 26 (2d Cir. 2021); Akshat Agrawal, *Andy Warhol Foundation v. Goldsmith: A Misnomer of a Debate*, PHIL IP & POL’Y (Oct. 22, 2022), <https://philipandpolicy.wordpress.com/2022/10/22/andy-warhol-foundation-v-goldsmith-a-misnomer-of-a-debate/> (providing a commentary of the Warhol case).

62. Syed & Bracha, *supra* note 5.

63. See Pamela Samuelson, *The Quest for a Sound Conception of Copyright’s Derivative Work Right*, 101 GEO. L.J. 1505, 1512–16, 1516 n.56 (2013); see also Berne Convention for the Protection of Literary and Artistic Works art. 12, Sept. 28, 1979, S. Treaty Doc. No. 99-27 (1979) [hereinafter Berne Convention].

64. Zechariah Chafee, Jr., *Reflections on Copyright Law*, 45 COLUM. L. REV. 503, 505 (1945).

existing works to compose their expression.⁶⁵ Specifically, cuttin' and scratchin', digital sampling, looping, and mashing up—all of which are integral parts of the hip hop music aesthetic, have been delegitimized.⁶⁶

Digital sampling is one of the integral “minerals” of hip hop music production.⁶⁷ It is often referred to as the African and African American community's tapestry.⁶⁸ The process involves inserting a particular sound or audio segment from a pre-existing recording into a new segment, sometimes manipulating various elements like the pitch or tempo, to create new expressions with an alternate aesthetic. Its aesthetic purpose is to make the incorporated sound recognizable yet the output distinct. This practice is embedded in their musicking and compositional practices.⁶⁹

The way U.S. courts have treated music sampling infringement cases in respect of the derivative work right reveals the conceptual dissonance between cultural practice and copyright policy.⁷⁰ Until now, courts unanimously say sampling of recognizable music segments infringe the derivative right. In *Grand Upright Music Ltd. v. Warner Brother Records*, Judge Duffy held sampling clearance to be a norm, and indicated sampling to be theft unless licenses were sought.⁷¹ In *Bridgeport Music Inc. v. Dimension Films*,⁷² the court pronounced

65. See Rosenblatt, *supra* note 23, at 606, 626, 629; see also Tonya M. Evans, *Sampling, Looping, and Mashing... Oh My!: How Hip Hop Music Is Scratching More Than the Surface of Copyright Law*, 21 FORDHAM INTELL. PROP. MEDIA & ENT. L.J. 843, 857 (2011); Chris Johnstone, *Underground Appeal: A Sample of the Chronic Questions in Copyright Law Pertaining to the Transformative Use of Digital Music in a Civil Society*, 77 S. CAL. L. REV. 397, 400 (2003); Jason H. Marcus, *Don't Stop That Funky Beat: The Essentiality of Digital Sampling to Rap Music*, 13 HASTINGS COMM. & ENT. L.J. 767, 790 (1990); Josh Norek, *You Can't Sing Without the Bling: The Toll of Excessive Sample License Fees on Creativity in Hip-hop Music and the Need for a Compulsory Sound Recording Sample License System*, 11 UCLA ENT. L. REV. 83, 102 (2004); Lauren Fontein Brandes, *From Mozart to Hip-Hop: The Impact of Bridgeport v. Dimension Films on Musical Creativity*, 14 UCLA ENT. L. REV. 93, 100 (2007).

66. Evans, *supra* note 65, at 2.

67. See SIVA VAIDYANATHAN, COPYRIGHTS AND COPYWRONGS: THE RISE OF INTELLECTUAL PROPERTY AND HOW IT AFFECTS CREATIVITY 145 (2001); see also Marcus, *supra* note 65.

68. Evans, *supra* note 65, at 46.

69. See Rosenblatt, *supra* note 23, at 626; see also Larisa K. Mann, *Decolonizing Copyright Law: Learning from the Jamaican Street Dance* 6, 42–44 (Fall 2012) (Ph.D. dissertation, University of California, Berkeley), <https://escholarship.org/content/qt7h8449q6/qt7h8449q6.pdf> (exploring how recognizable sampling is specifically embedded as an essential part of the compositional practice, in the context of Jamaican music).

70. Rosenblatt, *supra* note 23, at 629; see also Brandes, *supra* note 65.

71. 780 F. Supp. 182, 183 (S.D.N.Y. 1991); see also Rosenblatt, *supra* note 23, at 616, 638 (critiquing use of “plagiarists” to condemn cultural expressions and participants who rely on cultural norms).

72. 383 F.3d 390 (6th Cir. 2005); see also Brandes, *supra* note 65; Ponte, *supra* note 65.

looping a two-second portion of a guitar riff from the sound recording of the song “*Get Off Your Ass and Jam*” for use in “*I Got the Hook-Up*” to be infringing the plaintiff’s copyright.⁷³ The Court used § 114(b)—read with § 106—of the Copyright Act to hold that using an actual copy of any element of a sound recording, however insignificant it may be to the whole work, is impermissible.⁷⁴ It said—“*Get a license or do not sample.*”⁷⁵

These decisions gutted the ability of hip hop artists to use samples for downstream creation without adding to their costs.⁷⁶ The cost was not just the licensing fee, but also transactional costs associated with getting a license.⁷⁷ There was a severe decrease in the use of samples post-1991, and many up-and-coming artists in the hip hop genre had to change creative directions, because they could not afford licensing fees to clear samples.⁷⁸ Investors were reluctant to invest in artists who were using samples due to added cost and risk.⁷⁹ Entire styles had to be changed, which left no breathing space for

73. *Bridgeport Music*, 383 F.3d at 401–02.

74. *Id.*

75. *Id.* at 398.

76. See KEMBREW MCLEOD & PETER DI COLA, CREATIVE LICENSE: THE LAW AND CULTURE OF DIGITAL SAMPLING 27, 83, 105, 114–18, 137–44, 158–62 (2011); REBECCA GIBLIN & CORY DOCTOROW, CHOKEPOINT CAPITALISM 165–68 (2022); Evans, *supra* note 65, at 18–19; Rosenblatt, *supra* note 23, at 630–32; see generally Amanda Sewell, *How Copyright Affected the Musical Style and Critical Reception of Sample-based Hip-Hop*, 26 J. POPULAR MUSIC STUD. 295–320 (2014) (claiming that people other than the artists, such as producers or record labels, make financial decisions, requiring change or abandonment of music due to not being able to afford or clear the desired samples); Erik Nielson, *Did the Decline of Sampling Cause the Decline of Political Hip-Hop?*, ATLANTIC (Sept. 18, 2013), <https://www.theatlantic.com/entertainment/archive/2013/09/did-the-decline-of-sampling-cause-the-decline-of-political-hip-hop/279791/>; Mike Schuster, David Mitchel & Kenneth Brown, *Sampling Increases Music Sales: An Empirical Copyright Study*, 56 AM. BUS. L.J. 177, 200–01 (2019).

77. See GIBLIN & DOCTOROW, *supra* note 76, at 165–68; see also MCLEOD & DI COLA, *supra* note 76, at 158–62.

78. See Brandes, *supra* note 65, at 119 (citing VAIDYANATHAN, *supra* note 67, at 133, 140, 143); see also Rosenblatt, *supra* note 34, at 13 (citing KEMBREW MCLEOD, FREEDOM OF EXPRESSION 68 (1st ed. 2005)) (discussing how sampling-related copyright litigation led rap and hip hop creators to rely on fewer, more prominent samples rather than using a large number of less distinctive samples to create rich musical textures); see generally Cohen, *supra* note 34 (arguing copyright to substantially conform creativity and dictate what artists can and cannot do).

79. See KEMBREW MCLEOD, FREEDOM OF EXPRESSION 68 (1st ed. 2005); see also Brandes, *supra* note 65, at 123–25; Sewell, *supra* note 76, at 295–320.

creative self-determination.⁸⁰ The signifyin' rapper had lost its voice⁸¹ and was rather rendered a lazy thief.⁸²

3. *Is De Minimis Helpful? Dissonance with Compositional Logic*

In *VMG Salsoul, LLC v. Ciccone*, the Ninth Circuit minutely diverged from *Bridgeport Music*, and allowed copying sound recordings to the extent that the copied element is an unrecognizable trivial component of the original.⁸³ The court held that so long as a reasonable listener cannot identify appropriation, the use of a pre-existing sound recording would be *de minimis* and not infringe⁸⁴ (citing the Ninth's Circuit's logic in *Newton v. Diamond*,⁸⁵ which was in context of sampling musical compositions and not sound recordings).⁸⁶

This decision, however, is no victory for sampling and appropriation artists.⁸⁷ The *de minimis* threshold rejects *Bridgeport*, but as Oren Bracha shows, it provides an equally problematic maxim—"Get a license or never copy anything recognizable."⁸⁸ Any non-meager sample recognizable by ordinary members of the audience is not saved by this exception.⁸⁹

Rap artists often intentionally incorporate recognizable material to draw familiarity.⁹⁰ Sampling is a tool to talk-back to dense media portrayals that dominate the social environment.⁹¹ It is a discursive tactic to retell and recontextualize narratives.⁹² It is the very popularity of the sampled part of the song that makes it indulgent and provokes the impulse of recontextualizing or

80. See Kembrew McLeod, *How Copyright Law Changed Hip Hop: An Interview with Public Enemy's Chuck D and Hank Shocklee*, LITTLE VILLAGE (Oct. 17, 2011), <https://littlevillagemag.com/how-copyright-law-changed-hip-hop-an-interview-with-public-enemys-chuck-d-and-hank-shocklee/>.

81. See VAIDYANATHAN, *supra* note 67, at 143.

82. ANJALI VATS, *THE COLOR OF CREATORSHIP: INTELLECTUAL PROPERTY, RACE AND THE MAKING OF AMERICANS* (2020) (lamenting the racialization of human progress through labelling people of color as "lazy thieves" capable only of rote reproduction).

83. 824 F.3d 871, 881 (9th Cir. 2016).

84. *See id.*

85. *See id.* at 877; *Newton v. Diamond*, 388 F.3d 1189 (9th Cir. 2004).

86. 824 F.3d at 881; Oren Bracha, *Not De Minimis: (Improper) Appropriation in Copyright*, 68 AM. U. L. REV. 139, 156–57, 168 (2018).

87. Bracha, *supra* note 86, at 157, 183.

88. *Id.*

89. *See id.* at 165. These are still subject to the Fair Use exception which, however, is a cold comfort due to its unpredictability.

90. *See* Brandes, *supra* note 65, at 118; Wendy Gordon, *Reality as Artifact: From Feist to Fair Use*, 55 L. & CONTEMPORARY PROBS. 93, 98 n.25 (1992).

91. Keith Aoki, *Adrift in the Intertext: Authorship and Audience "Recoding" Rights*, 68 CHI.-KENT. L. REV. 805, 836 (1993).

92. *See* Sundar & Chander, *supra* note 13, at 619–621; Rosenblatt, *supra* note 23, at 643, 646.

producing an distinct construction with an alternate meaning, message or structure⁹³ to reimagine cultural expression.⁹⁴ The whole point of this cultural practice is to foster a discursive and dialogic community or to engage with repositories of social memory to enhance collective experience.⁹⁵ The *de minimis* rationale continues to undermine this common cultural practice of *glomming on* to parts of recognizable works⁹⁶ and does nothing to save such expressions. It continues to disrespect hip hop artists who make conscious aesthetic choices to recognizably sample from pre-existing works⁹⁷ by increasing their costs.

These decisions reflect the ethnocentric focus on culture while framing and interpreting contours of the derivative right, ignorant of borrowing as a normal cultural tendency and practice.⁹⁸

B. THE RIGHT TO EXCLUDE REPRODUCTIONS

1. *Contours*

The right to exclude reproductions encompasses much of the content of the derivative right and goes even beyond.⁹⁹ With the aim of compensating creators for substitution in their primary markets,¹⁰⁰ the reproduction right initially focused on identical copying or colourable changes by the defendant to evading complete copying.¹⁰¹ However, as Oren Bracha shows,¹⁰² pressures in the 19th century, triggered by industries realising the enormous benefit that accrues through broad exclusionary rights, significantly expanded its scope¹⁰³

93. Rosenblatt, *supra* note 23, at 618.

94. See VAIDYANATHAN, *supra* note 67, at 135; Brandes, *supra* note 65, at 118.

95. VAIDYANATHAN, *supra* note 67, at 137–38.

96. Bracha, *supra* note 86, at 185.

97. Olufunmilayo B. Arewa, *From J.C. Bach to Hip Hop: Musical Borrowing, Copyright and Cultural Context*, 84 N.C. L. REV. 547, 577–78 (2006).

98. See Olufunmilayo B. Arewa, *Copyright on Catfish Row: Musical Borrowing, Porgy and Bess, and Unfair Use*, 37 RUTGERS L.J. 277, 281, 332 (2006). Even U.S. courts have recognized borrowing to be essential to musical practice across genres. See, e.g., *Micro Star v. FormGen Inc.*, 154 F.3d 1107, 1110 (9th Cir. 1998) (stating that the derivative right is too broad because borrowing from known sources is all but necessary); *Gray v. Perry*, No. 2:15-CV-05642, 2018 WL 3954008 (C.D. Cal. Aug. 13, 2018) (stating that music “borrows and must necessarily borrow” from known and used works).

99. Michael Abramowicz, *A Theory of Copyright’s Derivative Right and Related Doctrines*, 90 MINN. L. REV. 317, 334 (2005).

100. Bracha & Syed, *Copyright Rebooted*, *supra* note 5.

101. *Stowe v. Thomas*, 23 F. Cas. 201 (C.C.E.D. Pa. 1853).

102. Bracha, *The Ideology of Authorship*, *supra* note 51.

103. *Id.* at 226.

to include copying even *elements of works* that have a similar “look and feel.”¹⁰⁴ Thus, over time, the scope of this right has bloated to cover “quite remote degrees of similarities under a very broad substantial similarity test.”¹⁰⁵

The baseline of copyright infringement in the primary market has changed from copying of the whole work or a large portion of it, to diminishing the value of *any part or element* of the original, and to free riding of “any” element from the labour of the original creator.¹⁰⁶ The judgment of similarity now rests upon a subjective and mystifying test where different circuits incorporate multiple different standards, focus on dissection of elements, and adjudicate upon the “similarity of feel” in the protectable elements of two works.¹⁰⁷ In other words, the scope of the protected “work” is now expanded to include its fragments, even when used or reproduced outside the context of the whole primary expression or its aesthetic appeal.¹⁰⁸

Due to more analysis of elemental similarity as against the overall expression, courts have, rightly, rendered certain building block elements of works as being *scenes a faire* and thus outside the scope of similarity analysis, given their use as stock inputs in multiple compositions. *Scenes a faire* is, thus, a limited saving grace. In *Skidmore v. Led Zeppelin*, the Ninth Circuit Court of Appeals held common or trite musical elements to not be subject to the substantial similarity analysis as no one person could claim ownership over them.¹⁰⁹ Recognizing arpeggios generally to be *scenes a faire* elements,¹¹⁰ the court refused to exclude or protect a combination of “a five-note descending chromatic scale in A minor; a sequence of half notes and whole notes in the

104. *See id.* at 227–28, 238–40; Craig, *Transforming “Total Concept and Feel,” supra* note 6, at 24.

105. Bracha & Syed, *Copyright Rebooted, supra* note 5.

106. Bracha, *The Ideology of Authorship, supra* note 51, at 228–32 (analyzing Justice Story’s contribution to dividing the “work” into “elements protectable” through *Folsom v. Marsh*, 9 F. Cas. 342 (C.C.D. Mass. 1841) to the extent that as long as the value of the original work is diminished, or the author’s labor are substantially appropriated, there can be copyright infringement even if there was no copying of the whole work).

107. *See* Bracha & Syed, *Copyright Rebooted, supra* note 5; Pamela Samuelson, *A Fresh Look at Tests for Nonliteral Copyright Infringement*, 107 NW. U. L. REV. 1821, 1823 (2013); Jeanne C. Fromer & Mark A. Lemley, *The Audience in Intellectual Property Infringement*, 112 MICH. L. REV. 1251, 1267–73 (2014); Mark A. Lemley, *Our Bizarre System for Proving Copyright Infringement*, 57 J. COPYRIGHT SOC’Y U.S.A. 719, 724–26 (2010).

108. *See* Bracha & Syed, *Copyright’s Atom, supra* note 6; *see also* Bracha, *The Ideology of Authorship, supra* note 51, at 234–35.

109. 952 F.3d 1051, 1069 (9th Cir. 2020).

110. *Id.* at 1070–71 (holding that “chromatic scales and arpeggios cannot be copyrighted by any particular composer”).

scale; a melody involving various arpeggios and note pairs; a rhythm of successive eighth notes; and a collection of pitches in distinct proportions.”¹¹¹

2. *Protection of Fragments: Joyful Noise?*

Gray v. Hudson is a recent case from the Ninth Circuit Court of Appeals that demonstrates how courts expand the scope of the reproduction right from the *overall work* itself, into its *fragments or elements*.¹¹²

In *Gray*, the plaintiff Marcus Gray (composer of the song “Joyful Noise”)¹¹³ alleged infringement over an eight-note ostinato against defendant-Katy Perry’s song¹¹⁴—“Dark Horse.”¹¹⁵ After consciously leaving “access” for the purposes of infringement unaddressed,¹¹⁶ the court analyzed substantial similarity through the two-part test of extrinsic and intrinsic similarity.¹¹⁷ While analyzing the extrinsic test, both the district court as well as the Ninth Circuit focused on figuring out whether any element of the plaintiff’s work was protected and objectively similar to any element of the defendant’s work.¹¹⁸ Both held the elements involved—for example chord progressions, tempos, recurring vocal phrases, repeating hook phrases, syncopation and arpeggios—to be common and trite elements that could not protectable.¹¹⁹ Rather than focusing on the overall aesthetic differences or similarities between the works, it broke compositions into parts and scrutinized protectability.¹²⁰

The court recognized that musical works generally do, and must, borrow from well-known elements used before.¹²¹ Importantly, the court reaffirmed the finding of the district court that elements ubiquitous in popular music and firmly rooted in a genre’s tradition, like chants, use of horns or glissando to

111. *Id.* at 1071–72.

112. *Gray v. Hudson*, 28 F.4th at 87 (affirming the district court’s decision in *Gray v. Perry*).

113. LilMeeker, *Joyful Noise-Flame ft. Lecrae*, YOUTUBE, (Feb. 15, 2009), https://www.youtube.com/watch?v=gWDutCdfS_s&ab_channel=LilMeeker.

114. Katy Perry (Official), *Dark Horse*, YOUTUBE (Feb. 20, 2014), https://www.youtube.com/watch?v=0KSOMA3QBU0&ab_channel=KatyPerryVEVO.

115. *Gray v. Hudson*, 28 F.4th at 92.

116. *Id.* at 96 (deciding not to address the access prong because the case may be resolved based on the “substantially similar” prong). The district court had already concluded the presence of access. *Gray v. Perry*, 2018 WL 3954008, at *5.

117. *Gray v. Hudson*, 28 F.4th at 96.

118. *Id.* at 96–98 (stating that the extrinsic test requires “breaking the works down into their constituent elements and comparing those elements for proof of copying as measured by substantial similarity” and emphasizing the importance of distinguishing between the protected and unprotected material in a plaintiff’s work).

119. *Id.* at 98–100.

120. *Id.* at 98 (holding that the elements plaintiff identified, instead of the whole work, were not copyrightable).

121. *Id.* at 99; *see also Gray v. Perry*, 2018 WL 3954008, at *6.

not be legally excludable because they are indispensable to cultural practice.¹²² Thus, the court rendered them *scenes a faire*, an idea that is not protected or excludable.¹²³ The court also held that sequences of notes, if commonplace to the genre, would not be protectable.¹²⁴ The rationale was that if the rules of the game only allow relatively few ways to express a combination of notes, given constraints of a particular musical convention and style, the same would not be excludable.¹²⁵ Further, while acknowledging combination of these unprotectable elements may be excludable when such combinations were original, the court recognized their excludability to be narrow and the protection thin—defined by the test of virtual identity.¹²⁶

These observations, although pro-copying, limiting the scope of the work, and cognizant of cultural norms, stem out of a misplaced focus on breaking down compositions into elements while adjudicating similarity, as against one focusing on overall aesthetic appeal of the two works as perceived by consumers. When there is no meaningful scrutiny of sufficient similarity on the level of the overall work and the sole focus is on filtering out unprotectable fragments, the outcome is an unfortunate distortion. Even if we would ideally think any two musical works to be different in terms of their aesthetic sensation, if certain protected elements are similar, it would be enough for the latter work to be infringing.

Such a fragmentary approach of dissecting elements and comparing works creates hierarchies in musical practice. Certain genres or kinds of musical composition necessarily involve use of similar elements due to strict compositional rules of the genre. They will, in the prevailing copyright regime, *either* inevitably constitute copyright infringement and thus will involve added costs of composition or investment, or will be less appropriable using exclusionary rights if they are deemed to constitute voluminous *scenes a faire* elements. I do not propose to argue that elements deemed to be *scenes a faire* ought to be protected. To the contrary, what I argue is that there is a need to tighten scope at the overall level of the work, in substantial similarity analysis by de-fragmenting it, as against relying on *scenes a faire* because the latter elemental approach unfortunately produces distortionary effects for works that voluminously involve elements that are rendered *scenes a faire* in law by reducing their potential of legal excludability and appropriability.

122. *Gray v. Perry*, No. 2:15-CV-05642, 2020 WL 1275221, at *5 (C.D. Cal. Mar. 16, 2020).

123. *Gray v. Hudson*, 28 F.4th at 98–99.

124. *Id.*

125. *Id.*

126. *Id.* at 101–02.

IV. INDIAN CLASSICAL MUSIC—EXPRESSION OR THIEVERY?

Classical musical forms significantly rely upon borrowing as a tool—for innovation through transformative imitation.¹²⁷ Extensive borrowing, inherent similarity of elements in two compositions, and a focus on performative improvisation in this tradition demonstrate dissonance with contemporary views of musical composition and copyright law. In this Part, I analyze an Indian (*Hindustani*) classical form of music and its compositional as well as performative practices, to show this dissonance.

Indian musical tradition reflects enormous influence of the *Raga* in its compositional and expressive practice. It is often referred to as the soul of the Indian music system.¹²⁸ The rules of the *Raga* system and its practice, however, are dissonant with copyright law's focus on broad exclusionary rights as a mode of inducing creation and investment. In Indian classical music, every curated composition inherently (1) involves desirable similarity of elements with other compositions in the same *Raga*, which the listeners and the performer can easily discern, (2) follows defined and strict rules of phrasing, sequencing, and performing compositions in that particular *Raga*, and (3) voluminously incorporates pre-existing expression that is critical to any composition in a particular *Raga*. Thus, substantial similarity, as is currently understood by courts, is inevitable. At the scope of copyright protection that currently persists, these expressions, due to their cultural traits, involve *either* high licensing costs, in case copyright law continues to protect all elements of the works in spite of the *inherence* of similarity, *or* high volumes of *scenes a faire* elements—in the case where copyright law renders motivic phrases, arpeggios, sequences and intonations, which must be followed in a *Raga*, as *scenes a faire*. Thus, such compositions are *either* costlier to produce, *or* relatively less excludable and appropriable than expressions from other genres of music. It is a lose-lose.

A. THE RAGA SYSTEM: RULES OF COMPOSITION

Every composition in Indian classical music is in a *Raga* or involves a perceivable amalgamation of multiple *Ragas*. A *Raga* is conceptualized as a “melodic mode/form or tonal matrix possessing a rigid and specific individual identity yet bearing immense potential for infinite improvisatory

127. Arewa, *supra* note 97, at 610.

128. See generally Santosh Kumar Pudaruth, *A Reflection on the Aesthetics of Indian Music, With Special Reference to Hindustani Raga-Sangita*, 6 SAGE OPEN 1 (2016), <https://journals.sagepub.com/doi/full/10.1177/2158244016674512>.

possibilities.”¹²⁹ It serves as a basic framework or a superstructure for composition and improvisation in Indian music (which is essentially melodic and monodic in nature).¹³⁰ Every *Raga* has strict compositional rules which play the role of an imaginary domain beyond which no composer can move.¹³¹

A *Raga* comprises defined note selection, confined to a single octave.¹³² Every *Raga* is distinct not just in the notes contained but also in the frequency of certain notes, volume of use, sequencing of ascending and descending segments, as well as signifiers.¹³³ These strict compositional subtleties permit intricate emotions to be expressed through the *Raga*'s perceptive individuality, represented in phrasings of its compositions.¹³⁴

Every *Raga* requires a certain minimum number of pitches, namely at least five notes out of the twelve recognized notes (“S”, “R”, “Rm” “G”, “Gm” “M”, “Mt”, “P”, “D”, “Dm”, “N”, “Nm”) within the Indian music tradition.¹³⁵ Various pitches can be expressly forbidden in particular *Raga* structures.¹³⁶ Incorrectly including impermissible notes alters the *Raga*, destroying its individuality.¹³⁷

The rules of composition in a *Raga* are strictly prescribed. Every *Raga* encapsulates an *aroha* and an *avroha*, the former signifying the notes and their sequence generally used in ascending parts of the composition, and the latter doing the same for descending parts.¹³⁸ These rules of note transition are mandatory while composing, phrasing, and performing expressions.¹³⁹ Phrasing requires specific focus on the peculiarities and rules of the *Raga*, including strict emphases on particular notes and intonation on specific

129. Sanchit Alekh, *Automatic Raga Recognition in Indian Classical Music* 1, ARXIV (Aug. 7, 2017), <https://arxiv.org/abs/1708.02322>.

130. Pudaruth, *supra* note 128.

131. MRIGANKA SEKHAR CHAKRABORTY, *INDIAN MUSICOLOGY: MELODIC STRUCTURE* 104 (1st ed. 1992).

132. NARENDRA KUMAR BOSE, *MELODIC TYPES OF HINDUSTAN: A SCIENTIFIC INTERPRETATION OF THE RAGA SYSTEM* 452 (1960).

133. *Id.*

134. Jeffrey M. Valla, Jacob A. Alapatt, Avantika Mathur & Nandini C. Singh, *Music and Emotion—A Case for North Indian Classical Music*, 8 *FRONT. PSYCHOL.* 2115 (2017).

135. *See id.*; *see generally* Christian Watson, *How Musicians Develop the Ability to Improvise: A Cross-cultural Comparison of Skill Development in the Egyptian, Hindustani Classical, and Jazz Traditions* (2012) (M.A. thesis, School of the Arts and Media University of New South Wales) (on file with the University of New South Wales Library).

136. *See* Watson, *supra* note 135.

137. *Id.*

138. *See* Suvarnalata Rao & Preeti Rao, *An Overview of Hindustani Music in the Context of Computational Musicology*, 43 *J. NEW MUSIC RES.* 24, 24–33 (2014); *see also* Pudaruth, *supra* note 128.

139. Pudaruth, *supra* note 128.

portions of the composition.¹⁴⁰ Every *Raga* involves a note that is supposed to be struck most frequently in all its phrasings—known as the king note, the *vadi swar*. The second most important and frequent note is the queen note, the *samvadi swar*. Finally, the third most prominent note is the *anuvadi swar*. These notes are supposed to predominate every composition in a particular *Raga*.¹⁴¹ The location of the king note illustrates whether the focus of compositions in the *Raga* has to be on ascending movements or descending movements.¹⁴² Most compositions or phrases conclude on the *nyasa* note of the *Raga*, which is its final prominent and resting note.¹⁴³ They are essential to evoke the mood and emotion associated with that individual *Raga*¹⁴⁴ and are also helpful in differentiating other *Ragas* which might involve similar notes but altered frequencies, or alternate gestures of performance¹⁴⁵ of the same notes.¹⁴⁶

Every composition in a *Raga* includes one or more motivic phrases, known as *pakad* or *vishistha taana*.¹⁴⁷ The *pakad* encapsulates the individuality of a *Raga* and is in fact the recognizable face of every composition.¹⁴⁸ These phrases help both the performer and audience grip the *Raga* and are considered crucial for conveying the peculiar feeling of the *Raga*. They are often present in the beginning of the compositions and repeated, as they are main clues and signifiers for the listeners to identify the *Raga* and distinguish it from other *Ragas*.¹⁴⁹ This characteristic phrase aims to stand out among other phrases drawing the mind to it over and over again, leaving a deep impression of the peculiarity of the *Raga* which is supposed to linger in memory even after the melody stops.¹⁵⁰ The *pakad* is crystallized in all compositions in the same *Raga* resulting in inevitable similarity across expressions.¹⁵¹

140. See Rao & Rao, *supra* note 138; see also BIMALAKANTA ROY CHAUDHARI, AESTHETICS OF NORTH INDIAN CLASSICAL MUSIC 25 (1st ed. 1993).

141. See CHAUDHARI, *supra* note 140.

142. ASHOK RANADE, KEYWORDS AND CONCEPTS IN HINDUSTANI CLASSICAL MUSIC 75 (1990).

143. See Rao & Rao, *supra* note 138; see also BOSE, *supra* note 132, at 455–56.

144. See Kunjal Gajjar & Mukesh Patel, *Computational Musicology for Raga Analysis in Indian Classical Music: A Critical Review*, 172 INT'L J. COMPUT. APPLICATIONS 42 (2017); see also BOSE, *supra* note 132, at 427.

145. CHAUDHARI, *supra* note 140.

146. Valla, Alapapatt, Mathur & Singh, *supra* note 134.

147. Gajjar & Patel, *supra* note 144; Rao and Rao, *supra* note 138.

148. BOSE, *supra* note 132, at 456.

149. See Pudaruth, *supra* note 128; see also Christian Watson, *supra* note 135.

150. BOSE, *supra* note 132, at 466.

151. Telephone Interview with Abhishek Mishra, Professor, Lalit Narayan Mithila University, Darbhanga (Nov. 16, 2022) (transcript on file with author).

B. IMPLICATIONS OF THESE RULES ON PHRASING

Due to these rigid rules of phrasing in *Ragas*, a certain level of perceivable aesthetic similarity in the minds of the lay listener is inherent to the multiple compositions produced within this same superstructure. This is, in fact, desirable, as every *Raga* embodies a unique musical idea or emotion known as *Ragabhava*.¹⁵² Compositions in a *Raga* are supposed to enable ready recognition due to them essentially consisting of these familiar recognizable patterns.¹⁵³ The song composed never overshadows the individuality and familiarity of the *Raga*.¹⁵⁴ Thus, each *Raga* has a distinct perceptible character of its own, perceived through compositions framed within its strict rules.¹⁵⁵

Sometimes multiple *Ragas* are mixed to produce a combination.¹⁵⁶ In the final expression, compositional rules of each *Raga* may be complied with or not, but either way, it will be easy to identify whether the expression is a combination of multiple *Ragas*.

Make no mistake, a *Raga* can be the basis of any number of compositions.¹⁵⁷ However, the melodic framework of each of these compositions represents a degree of similarity, easily discernable by the listener.¹⁵⁸ The difference between *Raga* music and non-*Raga*-based music is the strict loyalty associated with the definite structural arrangement of notes in the former.¹⁵⁹ The untrammelled freedom of both composers and vocalists is circumscribed within the four corners of the *Raga*.¹⁶⁰ This does not dismiss the possibility of limitless compositions within a *Raga*, but just makes a level of similarity of elements or fragments inevitable.¹⁶¹ The individuality of the *Raga* is marked to the extent that provokes identifiable similarity in its compositions.¹⁶²

152. *See id.*; *see also* Telephone Interview with Pt. Ashok Kumar Prasad, PhD. & M.A., Musicology (Indian Classical), Prayag Sangit Samiti, Allahabad (Nov. 17, 2022) (transcript on file with author).

153. BOSE, *supra* note 132, at 466.

154. *Id.*

155. PUDARUTH, *supra* note 128.

156. RANADE, *supra* note 142, at 75.

157. BOSE, *supra* note 132; Telephone Interview with Pt. Ashok Kumar Prasad, *supra* note 152.

158. MANJUSREE TYAGI, SIGNIFICANCE OF COMPOSITIONAL FORMS OF INDIAN CLASSICAL MUSIC, 185–86 (1997).

159. BOSE, *supra* note 132, at 377.

160. *See id.*; *see also* Telephone Interview with Pt. Ashok Kumar Prasad, *supra* note 152.

161. VIJAYA CHONDORKAR, COMPOSITIONAL FORMS OF HINDUSTANI MUSIC: A JOURNEY 1–2 (2012).

162. BOSE, *supra* note 132, at 455–56.

A *Raga* can be recognized through the compositions even if it appears in different rhythms or with different embellishments and styles of expressing its strict notational elements.¹⁶³ Creativity is in fact displayed by introducing variations and embellishments through pleasant ornamentations without disrupting the tone pattern of the *Raga*.¹⁶⁴

Importantly, compositions and performance of *Raga* music significantly involve voluminous use of *alankars* and *taans*, a concept similar to the idea of an arpeggio in western music.¹⁶⁵ *Alankars* and *taans* are rapid sequences of notes logically composed to create a meaningful melodic structure within rules of the *Raga*.¹⁶⁶ A song without an *alankar* is often referred to as a night without a moon, a river without water, and a creeper without a flower.¹⁶⁷ Various compositions involve similar *taans* which serve both as embellishments as well as ornamental essentials to convey the intended emotion of the *Raga*.¹⁶⁸ Only some of the varieties of *taans* can be used in a *Raga* due to its strict rule framework.¹⁶⁹ Allowing to exclude or monopolize any of these *alankars* or *taans*—which are essential elements of composition and performance in Indian classical music, can further limit the possibilities of expression over and above the limiting rules of the *Raga*. These are what can be referred to as *scenes a faire* elements, used voluminously while composing songs, thus making the composition less excludable and appropriable.

Indian classical music composition and performance uniquely focuses on oral transmission of knowledge through what is known as the *guru-shishya parampara*, where vocalists learn specific modalities of the notes through imitation.¹⁷⁰ This is significant as Indian classical music emphasizes intonation and performativeness, which cannot be transmitted through textual or visual modes.¹⁷¹ Performance fluency is often acquired through imitative vocalization and further internalized through memorization.¹⁷² Typically, the student observes (visually as well as auditorily) the teacher's performance and then attempts to emulate the phrase exactly as it sounded.¹⁷³ The idea is to

163. *Id.* at 346, 378.

164. ANUPAM MAHAJAN, RAGAS IN HINDUSTANI MUSIC: CONCEPTUAL ASPECTS 50 (2001).

165. Rao & Rao, *supra* note 138.

166. RANADE, *supra* note 142.

167. DR. SWATANTRA SHARMA, FUNDAMENTALS OF INDIAN MUSIC 9 (1996).

168. Rao & Rao, *supra* note 138.

169. MAHAJAN, *supra* note 164, at 48.

170. Watson, *supra* note 135.

171. *Id.*

172. *Id.*

173. Shyamal "Sony" Tiwari, *Oral Tradition and Musical Knowledge in Indian Composition Pedagogy* (May 2011) (M.A. thesis, New York University) (on file with author).

imitatively learn volumes of compositions in a particular *Raga* to get a grip of its character and be able to improvise and embellish during performance.¹⁷⁴ Excluding access and use of pre-existing material or composition is antithetical to this practice. Thus, this genre of music is relatively less excludable, as compared to those which rely less on voluminous exposure and use in downstream creations.

Further, Indian classical music resists the concept of romantic or individualistic authorship.¹⁷⁵ Most *Ragas* find their birth in social settings, often known as *gharanas*, which help group members perform compositions associated with these *Ragas* and ensure continuity of fundamental characteristics of the *Raga* system.¹⁷⁶ These *gharanas* are generally defined by locations of these social settings and is a metaphor representing the familial¹⁷⁷ collectivist tradition of performing arts.¹⁷⁸ Each *gharana*, as against an individual in the *gharana*, has developed distinctive performative features in various *Ragas*, deeply rooted in their common underlying tradition. *Ragas* and traditional compositions in Indian classical music have evolved in these social settings through long centuries of characteristic exhibition.¹⁷⁹

C. DISSONANCE WITH THE SCOPE OF RIGHTS

The peculiarities of the Indian classical music tradition show their clear dissonance with broad exclusionary rights prevalent in copyright policy.¹⁸⁰ My central claim is that compositions and elements of this form of music are inherently less excludable and appropriable, as well as relatively costlier due to the presence of a broad derivative right, due to (1) perceivable similarity in compositions, (2) defined and strict rules of composing and performing, (3) aural nature of transmission of knowledge through the *guru shishya parampara* relying on characteristically transforming voluminous pre-existing expression in an identical medium for an identical purpose, as well as (4) its inherent rejection of individuality of compositions.

174. TYAGI, *supra* note 158 at 40–41, 187.

175. Shyamal “Sony” Tiwari, *supra* note 173.

176. *Id.*

177. TYAGI, *supra* note 158, at 41.

178. *See* RANADE, *supra* note 142, at 62; *see also* Tiwari, *supra* note 173.

179. BOSE, *supra* note 132.

180. *See generally* Rajalakshmi Nadadur Kannan, *Performing Religious Music: Interrogating Karnatic Music Within a Postcolonial Setting* (Aug. 2013) (Ph.D. dissertation, School of Arts and Humanities, University of Stirling) (on file with author); *see also* Rajalakshmi Nadadur Kannan, *Copyright, Capitalism and a Postcolonial Critique of Karnatic Music*, FOCAALBLOG (Apr. 16, 2015), <https://www.focaalblog.com/2015/04/16/rajalakshmi-nadadur-kannan-copyright-capitalism-and-a-postcolonial-critique-of-karnatic-music/>. This is in context of a sub-genre of Indian music—Karnatic classical music.

Acknowledging this dissonance between law and cultural practice is important as cases asserting copyright infringement begin to come up. A recent example is the case of *Thaikkudam Bridge v. Hombale Films*, which was filed before the District Court of Kozhikode, Kerala, in India.¹⁸¹ Seeking a temporary as well a permanent injunction, the plaintiff, composer of the song “*Navarasam*,” asserted that the song “*Varaha Roopam*,” incorporated in the defendant’s film *Kantara*, infringed upon its copyright.¹⁸² The court granted the plaintiff’s application for preliminary injunction on the basis of there being a *prima facie* case, without analyzing any of the peculiarities of the songs involved and without any substantial reasoning.¹⁸³ Both *Navarasam* and *Varaha Roopam* are songs composed as a mixture of two *Ragas*—*Raga Panturavali* and *Raga Abir Bhairav*.¹⁸⁴ Most similarities in these songs are the result of characteristic phrases and compliance with *Raga* rules.¹⁸⁵ Upon filtering these characteristic phrases, it is clearly visible that the overall expressive characteristic and aesthetic appeal of the two songs can by no measure be *unmistakably* similar. There are significant lyrical and expressive differences as well as a difference in sequencing elements of *Raga* portrayal. Further, common *taans* and *alankars*, similar to arpeggios, are used in the pieces framed within the rules of these *Ragas*, evoking similarity that is in fact desirable in this form of music. By finding a *prima facie* case of infringement, the court has ignored these cultural elements peculiar to the form of music involved.¹⁸⁶

This instance clearly shows why the framings of global copyright law need to be altered to equally accommodate alternate cultural realities, that are often being estranged due to dissonant global norms. I claim that it would be dangerous to issue such injunctions as they could significantly limit expressions possible within classical Indian cultural practice.

This dissonance between the law and cultural practice furthers copyright law’s distortionary effects. It discourages investment in dissemination and

181. *Thaikkudam Bridge v. Hombale Films*, Unreported Judgment, Original Suit No. 14/2022, Principal District and Sessions Judge, Kozhikode District Court, Kerala.

182. *Thaikkudam Bridge v. Hombale Films*, Unreported Judgment, Order dt. Oct. 28, 2022 in IA No. 1/2022 in OS No. 14/2022. Order vacated on the date of writing this note by Kerala High Court on procedural grounds. *See Hombale Films v. Thaikkudam Bridge*, Unreported Judgment, FAO 147/2022, Order dt. Dec. 2, 2022.

183. *Id.*

184. *See* Telephone Interview with Abhishek Mishra, *supra* note 151; Karthik, *Navarasam (Music Review)*—Thaikkudam Bridge, MILLIBLOG (Nov. 1, 2015), <https://milliblog.com/2015/11/01/navarasam-music-review-thaikkudam-bridge/>.

185. *See* Akshat Agrawal, *Copyright and Classical Music: Not the Best Fusion*, SPICY IP (Nov. 25, 2022), <https://spicyip.com/2022/11/copyright-and-classical-music-not-the-best-fusion.html>.

186. *See id.* (providing a detailed analysis on this case and the order granting injunction).

circulation of such works due to involvement of relatively higher licensing costs *or* due to low relative potential of surplus appropriability as a result of inherent similarity between compositions, and voluminous use of elements that are *scenes a faire*.¹⁸⁷ It shows a perception of lower relative market value, in spite of the normative significance of these works in Indian culture.

The next part of this Note explains these distortionary effects.

V. DISSONANCE LEADS TO DISTORTION

Copyright's overt reliance on broad exclusionary rights exhibits a predictable bias for goods and expressions that generate the highest appropriable social value.¹⁸⁸ Conventional economic actors will only invest in distributing a work if they are able to sufficiently recoup social value by commodifying or selling them.¹⁸⁹ When presented with relative choices, most investors would invest in expressions that involve lower costs and offer high appropriability through exclusionary rights to fetch out the highest social and economic market value possible.¹⁹⁰ Thus, the current copyright system exerts enormous influence on the *kind* and content of expressions that receive enough investment to come into visible circulation.¹⁹¹ If probability of the highest possible return is diminished on a relative scale, investment and disseminative decisions are often *distorted away*, redirected to places with relatively higher return potential.¹⁹² Due to market liberalization of global cultural flows, these distortions have adverse ramifications on the *kind* of cultural expression that is globally visible and curated.

A. DISTORTIONARY EFFECTS

All information goods, Professors Kapczynski and Syed argue, exist on an excludability continuum.¹⁹³ Goods or expressions that are relatively non-

187. See Bracha & Syed, *supra* note 8, at 243–44.

188. Kapczynski & Syed, *supra* note 3, at 1905 (citing BRETT FRISCHMANN, *INFRASTRUCTURE* 109 (2012)).

189. See Kapczynski & Syed, *supra* note 3, at 1908; see also GIBLIN & DOCTOROW, *supra* note 76, at 258.

190. *Id.* at 1905–06, 1938; Lunney Jr., *supra* note 4, at 582 (discussing how investors consider risk and return of investment in deciding which works to financially support). Even if the work is highly popular, the possibility of appropriating value only through direct dissemination, as against direct dissemination as well as a licensing market, could potentially drive conventional investors away.

191. See Julie E Cohen, *Copyright as Property in the Post-Industrial Economy: A Research Agenda*, 2011 WIS. L. REV. 22 (2011).

192. See Kapczynski & Syed, *supra* note 3, at 1960; see also Lunney Jr., *supra* note 4, at 494–95.

193. Kapczynski & Syed, *supra* note 3, at 1920.

excludable in the continuum offer a diminished ratio of social value that is privately appropriable through copyrights. On the other hand, goods or expressions that are relatively easy to commodify and exclude, are on the higher end of the continuum of excludability and represent a larger ratio of privately appropriable social value.¹⁹⁴ In other words, the conventional investor considers works that are less easily copied, or more protected from copying, to be more economically valuable.¹⁹⁵

Contextualizing this analysis, works that inherently involve perceivable similarity of elements, *or* incorporate more *scenes a faire* elements, like compositions in Indian classical music,¹⁹⁶ would be relatively less excludable and hence potentially less appropriable through exclusionary rights. Thus, presumably, investing in the production of such content will offer relatively less social value. This significantly distorts investment decisions away from such cultural expressions weakening their cultural visibility and capability to shape tastes and preferences for autonomous, yet social, self-determination.¹⁹⁷

The cost of investment also plays a role in distorting away investments. Professors Kapczynski and Syed argue, under a given state of technology, norms, and institutions, some information will be more or less costly to exclude others from.¹⁹⁸ This is specifically due to the higher need to use pre-existing elements. Compositions in genres which ontologically rely on using pre-existing works would, in the current state of legal rules, involve higher licensing costs for production. This makes them less likely to attract investor interest. The fee demanded for licensing, as well as transactional costs involved, directly correlates to an increase in investment cost. For instance, record label representatives of sampled artists would require more and more of their works to be cleared through licenses under the current legal norms, increasing costs of investors in sampling artists. This exacerbates the already limited potential of appropriability from such expressions, driving and distorting investment away.¹⁹⁹

194. *Id.*

195. Lunney Jr., *supra* note 4, at 589 (explaining investment modelling results that suggest the investor will receive greater returns if investing in products more difficult to copy).

196. *See* discussion *supra* Section IV.C.

197. *See* Kapczynski & Syed, *supra* note 3, at 1947; *see generally* Fisher, *Reconstructing the Fair Use Doctrine*, *supra* note 20, at 1733–88.

198. Kapczynski & Syed, *supra* note 3, at 1919.

199. Talha Syed and Amy Kapczynski make a similar argument in context of patents and this distortion due to relative non-excludability of certain kinds of treatments like natural medication and checklist interventions which may be socially more valuable than modern medicine. *Id.* (“[P]atents will drive innovative effort and investments away from an optimally efficient allocation providing the greatest net social value and instead toward information goods that may provide lower net social value but higher private value owing to lower costs

Therefore, when norms of cultural practice relatively burden commodification or incorporate a cost that reduces the potential of appropriability through broad exclusionary rights, resources are often driven away from development of such expressions.²⁰⁰

Copyright law in its current state is complicit in undersupplying certain valuable expressions that either involve relatively higher costs of investment due to prevalent legal norms, or are relatively less excludable and appropriable due to cultural ontologies.²⁰¹ As a result, it does not prioritize enablement or development of these relatively less excludable expressions and instead actively works against them,²⁰² irrespective of their benefits of exposure. Works of equal or even higher social value may produce lower *enablement* for performers when copyright rules increase their cost of production, and the small share of their market value is capable of being internalized.

Depleting stock and visibility of such expressions diminishes their popular desire through what William Fisher calls the *sour grapes effect*, where aficionados of cultural expressions begin to persuade themselves that they did not really want to see the expressions they are unable to view, due to its purportedly lower value.²⁰³

Such distortion of resources entrenches an element of bias²⁰⁴ for intellectual expressions that generate the most appropriable value in consumer markets.²⁰⁵ Works which input pre-existing expressions to provide an alternate narrative, sometimes referred to as *heterodox* works, which normatively are essential to self-determination as they provide *meaningful variety* and are different from mainstream conceptions of works which have market value, often incorporate less control on secondary markets or are less consequential to recouping investment through derivative markets.²⁰⁶

or barriers to effective excludability.”); see also Arewa, *From J.C. Bach to Hip Hop*, *supra* note 97, at 639 (suggesting that a property rule implicitly assumes borrowing is not the norm, imposing extra costs because there needs to be consent to borrow, distorting the creation of music).

200. Cf. Kapczynski & Syed, *supra* note 3, at 1920 (arguing that investors have a limited ability to change social norms). I extend this argument to include cultural norms of practice embedded to a genre.

201. See Kapczynski & Syed, *supra* note 3, at 1938; see also Lunney Jr., *supra* note 4, at 483, 599, 655 (stating that the market will undersupply products that are more easily copied, while oversupplying products that are less easily copied).

202. Kapczynski & Syed, *supra* note 3, at 1941.

203. Fisher, *Reconstructing the Fair Use Doctrine*, *supra* note 20, at 1735–36, 1736 n.325, 326.

204. Kapczynski & Syed, *supra* note 3, at 1946.

205. *Id.* at 1946–47.

206. See Bracha & Syed, *supra* note 8, at 270–74 (discussing heterodox works).

On the other side, these works also involve a licensing cost due to their innate use of pre-existing output.²⁰⁷ For investors, due to the inherent focus on internalizing market value, they appear unattractive in spite of their significance to self-determination and need for equitable *enablement* of their creators. They also prove to incorporate an additional cost of production and provoke less income through derivative markets in the same medium. Thus, given wide derivative markets for other works, these works are disproportionately prejudiced as they provide less overall profits and involve a higher cost.²⁰⁸ The tendency of investors to minimize risk and earn maximum profit through derivative markets would thus crowd such works out. This is how the legal entitlement becomes systematically biased against such works.²⁰⁹ In a similar vein, expressions composed in Indian classical music tradition that inherently have similar elements as previous compositions are possibly deemed less profitable expressions deserving to be crowded out. Risk-averse creators and investors tend to get scared of investing in such works as the boundaries of protection in such works (and their elements) are often vague and uncertain.

Copyright law thus not only fails to enable investments in some socially beneficial expressions but can also affirmatively jeopardize the creation of such expressions.²¹⁰ It shapes deeper understandings and orientations of participants in the field evoking ideas around what kinds of expression are more desirable for drawing resources.²¹¹ It intimidates those composing expressions that inherently rely on borrowing, thus chilling cultural practices and next generation creativity in such genres of expression.²¹²

These distortions run directly contrary to its instrumental purpose of enabling or incentivizing multifarious lifestyles and ideas on public display for people to be able to develop their own mental and moral faculties.²¹³ The state

207. *Id.*

208. *Id.*

209. *Id.*; Kapczynski & Syed, *supra* note 3, at 1946–47.

210. See Kapczynski & Syed, *supra* note 3, at 1945; see also Fisher, *Reconstructing the Fair Use Doctrine*, *supra* note 20, at 1734–37; see generally Kapczynski, *The Cost of Price*, *supra* note 26.

211. Kapczynski & Syed, *supra* note 3, at 1947 (stating that the process of nonexcludable approaches repeatedly losing out to excludable ones may shape the understandings and orientations of various actors); see also Fisher, *Reconstructing the Fair Use Doctrine*, *supra* note 20, at 1736.

212. See Kapczynski & Syed, *supra* note 3, at 1945; see also Arewa, *From J.C. Bach to Hip Hop*, *supra* note 97, at 639–40 (“Even if legal standards do not impose absolute restrictions on borrowing, the current property rule standard has potential to create a chilling effect because many will be hesitant to borrow from existing material. Any such chilling effect is magnified by current practices of copyright holders that often focus on the strategic use of copyright to expand the scope of such rights. Such strategic uses often involve the use of threats of legal action or actual lawsuits, which may further intensify any chilling effect.”)

213. See discussion *supra* Sections II.A, II.B.

ought not to penalize expressive activity crucial to the preservation of diversity. Diversity must be nourished and rewarded. It is thus important to change the law and its interpretation to facilitate development of a diverse vocabulary of art, as against its creating asymmetrical market demand for, and enabling supply of, some highly excludable and appropriable expressions.²¹⁴

B. RELEVANCE IN THE GLOBAL POLITICAL ECONOMY

Why is this relevant for the United States? In the wake of global intellectual property systems which impose minimum standards of exclusionary rights, media industries selectively endorse highly commodifiable cultural identities, and thus distort cultural visibility, practice, indulgence, and exposure away from expressions that are less excludable. This shows a skew in the reality of the global political economy of cultural practice and is specifically influenced by prevalent understandings of copyright law in the United States, based on concepts underlying romantic creatorship that are central to Western philosophy and dissonant with other cultural practices.²¹⁵ Instruments like TRIPS and the American ideology of organizations like WIPO insert countries into what Amy Kapczynski refers to as a “transnational circuit.”²¹⁶ These are disciplined through use of politico-economic tools like the Special 301 United States Trade Representative reports which punish those refusing to comply with expansive exclusionary regimes.²¹⁷ Some have referred this as *neo-colonialism*²¹⁸ as it privileges a single objective reality of cultural consciousness

214. See Kapczynski & Syed, *supra* note 3, at 1905, 1947, 1950.

215. See Linda M.G. Zerill, *This Universalism Which Is Not One*, 28 *DIACRITICS* 2 (1998).

216. See Amy Kapczynski, *Harmonization and Its Discontents: A Case Study of TRIPS Implementation in India's Pharmaceutical Sector*, 97 *CALIF. L. REV.* 1571, 1645 (2009);

217. See Kapczynski, *supra* note 216, at 1636, 1645; see also Ewa Hemmungs Wirten, *Life, Liberty, and the Relentless Pursuit of Ownership: the “Americanization” of Intellectual Property Rights*, 35 *AM. STUD. SCANDINAVIA* 85, 85–93 (2003). Special 301 is a unilateral, abusive tactic through which the United States imposes maximalist IP norms at the behest of its industrial interests. See USTR, *SPICY IP*, <https://spicyip.com/tag/ustr> (last visited Mar. 31, 2023) (collecting a series of blog posts highlighting the unilateral nature of USTR Section 301 Special Reports); Christopher May, *Cosmopolitan Legalism Meets ‘Thin Community’: Problems in the Global Governance of Intellectual Property*, 39 *GOV'T & OPPOSITION* 393, 393–422; Ruth L. Okediji, *Legal Innovation In International Intellectual Property Relations: Revisiting Twenty-One Years of the TRIPS Agreement*, 36 *U. PA. J. INT'L L.* 191 (2014); CARLOS M. CORREA, *SPECIAL SECTION 301: US INTERFERENCE WITH THE DESIGN AND IMPLEMENTATION OF NATIONAL PATENT LAWS* (South Centre, Geneva, Research Paper No. 115, 2020), <https://www.econstor.eu/bitstream/10419/232238/1/south-centre-rp-115.pdf>.

218. See Andreas Rahmatian, *Neo-Colonial Aspects of Global Intellectual Property Protection*, 12 *J. WORLD INTELL. PROP.* 40, 74 (2009).

and reflects “a colonial imaginary that culturally impoverishes the self and orientalizes the other.”²¹⁹

Copyright’s standards are instruments of financing and organizing cultural production by global media and entertainment industries, driving the industries’ investment choices. These industries are mostly based out of the United States and have been strengthened by a series of mergers around the world from the late 1970s till today.²²⁰ Major players in global entertainment and media industries—which all develop, produce and distribute a plethora of disparate cultural products in many countries through countless corporate entities—influence the nature of cultural products that are disseminated based on potential of appropriating net surplus value through exclusionary rights.²²¹

Neoliberal conceptions of broad exclusionary rights—those protecting fragments of works and including wide derivative markets—do not accommodate alternate cultural expressions that resist these conceptions and are difficult to turn into equally excludable and appropriable commodity.²²² It affects the content of the cultural expression disseminated as well as the opportunity of people to participate in and access cultural discourse.²²³ Not only does this estrange indigenous cultural practices abroad from global visible circulation, but it also estranges contemporary cultural practices in the United States, like music sampling, that inherently re-work, derive, transform, or rely upon voluminous use of non-excludable expressions.

Thus, change in the United States is central as most (although not all) of the world’s biggest traders in culture are based here. If left unchecked, they pose a serious problem of cultural and expressive bias.

219. Rosemary J. Coombe, *Cultural and Intellectual Properties: Occupying the Colonial Imagination*, 16 POL. & LEGAL ANTHROPOLOGY REV. 8, 12 (1993); see also Rosemary Coombe, *Objects of Property and Subjects of Politics: Intellectual Property Laws and Democratic Dialogue*, 69 TEX. L. REV. 1853, 1860 (1991).

220. Rahmatian, *supra* note 218, at 59.

221. Natalie Fenton, *Bridging the Mythical Divide: Political Economy and Cultural Studies Approaches to the Analysis of the Media*, in MEDIA STUDIES: KEY ISSUES AND DEBATES 7, 12 (Eoin Devereux ed. 2007).

222. See Chakkri Chaipinit & Christopher May, *The Polanyian Perspective in the Era of Neoliberalism: The Protection of Global Intellectual Property Rights*, 19 J. POPULATION & SOC. STUD. 99, 111–12 (2010); see Alexander Peukert, *Fictitious Commodities: A Theory of Intellectual Property Inspired by Karl Polanyi’s “Great Transformation,”* 29 FORD. INTELL. PROP. MEDIA & ENT. L.J. 1151, 1184–88 (2019).

223. Fenton, *supra* note 221, at 12, 14–15 (arguing that cultural goods shifting from being public services to private commodities means that the groups able to access them are starkly different, because the “corporate machine” appropriates discourse that challenges the status quo).

VI. POLICY SIGNIFICANCE AND PRESCRIPTIONS

This Part prescribes preliminary structural revisions to copyright law and its interpretations in the United States to reduce these distortions. Before proceeding to prescriptions, however, I clarify the need to *resist* an expansionary impulse that may appear to be a solution for these distortionary effects.

A. RESIST EXPANSIONARY IMPULSE

An impulsive prescription to resolve distortions would be expanding the scope of exclusionary rights to increase appropriability of dissonant works. However, this solution ignores and, in fact, weakens cultural ontologies.²²⁴

Exclusionary rights themselves send distorted signals. Expanding their scope does not help get rid of these distortions.²²⁵ Let us imagine a world where the scope of copyright law was expanded to include *scenes a faire* elements voluminously used in certain cultural expressions. First, due to the scope of the derivative right as it currently stands, license fees for using pre-existing inputs would continue to exist. For example, if we increase the scope of copyright to protect *arpeggios* or *pakad* (the characteristic phrase of a *Raga*) to increase its appropriability, any other composition within the same *Raga* would further incur an additional licensing cost, significantly impacting the number of expressions that can be curated in a *Raga*. This increases distortion, apart from corrosively weakening cultural norms.²²⁶ Similar implications are visible in the *Thaikuddam Bridge* case discussed above, where the injunction effectively allowed exclusionary rights over signifiers and arpeggios inherent to the practice and performance of compositions in a *Raga*.²²⁷

Thus, expansion of exclusionary rights is corrosive and does nothing to remedy the underlying issue—bias and distortion of social value and resources, specifically due to the fundamentally irreconcilable nature of broad exclusionary rights and certain expressive practices. It would ameliorate one source of the balance, i.e., a lower internalization rate, only by exacerbating the other, i.e., creating high barriers in the form of costs of creation.²²⁸

B. STRUCTURALLY SCALE BACK EXCLUSIONARY RIGHTS

As a solution to copyright's *distortionary* effects, I suggest structurally scaling back exclusionary rights to a level where, for a particular kind of work that is

224. Kapczynski & Syed, *supra* note 3, at 1960.

225. *See id.* at 1943, 1960.

226. *See id.*

227. *See supra* notes 182–188 and accompanying text.

228. I am deeply grateful to Prof. Oren Bracha for this expositional framing.

protectable (such as a musical work), there is no distinction in the thickness of the right to exclude across different expressions. If, within the same superset of works, two kinds of expressions have distinct levels of excludability because of cultural norms, the work with lower levels of excludability or thinner protection would receive much less investment, resulting in much less supply.²²⁹ However if rights are optimally limited across the board to eradicate this fundamental distinction, these distortionary effects would potentially subside without demeaning cultural norms. When defining uniform contours of rights, the focus has to be on less legally excludable creative expressions,²³⁰ as against the highly excludable ones.

Although such an approach reduces the overall value that an owner can externalize out of a single highly excludable work, in parallel, it also expands the *kinds* of expressions that do not require investors to incur significant costs. It potentially increases the *breadth* or range of investment in, and dissemination of, cultural expressions. It also ensures that the social cost of freezing out expressions from various genres is well balanced with copyright's enablement function. The author or investor still gets exclusionary rights, however only to an extent that does not compromise diversity of cultural expressions.²³¹ It also enables a relatively egalitarian starting point for cultural speakers.

The concrete prescriptions I offer are nothing new. In fact, I use the analysis of copyright's *distortionary* effects to further bolster the case of some reformatory prescriptions that legal scholars have already offered. These prescriptions are (1) limiting the right to exclude derivative works only to adaptations in a *different form or medium of expression*²³² and (2) limiting the right to exclude reproductions to "works" (and not elements of works) that involve *unmistakable* overall similarity, or a similar overall aesthetic appeal, and will most probably substitute the original expression's primary market.²³³ Lack of

229. Fisher, *Reconstructing the Fair Use Doctrine*, *supra* note 20, at 1735.

230. See Alpana Roy, *Copyright: A Colonial Doctrine in a Post-colonial Age*, 26 COPYRIGHT REP. 112 (2008) (arguing for focusing law-making from the lens of cultural practices of those long subject to colonialism).

231. The concern that structurally scaling down rights would affect the autonomy of authors and reduce their contractual bargaining power with investors is a vertical issue that this paper does not address. The vertical issue subsists irrespective of these horizontal changes, and the vertical argument is in fact a distraction to weaken claims of structurally limiting exclusionary rights. The vertical issue of bargaining power has to be tackled on its own by providing contractual safeguards and is outside the scope of this Note.

232. Syed & Bracha, *Copyright Rebooted*, *supra* note 5.

233. See *id.*; see also Bracha & Syed, *Copyright's Atom*, *supra* note 6; cf. Craig, *supra* note 6 (providing proposal on similar lines, but minutely distinct on the kind of precedents they rely on and on the intricate filtering of unprotectable elements); Ann Bartow, *Copyright and Creative Copying*, 1 UNIV. OTTAWA L. & TECH. J. 75, 91 (2004).

an overly broad derivative right—one that includes expressions in the same *form*, market and medium as the primary work—would eradicate overt licensing costs in cultural compositions that inherently require re-mixing or use of pre-existing expression. Moreover, reducing the scope of an overly broad reproduction right—one that currently even protects elemental fragments—would eradicate any relative lack of appropriability that exists due to high volume of *scenes a faire* elements, so long as there is potential of aesthetic substitution of the primary market of the work.

I explain these prescriptions here.

1. *Limiting Derivative Markets to Alternate “Forms” of Representation*

The right to exclude derivative works must only cover those adaptations that are in a different medium **or** instrument of expression and perception. An instance of this is the adaptation of a book to a movie, as against another book. The former, **an** adaptation in a different medium of representation, constitutes part of the secondary market of the work which maps onto the derivative right. The latter, a version based on the previous work in the same medium of representation, constitutes part of the primary market which maps onto the reproduction right. Another example is that of a translation, where the content of the work is the same, but the work is presented in an alternate medium or form—a different language.²³⁴ Under this prescription, sequels or prequels, in the same form or medium of representation, are not part of the secondary market of the work.

This prescription can either be employed overtly by clarifying the limited scope of this right in the statute, or by interpreting the current framing properly.

The derivative right under § 106(2) is not supposed to allow excluding transformed uses of any kind, but only *transformed forms* of the original expression, where the same context is represented in a different medium of expression. The interpretation of ‘form’ ought to be limited to mean an alternate or different medium of representation—in other words, an alternate physical embodiment, alternate language, alternate way of presenting the same content, distinct from the medium of the original expression. Any other interpretation, pitting the scope of the derivative right against the

234. *But see* Samuelson, *supra* note 63, at 14–15. Prof. Samuelson argues translation, art reproduction, abridgement, condensation to be in the same medium of representation. However, I conceptualize them to be in a different medium of expression, as the market of a version that is an abridgement, condensation, translation, or art reproduction does not compete with the primary work in the same market. These represent the same content in an alternate medium.

transformative purpose and character exemption under the Copyright Act, seriously undermines First Amendment values that copyright aims to foster, not constrain.²³⁵ It also makes the co-existence of the derivative works right, reproduction right, and the purpose and character of use exemption to copyright infringement—which are all present in the same statute—completely incoherent. All the illustrations provided in the statute: *abridgement, translation, arrangement, condensation, etc.*,²³⁶ which are meant to guide the meaning of what the right is said to encompass,²³⁷ clearly point towards an alternate medium of representation of the same content, rather than an alternate and distinct work presented in the same medium. It also ensures that exclusionary rights do not extend beyond the content of the original work.

Clarifying this position allows expressive forms like digital sampling, Indian classical music, post-modern art, heterodox works and the like, which inherently rely on borrowing for expression, to be equally enabled, not inhibited, works. Looking at investment, it appears that limiting the derivative right ensures that access or licensing costs do not take a disproportionate toll on expressions which inherently involve borrowing pre-existing material or a level of perceptible similarity with another composition. Further, as these expressions are part of genres where similarity and borrowing are cultural norms, any control of a broad derivative market becomes less consequential to the recoupment of investment.²³⁸ Thus, limiting the derivative market of highly excludable works also brings relatively less excludable works on par with the potential of appropriability through exclusionary rights. It also allows ontological borrowing, which as a corollary helps advertise and enable a positive ripple effect on the distribution market of the primary work.²³⁹

An intuitive response to this prescription would be that it significantly reduces foreseeable incentives in broad derivative markets of highly excludable works. However, the incentives rationale justifying broad derivative rights is

235. Rebecca Tushnet, *Copy This Essay: How Fair Use Doctrine Harms Free Speech and How Copying Serves It*, 114 YALE L.J. 535, 555–60 (2004).

236. 17 U.S.C. § 101 (2016). Apart from musical arrangements and art reproduction which incorporate similar compositions presented in a different genre, no other supplementary works find place in the definition. Musical arrangements and art reproductions in a different genre often incorporate significantly distinct content which make them an imperfect adaptation, and hence whether techno versions of a rock song would be a part of this right is a proposition left open for further exploration.

237. See Samuelson, *supra* note 63, at 10.

238. Bracha & Syed, *supra* note 8, at 270–74 (discussing heterodox works).

239. See Schuster et al., *supra* note 76, at 219; see also Arewa, *supra* note 97.

highly overstated²⁴⁰ and subject to immense skepticism,²⁴¹ especially when pit against First Amendment concerns of downstream creation, self-determination, and cultural democracy. As Talha Syed and Oren Bracha argue, in cases of works that are highly successful in primary markets, the additional value that a broad derivative right provides is likely unnecessary to recoup investments or is beyond the optimal enablement that copyright should provide.²⁴² On the other hand, for relatively less successful works, the earning potential through a derivative market is highly unlikely to generate any additional enablement.²⁴³ Thus, the incentive “bang” earned for the access “buck” is unjustified in context of such broad rights that significantly hurt the diversity of downstream expression.²⁴⁴

This prescription, however, is incomplete and needs to be complemented by a second change that limits the scope of the right to exclude reproductions.

2. *De-Fragmenting the Work and Its Primary Market*

The reproduction right protects the primary market of the original work. While analyzing contours of this right, courts use the test of substantial similarity with an inward fragmentation approach²⁴⁵ that provokes a finding of infringement of this right even if merely some elements of the works are similar, despite the whole of the work being aesthetically different.²⁴⁶ Secondly, the test of substantial similarity is extremely vague, which constrains downstream creators who wish to use similar elements from pre-existing works.²⁴⁷ This is what we need to get rid of.

First, the right to exclude reproductions ought to be limited to what Talha Syed and Oren Bracha argue to be copyright’s “Atom”—the perception of the overall work, not its fragments.²⁴⁸ When analyzing substantial similarity courts

240. See Samuelson, *supra* note 63, at 22; see also Lydia Pallas Loren, *The Pope’s Copyright? Aligning Incentives with Reality by Using Creative Motivation to Shape Copyright Protection*, 69 LA. L. REV. 1, 8 (2008); Rebecca Tushnet, *Economies of Desire: Fair Use and Marketplace Assumptions*, 51 WM. & MARY L. REV. 513, 526–27 (2009); Syed & Bracha, *supra* note 26, at 1884–86.

241. Diane Leenheer Zimmerman, *Copyrights as Incentives: Did We Just Imagine That?*, 12 THEORETICAL INQUIRIES L. 29, 30–31 (2011).

242. See Syed & Bracha, *supra* note 26, at 1843.

243. *Id.* at 1907.

244. *Id.*

245. Syed & Bracha, *Copyright’s Atom*, *supra* note 6.

246. Samuelson, *supra* note 107.

247. *But see id.* at 1840–42 (analyzing all the different tests of substantial similarity that prevail in various circuits and showing how circuits mix up dissecting and a holistic comparison of the work, suggesting that dissecting elements ought to precede holistic comparison). This Note argues to the contrary.

248. Syed & Bracha, *Copyright’s Atom*, *supra* note 6.

ought to avoid disintegrating the whole of the work to figure out whether any of its elements have been copied and whether the allegedly infringing work incorporates similar elements.²⁴⁹ This elemental focus significantly expands the meaning of the “work” in the primary market of the original owner to include even independent parts of it.²⁵⁰ The author’s originality is the whole work in context—and does not extend to its parts when they are employed out of the context of the whole expression. Extending protection to such fragments precludes downstream creators from utilizing such fragments as building blocks to re-contextualize and portray an alternate aesthetic vision.

Secondly, the scope of the substantial similarity test ought to be tightened by asking whether the allegedly infringing work potentially substitutes²⁵¹ the overall aesthetic perception of the original work,²⁵² provoking consumers to buy it to have the same aesthetic experience. This is the stage of holistic comparison. If yes, then the court ought to figure out whether parts of the asserted original work are even original or involve volumes of *scenes a faire* elements. If they involve voluminous *scenes a faire* elements, the test of virtual identity replaces the test of substantial similarity.²⁵³ However, if substantial differences are discernable on a holistic comparison, the enquiry must stop right there.²⁵⁴

While analyzing holistic similarity, the court ought to ask whether the defendant’s work is a transparent rephrasing of the original or whether the defendant’s work comes so near the plaintiff’s overall expression to suggest it to be the same to the mind of almost every person seeing it. The question ought to be whether a reasonable spectator or viewer would have the *unmistakable* impression of the subsequent work to be a copy of the original.²⁵⁵

249. *Id.*

250. Bracha, *The Ideology of Authorship*, *supra* note 51, at 238–46.

251. Bracha & Syed, *Copyright Rebooted*, *supra* note 5; To clarify, by substitution, I do not mean market substitution from an antitrust law sense, which focusses more on the economics of a broader market. What I mean is perceptive substitution, where a reasonable consumer of cultural goods would rather buy the alternate representation of the work, because it does not change its aesthetic perception.

252. Rosenblatt, *supra* note 23, at 654–58.

253. *See* Craig, *supra* note 6, at 36–40 (proposing a similar test where the focus of infringement analysis is first on holistically comparing works for total similarity, and only if the works are holistically similar, without substantial differences, then dissect to filter unprotectable elements while comparing works); *see also* Lemley, *Our Bizarre System for Proving Copyright Infringement*, *supra* note 107, at 740 (arguing that the extrinsic-intrinsic test is backwards and in fact should be applied in the opposite order).

254. *See* Craig, *supra* note 6, at 36–40; *see* Rosenblatt, *supra* note 23, at 658.

255. A similar approach of focusing on *unmistakable* similarity as a whole, as against similarity of elements has been adopted by the Supreme Court of India in *R.G. Anand v. Deluxe Films*, (1978) 4 SCC 118, 140–41 (India) (“One of the surest and the safest test to

If there are broad dissimilarities which negate this, it shall not be infringement.²⁵⁶

How does this remedy copyright's *distortionary* effects? First, it avoids inward fragmentation and second, it focuses on overall aesthetic substitutability of the work, which ensures that cultural expressions which inevitably involve perceptive levels of fragmentary similarity due to cultural norms are not burdened *either* by licensing costs or by relative lack of appropriability. If the unit of protection is only the whole of the work, works which inherently contain perceivably similar elements, like compositions in Indian classical music, get the same amount of excludability as other works based on aesthetic substitutability of the whole—significantly erasing distortions that exist due to relatively lower potential of appropriability. This is because the excludability is structurally lowered across all works in that particular category—for instance, musical works. Removing protection of elements or fragments of works ensures that in spite of a work having *scènes a faire* elements, the focus of the inquiry is always on the substitution of its creative aesthetic appeal, *as a whole*.

An additional benefit of limiting the contours of excludability to the levels proposed is that it reduces not only internal distortions but also structural external distortions provoked by an expanding copyright regime. It is undisputed that copyright law has pervasively expanded over time.²⁵⁷ Jessica Litman has referred to this tendency as being similar to the “billowing white goo” which attempts to cover everything possible within its scope.²⁵⁸ Limiting the scope of rights in the way proposed could ensure that resources that are overtly employed into production of homogenous and highly excludable copyrightable works are better put into use where they would otherwise be more socially and economically valuable.²⁵⁹ In other words, resources invested in copyrighted works sometimes might produce greater social returns if invested elsewhere in the economy.²⁶⁰ Limiting the scope of rights could potentially contribute to remedying these external distortions as well.

determine whether or not there has been a violation of copyright is to see if the reader, spectator or the viewer after having read or seen both the works is clearly of the opinion and gets an *unmistakable* impression that the subsequent work appears to be a copy of the original.”)

256. *Id.* at 141.

257. Robert P. Merges, *One Hundred Years of Solicitude: Intellectual Property Law 1900-2000*, 88 CALIF L. REV. 2187, 2188 (2000).

258. Jessica D. Litman, *Billowing White Goo*, 31 COLUM. J.L. & ARTS 587 (2008).

259. Lunney Jr., *supra* note 4, at 655.

260. Abramowicz, *supra* note 99, at 320.

VII. CONCLUSION

The neoliberal discourse around expansive copyrights, through its dissonance with and ignorance of diverse cultural practices, proves its inelasticity in accommodating forms of alterity that are less marketable.²⁶¹ Forms of cultural expression that are ontologically dissonant have a choice to comply and de-legitimize their cultural norms, or be distorted from a distributional space that chooses on the basis of legal excludability and appropriability. Appreciating these distortionary effects justifies fine-tuning copyright doctrine to better enable production and dissemination of diverse cultural expressions. It shows how the law is not a neutral conduit when it comes to enabling social value that can be derived from expressive output.

Musical practices explored above—digital sampling in hip hop musical tradition, perceivable similarity in compositions in the *Raga* system of Indian classical music, and many similar cultural norms that are prevalent across mediums of expression—embrace downstream use and similarity while dismissing exclusion. Either by imposing dissonant legal norms on these practices or by ignoring them while making the law, the market around culture has globally been disembedded from many social and cultural practices.²⁶² This creates inequality in opportunity to participate in cultural discourse and hurts autonomous self-determination—an important purposive end goal of copyright policy.

Thus, limiting the scope of exclusionary rights under the U.S. Copyright Act through statutory and interpretive proposals offered could significantly ensure that copyright, as a legal tool, continues to equitably remain a means to the end of participation and optimal enablement of diverse expressions. By eliminating these distortionary effects, such reform can help ensure that we do not disenfranchise humans by subordinating the cultural consciousness of society to the rules of the market and interests of marketers.²⁶³

261. Rosemary Coombe, *Legal Claims to Culture in and Against the Market: Neoliberalism and the Global Proliferation of Meaningful Difference*, 1 L., CULTURE & HUMANITIES 35, 37–40 (2005).

262. Kannan, *supra* note 180.

263. Timothy Macneill, *The End of Transformation? Culture as the Final Fictitious Commodity*, 12 PROBLÉMATIQUE 17 (2010).

UNDERSTANDING *UNICOLORS*: MISTAKES OF LAW DON'T NECESSARILY INVALIDATE COPYRIGHT REGISTRATION CERTIFICATES

Samantha Cox-Parra[†]

TABLE OF CONTENTS

I.	INTRODUCTION	1249
II.	BACKGROUND.....	1252
	A. COPYRIGHT FORMALITIES AND THE PURPOSE OF COPYRIGHT REGISTRATION.....	1252
	B. SECTION 411.....	1255
III.	<i>UNICOLORS V. H&M</i> CASE SUMMARY.....	1258
	A. FACTUAL AND PROCEDURAL HISTORY	1258
	B. THE SUPREME COURT'S RULING.....	1262
IV.	MISTAKE OR MISUSE?	1265
	A. THE COPYRIGHT TROLL PROBLEM	1265
	B. PROPOSED EFFORTS TO DETER THE TROLL.....	1268
	1. <i>Congress</i>	1269
	2. <i>The Courts</i>	1270
	3. <i>The Copyright Office</i>	1271
V.	CONCLUSION.....	1272

I. INTRODUCTION

Can mistakes of fact or law in an application for copyright registration invalidate a copyright certificate? As many lawyers love to say, it depends. And in this inquiry, the validity of the copyright certificate depends on the copyright owner's mental state when including the inaccuracy on her application for copyright registration.

DOI: <https://doi.org/10.15779/Z38HH6C716>

© 2023 Samantha Cox-Parra.

[†] J.D. 2023, University of California, Berkeley, School of Law. Thank you, Professor Pamela Samuelson and Professor Talha Syed, for your advice and feedback.

In the only intellectual property decision of the October 2021–2022 term, the Supreme Court held in *Unicolors v. H&M* that the Copyright Act’s § 411(b) safe-harbor provision protects the validity of a copyright certificate against good faith mistakes of either fact or law made by the copyright holder in the application for copyright registration.¹ In a relatively short opinion authored by Justice Breyer,² the majority adopted a purposivist interpretation of the statutory language and legislative history of § 411(b) to clarify that exclusion from the safe harbor requires “knowledge,” which means “actual, subjective awareness of both the facts and the law,” rather than a higher “intent-to-defraud” standard.³ While ultimately finding in favor of the copyright owner, Unicolors, Justice Breyer provided an important carve-out that allows courts to look for willful blindness or constructive knowledge of inaccuracies in place of actual knowledge to exclude copyright holders from § 411(b)’s safe harbor protection.⁴

In 2016, Plaintiff Unicolors sued H&M for copyright infringement alleging that H&M’s “Xue Xu” design printed on jackets and skirts infringed Unicolors’ “EH101” textile copyright.⁵ The matter went to trial and the jury found H&M liable for copyright infringement.⁶ However, at the close of trial, H&M asked the court to grant it judgment as a matter of law that Unicolors’ registration should be referred to the Register of Copyrights under § 411(b)(2) to determine the validity of Unicolors’ copyright because Unicolors’ registration contained inaccurate information regarding publication.⁷

The District Court denied H&M’s motion, finding that there was no evidence of intent-to-defraud, and rather that Unicolors simply did not know

1. *Unicolors, Inc. v. H&M Hennes & Mauritz, L.P.*, 142 S. Ct. 941 (2022).

2. *Unicolors v. H&M* was Justice Breyer’s last intellectual property decision before retiring from the Supreme Court. Over the course of his time on the bench, Justice Breyer has shaped copyright law. *See* *Kirtsaeng v. Wiley*, 568 U.S. 519 (2013); *ABC v. Aero*, 573 U.S. 431 (2014); *Google v. Oracle*, 141 S. Ct. 1183 (2021); *MGM v. Grokster*, 545 U.S. 913 (2005) (Breyer, J., joining unanimous opinion by Justice Souter); *Allen v. Cooper*, 140 S. Ct. 994 (2020) (Breyer, J., joining majority opinion); *Eldred v. Ashcroft*, 537 U.S. 186 (2003) (Breyer, J., dissenting); *Golan v. Holder*, 565 U.S. 302 (2012) (Breyer, J., dissenting); *Petrella v. MGM*, 572 U.S. 663 (2014) (Breyer, J., dissenting); *Star Athletica v. Varsity Brands*, 137 S. Ct. 1002 (2017) (Breyer, J., dissenting); *Georgia v. Public Resources*, 140 S. Ct. 1498 (2020) (Breyer, J., joining both Justice Thomas’ and Justice Ginsburg’s dissents); *see also* Stephen Breyer, *The Uneasy Case for Copyright: A Study of Copyright in Books, Photocopies, and Computer Programs*, 84 HARV. L. REV. 281 (1970).

3. *Unicolors v. H&M*, 142 S. Ct. at 947.

4. *Id.* at 948.

5. *Unicolors, Inc. v. H&M Hennes & Mauritz, L.P.*, 959 F.3d 1194, 1195 (9th Cir. 2020).

6. *Unicolors, Inc. v. H&M Hennes & Mauritz, L.P.*, 2018 WL 10307045, at *1 (C.D. Cal. Aug. 1, 2018).

7. *Id.*

that it failed the same unit of publication requirement.⁸ H&M appealed to the Ninth Circuit, which clarified that there was no intent-to-defraud requirement for denial of the safe harbor but found that Unicolors did not benefit from the safe harbor because § 411(b) excuses only mistakes of *facts* not mistakes of *law*.⁹ Unicolors sought certiorari and the Supreme Court granted the petition to address whether mistakes of law were protected by the safe harbor and to clarify the requisite mental state required for denial of the safe harbor.¹⁰

While some scholars argue that *Unicolors v. H&M* was not a monumental Supreme Court case since it resolved a narrow question in which there was not much, if any, disagreement in the lower courts,¹¹ the underlying facts of the case draw attention to the copyright trolling issue lurking in the background of much of copyright litigation. As H&M pointed out, Unicolors has filed hundreds of similar copyright infringement suits against numerous fashion brands.¹² In oral arguments, Justice Sotomayor was the only Justice to raise the copyright troll issue, and in addressing Unicolors' counsel, she asked, "how do I describe a truly innocent mistake of law from one in which a sophisticated party with the capacity to confer with lawyers makes a mistake that they could have easily checked?"¹³ Despite briefing in the case, largely by amici in support of H&M,¹⁴ the Court's opinion does not directly address how sophisticated, professional plaintiffs—or copyright trolls—evade responsibility for their misuse of the copyright registration system, but rather focuses instead on the impact of § 411(b) on relatively unprofessional plaintiffs.

This Note addresses the tension between copyright's fundamental goal to promote accessibility of copyright protections to non-lawyer creatives and the exploitation of the copyright system by copyright trolls. Part II of this Note

8. *Id.*

9. *Unicolors v. H&M*, 959 F.3d at 1200.

10. *Unicolors, Inc. v. H&M Hennes & Mauritz, L.P.*, 141 S. Ct. 2698 (2021).

11. See Jasper L. Tran, Response, *Copyright's Legal Mistake*, GEO. WASH. L. REV. ON DOCKET (Mar. 22, 2022), <https://www.gwlr.org/copyrights-legal-mistake>; see also Ronald Mann, *Justices Require Actual Knowledge That Application Was Erroneous to Invalidate Copyright Filing*, SCOTUSBLOG (Feb. 24, 2022), <https://www.scotusblog.com/2022/02/justices-require-actual-knowledge-that-application-was-erroneous-to-invalidate-copyright-filing/> ("As I noted above, *Unicolors* resolves a narrow question. Indeed, as the dissent points out, the question that the court answers is not even one on which there is any disagreement in the lower courts. There is every reason to think that *Unicolors* will fade from view in the not-so-distant future.")

12. Brief for Respondent at 7, *Unicolors, Inc. v. H&M Hennes & Mauritz, L.P.*, 142 S. Ct. 941 (2022) (No. 20-915).

13. Transcript of Oral Argument at 17, *Unicolors, Inc. v. H&M Hennes & Mauritz, L.P.*, 142 S. Ct. 941 (2022) (No. 20-915).

14. See generally Brief for Professors of Copyright Law as Amici Curiae Supporting Respondent, *Unicolors, Inc. v. H&M Hennes & Mauritz, L.P.*, 142 S. Ct. 941 (2022) (No. 20-915).

focuses on the historical and legal background necessary to understand the Supreme Court's ruling and reasoning. Part III provides a deeper understanding of the Supreme Court's *Unicolors* decision. Lastly, Part IV identifies the underlying copyright troll problem and contemplates some proposed efforts that Congress, the courts, and the Copyright Office can adopt to deter the troll.

II. BACKGROUND

A. COPYRIGHT FORMALITIES AND THE PURPOSE OF COPYRIGHT REGISTRATION

Copyright law in the United States aims to incentivize creation by providing exclusive rights to copyright holders while striking a balance with the public interest of access to these works.¹⁵ The Intellectual Property Clause of the United States Constitution authorizes Congress to “promote the Progress of Science and useful Arts, by securing for limited Times to Authors and Inventors the exclusive Right to their respective Writings and Discoveries.”¹⁶

Historically, Congress sought to fulfill the goals of copyright law by enforcing compliance with the copyright formalities of registration, renewal, notice, and deposit. As explained by legal scholar Christopher Sprigman, “For most of our history, U.S. copyright included a system of procedural mechanisms, referred to collectively as ‘copyright formalities,’ that helped to maintain copyright’s traditional balance between providing private incentives to authors and preserving a robust stock of public domain works from which future creators could draw.”¹⁷ In exchange for the market-based incentives provided to authors to create works, copyright formalities—namely registration, renewal, notice, and deposit—have functioned to inform the public and ensure preservation of the work for future public use.

In 1790, Congress enacted the first copyright statute,¹⁸ which required authors to comply with copyright formalities.¹⁹ Importantly, failure to comply

15. *See Sony Corp. of Am. v. Universal City Studios*, 464 U.S. 417, 429 (1984) (“The monopoly privileges that Congress may authorize are neither unlimited nor primarily designed to provide a special private benefit. Rather, the limited grant is a means by which an important public purpose may be achieved. It is intended to motivate the creative activity of authors and inventors by the provision of a special reward, and to allow the public access to the products of their genius after the limited period of exclusive control has expired.”).

16. U.S. CONST. art. I, § 8, cl. 8.

17. Christopher Sprigman, *Reform(aliz)ing Copyright*, 57 STAN. L. REV. 485, 487 (2004).

18. Copyright Act of May 31, 1790, ch. 15, 1 Stat. 124 (repealed 1947).

19. *See Sprigman, supra* note 17, at 487.

resulted in termination of the copyright.²⁰ In 1886, various countries adopted the Berne Convention for the Protection of Literary and Artistic Works, which made formalities largely voluntary.²¹ To comply with international norms and make copyright less onerous and more obtainable, Congress deformedalized copyright over a series of reforms and legislation. Beginning with the Copyright Act of 1976²² and culminating with the Berne Convention Implementation Act of 1988,²³ the Copyright Renewal Act of 1992,²⁴ and the Sonny Bono Copyright Term Extension Act of 1998,²⁵ “Congress pared back, and in some instances entirely discarded, copyright formalities.”²⁶

Under current copyright law, a work of authorship is protected by copyright from the moment it is created so long as the work is original and fixed in a tangible medium of expression.²⁷ While copyright registration is no longer required for a work to be protected by copyright, registration provides several important benefits to both copyright holders and the public.²⁸ As explained by Robert J. Kasunic, the Associate Register of Copyrights and Director of Registration Policy and Practice for the U.S. Copyright Office: “Although copyright registration as a condition for copyright protection has gradually been eliminated over the past two centuries, the importance of registration and the benefits it bestows has increased.”²⁹ To incentivize registration, Congress created various benefits to authors to register their works with the Copyright Office.³⁰ Timely copyright registration provides

20. *See* *Wheaton v. Peters*, 33 U.S. 591, 593 (1834) (“The security of a copyright to an author, by the acts of congress, is not a technical grant of precedent and subsequent conditions. All the conditions are important: the law requires them to be performed, and, consequently, their performance is essential to a perfect title.”).

21. Berne Convention for the Protection of Literary and Artistic Works, Sept. 9, 1886, as revised at Paris on July 24, 1971 and amended in 1979. S. TREATY DOC. NO. 99-27 (1986).

22. Copyright Act of 1976, Pub. L. No. 94-553, 90 Stat. 2541 (1967) [hereinafter 1976 Act].

23. Berne Convention Implementation Act of 1988, Pub. L. No. 100-568, 102 Stat. 2853 (1988).

24. Copyright Renewal Act of 1992, Pub. L. No. 102-307, 106 Stat. 264 (1992).

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

26. Sprigman, *supra* note 17, at 487. This shift from a “conditional” copyright system to an “unconditional” system in which formalities are largely voluntary, has put a strain on copyright law and the balance it seeks to achieve. Consequently, a movement to re-formalize copyright formalities in such a way that embraces modern technology has emerged.

27. *See* 17 U.S.C. §§ 102(a), 408(a).

28. U.S. COPYRIGHT OFF., COMPENDIUM OF U.S. COPYRIGHT OFFICE PRACTICES § 202 (3d ed. 2021).

29. Robert J. Kasunic, *The Benefits of Registration*, 68 J. COPYRIGHT SOC’Y U.S.A. 83, 85–86 (2020–2021).

30. *Id.* at 84.

copyright holders with the ability to file suit for infringement,³¹ to claim statutory damages and attorney's fees,³² and to use registration as *prima facie* evidence of the validity of the copyright and the facts stated in the certificate.³³

Copyright registration, in addition to filtering out meritless copyright claims that would otherwise overwhelm courts,³⁴ also serves to provide the public with information about the copyrighted work and put potential infringers on notice via the Copyright Office's public records database.³⁵

31. See 17 U.S.C. § 411(a); see also *Fourth Estate Pub. Benefit Corp. v. Wall-Street.com*, 139 S. Ct. 881, 892 (2019) (“[T]he Copyright Act safeguards copyright owners, irrespective of registration, by vesting them with exclusive rights upon creation of their works and prohibiting infringement from that point forward.”).

32. See 17 U.S.C. §§ 412(c), 504–05.

33. See *id.* § 410(c).

34. See George Thuronyi, *The Fourth Estate Decision and Copyright Registration*, LIBR. CONGRESS: COPYRIGHT CREATIVITY AT WORK (Mar. 14, 2019) (“The registration approach is part of Congress’s considered scheme to filter copyright claims through the Copyright Office, resulting in an improved record for the courts as well as the public at large to rely upon.”) <https://blogs.loc.gov/copyright/2019/03/the-fourth-estate-decision-and-copyright-registration/>.

35. See Kasunic, *supra* note 29, at 91–92. In *Unicolors*, amici in support of H&M highlight some shortcomings of the Copyright Office’s public databases, especially as it concerns visual art. See Brief of Amici Curiae California Fashion Association in Support of Respondent at 10, *Unicolors, Inc. v. H&M Hennes & Mauritz, L.P.*, 142 S. Ct. 941 (2022) (No. 20-915) (“The Copyright Office’s online catalog does not include pictures or copies of the registered work itself; only the written information submitted on the application form itself. . . . Theoretically, the only way to be sure that a given fabric design or other work of authorship was not previously registered by someone else would be to put the design in question firmly in mind, and then go through and search the entire Library of Congress collection of deposit copies to look for the proverbial matching needle in one of the world’s largest haystacks.”). Even legal scholar Jane Ginsburg argues, “formalities that condition the existence or enforcement of copyright on supplying information about works of authorship should enable effective title searching, thus furthering the economic interests both of copyright owners and of potential exploiters.” See Jane C. Ginsburg, *The U.S. Experience with Copyright Formalities: A Love/Hate Relationship*, 33 COLUM. J.L. & ARTS 311, 312–13 (2010). Excitingly, the Copyright Office agrees and is currently undertaking various modernization efforts to improve the copyright registration system and redesign the public records database system with an ability to utilize APIs. See Registration Modernization, 85 Fed. Reg. 12704, 12709 (Mar. 3, 2020) (“A copyright system of the twenty-first century demands flexibility, agility, and adaptability to technological advancement The Office believes that the use of APIs—interfaces that permit communication between two systems or software programs—could improve the registration system by enabling programs used in the process of creating works to submit copyright registration applications or extract data from the online public record.”). Considering the differing needs and wants of various types of copyright holders, APIs will allow for copyright holders to utilize the Copyright Office’s records to best satisfy their differing needs. Regarding textile copyrights, “[d]atabases could help a company avoid infringing existing designs or recognize a supplier didn’t create a fabric. Once infringement is alleged, a database could also help attorneys find similar designs to show claimed originality wasn’t actually that creative and

Copyright registration is able to do so because it creates data about a copyrighted work.³⁶ A certificate of registration creates a public record of key facts relating to a work, “including the title of the work, the author of the work, the name and address of the claimant or copyright owner, the year of creation, and information about whether the work is published, has been previously registered, or includes preexisting material.”³⁷

In the registration process, the putative copyright owner provides various facts relevant to the work she is seeking to register in an application for registration.³⁸ Unlike applications for patents or trademarks, the Copyright Office accepts many of the facts stated in the application at face value without investigation or verification.³⁹ Rather, an examiner at the Copyright Office uses her knowledge and expertise to spot problems and inconsistencies that can be addressed prior to issuing a certificate, often corresponding with applicants to remedy such issues.⁴⁰ Notwithstanding, applications with mistakes—intentional or otherwise—may be issued a valid certificate of copyright registration.⁴¹

B. SECTION 411

Under § 411(a), a copyright holder must register her work with the Copyright Office in order to bring a civil action for infringement.⁴² As Justice Ruth Bader Ginsburg explained in *Fourth Estate*, “[i]n enacting 17 U.S.C. § 411(a), Congress both reaffirmed the general rule that registration must precede an infringement suit, and added an exception in that provision’s second sentence to cover instances in which registration is refused.”⁴³ Copyright registration triggers the right to sue.

deserves thin protection[.]” See Kyle Jahner, *Textile Design Copyrights Remain Tricky After High Court Ruling*, BLOOMBERG L. (Feb. 28, 2022), <https://news.bloomberglaw.com/ip-law/textile-design-copyrights-remain-tricky-after-high-court-ruling>.

36. Sprigman, *supra* note 17, at 487.

37. U.S. COPYRIGHT OFF., CIRCULAR 1: COPYRIGHT BASICS 5 (2021), <https://www.copyright.gov/circs/circ01.pdf>.

38. See generally U.S. COPYRIGHT OFF., CIRCULAR 2: COPYRIGHT REGISTRATION (2021), <https://www.copyright.gov/circs/circ02.pdf>.

39. Kasunic, *supra* note 29, at 89.

40. *Id.* at 90–93 (“Given the 25% correspondence rate annually, the interaction with examiners as intermediaries in the registration process plays an important role in improving the public record and resolving many issues prior to the issuance of a certificate of registration.”).

41. See generally U.S. COPYRIGHT OFF., *supra* note 38.

42. Registration, or a refusal of registration, by the Copyright Office is a prerequisite to filing a lawsuit for copyright infringement involving a U.S. work. See 17 U.S.C. § 411(a).

43. *Fourth Estate Public Benefit Corp. v. Wall-Street.com*, 139 S. Ct. 881, 890–91 (2019) (citing to H.R. REP. NO. 94-1476, at 157 (1976)); see also *Petrella v. Metro-Goldwyn-Mayer*,

Congress chose to relax the registration formality of § 411 by enacting § 411(b) as a safe harbor to protect copyright holders.⁴⁴ In 2008, Congress enacted the Prioritizing Resources and Organization for Intellectual Property Act⁴⁵ (Pro-IP Act) with the intention of making “a number of changes to copyright and trademark law that [would] enhance the ability of intellectual property rights holder to enforce their rights.”⁴⁶ The Pro-IP Act amended the Copyright Act to add § 411(b) as a rule of “Harmless Error,” which provides a safe harbor for copyright holders against invalidation from mistakes made on an application for registration.⁴⁷ The safe harbor amendment was largely motivated by Congress’s desire “to prevent intellectual property thieves from exploiting [a] potential loophole” by “argu[ing] in litigation that a mistake in the registration documents, such as checking the wrong box on the registration form, renders a registration invalid[.]”⁴⁸

Under the § 411(b)(1) safe harbor, a copyright holder’s certificate of registration is valid regardless of whether the certificate contains any inaccurate information, unless:

(A) the inaccurate information was included on the application for copyright registration *with knowledge that it was inaccurate*; and

(B) the inaccuracy of the information, if known, would have caused the Register of Copyrights to refuse registration.⁴⁹

If it is alleged that the copyright holder knowingly included the inaccuracy on the application, § 411 (b)(2) is triggered. Under § 411(b)(2):

[T]he court shall request the Register of Copyrights to advise the court whether the inaccurate information, if known, would have caused the Register of Copyrights to refuse registration.⁵⁰

Inc., 132 S. Ct. 1962, 1977 (2014) (“Although registration is ‘permissive,’ both the certificate and the original work must be on file with the Copyright Office before a copyright owner can sue for infringement.”); *Alaska Stock, LLC. v. Houghton Mifflin Harcourt Publ’g Co.*, 747 F.3d 673, 67 (9th Cir. 2014) (“Though an owner has property rights without registration, he needs to register the copyright to sue for infringement.”).

44. Grace Pyun, *2008 Pro-IP Act: The Inadequacy of the Property Paradigm in Criminal Intellectual Property Law and Its Effect on Prosecutorial Boundaries*, 19 DEPAUL J. ART TECH. & INTELL. PROP. L. 355, 375 (2009).

45. Prioritizing Resources and Organization for Intellectual Property Act of 2008, Pub. L. No. 110-403, 122 Stat. 4256 (2008).

46. H.R. REP. NO. 110-617, at 23 (2008).

47. Pro-IP Act § 101(a).

48. H.R. REP. NO. 110-617, at 24, n.15 (2008) (citing *In re Napster, Inc.*, 191 F. Supp. 2d 1087, 1099 (N.D. Cal. 2002)).

49. 17 U.S.C. § 411(b)(1) (emphasis added).

50. *Id.* § 411(b)(2).

Put simply, if a copyright holder made a mistake, her registration is protected by the safe harbor. However, if she did not make a mistake, her registration is at risk of invalidation.

What is the difference between a mistake and a non-mistake? In other words, what mental state would satisfy the “with knowledge that it was inaccurate” requirement: fraud, intentional deception, actual knowledge, constructive knowledge, or willful blindness? Fraud and intentional deception would require that the copyright holder be aware of the inaccuracy and choose to include the inaccuracy on the application for copyright registration in order to trick the Copyright Office into granting a copyright that likely should not be granted. Actual knowledge would require that the copyright holder was aware of the inaccuracy and chose to include it in the application but did not have any sort of intention to game the Copyright Office. Constructive knowledge, on the other hand, would not require that the copyright holder was aware of the inaccuracy, but rather that she should have been aware. Like constructive knowledge, willful blindness would not require that the copyright holder be aware of inaccuracy, but rather that she chose to keep herself unaware and in the dark as to whether the inaccuracy was inaccurate or not.

Some courts, scholars, and even the Register of Copyrights interpreted the “with knowledge that it was inaccurate” requirement of § 411(b)(1)(A) to mean that the Pro-IP Act amended § 411 of copyright law to codify the doctrine of “fraud on the Copyright Office”⁵¹ in the registration process.⁵² Under this assumption, a showing of intentional deception or fraud would be required for exclusion from the safe harbor, thereby risking invalidation of the copyright registration by the Register of Copyrights. Yet, there remained an apparent split among scholars and the circuit courts as to the request mental state required by § 411(b)(1)(A) for exclusion of the safe harbor protections. According to the Eleventh Circuit in *Roberts v. Gordy*, § 411(b)(1)(A) requires “deceptive intent,” whereas the Ninth Circuit held in *Gold Value International*

51. The doctrine of fraud on the Copyright Office developed by way of cases that upheld the validity of copyright registrations where inadvertent or immaterial errors were made. For example, in *Advisers, Inc. v. Wiesen-Hart, Inc.*, the Sixth Circuit held that “an innocent misstatement, or a clerical error, in the affidavit and certificate of registration, *unaccompanied by fraud* or intent to extend the statutory period of copyright protection, does not invalidate the copyright, nor is it thereby rendered incapable of supporting an infringement action.” 238 F.2d 706, 708 (6th Cir. 1956) (emphasis added). Overtime, the converse proposition that a registration may be invalidated by fraud birthed the doctrine and defense of fraud on the Copyright Office.

52. U.S. COPYRIGHT OFF., ANNUAL REPORT OF THE REGISTER OF COPYRIGHTS 13 (2008), <https://www.copyright.gov/reports/annual/2008/ar2008.pdf>.

Textile v. Sanctuary Clothing that “there is no such intent-to-defraud requirement.”⁵³

If intentional deception or fraud is not required for exclusion from the safe harbor, then what lesser mental state is sufficient to exclude a copyright holder from the safe harbor? The clarification of the requisite mental state in § 411(b)(1)(A) is crucial because it triggers § 411(b)(2), in which a court shall refer the certification issue to the Register of Copyrights.⁵⁴ If the Register of Copyrights determines that, had she known of the inaccuracy at the time, she would still have granted registration, the certification of registration would be protected by the safe harbor. However, if the Register of Copyrights determines that she would have refused registration, the certification of registration is invalidated, likely allowing the defendant to escape liability for copyright infringement.⁵⁵ The lower the mental state required, the easier it is for infringing defendants to argue that a copyright holder’s mistake in her registration should not be protected by the safe harbor and her application should be reexamined by the Register of Copyrights to determine whether the copyright certificate should be invalidated. The higher the mental state requirement, the more difficult it would be to ultimately invalidate a copyright holder’s registration certificate.

III. *UNICOLORS V. H&M* CASE SUMMARY

A. FACTUAL AND PROCEDURAL HISTORY

Unicolors and H&M are no strangers to the courtroom, and neither party is particularly sympathetic.⁵⁶ Unicolors is a Los Angeles, California-based company in the business of creating, purchasing, and obtaining copyrights to graphic artworks that are printed on fabrics and sold to fashion brands.⁵⁷ As H&M noted:

53. *See* *Roberts v. Gordy*, 877 F.3d 1024, 1030 (11th Cir. 2017); *Gold Value Int’l Textile, Inc. v. Sanctuary Clothing, L.L.C.*, 925 F.3d 1140, 1147 (9th Cir. 2019).

54. *See* 17 U.S.C. § 411.

55. This determination by the Register of Copyrights is a bit more complicated than Congress may have anticipated. In practice, an examiner at the Copyright Office communicates with applicants to remedy any issues in their application to approve certification rather than outright refusing the application as is. In practice, an examiner refuses registration if the applicant does not respond in a timely manner or refuses to correct the inaccuracy. *See* Kasunic, *supra* note 29, at 90.

56. *Unicolors v. H&M: A Fast Fashion Copycat and Alleged ‘Copyright Troll’ Go to Trial*, FASHION L. (Dec. 7, 2017), <https://www.thefashionlaw.com/a-fast-fashion-copycat-and-an-alleged-copyright-troll-are-currently-at-trial/>.

57. *Unicolors v. H&M*, 959 F.3d at 1195.

“A cursory PACER search shows that Unicolors has filed literally hundreds of these lawsuits. Its victims include Amazon, Bass Pro, Bloomingdale’s, Burlington Stores, Century 21, Dillard’s, The Dress Barn, JCPenney, Kmart, Kohl’s, Lord & Taylor, Macy’s, Neiman Marcus, Nordstrom, Ross, Saks, Sears Roebuck, The TJX Companies, Urban Outfitters, and Walmart.”⁵⁸

Very few of Unicolors’ cases have gone to trial, and most have settled out of court, ahead of trial, “largely because it tends to be much cheaper and more expeditious to settle a case than it is to finance and fight through a trial.”⁵⁹ Unsurprisingly, Unicolors has been called a copyright troll.⁶⁰ H&M, on the other hand, is a large international clothing retailer that is widely considered a fast-fashion⁶¹ copycat.⁶² H&M has been on the receiving end of numerous copyright infringement suits and allegations, largely from fashion designers and labels.⁶³

In 2011, Unicolors registered thirty-one fabric designs in a single registration, the ’400 Registration, under the Copyright Office’s same unit of publication exception.⁶⁴ The same unit of publication exception allows for works packaged together and published on the same date to be registered in a single application rather than individual applications.⁶⁵ In their application for registration, Unicolors listed January 15, 2011, as the publication date for all of

58. Brief for Respondent at 7, *Unicolors, Inc. v. H&M Hennes & Mauritz, LP.*, 142 S. Ct. 941 (2022) (No. 20-915).

59. *A Fast Fashion Copycat*, *supra* note 56.

60. Brief for Respondent at 6, *Unicolors v. H&M.*, 142 S. Ct. 941 (2022) (No. 20-915); *A Fast Fashion Copycat*, *supra* note 56; Noah Smith, *Are Copyright Trolls Taking Over the Fashion Industry?*, *FORTUNE* (Oct. 7, 2015), <https://fortune.com/2015/10/07/patent-trolls-fashion/>.

61. A class action lawsuit was recently filed against H&M, alleging that H&M is “greenwashing” or engaging in false and misleading marketing regarding the sustainability of its clothing. *See generally* Complaint, *Commodore v. H&M Hennes & Mauritz L.P.*, 7:2022-cv-06247, (S.D.N.Y. filed July 22, 2022).

62. *See* Reply Brief for Petitioner at 2, *Unicolors, Inc. v. H&M Hennes & Mauritz, L.P.*, 142 S. Ct. 941 (2022) (No. 20-915).

63. Fashion clothing designs are not protectable under U.S. copyright laws. However, there have been attempts to secure stronger legal protection for fashion. For example, the Council of Fashion Designers of America tried to pass the Innovative Design Protection Act of 2012, which sought to provide designers a three-year period during which designs could be protected, so long as they went through a rigorous process to prove they were “novel” and had never existed before; the bill was never brought to a vote. *See* S.3523, 112th Cong. (2012).

64. Registration No. VA 1-770-400 (“the ’400 Registration”).

65. 37 C.F.R. § 202.3(b)(4) (2020) (“For the purpose of registration on one application and upon the payment of one filing fee, the following shall be considered one work: In the case of published works, all copyrightable elements that are otherwise recognizable as self-contained works, that are included in the same unit of publication, and in which the copyright claimant is the same.”).

the thirty-one works.⁶⁶ When questioned about the '400 Registration at trial, Unicolors' President Nader Pazirandeh testified that Unicolors submits collections of works in a single copyright registration "for saving money."⁶⁷ Within the '400 Registration, Unicolors included the "EH101" design,⁶⁸ which it claims was a work for hire made by Hannah Lim, a designer for Unicolors.⁶⁹

In 2016, Unicolors filed suit against H&M for copyright infringement alleging that H&M's "Xue Xu" design printed on jackets and skirts infringed Unicolors' EH101 copyright.⁷⁰ The matter went to trial and the jury awarded Unicolors \$817,920 in disgorgement damages and \$28,800 in lost profit damages.⁷¹ At the close of trial, H&M asked the District Court to grant it judgment as a matter of law that Unicolors' registration should be referred to the Register of Copyrights because it contained inaccurate publication information.⁷² Specifically, the '400 Registration contained inaccurate publication information because only twenty-two of the thirty-one designs were made available to the public on January 15, 2011, while the remaining

66. *Unicolors v. H&M*, 959 F.3d at 1196.

67. *Id.*

68. The EH101 design was one of the twenty-two designs that was made available to the public. Although the referral to the Register of Copyrights was stayed, this fact could have potentially been key in the determination of whether the Register of Copyrights would have refused or granted the registration.

69. Joint Appendix at 21, *Unicolors, Inc. v. H&M Hennes & Mauritz, LP.*, 142 S. Ct. 941 (2022) (No. 20-915). Interestingly, in 2015, H&M acquired a Chinese copyright for the allegedly infringing "Xue Xu" design from Shaoxing County DOMO Apparel Co., Ltd. with a publication date of June 18, 2014. *Id.* at 33. H&M's Chinese "Xue Xu" copyright was the same design claimed in Unicolors' EH101 copyright. In its motion for judgment as a matter of law, H&M contended that while the court granted judicial notice of the Chinese copyright registration, the jury was not instructed properly on the presumptions that flow from the copyright registrations for Xue Xu. *Id.* at 94. Ultimately, the District Court held, "Even if the Chinese copyright registration did create a presumption of originality, the Court still did not err in refusing to instruct the jury about such a presumption because H&M LP failed to establish a connection between the Chinese copyright registration and its own garments. The Court took judicial notice of the Chinese Xue Xu registration, which established the registration's existence. But H&M LP presented no testimony or documents whatsoever about the origin of the design on its own garments H&M LP's failure to establish any connection between the Chinese Xue Xu registration and its own design rendered the registration irrelevant." *Id.* at 187.

70. *Unicolors v. H&M*, 959 F.3d at 1195.

71. *Unicolors v. H&M*, 2018 WL 10307045, at *1.

72. It appears that H&M requested referral to the Register of Copyrights as a last-ditch effort to escape copyright infringement liability in the hopes that the Register of Copyright would find that Unicolors' '400 Registration should have been refused. Consequently, this would strike down the jury's finding of copyright infringement and damages award against H&M. This is exactly the kind of action that Congress intended to protect copyright holders against when enacting the § 411(b) safe harbor.

nine were only presented to Unicolors' salespeople, not the public.⁷³ A work is published when it is offered to the public.⁷⁴ The '400 Registration contained published and unpublished works in violation of the same unit of publication exception.

The District Court denied H&M's motion for judgment as a matter of law, finding that the publication inaccuracy did not invalidate Unicolors' registration since there was no evidence that Unicolors intended to defraud the Copyright Office.⁷⁵ Rather, the District Court reasoned that Unicolors did not know that it had failed to satisfy the same unit of publication requirement because it provided the inaccurate information in the application *without* "knowledge that it was inaccurate."⁷⁶

H&M then appealed to the Ninth Circuit. The Ninth Circuit clarified that there was no intent-to-defraud requirement for denial of the safe harbor protection given its recent ruling in *Gold Value*.⁷⁷ Under a view that § 411(b) excused only good-faith mistakes of facts, not mistakes of *law*, the Ninth Circuit held that Unicolors' mistake as to the same unit of publication requirement was not protected by the safe harbor.⁷⁸ The Ninth Circuit struck down the infringement claims and damages awards against H&M and remanded to the District Court to submit an inquiry to the Register of Copyrights asking whether she would have refused the registration if she had known of the inaccuracy at the time the application for registration was submitted.⁷⁹

Following the Ninth Circuit's ruling, the District Court referred the certification issue to the Register of Copyrights for advice, but Unicolors successfully requested a stay of the response from the Copyright Office in light of the Supreme Court's decision to review the case.⁸⁰ The Register of Copyrights did not issue a response to the request.

73. *Unicolors v. H&M*, 959 F.3d at 1196.

74. See 17 U.S.C. § 101 (defining publication as "the distribution of copies or phonorecords of a work to the public by sale or other transfer of ownership, or by rental, lease, or lending").

75. *Unicolors v. H&M*, 2018 WL 10307045, at *3–4.

76. *Id.*

77. *Unicolors v. H&M*, 959 F.3d at 1198; see *Gold Value*, 925 F.3d at 1147.

78. *Unicolors v. H&M*, 959 F.3d at 1200 (emphasis added).

79. *Id.* at 1200–01.

80. Joint Appendix at 220, *Unicolors, Inc. v. H&M Hennes & Mauritz, L.P.*, 142 S. Ct. 941 (2022) (No. 20-915).

B. THE SUPREME COURT'S RULING

Unicolors sought certiorari and the Supreme Court granted the petition to address the circuit split regarding the requisite mental state required in § 411(b)(1)(A).⁸¹ As aforementioned, under the Eleventh Circuit's view in *Roberts v. Gordy*, § 411(b)(1)(A) requires "deceptive intent," such that an accidental or innocent mistake of fact or law included in an application for registration would not invalidate the copyright.⁸² In comparison, under the Ninth Circuit's new interpretation in *Unicolors*, § 411(b)(1)(A) does not require an intent-to-deceive and mistakes of *fact* may be protected by the safe harbor, but any mistake of *law* on the application for registration would invalidate the copyright.⁸³ Initially, Unicolors posed the question presented to the Court as:

Did the Ninth Circuit err in breaking with its own prior precedent and the findings of other circuits and the Copyright Office in holding that 17 U.S.C. § 411 requires referral to the Copyright Office where there is no indicia of fraud or material error as to the work at issue in the subject copyright registration?⁸⁴

However, the question was refined and briefed on the merits as:

Whether that 'knowledge' element precludes a challenge to a registration where the inaccuracy resulted from the applicant's good-faith misunderstanding of a principle of copyright law?⁸⁵

Although Justice Breyer stated that the question whether "knowledge" required an "indicia of fraud" was a "subsidiary question fairly included" in the question presented to the Court,⁸⁶ the dissent, written by Justice Thomas, disagreed and argued that the case should have been dismissed as improvidently granted.⁸⁷

Notwithstanding, the Supreme Court held that a copyright registration is protected by the § 411(b) safe harbor if the copyright holder did not have actual knowledge that she included inaccurate information in her registration, regardless of whether the inaccuracy stemmed from either a mistake of fact or

81. *Unicolors v. H&M*, 142 S. Ct. at 945.

82. *See Roberts v. Gordy*, 877 F.3d 1024, 1030 (11th Cir. 2017).

83. *Unicolors v. H&M*, 959 F.3d at 1200 (emphasis added).

84. Eileen McDermott, *Justices Express Frustration Over Question Presented in Unicolors v. H&M, But Lean Toward Preserving Copyright Registrations*, IP WATCHDOG (Nov. 8, 2021), <https://ipwatchdog.com/2021/11/08/justices-express-frustration-question-presented-unicolors-v-hm-lean-toward-preserving-copyright-registrations/id=139686/>.

85. Unicolors Brief at I, *Unicolors, Inc. v. H&M Hennes & Mauritz, LP.*, 142 S. Ct. 941 (2022) (emphasis added).

86. *Unicolors v. H&M*, 142 S. Ct. at 949.

87. *Id.* at 952 (Thomas, J., dissenting).

of law.⁸⁸ Justice Breyer employed a purposivist interpretation of § 411(b)(1) to hold that “knowledge” refers to “actual, subjective awareness of both the facts and law.”⁸⁹ Relying on the House Report, Justice Breyer explained that “Congress enacted § 411(b) to make it easier, not more difficult, for nonlawyers to obtain valid copyright registrations.”⁹⁰ Congress intended to “eliminat[e] loopholes that might prevent enforcement of otherwise validly registered copyrights,” to prevent copyright infringers from escaping liability based on a technicality.⁹¹ Considering this history, the Court found that “it would make no sense if § 411(b) left copyright registrations exposed to invalidation based on applicants’ good-faith misunderstandings of the details of copyright law.”⁹²

To illustrate his reasoning, Justice Breyer provided a helpful birdwatching analogy:

A brief analogy may help explain the issue we must decide. Suppose that John, seeing a flash of red in a tree, says, “There is a cardinal.” But he is wrong. The bird is not a cardinal; it is a scarlet tanager. John’s statement is inaccurate. But what kind of mistake has John made?

John may have failed to see the bird’s black wings. In that case, he has made a mistake about the brute facts. Or John may have seen the bird perfectly well, noting all of its relevant features, but, not being much of a birdwatcher, he may not have known that a tanager (unlike a cardinal) has black wings. In that case, John has made a labeling mistake. He saw the bird correctly, but does not know how to label what he saw. Here, Unicolors’ mistake is a mistake of labeling. But unlike John (who might consult an ornithologist about the birds), Unicolors must look to judges and lawyers as experts regarding the proper scope of the label “single unit of publication.” The labeling problem here is one of law. Does that difference matter here? . . . We think it does not.⁹³

In other words, Unicolors’ misunderstanding of the legal requirements of publication meant that Unicolors did not include the publication inaccuracy in its application “*with knowledge that it was inaccurate*,” and as such Unicolors was protected by the safe harbor and its registration remained valid.⁹⁴

88. *Id.* at 944.

89. *Id.* at 947.

90. *Id.* at 948.

91. *Id.* (citing H.R. REP. NO. 110-617, at 20 (2008)).

92. *Id.*

93. *Id.* at 946 (internal citation omitted).

94. *Id.* at 947 (quoting § 411(b)(1)(A)).

In addressing arguments that copyright holders will too easily claim lack of knowledge to avoid the consequences of inaccurate applications, Justice Breyer provided an important carve-out in which willful blindness or constructive knowledge of inaccuracies can satisfy the “with knowledge that it was inaccurate” requirement of § 411(b).⁹⁵ Justice Breyer explained that “courts need not automatically accept a copyright holder’s claim that it was unaware of the relevant legal requirements of copyright law.”⁹⁶ Courts may look at “[c]ircumstantial evidence, including the significance of the legal error, the complexity of the relevant rule, the applicant’s experience with copyright law, and other such matters,” to find that an applicant was actually aware of, or willfully blind to, the legal inaccuracy.⁹⁷

Interestingly, the Supreme Court asked, “But did Unicolors *know* about this inaccuracy?”⁹⁸ Although Justice Breyer did not directly address the copyright troll issue in his opinion, perhaps H&M served as an unsympathetic sacrificial lamb that allowed him to inexplicitly address the tension between copyright’s fundamental goals to promote accessibility of copyright protections to non-lawyer creatives and the exploitation of the copyright system by copyright trolls. The Court could have made it harder for defendants to exclude copyright holders from safe harbor protections, but it chose not to. The Court rejected the higher intent-to-defraud the Copyright Office standard in favor of a lower actual knowledge standard to exclude a copyright holder from the § 411(b) safe harbor while basing its reasoning in Congress’ desire to make it easier for non-lawyer creatives to obtain valid copyrights. This appears counterintuitive, but an actual knowledge standard with a constructive knowledge or willful blindness carve-out likely will not harm the non-lawyer

95. *Id.* at 948.

96. *Id.*

97. *Id.* (citing to *Intel Corp. Inv. Pol’y Comm. v. Sulyma*, 140 S. Ct. 768, 778–79 (2020)).

98. *Id.* at 946. On remand the Ninth Circuit stated that “under the correct reasoning of the safe-harbor provision in light of the Supreme Court’s ruling . . . a court’s § 411(b) finding regarding a registrant’s lack of intent to defraud is also a § 411(b) finding regarding the registrant’s lack of knowledge that his copyright application contained inaccuracies—factual or legal.” *See Unicolors, Inc. v. H&M Hennes & Mauritz, L.P.*, 52 F.4th 1054, 1069 (9th Cir. 2022). As a result, the Ninth Circuit relied on the district court’s express conclusion that H&M failed to make “any showing that Unicolors intended to defraud the Copyright Office,” which meant that “the district court determined that Unicolors lacked knowledge that it submitted inaccuracies with its application and as a result that its ’400 Registration is entitled to the safe-harbor provision’s protection.” *Id.* Despite Unicolors’ familiarity with the copyright registration system, experience with copyright law, and initiation of hundreds of suits alleging infringement, the Ninth Circuit did not make an inquiry into whether Unicolors was willfully blind to the legal requirements of the same unit of publication exception, because “this case does not present a context where Unicolors has taken a legal position that egregiously misapplies a clear statute.” *Id.*

creative earnestly applying for and asserting her rights, however it does serve to protect alleged infringers from claims that likely should not have been brought in the first place by copyright trolls.

IV. MISTAKE OR MISUSE?

A. THE COPYRIGHT TROLL PROBLEM

While patent trolls have received attention from intellectual property scholars, policy makers, and courts for at least two decades, copyright trolls have lurked in the background of the majority of intellectual property discussions until fairly recently.⁹⁹ “Recent empirical studies show the field of copyright litigation is increasingly being overtaken by ‘copyright trolls’”¹⁰⁰ Likewise, courts have begun to recognize “the challenge in administering intellectual property law to discourage so-called intellectual property ‘trolls’ while protecting genuine creativity.”¹⁰¹

Defining what makes a copyright holder a troll is inevitably controversial.¹⁰² While a clear definition of what or who is a copyright troll as opposed to an overzealous copyright owner may be appealing, a bright line rule or rigid definition is too restrictive as the business model and strategy of copyright trolls evolves with changes to the law and copyright system. Rather, a more fluid and descriptive concept of copyright trolls and trolling draws attention to the different methods and strategies that are employed in practice.¹⁰³ Notwithstanding, if there is any unifying characteristic of a troll, it is that they are “systematic opportunists.”¹⁰⁴

The concept of copyright trolling generally involves an opportunistic plaintiff that is more focused on the business of litigation to enforce copyrights than the creation or licensing of works.¹⁰⁵ Copyright trolls “recognize[d] the existence of a potentially lucrative business model hidden within the contours

99. Matthew Sag, *Copyright Trolling, An Empirical Study*, 100 IOWA L. REV. 1105, 1107 (2014) (providing a detailed empirical and doctrinal study of copyright trolling showing the astonishing rate of growth of multi-defendant John Doe litigation associated with allegations of infringement concerning BitTorrent file sharing and pornographic films).

100. *Malibu Media, L.L.C. v. Doe*, 2015 WL 4092417, at *2 (S.D.N.Y. July 6, 2015).

101. *Design Basics L.L.C. v. Lexington Homes, Inc.*, 858 F.3d 1093, 1096 (7th Cir. 2017).

102. Sag, *supra* note 99, at 1108.

103. There are various subsets of copyright trolls, such as multi-defendant John Doe file sharing trolls, paparazzi trolls, music sampling trolls, fabric textile trolls, and more. For a brief history and evolution of the copyright troll concept, see James DeBriyn, *Shedding Light on Copyright Trolls: An Analysis of Mass Copyright Litigation in the Age of Statutory Damages*, 19 UCLA ENT. L. REV. 79, 86–90 (2012).

104. Sag, *supra* note 99, at 1113.

105. *Id.* at 1107.

of the Copyright Act,” and crept into U.S. copyright law.¹⁰⁶ The copyright registration system is a lenient honor system that trusts the accuracy of the applicant.¹⁰⁷ As clarified by the Court in *Unicolors*, the Copyright Act’s § 411(b) safe-harbor provision protects copyright holders from having their registration invalidated due to innocent mistakes of fact or law. However, the congressional intent to make valid copyrights more accessible to non-lawyer creatives unintentionally opened the door to exploitation by copyright trolls.

Much like the business model of patent trolls, copyright trolls typically operate by creating or acquiring copyrights, seeking out actual or potential infringers, and commencing or threatening litigation with the goal of forcing monetary settlement or an award of damages.¹⁰⁸ As explained by legal theorist Shyamkrishna Balganesh, “To the copyright troll, the substance of an

106. Shyamkrishna Balganesh, *The Uneasy Case Against Copyright Trolls*, 86 S. CAL. L. REV. 723, 738 (2013). Harry Wall, an Englishman from the 1800s, is often named as “the world’s first ‘copyright troll.’” In 1842, the United Kingdom’s 1833 Dramatic Copyright Act was extended to cover musical compositions. As the myth goes, Wall, husband to comic singer Annie Adams, quickly understood and took advantage of the 1842 Act’s new opportunities and developed a business in which he obtained a power of attorney from the assignee of deceased composers to collect statutory fees for unauthorized performances of songs by sending demand letters to performers. It was reported that “ladies would be discouraged from singing songs in public for fear of receiving letters from agents such as Wall demanding money.” Unsurprisingly, the U.K. musical scene of the late 1800s and the Royal Copyright Commission were outraged by Wall’s exploitation of statutory penalties. Music publishers raised the matter before the Royal Copyright Commission, to which Thomas Chappell said he had refused to deal with Wall because he ‘did not like the character of the man or the character of the proceedings,’ later adding that such things were done by ‘people who do not care anything for the work or anything else, all they want is the money they can get.’ See Isabella Alexander, *Neither Bolt nor Chain, Iron Safe nor Private Watchman, Can Prevent the Theft of Words: The Birth of the Performing Right in Britain*, in PRIVILEGE AND PROPERTY: ESSAYS ON THE HISTORY OF COPYRIGHT 321, 339 (Ronan Deazley et al. eds., 2010) (quoting Royal Commission on Laws and Regulations relating to Home, Colonial and Foreign Copyrights (1878) 24 Parliamentary Papers [C.2036] at pp. 106, 109 & 115).

107. See generally Kasunic, *supra* note 29, at 83 (discussing the benefits of copyright registration as well as the process by which Copyright Office examiners review applications for registration).

108. See Balganesh, *supra* note 106, at 732–33. In the early 2010s, a Nevada-based company called Righthaven LLP discovered a fatally fruitful business model that followed three simple steps: (1) recruit content owners, principally newspapers; (2) identify plausible cases of copyright infringement, such as the reposting of newspaper articles on blogs; and (3) acquire a partial assignment of copyright that is tailored precisely to the infringement identified in step two. This model worked well for some time, but those “assignments” were subject to a secret “Strategic Alliance Agreement,” meaning that Righthaven possessed nothing more than a right to sue. Righthaven flew too close to the sun because an agreement transferring the right to sue without any of the copyright owner’s other exclusive rights is ineffectual. Ultimately, Righthaven’s suits were dismissed and the firm went under from legal fees. See Sag, *supra* note 99, at 1111–13.

individual claim matters much less than its aggregate returns from the enforcement of multiple claims. This explains why it is able to settle each claim for amounts much lower than the damages it seeks. Yet, when aggregated together, the settlements prove to be beneficial.”¹⁰⁹ Rather than pursue actual damages—which are often difficult to calculate—the Copyright Act allows copyright owners to pursue statutory damages in an infringement action.¹¹⁰

More often than not, a copyright troll’s business model complies with all of copyright’s formal rules as a matter of law, which places courts in muddy waters when trying to balance the core aims of copyright law and policy.¹¹¹ As discussed above, post-Berne, copyright formalities are largely voluntary; however, Congress still sought to encourage registration by providing legal and market-based benefits.¹¹² While there is debate whether copyright’s fundamental theory of market incentives does in fact drive creativity or not,¹¹³ in practice, copyright law provides economic incentives, namely registration, to which copyright trolls have latched on. Copyright trolls are opportunistic and their business strategies exploit loopholes and leniencies in the law.¹¹⁴

In the early 2000s, copyright trolls “found their way into the fashion industry after their predecessors encountered success in bringing ‘strike suits’ in other fields.”¹¹⁵ Copyright infringement suits over print designs on garments have been brought by textile companies largely based in Los Angeles, California.¹¹⁶ This model of copyright textile trolling is particularly attractive “[b]ecause U.S. copyright law allows for copyright claims to target all parties

109. Balganes, *supra* note 106, at 765; *see also* Sag, *supra* note 99, at 1108 (“The paradigmatic troll plays a numbers game in which it targets hundreds or thousands of defendants, seeking quick settlements priced just low enough that it is less expensive for the defendant to pay the troll rather than defend the claim.”).

110. 17 U.S.C. § 504(c)(1).

111. Balganes, *supra* note 106, at 780.

112. Dotan Oliar, Nathaniel Pattison, & K. Ross Powell, *Copyright Registrations: Who, What, When, Where, and Why*, 92 TEX. L. REV. 2211 (2014) (presenting a systematic study of the registration records at the U.S. Copyright Office that used an original data set containing all 2.3 million registrations from 2008 to 2012 focused on understanding who is registering what, where, when, and why.).

113. *See generally* Diane Leenheer Zimmerman, *Copyrights as Incentives: Did We Just Imagine That?*, 12 THEORETICAL INQUIRIES L. 29 (2011) (challenging the notion that copyright protections are needed to provide authors with the necessary economic incentives to create).

114. Whether or not copyright litigation as a content-independent revenue stream results in frivolous lawsuits or helps plaintiffs assert their rights is a hotly contested debate. *See* Shyamkrishna Balganes & Jonag B. Gelbach, *Debunking the Myth of the Copyright Troll Apocalypse*, 101 IOWA L. REV. 43, 48–49 (2016).

115. *Id.*; Charles E. Colman, *The History and Doctrine of American Copyright Protection for Fashion Design: Managing Mazer*, 7 HARV. J. SPORTS & ENT. L. 151, 179 (2016).

116. Smith, *supra* note 60.

involved in the production and sales process, from retailers to labels to textile manufacturers,” meaning “misappropriated patterns could represent millions of dollars to a successful plaintiff.”¹¹⁷ Since textile designs are often composed of “redundant-by-definition patterns consisting of finite elements and shapes,” this further complicates infringement analysis.¹¹⁸ When textile designs are registered, the presumption of validity in registration often forces defendants to settle.¹¹⁹ According to Doug Lipstone, partner at Weinberg Gosner whose clients are targets of textile copyright trolls, “This is not about copyright. This is about legalized extortion, it is shakedowns under the presumption of validity you can get from a copyright registration It is an absolute tax.”¹²⁰

B. PROPOSED EFFORTS TO DETER THE TROLL

There is much debate as to whether Congress, the courts, or the Copyright Office is best positioned administratively to address the copyright troll problem. While certain efforts can be made by Congress, the courts and the Copyright Office are better equipped to deter trolling.

117. *Id.*

118. Jahner, *supra* note 35.

119. *Id.* Stephen Doniger and Scott Alan Burroughs are credited to have “pioneered these lawsuits over the past decade,” in which they have filed over 700 suits for copyright infringement that have “targeted garments which feature designs they claim are purloined from their clients, who are mostly large textile converters and importers, such as L.A. Printex Industries, Star Fabrics, Unicolors, and UFI.” See Smith, *supra* note 60; see also Michael Goodyear, *A Shield or a Solution: Confronting the New Copyright Troll Problem*, 21 TEX. REV. ENT. & SPORTS L. 77 (2020) (examining the motivations and dangers of the attorney as a copyright troll focused on extorting money over pursuing legitimate claims to the detriment of the spirit of copyright); *Usherson v. Bandshell Artist Mgmt.*, 2020 WL 3483661 (S.D.N.Y. June 26, 2020), *aff’d in part sub nom. Liebowitz v. Bandshell Artist Mgmt.*, 858 F. App’x 457 (2d Cir. 2021), and *aff’d sub nom. Liebowitz v. Bandshell Artist Mgmt.*, 6 F.4th 267 (2d Cir. 2021) (“Richard Liebowitz, who passed the bar in 2015, started filing copyright cases in this District in 2017. Since that time, he has filed more cases in this District than any other lawyer: at last count, about 1,280; he has filed approximately the same number in other districts. In that same period, he has earned another dubious distinction: He has become one of the most frequently sanctioned lawyers, if not *the* most frequently sanctioned lawyer, in the District. Judges in this District and elsewhere have spent untold hours addressing Mr. Liebowitz’s misconduct, which includes repeated violations of court orders and outright dishonesty, sometimes under oath. He has been called ‘a copyright troll,’ . . . ‘a clear and present danger to the fair and efficient administration of justice,’ . . . a ‘legal lamprey[],’ . . . and an ‘example of the worst kind of lawyering,’ . . .”) (internal citations omitted).

120. Smith, *supra* note 60. For defense strategies against copyright trolls, see generally Matthew Sag & Jake Haskell, *Defense Against the Dark Arts of Copyright Trolling*, 103 IOWA L. REV. 571 (2018) (analyzing the practices of copyright trolls in bringing infringement suits).

1. Congress

A popular proposed solution to deterring the copyright troll has been to reduce or even eliminate statutory damages, a driving incentive in the copyright troll's business model.¹²¹ Statutory damages allow the court to award the copyright holder a fixed fee per work infringed, as it "considers just," between \$750 and \$30,000 per work.¹²² The Copyright Act also empowers courts to raise the award to as much as \$150,000 per work when a plaintiff succeeds in establishing willful infringement.¹²³ These statutory damages are determined independent of any harm, which effectively assures the troll of a worthwhile recovery and induces potential defendants to settle their claims with the troll in advance of a court's decision.¹²⁴

While appealing, either reducing or eliminating statutory damages are aggressive solutions that would likely cause significant harm to other areas of copyright law.¹²⁵ This is undesirable as statutory damages were implemented because of the difficulties copyright holders face when seeking to prove damage.¹²⁶ Changes to statutory damages would deter copyright trolls, however, given the impact on copyright holders and copyright law, this would be inconsistent with the overall aims of copyright. Notwithstanding, courts ought to scrutinize the election for statutory damages more closely to ensure that the compensatory purpose of statutory damages is not lost in overly punitive awards.¹²⁷

Considering how Congress sought to close the loophole of defendant infringers invalidating registrations based on a technicality via § 411(b), Congress could consider some larger legislative solution to the troll problems,

121. See Luke S. Curran, *Copyright Trolls, Defining the Line Between Legal Ransom Letters and Defending Digital Rights: Turning Piracy into a Business Model or Protecting Creative from Internet Lawlessness?*, 13 J. MARSHALL REV. INTELL. PROP. L. 170, 195 (2013) (citing to Pamela Samuelson & Tara Wheatland, *Statutory Damages in Copyright Law: A Remedy in Need of Reform*, 51 WM. & MARY L. REV. 439, 510 (2009) ("As part of a more general revision of copyright law, Congress might even reconsider whether statutory damages serve a desirable purpose in copyright law The compensatory purpose of statutory damages continues to be important, but, owing to the 1976 Act's creation of an enhanced level of authorized statutory damages for willful infringements, and the lack of principles to guide jury or judicial deliberations on statutory damages, awards have too often been arbitrary and inconsistent, and sometimes grossly excessive."). For a view that statutory damages should be eliminated to deter copyright trolls, see generally James DeBriyn, *Shedding Light on Copyright Trolls: An Analysis of Mass Copyright Litigation in the Age of Statutory Damages*, 19 UCLA ENT. L. REV. 79 (2012).

122. 17 U.S.C. § 504(c)(1).

123. *Id.* § 504(c)(2).

124. Balganes, *supra* note 106, at 737.

125. See Curran, *supra* note 121, at 196.

126. See Samuelson & Wheatland, *supra* note 121, at 499.

127. See Balganes, *supra* note 106, at 774.

but this is beyond the scope of this Note. While well-intentioned, Congress simply did not realize that its amendment to § 411 would facilitate trolling. Congress could also seek to address the copyright troll through statutory amendment, perhaps of § 411 itself. However, due to the evolving nature and difficulty of defining a copyright troll, such efforts may be ineffective in deterring the litigation-based business model of a copyright troll.

2. *The Courts*

If copyright trolling is suspected, courts should utilize the willful blindness carve-out that the Supreme Court provided. Like Justice Sotomayor questioned during oral arguments, the difference between a truly innocent mistake and an alleged mistake on the part of a sophisticated party with the capacity and wherewithal to know better can be difficult to differentiate.¹²⁸ As counsel for Unicolors themselves replied to the inquiry in oral arguments, “that’s the beauty of willful blindness.”¹²⁹

In place of actual knowledge, the Supreme Court expressly stated, “courts need not automatically accept a copyright holder’s claim that it was unaware of the relevant legal requirements of copyright law.”¹³⁰ The Supreme Court provided three relevant factors to aid courts in determining willful blindness based on circumstantial evidence in place of actual knowledge: (1) the significance of the legal error; (2) the complexity of the relevant law or rule; and (3) the applicant’s experience with copyright law.¹³¹ Based on the *Unicolors* decision, it appears that the Supreme Court does not intend for these factors to be exhaustive, but rather indicative of willful blindness. Likewise, the Court does not place any assignment of weight to the factors and suggests that the determination should be based on the totality of the circumstances at the court’s discretion.¹³² Other such factors that courts can, and should, look to include in their determination are: (1) the applicant’s litigation history; (2) any financial motivations that could explain the mistake; and (3) the line of work or business of the applicant.

While bright line rules and clear definitions are easy for courts to administer, the complexity of copyright law and the concept of copyright trolling requires courts to carefully balance competing aims. If courts utilize the *Unicolors* willful blindness carve-out and address the relevant factors,

128. Transcript of Oral Argument at 17, *Unicolors, Inc. v. H&M Hennes & Mauritz, L.P.*, 142 S. Ct. 941 (2022) (No. 20-915).

129. *Id.* at 18.

130. *Unicolors v. H&M*, 142 S. Ct. at 948.

131. *Id.* (citing to *Intel Corp. Inv. Pol’y Comm. v. Sulyma*, 140 S. Ct. 768, 778–79 (2020)).

132. *Id.* (stating that “other such matters . . . may also lead a court to find that an applicant was actually aware of, or willfully blind to, legally inaccurate information.”).

copyright trolls may be deterred from taking advantage of leniencies in copyright law. At worst, it will force copyright trolls to ensure their registrations are accurate.

3. *The Copyright Office*

Lastly, the Copyright Office, in order to protect the integrity of the registration system, should also consider certain efforts to deter the copyright troll from disturbing, and even diluting, the rights of all copyright holders.¹³³ Most importantly, the Copyright Office needs to provide some clarity as to how it determines referrals under § 411(b)(2).

Although the Supreme Court has now offered some clarity on how to determine if an inaccuracy was knowingly included in an application for registration under § 411(b)(1), it remains unclear how the Register of Copyrights retroactively determines whether a registration would have been refused when asked for advice from courts under § 411(b)(2). As with most areas of law, there is a tension between the law on the books and the law in action. According to the plain language of the statute, it appears that the phrase, “the inaccurate information, *if known*, would have *caused* the Register of Copyrights to refuse registration,” implies a determination of the application as submitted.¹³⁴ However, in practice at the moment, if an examiner had known of the inaccuracy on the application, she likely would have corresponded with the applicant to remedy the mistake rather than automatically refuse registration.¹³⁵ Based on the Copyright Office’s recent replies to § 411(b)(2) referrals, there does not appear to be a consistent approach.¹³⁶ Considering the Court’s holding in *Unicolors*, the Copyright Office should put forth a clear standard, of which an “as submitted” determination seems most appropriate.

133. While some may argue that the Copyright Office should reform their review of applications for registration to include some sort of prior art search, like the Patent and Trademark Office practices, this is an unrealistic burden on the office. In 2021 alone, the Copyright Office registered 403,593 claims for registration involving millions of works and recorded 11,625 documents containing titles of 961,291 works. See U.S. COPYRIGHT OFF., U.S. COPYRIGHT OFFICE ANNUAL REPORT FISCAL YEAR 2021, at 3 (2021), <https://www.copyright.gov/reports/annual/2021/ar2021.pdf>.

134. 17 U.S.C. § 411(b)(2) (emphasis added).

135. Kasunic, *supra* note 29, at 90.

136. See generally *Copyright Office Filings Under Section 411*, U.S. COPYRIGHT OFF., <https://www.copyright.gov/rulings-filings/411/> (last visited Dec. 20, 2023).

V. CONCLUSION

The underlying facts of the recent Supreme Court case *Unicolors v. H&M* draw attention to the copyright troll issue that is plaguing much of copyright litigation. In holding that actual knowledge, rather than fraud, is required to disqualify a copyright holder from the § 411(b) safe harbor, Justice Breyer's majority opinion is consistent with Congress's intention to protect and promote the rights of non-lawyer copyright holders against efforts to invalidate their copyright certificate by infringing defendants seeking to escape liability for their infringing acts. At the same time, Justice Breyer indirectly struck a balance between the tension of protecting non-lawyer creatives and deterring copyright trolls from exploiting the leniencies of copyright law and the copyright registration system. To effectively deter copyright trolls, courts should approach a copyright holder's claim that she was unaware of the relevant legal requirements with skepticism. Where appropriate, courts should feel empowered to apply the willful blindness carve-out to determine whether, in the totality of the circumstances, a copyright holder truly made a mistake or knew what she was doing. Likewise, the Copyright Office should provide the courts with clarity on how the Register of Copyrights retroactively determines whether the inaccuracy, if known at the time, would have caused her to refuse registration. The risk of invalidating a copyright troll's certificate of registration and the right to sue it provides is a potent deterrent, especially given that the right to bring suit is a foundational aspect of a copyright troll's business model.

DANGLING THE CARROT OF TRADEMARK REGISTRATION

Brigitte Desnoes[†]

TABLE OF CONTENTS

I.	INTRODUCTION	1274
II.	IN RE ELSTER BACKGROUND	1276
	A. INITIAL PTO REFUSAL.....	1276
	B. TTAB OPINION.....	1278
	1. Purpose of § 2(c).....	1278
	2. In re ADCO Decision.....	1282
	C. FEDERAL CIRCUIT OPINION.....	1283
III.	TRADEMARK REGISTRATION AND THE FIRST AMENDMENT	1284
	A. TRADEMARK REGISTRATION OVERVIEW.....	1284
	B. TAM AND BRUNETTI.....	1287
	1. Matal v. Tam.....	1287
	2. Iancu v. Brunetti.....	1289
IV.	FIRST AMENDMENT FRAMEWORKS	1290
	A. COMMERCIAL SPEECH.....	1293
	1. Overview.....	1293
	2. Analogy to Trademark.....	1296
	B. GOVERNMENT SUBSIDIES AND GOVERNMENT PROGRAMS.....	1298
	1. Overview.....	1298
	2. Analogy to Trademark.....	1299
	C. LIMITED PUBLIC FORUM.....	1299
	1. Overview.....	1299

DOI: <https://doi.org/10.15779/Z38J678Z0T>

© 2023 Brigitte Desnoes.

† J.D. 2024, University of California, Berkeley, School of Law. Thank you to Professor Talha Syed, Professor Sonia Katyal, and my fellow students in the 2022 Law & Technology Writing Workshop. Since this Note was completed in December 2022, the Office of the Solicitor General filed a Petition for Writ of Certiorari with the Supreme Court on January 27, 2023. The Petition was granted on June 5, 2023, and oral argument was heard on November 1, 2023. During oral argument, Justice Kavanaugh and Justice Kagan inquired about and seemingly expressed support for the application of the limited public forum framework to trademark registration. Transcript of Oral Argument, *Vidal v. Elster*, 143 S. Ct. 2579 (2023).

2. <i>Analogy to Trademark</i>	1302
V. PARODIC USE CARVEOUT?	1304
VI. CONCLUSION	1306

I. INTRODUCTION

At the core of the free speech clause of the First Amendment is the idea that the government may not restrict expression that it finds offensive or disagreeable.¹ Such expression includes the Westboro Baptist Church’s picketing at military funerals,² an individual’s wearing a jacket that says “Fuck the Draft,”³ and the burning of the American flag.⁴ Yet the First Amendment does not prevent a state DMV from refusing to approve a specialty license plate that might be offensive to others,⁵ or a city from prohibiting posterage on utility poles.⁶ At the same time, it does prevent a government’s banning tobacco advertising within one thousand feet of schools.⁷ The thread between these cases lies in a convoluted maze of categories and hierarchies that the Supreme Court has developed over the past century in its attempt to balance legitimate governmental and societal interests in speech regulation with the overarching ethos of free expression.

Speech “regulation” that occurs through the process of trademark registration is in limbo in this maze. The United States Patent and Trademark Office (PTO) reviews trademark applications and can refuse registration to marks if they do not meet certain requirements as laid out by Congress in the Lanham Act. How much should we scrutinize the justifications for registration refusal? The Supreme Court has held that the PTO’s refusals to register both “disparag[ing]” and “immoral or scandalous” trademarks were unconstitutional violations of the First Amendment.⁸ Most recently, the Federal Circuit held that the PTO’s refusal to register the trademark TRUMP

1. *Texas v. Johnson*, 491 U.S. 397, 414 (1989) (“If there is a bedrock principle underlying the First Amendment, it is that the Government may not prohibit the expression of an idea simply because society finds the idea itself offensive or disagreeable.”).

2. *See Snyder v. Phelps*, 562 U.S. 443 (2011) (holding that the picketing church members had a First Amendment right to speak on matters of public concern).

3. *See Cohen v. California*, 403 U.S. 15 (1971) (holding the jacket language protected under the First Amendment).

4. *See Johnson*, 491 U.S. at 414 (holding the right to burn the American flag was protected speech under the First Amendment).

5. *See Walker v. Tex. Div., Sons of Confederate Veterans, Inc.*, 576 U.S. 200 (2015).

6. *See City Council of Los Angeles v. Taxpayers for Vincent*, 466 U.S. 789 (1984).

7. *See Lorillard Tobacco Co. v. Reilly*, 533 U.S. 525 (2001).

8. *Matal v. Tam*, 137 S. Ct. 1744 (2017); *Iancu v. Brunetti*, 139 S. Ct. 2294 (2019).

TOO SMALL was unconstitutional as well.⁹ According to the trademark applicant, Steve Elster, the mark was political commentary not only about Senator Marco Rubio’s “small hands” insult at a 2016 presidential debate, but also about the “smallness of Donald Trump’s overall approach to governing as president.”¹⁰ The PTO rejected his trademark application under a relatively straightforward provision of the Lanham Act, § 2(c), which states that a trademark shall be rejected if it “identif[ies] a particular living individual except by his written consent.”¹¹ Despite this provision, the court found that the protection of political speech outweighed any government interest in the right of publicity or right of privacy, not least because of Trump’s status as a deeply public figure.¹²

Yet, the Federal Circuit court failed to do what the Supreme Court has painstakingly done for decades in crafting its modern free speech jurisprudence—carefully consider the context of the speech. In *Elster*, the speech in question was not expressed and restricted in a public forum, but in a forum for federal trademark registration. A trademark refused federal registration is not deprived of its status as a trademark, because trademarks arise from use in commerce and exist at common law with or without registration. Nor is the applicant stripped of the ability to use the phrase “Trump Too Small” in any capacity, on merchandise or otherwise. Trademark registration merely provides certain benefits to a trademark owner, primarily through an evidentiary presumption of validity when a mark owner is either defending against or alleging trademark infringement. In practice, the trademark registration system looks a lot like a limited public forum, one of the First Amendment “contexts” in which restricted speech is merely denied certain benefits of occurring in a government provided forum, but not restricted altogether. Examples of forums given this title by the Court include a student-organization forum upon registration with a law school¹³ and an internal school district mail system.¹⁴ In the former case, the Court emphasized that by offering benefits from registering an organization with the school and

9. *In re Elster*, 26 F.4th 1328, 1330 (Fed. Cir. 2022).

10. Response to Office Action of February 19, 2018, U.S. Trademark Application Serial No. 87/749,230 (filed July 7, 2018).

11. 15 U.S.C. § 1052(c).

12. *See In re Elster*, 26 F.4th at 1335. (“With respect to privacy, the government has no legitimate interest in protecting the privacy of President Trump, ‘the least private name in American life,’ from any injury to his ‘personal feelings’ caused by the political criticism that Elster’s mark advances.”).

13. *See* Christian Legal Soc’y Chapter of Univ. of Cal., Hastings Coll. of L. v. Martinez, 561 U.S. 661 (2010).

14. *See* Perry Educ. Ass’n v. Perry Loc. Educators’ Ass’n, 460 U.S. 37 (1983).

complying with the school's policies, the school was "dangling the carrot of subsidy, not wielding the stick of prohibition."¹⁵

With *Elster* and § 2(c) as a backdrop, this Note delves into the theoretical and policy underpinnings of the limited public forum, as well as two other First Amendment frameworks that have been proposed as means through which to analyze trademark registration regulation: regulation of commercial speech and regulation of speech through a government subsidy. First, Part II details the procedural history and trademark issues leading to *In re Elster*. Part III describes the nature of trademark registration and the two cases in which the Supreme Court considered trademark registration in the context of the First Amendment: *Matal v. Tam*¹⁶ and *Iancu v. Brunetti*.¹⁷ Part IV analyzes the aforementioned First Amendment frameworks in the context of trademark registration, landing on the limited public forum as most analogous to registration. Part V then considers and rejects a parodic use carveout.

II. *IN RE ELSTER* BACKGROUND

In this Part, Section II.A sets the stage for the Federal Circuit decision, detailing the PTO's initial refusal of Steve Elster's TRUMP TOO SMALL trademark registration. Section II.B details the TTAB's opinion of Elster's appeal, with sub-Sections II.B.1 and II.B.2 highlighting two foundations of the TTAB's position: the purpose of § 2(c), and an earlier decision, *In re ADCO Industries-Technologies, L.P.*,¹⁸ which dealt with substantially the same issue. Finally, Section II.C details the Federal Circuit's opinion in *Elster*.

A. INITIAL PTO REFUSAL

On January 10, 2018, Steve Elster filed a trademark application with the PTO for the mark TRUMP TOO SMALL for use on shirts.¹⁹ A little over a month later, the PTO responded with a straightforward notice of refusal pursuant to § 2(c) of the Lanham Act.²⁰ Section 2(c) bars the registration of a trademark which "[c]onsists of or comprises a name, portrait, or signature

15. *Christian Legal Soc'y*, 561 U.S. at 683.

16. 137 S. Ct. 1744 (2017).

17. 139 S. Ct. 2294 (2019).

18. *In re ADCO Indus.-Techs., L.P.*, 2020 U.S.P.Q.2d 53786 (I.T.A.B. 2020).

19. TEAS RF New Application, U.S. Trademark Application Serial No. 87/749,230 (filed Jan. 10, 2018).

20. The Trademark Act of 1946—also known as the Lanham Act—is the federal statute that governs federal trademark registration. As will be discussed *infra* Part III.A, registration of a mark pursuant to the Lanham Act does *not* create the existence of trademark, in the way registration of an invention creates the existence of a patent. J. THOMAS MCCARTHY, 2 MCCARTHY ON TRADEMARKS AND UNFAIR COMPETITION § 6:8 (5th ed. 2019).

identifying a particular living individual except by his written consent.”²¹ The PTO determines that a name in a mark “identifies” an individual if (1) the person is so well known that the public would reasonably assume a connection between the person and the goods or services, or (2) the individual is publicly connected with the business in which the mark is used.²² In this case, the PTO noted that the word TRUMP “clearly references” Donald Trump, “the subject of frequent media attention,” and his written consent was not on record.²³

In his initial response to the rejection, Elster stated that the mark was political commentary both about Donald Trump’s refutation of Marco Rubio’s “insinuation that Donald Trump has a small penis” during a March 3, 2016 Republican presidential debate, and the “smallness of Donald Trump’s overall approach to governing as president of the United States.”²⁴ And, in fact, because Trump tried to “repudiate the assertion that his penis is small,” no consumer would reasonably think he sponsored or endorsed the goods accompanied by the trademark.²⁵ After the PTO again straightforwardly refused registration under § 2(c), Elster responded with the argument that the trademark was “core political speech about a political figure.”²⁶ Elster cited *Cardtoons, L.C. v. Major League Baseball Players Ass’n*, a case in which the Tenth Circuit upheld the production of parody baseball cards caricaturing active and former baseball players in spite of the MLB’s claim of players’ rights of publicity.²⁷ The PTO denied the request for reconsideration on the same grounds as prior refusals, adding that the *Cardtoons* case was “misplaced.”²⁸ Elster appealed to the Trademark Trial and Appeal Board (TTAB), after which the PTO amended the refusal on an additional ground under the false

21. 15 U.S.C. § 1052(c).

22. *In re Nieves & Nieves L.L.C.*, 113 U.S.P.Q.2d 1639, 1650 (T.T.A.B. 2015).

23. Office Action Outgoing, U.S. Trademark Application Serial No. 87/749,230 (filed Feb. 19, 2018).

24. Response to Office Action of February 19, 2018, *supra* note 10. See Gregory Krieg, *Donald Trump Defends the Size of His Penis*, CNN (Mar. 4, 2016), <https://www.cnn.com/2016/03/03/politics/donald-trump-small-hands-marco-rubio/index.html> (reporting that Donald Trump stated in the debate, “[Rubio], he referred to my hands, if they’re small, something else must be small. I guarantee you there’s no problem. I guarantee.”); see also Emily Shapiro, *The History Behind the Donald Trump ‘Small Hands’ Insult*, ABCNEWS (Mar. 4, 2016), <https://abcnews.go.com/Politics/history-donald-trump-small-hands-insult/story?id=37395515>.

25. Response to Office Action of February 19, 2018, *supra* note 10.

26. Request for Reconsideration after Final Action & Response to Office Action of July 30, 2018, U.S. Trademark Application Serial No. 87/749,230 (filed Jan. 29, 2019).

27. *Id.* (citing *Cardtoons, L.C. v. Major League Baseball Players Ass’n*, 95 F.3d 959, 969 (1996)).

28. Request for Reconsideration Denied, U.S. Trademark Application Serial No. 87/749,230 (filed Feb. 25, 2019).

association clause of § 2(a), which bars registration of marks which “falsely suggest a connection with persons, living or dead.”²⁹

Finally, Elster responded on explicit constitutional grounds.³⁰ He stated that the government’s refusal to register a certain trademark was a content-based regulation of private speech, meaning that it is subject to strict scrutiny under the First Amendment.³¹ Strict scrutiny requires the government to prove that a restriction of speech furthers a compelling interest and is narrowly tailored to achieve that interest.³² Elster stated that the government had not articulated a compelling interest in § 2(a), nor was the regulation narrowly tailored.³³

B. TTAB OPINION

On July 2, 2022, the TTAB affirmed the PTO’s § 2(c) refusal, finding it unnecessary to reach the issue of § 2(a)’s false association clause.³⁴ Rather, the TTAB reiterated what the PTO had pointed out—it is not required, as for refusal under § 2(a), that the public perceive an association with or endorsement by the individual identified in a mark.³⁵ Therefore, Elster’s argument that the mark is the “antithesis of what consumers would understand to be sponsored by” Trump was moot.³⁶ On the First Amendment question, the Board cited its own decision in *In re ADCO Industries-Technologies, L.P.*,³⁷ in which it addressed and rejected the same issue. Sub-Sections II.B.1 and II.B.2 delve further into each of these points.

1. Purpose of § 2(c)

The Lanham Act provides that no trademark application that can be distinguished from the goods of others can be denied registration unless it falls under one of the § 2 provisions, including: falsely suggesting a connection with persons, living or dead; consisting of deceptive matter; containing a geographical indication used in connection with wines or spirits which identifies a place other than the origin of the goods; and containing a flag or

29. Motion to Remand, U.S. Trademark Application Serial No. 87/749,230 (filed May 31, 2019); Office Action Outgoing, U.S. Trademark Application Serial No. 87/749,230 (filed June 24, 2019); 15 U.S.C. § 1052(a).

30. Response to Office Action of June 24, 2019, U.S. Trademark Application Serial No. 87/749,230 (filed Sept. 9, 2019).

31. *Id.* at 3.

32. *Id.* (citing *Reed v. Town of Gilbert*, 135 S. Ct. 2218, 2227 (2015)).

33. *Id.*

34. *In re Elster*, Serial No. 87/749,230 (T.T.A.B. July 2, 2020).

35. *Id.* at 5.

36. *Id.*

37. 2020 U.S.P.Q.2d 53786 (T.T.A.B. 2020).

coat of arms of the United States, any state or municipality, or any foreign nation.³⁸ The Lanham Act defines a trademark as a designation used to “identify and distinguish” goods from those sold by others,³⁹ and the test for protection of a trademark is “likelihood of confusion” by consumers.⁴⁰ The false association and deception clauses of § 2(a) perhaps flow intuitively based on these purposes, but some § 2 provisions bar registration for reasons beyond preventing likelihood of confusion. For example, the bar on certain geographic indications for wines and liquors came from the Uruguay Round Agreements Act, which implemented the international trade agreement that created the World Trade Organization (WTO); the restrictions do not require proof that the trademark is false or misleading.⁴¹

In addition, § 2(c) captures rights outside the bounds of likelihood of confusion but recognized at common law—namely, the rights of privacy and publicity.⁴² The provision provides that registration of a trademark is allowed unless it “[c]onsists of or comprises a name, portrait, or signature identifying a particular living individual except by his written consent, or the name, signature, or portrait of a deceased President of the United States during the life of his widow, if any, except by the written consent of the widow.”⁴³ While drafting the Act, one House member noted of the President clause that “we would not want to have Abraham Lincoln gin,” while another agreed but added that they “would not say the use of G. Washington on coffee should not be permissible.”⁴⁴ Clearly, in 1939, it’s unlikely consumers would have

38. 15 U.S.C. § 1052.

39. 15 U.S.C. § 1127.

40. 2 MCCARTHY, *supra* note 20, § 23:1.

41. *Id.* § 14:40; *see also The Uruguay Round*, WTO, https://www.wto.org/english/thewto_e/whatis_e/tif_e/fact5_e.htm (last visited Dec. 17, 2022). TRIPS art. 23(1) provided that each member shall provide the legal means to “prevent use of a geographical indication identifying wines [or spirits] for wines [or spirits] not originating in the place indicated by the geographical indication . . . even where the true origin of the goods is indicated or the geographical indication is used in translation or accompanied by expressions such as ‘kind,’ ‘style,’ ‘imitation,’ or the like.” For example, it is not necessary that the geographic origin of a wine named ALASKA CHARDONNAY be believed or material to the consumer decision. Rebecca Tushnet, *The First Amendment Walks into a Bar: Trademark Registration and Free Speech*, 92 NOTRE DAME L. REV. 381, 407 n.99 (2016). Nevertheless, Tushnet notes that this “might still be justified on a prophylactic deception-avoidance rationale.” *Id.*

42. *See Univ. of Notre Dame Du Lac v. J.C. Gourmet Food Imports Co.*, 703 F.2d 1372 (Fed. Cir. 1983) (“[T]here may be no likelihood of such confusion as to the source of goods even under a theory of ‘sponsorship’ or ‘endorsement,’ and, nevertheless, one’s right of privacy, or the related right of publicity, may be violated.”) (discussing 2(a) false association but noting 2(c) “is also of this nature”).

43. 15 U.S.C. § 1052(c).

44. Hearings on H.R. 4744 Before the Subcomm. on Trademarks of the H. Comm. on Patents, 76th Cong. 19 (1939) (statement of Thomas E. Roberts, former Comm’r of Patents).

been that more confused about Abraham Lincoln's relationship with a distillery than George Washington's relationship with a coffee roastery. The representatives were articulating something different from likelihood of confusion—possibly what Jennifer Rothman and Robert Post have described as “diminishment.”⁴⁵ In contrast to the harm of confusion of an individual's sponsorship or participation in some commercial venture, which can injure the value of that person's interests whether or not they are famous, “the harm of *diminishment* applies primarily (perhaps exclusively) to those plaintiffs whose identities already possess goodwill in the market,” wherein their identities “may be distinctly vulnerable to damage through overexposure and tarnishment.”⁴⁶

In fact, § 2(c) codifies a long history of common law and statutory protections rooted in a “personality-based understanding of trademark law.”⁴⁷ For example, Rothman cites an 1873 treatise which describes a trademark as “carr[ying] the idea of a man's personality, like his ordinary autograph.”⁴⁸ Given the early use of trade names that were derived from one's identity, trademark was understood to encompass more than merely market-based interests; rather, unauthorized uses of another's name in trade were “also understood as an affront to a person's autonomy interests, their dignity, and their natural right to the fruits of their own labor.”⁴⁹ In 1898, for example, the predecessor to the PTO refused to register “Dewey's Chewies” for confections because George Dewey, a famous Admiral, “was a ‘living celebrity’ who was ‘entitled to protection from the ordinary trader,’” regardless of a prospective consumer's confusion.⁵⁰ The idea of ownership in oneself ultimately led to the common law adoption of the standalone rights of privacy and publicity, which most states recognize today.⁵¹

45. Jennifer Rothman & Robert Post, *The First Amendment and the Right(s) of Publicity*, 130 YALE L.J. 86, 111–12 (2020).

46. *Id.*

47. See Jennifer Rothman, *Navigating the Identity Thicket: Trademark's Lost Theory of Personality, The Right of Publicity, and Preemption*, 135 HARV. L. REV. 1272, 1307 n.167 (2022) (citing Trade-Mark Act of 1905, ch. 592, § 5(b), 33 Stat. 724, 726 (“[N]o portrait of a living individual may be registered as a trade-mark, except by the consent of such individual, evidenced by an instrument in writing . . .”) and the California Act of April 4, 1941, ch. 58, § 14242(g), 1941 Cal. Stat. 703, 705 (precluding registration of a mark that consists of “[t]he portrait of a living person except by consent of the person evidenced by an instrument in writing”).

48. Rothman, *supra* note 47, at 1295–96 (citing WILLIAM HENRY BROWNE, A TREATISE ON THE LAW OF TRADE-MARKS AND ANALOGOUS SUBJECTS § 90 (Boston, Little, Brown & Co. 1873)).

49. *Id.* at 1296.

50. *Id.* at 1308–09 (citing *Ex parte McInnerney*, 85 Off. Gaz. Pat. Off. 148, 149 (1898)).

51. *Id.* at 1297; see also J. THOMAS MCCARTHY, RIGHTS OF PUBLICITY AND PRIVACY § 6:3 (2d ed. 2016).

The right of privacy broadly includes: an intrusion upon one's seclusion or solitude, public disclosure of private facts, publicity which places one in false light, and appropriation of one's name or likeness.⁵² This fourth right has developed into a standalone right of publicity.⁵³ While the former generally is the right "to be let alone," and encompasses a kind of *psychic* protection, the latter protects a famous person's *pecuniary* interest in the commercial exploitation of his identity.⁵⁴

Because a right of privacy or publicity violation is directly tied to one's identity, § 2(c) requires "identification."⁵⁵ If a person is neither "generally known," nor publicly connected to the field relating to the business concerned, then the mark cannot be said to constitute "identification" of a particular person.⁵⁶ The PTO accordingly excludes protection for someone who "coincidentally bears an applied-for name,"⁵⁷ but also recognizes that identification of a person can occur even if it is a name shared by others⁵⁸ or only part of their full name.⁵⁹ In any scenario, an individual who is not well known would have to show that the consuming public connects them with the product, whereas "well-known individuals such as celebrities and world-famous political figures are entitled to the protection of § 2(c) without having to evidence a connection with the involved goods or services."⁶⁰ This results in sometimes obvious work on the PTO's part, for example, to point to news articles to in support of the finding that Donald Trump is "well known by the public."⁶¹ In rejecting a trademark application for OBAMA PAJAMA, the

52. *Carson v. Here's Johnny Portable Toilets, Inc.*, 698 F.2d 831 (6th Cir. 1983) (citing William L. Prosser, *Privacy*, 48 CALIF. L. REV. 383, 389 (1960)).

53. *Id.*

54. *Id.*; see also RESTATEMENT (THIRD) OF UNFAIR COMPETITION § 46, cmt. b (1995) ("The distinction between the publicity and privacy actions . . . relates primarily to the nature of the harm suffered by the plaintiff; similar substantive rules govern the determination of liability").

55. *Martin v. Carter Hawley Hale Stores, Inc.*, 206 U.S.P.Q. 931, 932–33 (T.T.A.B. 1979).
56. See Trademark Manual of Examining Procedure § 1206.03 (November 2023).

57. *In re Richard M. Hoefflin*, U.S. Trademark Application Serial No. 77/632,391, 97 U.S.P.Q.2d 1174, 1175 (T.T.A.B. 2010).

58. See *In re Steak & Ale Rests. Am., Inc.*, 185 U.S.P.Q. 447 (T.T.A.B. 1975) (affirming refusal to register PRINCE CHARLES for meat, since it identified Prince Charles, a member of the English royal family).

59. See *Ross v. Analytical Tech. Inc.*, 51 U.S.P.Q.2d 1269 (T.T.A.B. 1999).

60. *In re Hoefflin*, 97 U.S.P.Q.2d at 1177.

61. Office Action Outgoing, U.S. Trademark Application Serial No. 87/749,230 (filed July 30, 2018). In *In re Hoefflin*, the TTAB noted that "[t]he Trademark Examining Attorney has done an excellent job marshalling a variety of press excerpts to demonstrate the obvious—namely, that President Barack Obama is extremely well known." *In re Hoefflin*, Serial No. 77/632,391, 97 U.S.P.Q.2d at 8.

TTAB noted that despite the applicant's argument that no one would conclude that President Obama was connected with the brand's pajamas, sleepwear, and underwear, it was "because he is the President of the United States" that § 2(c) applied.⁶² Accordingly, the application was rejected.⁶³ At the same time, although the mark ARNOLD BRAND for fresh tomatoes encompassed the name of an individual named Arnold Brand, a relatively well-known attorney, Brand had not attained recognition in the field of business in which the mark was used, and his attempt to cancel the trademark registration under § 2(c) failed.⁶⁴

For a famous person, the only harms that a court will likely recognize in the face of a § 2(c) violation are the pecuniary and goodwill harms attendant to the right of publicity, rather than the right of privacy. For example, in one case, the Sixth Circuit found that no right of privacy interests were invaded simply because the plaintiff (Johnny Carson) was "embarrassed by" the defendant's product (portable toilet seats named "Here's Johnny").⁶⁵ Right of publicity interests, however, were.⁶⁶ The court noted that the theory of the right is that "a celebrity's identity can be valuable in the promotion of products, and the celebrity has an interest that may be protected from the unauthorized commercial exploitation of that identity."⁶⁷

2. *In re ADCO Decision*

In *In re ADCO*,⁶⁸ the Court considered a claim similar to that in *In re Elster*.⁶⁹ The proposed mark in that case consisted of the phrase TRUMP-IT for a package opener.⁷⁰ As in *Elster*, the PTO rejected the mark under both §§ 2(a) and 2(c), which the applicant claimed was unconstitutional under the First Amendment.⁷¹ The TTAB noted that as a threshold matter, the applicant erred in treating provisions of the Lanham Act as akin to direct restrictions on free speech.⁷² The Board stated that "Section 2 does not prevent an applicant from using any slogan of its choice on its merchandise or from advertising that

62. *In re Hoeflin*, 97 U.S.P.Q.2d at 1177.

63. *Id.* at 1178.

64. *Martin v. Carter Hawley Hale Stores*, 206 U.S.P.Q. at 933 (citing *Brand v. Fairchester Packing Co.*, 84 U.S.P.Q. 97 (Comm'r Pat. 1950)).

65. *Carson v. Here's Johnny Portable Toilets, Inc.*, 698 F.2d 831, 834 (6th Cir. 1983).

66. *Id.* at 835.

67. *Id.*

68. *In re ADCO Indus.-Techs., L.P.*, 2020 U.S.P.Q.2d 53786 (I.T.A.B. 2020).

69. *In re Elster*, 26 F.4th at 1330.

70. *In re ADCO Indus.-Techs.*, 2020 U.S.P.Q.2d 53786, at *1.

71. *Id.* at *21.

72. *Id.* at *25.

merchandise through any advertising message of its choosing.”⁷³ In addition, the provisions do not call for rejecting trademarks based on viewpoint, but rather on viewpoint neutral criteria.⁷⁴ Finally, even if the challenged provisions were evaluated as outright restrictions on speech, rather than on registration, they are within Congress’s authority to make decisions for the sake of the public; both of the statutes “recognize[] the right of privacy and publicity that a living person has in his or her identity and protects consumers against source deception.”⁷⁵ Certainly, as the TTAB added in *Elster*, § 2(c) is narrowly tailored, since it “consistently and reliably applies to any mark that consists of or comprises a name, portrait or signature identifying a particular living individual.”⁷⁶

C. FEDERAL CIRCUIT OPINION

Finally, the Federal Circuit heard the case, ultimately agreeing with *Elster* and finding the application of § 2(c) to reject registration of the mark unconstitutional.⁷⁷ The court declined to decide how best to analyze the nature of trademark registration rejections—whether through strict or intermediate scrutiny as expressive or commercial speech.⁷⁸ Under either conceptualization, the court held, the government had no sufficient right of publicity or right of privacy interests to overcome the powerful First Amendment protections of the “political criticism” embodied in trademark.⁷⁹

First, there was no “plausible” claim of Trump’s right of privacy from criticism in the absence of actual malice.⁸⁰ In fact, there is no right of privacy at all when the government restricts speech that comments on or criticizes public officials.⁸¹ Moreover, the court stated that no right of publicity existed because no claim was made that Trump’s name was misappropriated in a way that exploited his commercial interests or diluted the commercial value of his name.⁸² Moreover, “no plausible claim” could be made that the disputed mark suggests President Trump has endorsed *Elster*’s product.⁸³ Broadly, as with the right of privacy, there is no “substantial” interest in a right of publicity claim

73. *Id.* at *27

74. *Id.* at *27–28.

75. *Id.* at *28.

76. *In re Elster*, Serial No. 87/749,230, at *11.

77. *In re Elster*, 26 F.4th at 1330.

78. *Id.* at 1338–39.

79. *Id.* at 1338.

80. *Id.* at 1336.

81. *Id.* at 1335.

82. *Id.* at 1336.

83. *Id.*

in the context of criticism of a public official.”⁸⁴ Despite the government’s claim that Congress’s enacting § 2(c) was a targeted effort to preclude registration that facilitates a type of commercial behavior already banned in most states, the court’s “review of state-law cases revealed no authority holding that public officials may restrict expressive speech to vindicate their publicity rights.”⁸⁵

III. TRADEMARK REGISTRATION AND THE FIRST AMENDMENT

The Federal Circuit decision reflects a fundamental misunderstanding of trademark and trademark registration, as well as how best to conceptualize such speech “regulation” in the context of First Amendment jurisprudence. In this Part, Section III.A explains trademark registration. Section III.B details the two Supreme Court cases preceding *In re Elster*, in which the Court first considered the constitutionality of federal trademark registration provisions.

A. TRADEMARK REGISTRATION OVERVIEW

A trademark is a designation of the source of goods and services used to “identify and distinguish” the source from those manufactured or sold by others.⁸⁶ Early trademark protection evolved out of the common law tort of fraud and deceit, in which the fraudulent intent to deceive consumers through the use of another’s trademark was the key inquiry, rather than consumers’ confusion itself.⁸⁷ Today, while some disagree about the primacy of one goal over another, the protection of trademarks broadly has two goals: (1) to protect consumers from deception and confusion over trademarks, and (2) to protect the owner of trademark from misappropriation by others.⁸⁸

Though often included under the umbrella of intellectual property, trademarks are fundamentally different from patents or copyrights.⁸⁹ First, as

84. *Id.*

85. *Id.* at 1338.

86. 15 U.S.C. § 1127.

87. 2 MCCARTHY, *supra* note 20, § 5:2.

88. *Id.* § 2:2 (citing S. REP. NO. 133, at 3 (1946)). For example, Professor McKenna has argued that trademark law is rooted in unfair competition law and was never traditionally intended to protect consumers. Mark P. McKenna, *The Normative Foundations of Trademark Law*, 82 NOTRE DAME L. REV. 1839, 1841 (2007). On the other hand, Professor Tushnet has stated that “[p]rotection against consumer confusion is the rhetorical core of modern trademark law.” Rebecca Tushnet, *Registering Discontent: Registration in Modern American Trademark Law*, 130 HARV. L. REV. 867 (2017).

89. 1 MCCARTHY, *supra* note 20, § 6:1 (describing “a tremendous amount of confusion in the mind of the public and even the practicing bar as to the fundamental differences between patents, trademarks and copyrights”).

indicated by the twin goals of trademark law, trademark protection does not exist to foster or reward innovation.⁹⁰ Second, trademark rights do not exist for a given duration upon registration with the U.S. Government. Rather, the exclusive right to a trademark “grows out of its *use*, and not its mere adoption.”⁹¹ Without continual use in commerce, trademarks are meaningless. These differences track to the constitutional basis for the federal power to regulate trademarks. While Congress has the power to regulate patents and copyrights under the patent and copyright clause of the Constitution—which grants Congress the power to “promote the progress of sciences and the useful arts, by securing for limited times to authors and inventors, the exclusive right to their respective writings and discoveries”—trademarks are excluded from such a grant.⁹² Because trademarks generally grow out of “a considerable period of use, rather than a sudden invention,” a clause concerning authors and inventors is inapplicable.⁹³ Rather, Congress has power to regulate trademarks only under its power to regulate commerce in the Commerce Clause.⁹⁴

The Lanham Act was enacted in 1946 to establish a federal system of national trademark registration. Importantly, the Act was not intended to change the common law of trademark that had developed up to its enactment.⁹⁵ While federal registration of a patent or copyright confers property-like exclusive rights to use and protection of the creation, federal registration of trademark requires proof of “use in commerce” and is, as in common law, meaningless absent continual use of the mark.⁹⁶ There is an opportunity provided in the Lanham Act to file an application for registration based on a good faith “intent to use,” but only after the mark is used in commerce is the registration actually issued.⁹⁷

There are benefits, however, to federal registration of a trademark. First, registration provides constructive national notice of ownership.⁹⁸ Second, if an owner of a registered trademark were to file an infringement action, the fact

90. *See In re Trade-Mark Cases*, 100 U.S. 82, 93–94 (1879) (noting that a trademark does not depend upon “novelty, invention, discovery, or any work of the brain”).

91. *Id.*

92. *Id.*

93. *Id.*

94. *Id.*

95. “The Lanham Act did not supplant the state common law of trademarks . . . [i]n fact, Section 15 of the Lanham Act grants federally registered marks the right to exclusive use of the mark only insofar as they do not conflict with any pre-existing rights acquired under state law.” *Dorpan v. Hotel Meliá, Inc.*, 728 F.3d 55, 62 (1st Cir. 2013) (citing 15 U.S.C. § 1065).

96. 2 MCCARTHY, *supra* note 20, § 19:1.25.

97. *Id.*

98. 15 U.S.C. § 1072.

of registration can be introduced as prima facie evidence of the mark's validity and the plaintiff's ownership.⁹⁹ Third, trademark registration provides the opportunity to file incontestability. Between the fifth and sixth anniversary of registering a trademark, if the mark has been consecutively used in commerce for five years, a trademark owner can file for "incontestable status."¹⁰⁰ A defendant defending against an incontestable trademark owner's infringement claim is left with limited defenses.¹⁰¹ There are also some non-governmentally conferred advantages to registration. For example, Amazon requires trademark registration or a pending trademark registration application for inclusion of the relevant product in the Amazon Brand Registry.¹⁰²

Fundamentally, though, while federal trademark registration confers some procedural and substantive benefits, it does not change the nature of a trademark that exists, registered or unregistered, through use in business.¹⁰³ There are also trademark uses protected in the common law but not federally registerable, such as commercial trade names, and trademarks not used in interstate or foreign commerce.¹⁰⁴ Moreover, unregistered trademarks are enforceable under § 43(a) of the Lanham Act, which creates a federal cause of action for infringement of unregistered trademarks.¹⁰⁵

99. 15 U.S.C. § 1115(a). A defendant can still raise various legal and equitable defenses.

100. 15 U.S.C. § 1065.

101. 15 U.S.C. § 1115(b).

102. *See Get Started in Three Steps*, AMAZON, <https://brandservices.amazon.com/brandregistry/eligibility> (last visited Dec. 15, 2023) (describing relevant eligibility requirements to enroll in the program). The program offers a number of exclusive programs to "build and protect your brand," such as a tool to report IP infringement. *Brand Protection Quick Start Guide*, AMAZON, <https://brandservices.amazon.com/protect-brand> (last visited Dec. 15, 2023).

103. *See B & B Hardware, Inc. v. Hargis Indus., Inc.*, 575 U.S. 138, 142 (2015) (noting that "federal law does not create trademarks"); *see also San Juan Products, Inc. v. San Juan Pools of Kansas, Inc.*, 849 F.2d 468, 474 (10th Cir. 1998) (cited in 2 MCCARTHY, *supra* note 20, § 19:3) ("Unlike the registration of a patent, a trademark registration of itself *does not create the underlying right to exclude*. Nor is a trademark created by registration. While federal registration triggers certain substantive and procedural rights, the absence of federal registration does not unleash the mark to public use. The Lanham Act protects unregistered marks as does the common law.").

104. 2 MCCARTHY, *supra* note 20, § 19:8.

105. *See Two Pesos, Inc. v. Taco Cabana, Inc.*, 505 U.S. 763, 768 (1992) ("Section 43(a) prohibits a broader range of practices than does § 32, which applies to registered marks, but it is common ground that § 43(a) protects qualifying unregistered trademarks.").

B. TAM AND BRUNETTI

1. Matal v. Tam

In *Matal v. Tam*, the Court unanimously decided that the “disparagement clause” of the Lanham Act, § 2(a), was facially unconstitutional.¹⁰⁶ The case marked the first time a federal intellectual property statute was invalidated on constitutional grounds since 1879.¹⁰⁷ Section 2(a) prohibited the registration of a trademark “which may disparage . . . persons, living or dead, institutions, beliefs, or national symbols, or bring them into contempt, or disrepute.”¹⁰⁸ Based on this clause, the PTO rejected Simon Tam’s trademark registration of his band name “THE SLANTS,” citing the fact that the name, a derogatory term for persons of Asian descent, “had been found offensive numerous times.”¹⁰⁹ The Court held that such a provision violated the “bedrock First Amendment principle” that the government cannot discriminate against “ideas that offend.”¹¹⁰

One key question courts ask in deciding how to scrutinize a given restriction on speech is whether it is viewpoint-discriminatory. The test asks whether the government has singled out and disfavored a subset of messages based on the views expressed.¹¹¹ A Texas law prohibiting flag desecration, for example, was clearly viewpoint discriminatory by prohibiting one’s attitude toward the American flag.¹¹² Of all forms of speech regulation, the Court is most skeptical of viewpoint-discriminatory speech, since the “bedrock” principle of the First Amendment is that the government may not prohibit expression of an idea.¹¹³ Accordingly, such restrictions are subject to “the most exacting scrutiny” and are presumptively unconstitutional.¹¹⁴ In *Tam*, the clause was viewpoint discriminatory since “giving offense is a viewpoint,”¹¹⁵

106. 137 S. Ct. 1744 (2017).

107. The Court struck down federal trademark legislation in *The Trade-Mark Cases*, 100 U.S. 82 (1879).

108. 15 U.S.C. § 1052(a).

109. *Tam*, 137 S. Ct. at 1754.

110. *Id.* at 1751.

111. *Id.* at 1750.

112. See *Texas v. Johnson*, 491 U.S. 397, 413, n.9 (“[I]f Texas means to argue that its interest does not prefer *any* viewpoint over another, it is mistaken; surely one’s attitude toward the flag and its referents is a viewpoint.”).

113. *Id.* at 414. Viewpoint discrimination is a “form of speech suppression so potent that it must be subject to rigorous constitutional scrutiny.” *Tam*, 137 S. Ct. at 1765. (Kennedy, J., concurring in part).

114. *Johnson*, 491 U.S. at 412.

115. *Tam*, 137 S. Ct. at 1763.

and the Government's proffered justifications were insufficient under such exacting scrutiny.¹¹⁶

The problem with the Court's straightforward application of such rigorous viewpoint-discrimination scrutiny is that it allowed the Court to avoid the question of what kind of speech regulation restrictions on trademark registration are. The Court acknowledged the "debate" about whether trademarks are commercial speech, rather than expressive speech, but noted that it "need not resolve this debate . . . because the disparagement clause cannot withstand even [the lesser commercial speech scrutiny]."¹¹⁷ The Court *did*, however, decide that trademarks are private, not government speech.¹¹⁸ In so doing, it distinguished trademark registration from a state's specialty license plate approval process.¹¹⁹ Whereas license plates are subject to direct state control, are closely identified with the State in the public mind, and have been used to convey State messages, trademarks have not historically conveyed Government messages, nor does the public associate the contents of a trademark with the Government.¹²⁰

Yet the label of "private speech" provides little insight into the relevant First Amendment framework, beyond signaling that the First Amendment has restrictive import in ways it doesn't for purely government speech.¹²¹ And while the First Amendment can be broadly thought of as a protection of speech, it more practically is "a bundle of different but interrelated concepts, joined together under the oversimplifying rubric of 'freedom of speech.'"¹²² For better or worse, this has come to mean that categorization of speech

116. *Id.* at 1765.

117. *Id.* at 1764.

118. *Id.* at 1760.

119. *Id.* (distinguishing *Walker v. Tex. Div.*, 576 U.S. at 200, in which the Court held that Texas did not violate an organization's free speech in rejecting its application for a specialty license plate with a confederate battle flag).

120. *Id.* The Court also distinguished *Johanns v. Livestock Mktg. Ass'n*, 544 U.S. 550 (2005) (holding that federally mandated beef advertisements were government speech), and *Pleasant Grove City, Utah v. Summum*, 555 U.S. 460 (2009) (holding that a small city's selection of a permanent monument in a public park was government speech).

121. *See, e.g., Pleasant Grove City, Utah v. Summum*, 555 U.S. 460, 467 (2009) ("The Free Speech Clause restricts government regulation of private speech; it does not regulate government speech.").

122. Frederick Schauer, "Private" Speech and the "Private Forum": *Givhan v. Western Line School District*, 1979 SUP. CT. REV. 217, 217-18 (1979). More critically, Professor Schauer has posited that "if there exists a single theory that can explain the First Amendment's coverage, it has not yet been found." Frederick Schauer, *The Boundaries of the First Amendment: A Preliminary Exploration of Constitutional Salience*, 117 HARV. L. REV. 1765, 1786 (2004).

matters significantly in determining the extent of the First Amendment's protection.¹²³ The decision, then, was left for another day.

2. *Iancu v. Brunetti*

Just two years after *Tam*, the Court struck down another clause of § 2(a) of the Lanham Act as facially unconstitutional.¹²⁴ The clause prohibited registration of “immoral[] or scandalous” trademarks, and was used to reject Erik Brunetti's registration of the trademark FUCT, the name of his clothing line.¹²⁵ The Court again provided no framework through which trademark registration should be analyzed. As in *Tam*, the “key question” was simply whether the criterion was viewpoint-neutral or viewpoint-based,¹²⁶ which it was, since the clause “disfavors certain ideas.”¹²⁷ Importantly, the unanimous opinion reflected only the Court's determination that the bar on registration of “immoral” trademarks was viewpoint-discriminatory. Three Justices—Chief Justice Roberts, Justice Breyer, and Justice Sotomayor—dissenting in part, argued that the bar on registration on “scandalous” trademarks could be a narrowly construed, viewpoint-neutral, and reasonable regulation of expressive activity.

Chief Justice Roberts noted that refusal to register “obscene” or “vulgar” marks encompassed by a narrow reading of “scandalous” would not offend the First Amendment—“regardless of how exactly the trademark registration system is best conceived . . . a question we left open in *Tam*.”¹²⁸ Notably, he distinguished the stakes of trademark speech from that of more traditional

123. “[I]dentifying the category of speech at issue (e.g., commercial speech, obscenity) is an important step in determining what First Amendment standards, including what level of judicial scrutiny, a court might apply to the law.” VICTORIA L. KILLION, CONG. RESEARCH SERV., IF11072, *THE FIRST AMENDMENT: CATEGORIES OF SPEECH* 1 (2019); see also Joseph Blocher, *Categoricalism and Balancing in First and Second Amendment Analysis*, 84 N.Y.U. L. REV. 375, 397 (“Rather than fully embracing categorization or balancing at all levels of analysis, First Amendment doctrine generally combines the two, for example by using balancing or other standard-like tests to establish the borders of constitutional coverage and then applying categorical rules to speech in certain subcategories [or vice versa].”). In oral argument in one case, Chief Justice Roberts noted that such balancing standards “just kind of developed over the years as sort of baggage that the First Amendment picked up.” *Id.* (internal citation omitted). This Note does not address the many critiques of First Amendment category delineation and tiers of scrutiny, and instead describes a solution that fits within such “baggage,” i.e., the Court's established doctrine.

124. *Brunetti*, 139 S. Ct. at 2297.

125. *Id.*

126. *Id.* at 2299.

127. *Id.* at 2297. The PTO examining attorney determined that FUCT was a “total vulgar” and “therefore . . . unregistrable” mark. *Id.*

128. *Id.* at 2303 (Roberts, C.J., concurring in part and dissenting in part).

speech, writing: “no speech is being restricted; no one is being punished. The owners of such marks are merely denied additional benefits.”¹²⁹ Justice Sotomayor, too, downplayed the stakes of trademark speech restriction, noting that they are “far removed” from a situation threatening Brunetti’s liberty, or even his right to use or enforce his trademark.¹³⁰ She likened the registration system either to a government subsidy or a limited public (or nonpublic) forum, two conceptions of forums for speech different from traditional expressive speech; under either framework, “reasonable viewpoint-neutral content discrimination is generally permissible.”¹³¹ Finally, Justice Breyer wrote that the trademark statute fits into no First Amendment categories, and should be subjected to a “proportionality” analysis.¹³²

IV. FIRST AMENDMENT FRAMEWORKS

Tam and *Brunetti* were straightforwardly decided because the provisions invoked the most “egregious form of content discrimination”¹³³ — discrimination based on viewpoint. However, a restriction that is viewpoint-neutral but nevertheless discriminates based on the content of speech is still subject to strict scrutiny.¹³⁴ In other words, such content discrimination is presumptively unconstitutional and justified only if the government proves such restrictions are narrowly tailored to serve compelling state interests (a standard only somewhat less demanding than “exacting” scrutiny).¹³⁵

To determine whether a provision is content-neutral, the Court asks whether it can be justified without reference to the content of the regulated speech.¹³⁶ As an illustration, consider the difference between two sign regulation cases that reached the Court. In *City Council of Los Angeles v. Taxpayers for Vincent*, a political sign service company challenged the City of Los

129. *Id.*

130. *Id.* at 2312 (Sotomayor, J., concurring in part and dissenting in part).

131. *Id.* at 2317.

132. *Id.* at 2306 (Breyer, J., concurring in part and dissenting in part). In *Reed v. Town of Gilbert*, 576 U.S. 155, 175 (2015), Justice Breyer similarly expressed his frustration with First Amendment categories. He wrote, “I believe that categories alone cannot satisfactorily resolve the legal problem before us. The First Amendment requires greater judicial sensitivity both to the Amendment’s expressive objectives and to the public’s legitimate need for regulation than a simple recitation of categories, such as ‘content discrimination’ and ‘strict scrutiny’ would permit. In my view, the category “content discrimination” is better considered in many contexts, including here, as a rule of thumb, rather than as an automatic ‘strict scrutiny’ trigger, leading to almost certain legal condemnation.”

133. *Rosenberger v. Rector & Visitors of Univ. of Va.*, 515 U.S. 819, 829 (1995).

134. *Reed*, 576 U.S. at 2227.

135. *Id.*

136. *Id.*

Angeles's removal of the group's political signs from utility poles pursuant to an ordinance prohibiting the posting of signs on various forms of public property.¹³⁷ The Court found that such an ordinance was content-neutral and impartially applied, impacting all forms of speech in the same way.¹³⁸ In contrast, in *Reed v. Town of Gilbert*, a church challenged a city's code which provided differing restrictions on signs based on their categories as, for example, an "ideological sign" or "political sign."¹³⁹ The Court found the code was content-discriminatory, and ultimately unconstitutional.¹⁴⁰

Section 2(c) and the other § 2 provisions clearly discriminate on the basis of content.¹⁴¹ For example, § 2(c) looks to the content of a trademark to ask: does this mark identify an individual without their consent? Without a conception of such speech "regulation" that subjects the § 2 provisions to a lower level of scrutiny, they are potentially all at risk of being struck down as, at a minimum, unconstitutional as-applied in certain scenarios.¹⁴² The Court's selection of a framework would provide guidance to Congress, trademark examiners, trademark registrants, and courts about the Government's ability to restrict trademarks from registration by articulating the level of scrutiny with which to compare the regulation to countervailing interests, such as protection of political speech.

In their partial dissents in *Brunetti*, the three Justices stated that the "scandalous" provision could be saved if it were construed in a viewpoint-neutral way. In so doing, they implied that trademark registration is different than other content-based restrictions; they noted that "scandalous" trademarks could comport with the reasonable or general interests of the Government without reference to "compelling governmental interests" or "narrow tailoring."¹⁴³ In other words, they implied that trademark registration is subject

137. 466 U.S. at 793.

138. *Id.* at 817.

139. 576 U.S. at 2227.

140. *Id.* at 2232.

141. *See Brunetti*, 139 S. Ct. at 2317 n.12 (Sotomayor, J., concurring in part and dissenting in part) ("Though I do not address the constitutionality of provisions not before the Court, I note as well that the 'scandalous' bar in § 1052(a) is hardly the only provision in § 1052 that could be characterized as content discriminatory. *See, e.g.*, § 1052(b) (no flags or insignias); § 1052(c) (no unapproved markers of deceased U.S. Presidents during the lives of their spouses)."); Tushnet, *supra* note 41, at 382 ("Section 2 is almost nothing but content-based.").

142. *See In re Elster*, 26 F.4th at 1331 ("Neither *Tam* nor *Brunetti* resolves the constitutionality of section 2(c). Both holdings were carefully cabined to the narrow, 'presumptive[] unconstitutional[ity]' of section 2(a)'s viewpoint-based restrictions.") (internal citations omitted).

143. *See Brunetti*, 139 S. Ct. at 2308 (Sotomayor, J., concurring in part and dissenting in part) ("Such a narrowing construction would save that duly enacted legislative text by rendering it a reasonable, viewpoint-neutral restriction on speech that is permissible in the

to a lower level of scrutiny than expressive speech (such as rational basis review).

This Part will explore the First Amendment frameworks that the Justices considered in both *Tam* and *Brunetti*—trademark registration as a regulation of commercial speech, as a government subsidy and as a limited public forum. Though some Justices expressed a preference for the appropriate framework—Justice Sotomayor for a non-cash government program or limited public forum¹⁴⁴ and Justice Alito for a limited public forum¹⁴⁵—all explicitly left open the question for a future time.¹⁴⁶

The *Elster* court also declined to select a framework, though it only conceptualized trademarks as either expressive political speech or commercial speech.¹⁴⁷ In support of the protection of expressive speech, the *Elster* court cites, among other cases, *Cohen v. California*, in which a court upheld First Amendment protection of a jacket with the words “Fuck the Draft.”¹⁴⁸ That trademark registration is not a restriction on expressive speech should be clear from Section II.A. Preventing someone from wearing a jacket with certain words is an imposition on free expression, whereas refusing to register a trademark prevents no expression from being spoken or sold.¹⁴⁹ However,

context of a beneficial governmental initiative like the trademark-registration system.”), 2307 (Breyer, J., concurring in part and dissenting in part) (“The Government has at least a reasonable interest in ensuring that it is not involved in promoting highly vulgar or obscene speech.”), 2303 (Roberts, C.J., concurring in part and dissenting in part) (“The Government . . . has an interest in not associating itself with trademarks whose content is obscene, vulgar, or profane.”).

144. *See id.* at 2316 n.10 (Sotomayor, J., concurring in part and dissenting in part) (“Trademark registration differs [from cash-subsidy government programs] because any ‘subsidy’ comes in the form of a noncash benefit, but that difference does not foreclose understanding the registration system as a beneficial, noncash governmental program. No Justice, meanwhile, rejected the limited-public-forum analogy . . . and scholars have noted arguments for adopting it.”).

145. *Tam*, 137 S. Ct. at 1763 (noting that trademark registration is “potentially more analogous” to a limited public forum than government programs or subsidies, the latter of which he deemed “nothing like” the programs at issue). Chief Justice Roberts, and Justices Thomas and Breyer also joined Justice Alito in this opinion.

146. *See Brunetti*, 139 S. Ct. at 2302 (declining to “say anything about how to evaluate viewpoint-neutral restrictions on trademark registration.”).

147. *See In re Elster*, 26 F.4th at 1338–39.

148. *Id.* at 1333.

149. *See* Sonia K. Katyal, *Trademark Intersectionality*, 57 UCLA L. REV. 1601, 1646 (2010) (“The phrase in *Cohen* involved pure expression in the marketplace of ideas, whereas the same phrase in the Lanham Act context involves federal registration of a commodity as a brand in the marketplace of goods.”). Professor Katyal also noted that an additional difference is that *Cohen* involved a “criminal prohibition on speech” rather than a refusal to registration, which involves no speech prohibition. *Id.*

there is more academic support for trademark registration as a regulation of commercial speech,¹⁵⁰ so, though various Justices have seemed to express more support for either the limited public forum or government subsidy framework, it is worth exploring more fully.

A. COMMERCIAL SPEECH

1. Overview

Commercial speech is evidence of the moving target of First Amendment protections.¹⁵¹ Though rejected as outside the bounds of the First Amendment in 1942,¹⁵² it was accepted as protectable speech in 1976 in *Virginia State Board of Pharmacy v. Virginia Citizens Consumer Council, Inc.*¹⁵³ There, the Court considered a restriction prohibiting pharmacists from advertising the truthful price of their drugs.¹⁵⁴ Virginia had reasoned that if pharmacists could cut corners to offer and advertise low prices, they might be able to trick consumers into their low-quality services and harm them.¹⁵⁵ The Court was skeptical of this “highly paternalistic” approach—outright banning truthful speech—and found that in the absence of any tangible evidence of deception, the restriction was unconstitutional under the First Amendment.¹⁵⁶ The Court made it clear that commercial speech was different from expressive speech, and a lower degree of protection of the speech was needed to ensure an unimpaired flow of “truthful and legitimate commercial information.”¹⁵⁷ The Court articulated a balancing test just four years later in *Central Hudson Gas & Electric Corp. v. Public Service Commission of New York*.¹⁵⁸ The first prong (1) asks whether the commercial speech concerns lawful activity and is not misleading. If the speech is commercial in nature but misleading or unlawful, “there can be no constitutional objection to the suppression of commercial messages that do

150. See, e.g., Ned Snow, *Denying Trademark for Scandalous Speech*, 51 U.C. DAVIS L. REV. 2331, 2363 n.125 (“It seems likely that the test would apply given that trademark rights do not exist unless there is a bona fide use in commerce of the mark.”).

151. See Amanda Shanor, *First Amendment Coverage*, 93 N.Y.U. L. REV. 318, 326 (noting that commercial speech was “once explicitly excluded from First Amendment coverage altogether, but was several decades later swept within the First Amendment’s ambit,” and that “[t]he scope of the First Amendment is dynamic, not static”).

152. See *Valentine v. Chrestensen*, 316 U.S. 52, 54 (1942) (holding that in contrast to communicating information and opinions in the street, “the Constitution imposes no such restraint on government as respects purely commercial advertising”).

153. 425 U.S. 748, 770 (1976).

154. *Id.*

155. *Id.* at 769.

156. *Id.* at 770.

157. *Id.* at 771 n.24.

158. 447 U.S. 557, 563 (1980).

not accurately inform the public about lawful activity.”¹⁵⁹ The next three prongs of the test articulate the “intermediate” scrutiny a court should apply, and ask whether: (2) the government interest is substantial, (3) the regulation directly advances the governmental interest asserted, and (4) the regulation is not more extensive than necessary to serve that interest.¹⁶⁰

Typically, the central inquiry in commercial speech cases is into the consumer harm which such commercial speech restrictions seek to alleviate.¹⁶¹ This inquiry often depends on the nature of the audience of certain commercial speech.¹⁶² For example, in *Edenfield v. Fane*, the Court found that a restriction on direct solicitations by CPAs was a violation of the First Amendment, distinguishing the decisions from the Court’s upholding of a ban on solicitations by attorneys, because “[t]he typical client of a CPA is far less susceptible to manipulation.”¹⁶³ And in *Lorillard Tobacco Co. v. Reilly*, the Court found that despite the Government’s interest in protecting children, a regulation prohibiting outdoor tobacco advertising near schools was unduly restrictive because “tobacco retailers and manufacturers have an interest in conveying truthful information about their products to *adults*, and adults have a corresponding interest in receiving truthful information about tobacco products.”¹⁶⁴

Such audience orientation, however, can nevertheless create tension with what the Court often perceives as paternalistic government measures. In *Virginia State Board of Pharmacy*, the Court questioned the state Board of Pharmacy’s approach of banning the advertisement of drug prices as “highly paternalistic.” In considering the choice between suppression of advertisements and the danger of misuse of advertisement information, the Court stated that the First Amendment makes the choice for the Court—to “open the channels of communication.”¹⁶⁵ This anti-paternalistic approach has taken on such a role in recent commercial speech opinions that the scrutiny of regulation of commercial speech has arguably been elevated to that of

159. *Id.*

160. *Id.*

161. *See* *Edenfield v. Fane*, 507 U.S. 761, 770–71 (1993) (“[A] governmental body seeking to sustain a restriction on commercial speech must demonstrate that the harms it recites are real and that its restriction will in fact alleviate them to a material degree.”).

162. *See* Robert Post, *The Constitutional Status of Commercial Speech*, 48 UCLA L. REV. 1, 14 (2000) (describing commercial speech as “sharply audience oriented”).

163. 507 U.S. at 775.

164. 533 U.S. 525, 564 (2001) (emphasis added).

165. *Va. State Bd. of Pharmacy v. Va. Citizens Consumer Council*, 425 U.S. 748, 770 (1976).

expressive speech, beyond the “intermediate” scrutiny of *Central Hudson*.¹⁶⁶ In *Thompson v. Western States Medical Center*, for example, the Court stated that even if the government had argued that it prohibited certain drug advertisements because of a fear that such advertisements put people at risk, “this concern amounts to a fear that people would make bad decisions if given truthful information about compounded drugs.”¹⁶⁷

Nearly all the regulations that the Court has recently considered under the commercial speech doctrine involve outright prohibitions of certain forms of speech. Such cases include an FDA regulation prohibiting manufacturers from advertising compound drugs;¹⁶⁸ a prohibition on advertising the price of alcohol;¹⁶⁹ a prohibition on the sale of pharmacy records that reveal prescribing practices of individual doctors,¹⁷⁰ and a federal ban on stating the alcohol content on beer labels.¹⁷¹ In each case, the court rejected the government’s attempts to restrict truthful, factual information as violative of the First Amendment given the lack of fit between the harm in question and the government’s approach to alleviating such harm.¹⁷² In addition, the Court has suggested that disclosure and disclaimer requirements are preferable to pure speech suppression.¹⁷³ Alternatively, government speech on its own is preferable to speech suppression in order to accomplish the government’s policy or consumer protection goals.¹⁷⁴

166. See, e.g., *Sorrell v. IMS Health Inc.*, 564 U.S. 552, 577 (2011) (noting that the “fear that people would make bad decisions if given truthful information” cannot justify content-based burdens on speech). The Court also stated that the First Amendment “directs us to be especially skeptical of regulations that seek to keep people in the dark for what the government perceives to be their own good.” *Id.* (citing *44 Liquormart, Inc. v. Rhode Island*, 517 U.S. 484, 503 (1996) (opinion of Stevens, J.)).

167. 535 U.S. 357, 374 (2002).

168. *Id.*

169. *44 Liquormart, Inc. v. Rhode Island*, 517 U.S. 484 (1996).

170. *Sorrell v. IMS Health Inc.*, 564 U.S. 552 (2011).

171. *Rubin v. Coors Brewing Co.*, 514 U.S. 476 (1995).

172. That each case restricted truthful information was crucial to the Court. See *Sorrell*, 564 U.S. at 579 (“The State nowhere contends that detailing is false or misleading within the meanings of this Court’s First Amendment precedents.”).

173. Rebecca Tushnet, *Trademark Law as Commercial Speech Regulation*, 58 S.C. L. REV. 737, 748 (2006) (citing *Zauderer v. Off. of Disciplinary Counsel*, 471 U.S. 626, 651 (1985)); see also *In re R. M. J.*, 455 U.S. 191, 201 (1982) (“[A] warning or disclaimer might be appropriately required . . . in order to dissipate the possibility of consumer confusion or deception.”) (internal citations omitted).

174. See *Sorrell*, 564 U.S. at 578 (“Vermont may be displeased that details who use prescriber-identifying information are effective in promoting brand-name drugs. The State can express that view through its own speech.”); see also *44 Liquormart*, 517 U.S. at 498 (“[T]he remedy to be applied is more speech, not enforced silence. Only an emergency can justify repression.”) (plurality opinion) (internal citations omitted).

2. *Analogy to Trademark*

Trademark registration should not be characterized as regulation of commercial speech. As an initial matter, that a slogan or form of expression is embroidered on merchandise or generally for sale does not necessarily transform it into commercial speech. However, even if all marks are conceived of as commercial speech, trademark registration decisions do not take on the form of commercial speech regulation as an analytical framework. First, trademark registration decisions do not involve outright prohibitions of speech, but rather the denial of access to the benefits of registration. Second, while some registration provisions under § 2 bar registration if a mark is deceptive or misleading, others, including § 2(c), facilitate the protection of third party rights already provided by common law. In other words, there is no tension between the “right to information” and paternalistic goals, because consumers are not deprived of information, nor is the purpose of a provision like § 2(c) to protect consumers.

As Chief Justice Roberts noted in his opinion in *Brunetti*, in the face of trademark registration denial, “[n]o speech is being restricted; no one is being punished.”¹⁷⁵ Whether a mark is registered does not prevent its use in business or commerce. Steve Elster can sell any number of products with the phrase “Trump too small” absent a trademark registration. In fact, the PTO, in its early rejections of Elster’s application, noted multiple items for sale without trademark registration such as a t-shirt with the phrase “Can’t Build a Wall If Your Hands Are Too Small”¹⁷⁶ and a soap called “Trump’s Small Hand Soap.”¹⁷⁷ The only way Elster could be hampered in a material way from selling his envisioned t-shirts would be if Donald Trump himself registered (or gave consent to another to register) the mark and made sufficient use of the mark to satisfy the requirement for “use in commerce.”¹⁷⁸

175. *Brunetti*, 139 S. Ct. at 2303.

176. *Can’t Build Wall Hands Too Small Justice Baby Trump T-Shirt*, AMAZON, <https://www.amazon.com/BUILD-HANDS-SMALL-JUSTICE-TRUMP^=dp/B07FLTBVYZ> (last visited Dec. 18, 2023).

177. *Trump’s Small Hands Soap – Republican and Democrat – Made in the USA*, AMAZON, <https://www.amazon.com/Trumps-Small-Hands-Soap-Republican/dp/B076JKNJ41> (last visited Dec. 18, 2023).

178. Samuel Ernst considered this in a recent Article, noting the unlikelihood of Trump selling shirts informing the public that his policies and body parts were “TOO SMALL,” and adding that even if Trump were to do that, it would be hard for him to prove likelihood of confusion. See Samuel Ernst, *Trump Really Is Too Small: The Right To Trademark Political Commentary*, 88 BROOK. L. REV. 839, 872 (2023). Professor Ernst does, however, argue that § 2(c) is unconstitutional.

In addition, § 2(c), and other provisions of § 2, have a purpose beyond preventing consumer deception or protecting consumers in the vein of an FDA regulation, or tobacco advertising restriction. In contrast, the commercial speech analysis hinges on whether the speech being policed is misleading or truthful. Rebecca Tushnet has noted that the goal of restricting commercial speech in fostering access to truthful information “depends on truth and falsity being pure binaries.”¹⁷⁹ Yet, as discussed in Section III.C, § 2(c) has purpose beyond assisting the policing of deceptive trademarks. Rather, it tackles a kind of dilution-like interest of a celebrity or known person in preventing “diminishment,” as acknowledged by common law rights of privacy and publicity.¹⁸⁰ Even if one thinks such interests are not sufficiently worthy of preservation (though there is certainly support for a federal right of publicity),¹⁸¹ leaving room for Congress to be able to enact legislation with meaningful purpose beyond the bounds of falsity is.

In contrast, in *In re Elster*, the Federal Circuit stated that “no plausible claim” could be made that President Trump has endorsed the shirt, and therefore he couldn’t possibly have a right of publicity claim.¹⁸² Under this interpretation, any PTO rejection of a trademark based on § 2(c) without false endorsement would not survive commercial speech scrutiny. And even if some restrictions could survive in certain fact patterns that suggests false endorsement, the purpose of § 2(c), or any restrictions which Congress may want to enact that reach beyond deception, may be lost. For example, in *Native American Arts, Inc. v. Waldron Corp.*, Judge Posner reversed the district judge’s finding that the Indian Arts and Crafts Act (IACA), which forbids selling a good “in a manner that falsely suggests it is . . . an Indian product” is unconstitutional.¹⁸³ However, he affirmed the district court’s judgement for the defendant despite plaintiffs’ contention that the instructions to the jury did not sufficiently capture the purpose of the IACA.¹⁸⁴ The plaintiffs argued that there must be a qualifier or disclaimer along with the use of a misleading word, since the statute forbids “unqualified use” specifically. However, Judge Posner

179. Tushnet, *supra* note 173, at 748.

180. *See supra* Section II.B.1.

181. *See, e.g.*, Kevin L. Vick & Jean-Paul Jassy, *Why a Federal Right of Publicity Statute is Necessary*, 28 COMM. LAW. 14 (2011); Brittany Lee-Richardson, *Multiple Identities: Why the Right of Publicity Should be a Federal Law*, 20 UCLA ENT. L. REV. 190 (2013).

182. *In re Elster*, 26 F.4th at 1335.

183. 399 F.3d 871, 874 (7th Cir. 2005) (“A non-Indian maker of jewelry designed to look like jewelry made by Indians is free to advertise the similarity but if he uses the word ‘Indian’ he must qualify the usage so that consumers aren’t confused and think they’re buying not only the kind of jewelry that Indians make, but jewelry that Indians in face made. There is no constitutional infirmity.”).

184. *Id.* at 875.

found that a context-based false association finding by the jury was sufficient, since “[w]e expect the jury would have been confused by such a regulation.”¹⁸⁵ The court thus narrowed the IACA to be a regulation rooted in consumer protection from deception alone, and eschewed the broader interests of and protections for Native American cultural heritage that formed the foundation of the Act.¹⁸⁶ As Rebecca Tushnet noted, “[t]his is part of a broader dynamic that encourages a regulation’s defenders to define their goal as avoiding falsity, since that is the only aim for which current First Amendment commercial speech doctrine has any sympathy.”¹⁸⁷

B. GOVERNMENT SUBSIDIES AND GOVERNMENT PROGRAMS

1. *Overview*

Government subsidies and government programs are an outgrowth of government speech doctrine. In both situations, the Court is deferent to the government’s decisions to selectively speak, recruit private speakers to speak on its behalf, or selectively provide subsidies, even if such decisions include viewpoint-discrimination. Two key theories underlie this deference. First, where the government speaks, either on its own or through a private entity, it is accountable to the electorate, and the citizenry has the ability to elect new officials if it disagrees with the substance of such speech.¹⁸⁸ This doctrine “reflects the fact that it is the democratic electoral process that first and foremost provides a check on government speech.”¹⁸⁹ Second, the Constitution does not confer an entitlement to funds, even if a lack of such funds prevents a full expression of free speech. In other words, in government program cases, “a decision not to subsidize the exercise of a fundamental right does not infringe the right, and thus is not subject to strict scrutiny.”¹⁹⁰

In government program cases, the government appropriates public funds to establish a program, and “is entitled to broadly define that program’s limits.”¹⁹¹ In other words, discriminating based on viewpoint is reframed as “simply insisting that public funds be spent for the purpose for which they are authorized.”¹⁹² The Court has upheld government program restrictions such as a federal act prohibiting public libraries from receiving certain federal

185. *Id.*

186. Tushnet, *supra* note 173, at 753.

187. *Id.*

188. *Legal Servs. Corp. v. Velazquez*, 531 U.S. 533, 541 (2001).

189. *Walker v. Texas Div.*, 576 U.S. at 207.

190. *Regan v. Tax’n with Representation of Wash.*, 461 U.S. 540, 549 (internal citations omitted).

191. *United States v. Am. Libr. Ass’n*, 539 U.S. 194, 196 (2003).

192. *Id.*

assistance unless they install software blocking obscene or pornographic images,¹⁹³ and a federal regulation that prohibited a medical professional receiving Title X funding from providing abortion counseling.¹⁹⁴

In subsidy cases too, the Court gives Congress “wide latitude” to set its spending priorities.¹⁹⁵ More explicitly, the government “may allocate competitive funding according to criteria that would be impermissible were direct regulation of speech or a criminal penalty at stake.”¹⁹⁶ So, in *National Endowment for the Arts v. Finley*, the Court upheld an act which required a consideration of “decency and respect” in awarding artistic grants.¹⁹⁷

2. *Analogy to Trademark*

Restrictions on government subsidy and government speech are not analogous to trademark registration because trademarks are private, not government speech.¹⁹⁸ As the Court noted, “[t]he Government does not dream up these marks, and it does not edit marks submitted for registration.”¹⁹⁹ Moreover, registration of a mark does not constitute approval of the mark. Finally, the Court held it is unlikely that “more than a tiny fraction of the public has any idea what federal registration of a trademark means.”²⁰⁰ Accordingly, the “accountability” which makes deference to government speech palatable is nearly non-existent in the trademark context.

C. LIMITED PUBLIC FORUM

1. *Overview*

When private speech is restricted on government property, Courts ask whether the property falls within one of a three types of “fora”: a traditional public forum,²⁰¹ a designated public forum, or a limited public forum.²⁰²

193. *Id.* at 195.

194. *Rust v. Sullivan*, 500 U.S. 173 (1991).

195. *Nat'l Endowment for the Arts v. Finley*, 524 U.S. 569, 587 (1998).

196. *Id.* at 571.

197. *Id.*

198. *Tam*, 137 S. Ct. at 1757.

199. *Id.* at 1758.

200. *Id.* at 1759.

201. “When government regulation discriminates among speech-related activities in a public forum, the Equal Protection Clause mandates that the legislation be finely tailored to serve substantial state interests.” *Carey v. Brown*, 447 U.S. 455, 461 (1980). Examples of traditional public forum include: streets and parks that since “time out of mind, have been used for purposes of assembly, communicating thoughts between citizens, and discussing public questions.” *Perry*, 460 U.S. at 45.

202. In first introducing the tripartite framework, the Court labeled the third type of forum as a “nonpublic forum.” *Perry*, 460 U.S. at 45. The Court has since used “nonpublic forum” and “limited public forum” interchangeably. *See, e.g.*, *Minn. Voters Alliance v. Mansky*,

Government property can mean both physical bounds of government ownership, such as government-owned airport terminals²⁰³ and school district mail systems,²⁰⁴ but also property in a “metaphysical” sense, such as a student activities fund²⁰⁵ or a public television broadcast.²⁰⁶ In these cases, the same type of speech restrictions may be valid or invalid depending on the location (metaphysical or otherwise) of the speaker.²⁰⁷

Examples of traditional public forums include streets and parks, which, since “time out of mind,” have been used “for purposes of assembly, communicating thoughts between citizens, and discussing public questions.”²⁰⁸ The state may enforce restrictions on a traditional public forum only if it is content-neutral, narrowly tailored to serve a significant government interest, and “leave[s] open ample alternative channels of communication.”²⁰⁹ Importantly, the Court has held that the government does not create a public forum by inaction, but by intentionally opening up the forum for public discourse.²¹⁰ A government entity can create a *designated* public forum if government property that has not traditionally been regarded as a public forum is intentionally opened up for that purpose,²¹¹ whereby restrictions on speech in such a forum are subject to the same levels of scrutiny as a traditional public forum.²¹²

Finally, the *limited* public forum is a government property (which, again, can be “metaphysical” in nature) that the government has reserved for a specific purpose. “Implicit” in this concept is “the right to make distinctions

138 S. Ct. 1876, 1885 (“Generally speaking, our cases recognize three types of government-controlled spaces: traditional public forums, designated public forums, and nonpublic forums.”); *Christian Legal Soc’y*, 561 U.S. at 679 n.10 (“In conducting forum analysis, our decisions have sorted government property into three categories. First, in traditional public forums Second, [when] governmental entities create designated public forums Third [when] governmental entities establish limited public forums.”).

203. *Int’l Soc’y for Krishna Consciousness, Inc. v. Lee*, 505 U.S. 672 (1992).

204. *Perry Ed. Ass’n v. Perry Local Educators’ Ass’n*, 460 U.S. 37 (1983).

205. *Rosenberger v. Rector & Visitors of Univ. of Va.*, 515 U.S. 819, 830 (1995) (“The [Student Activities Fund] is a forum more in a metaphysical than in a spatial or geographic sense, but the same principles are applicable.”).

206. *Arkansas Educ. Television Comm’n v. Forbes*, 523 U.S. 666 (1998).

207. See Daniel A. Farber & John E. Nowak, *The Misleading Nature of Public Forum Analysis: Content and Context in First Amendment Adjudication*, 70 VA. L. REV. (1984) (“Public forum analysis might well be called the ‘geographical’ approach to first amendment law, because results often hinge almost entirely on the speakers’ location.”).

208. *Perry*, 460 U.S. at 45.

209. *Id.*

210. *Cornelius v. NAACP Leg. Def. Fund*, 473 U.S. 788, 802 (1985).

211. *Id.*

212. *Id.*

in access on the basis of subject matter and speaker identity.”²¹³ These distinctions may be impermissible in a public forum but are inherent and inescapable in the process of limiting a nonpublic forum to activities compatible with the intended purpose of the property.”²¹⁴ The core of this concept is the holding that “government has much more flexibility to craft rules limiting speech,”²¹⁵ but also that the Constitution does not require free access “on every type of Government property without regard to the nature of the property or to the disruption that might be caused by the speaker’s activities.”²¹⁶ As such, as long as the regulation of speech is reasonable for the forum’s intended purpose, and not intended to suppress viewpoints contrary to a public official, it is not in violation of the First Amendment.²¹⁷

For example, a ban on wearing a political badge, button, or political insignia “plainly restricts a form of expression within the protection of the First Amendment.”²¹⁸ However, because a polling place on election day is a nonpublic forum, set aside for the sole purpose of voting, the restriction is only subject to a “reasonable” review in light of the purpose of voting.²¹⁹

To ascertain the creation of such a forum, the Court has looked to (1) the policy and practice of the government and (2) the nature of the property and its compatibility with expressive activity.²²⁰ On (1), the Court uses policy and practice indicia to “ascertain whether [the government] intended to designate a place not traditionally open to assembly and debate as a public forum.”²²¹ On (2), the Court asks whether the forum is clearly a space for expressive activity such that it should be deemed a public forum. That a forum may be used for communication of information and ideas is *not* sufficient to transform the space into a public forum.²²² Moreover, the existence of alternative channels is important to the determination of the forum—“[r]arely will a

213. *Perry*, 460 U.S. at 49.

214. *Id.* at 46.

215. *Id.*

216. *Cornelius*, 473 U.S. at 800.

217. *Id.*

218. *Minn. Voters Alliance*, 138 S. Ct. at 1885.

219. *Id.*

220. *Cornelius*, 473 U.S. at 802.

221. *Id.*

222. *See* *United States Postal Serv. v. Council of Greenburgh Civic Ass’ns*, 449 U.S. 1076 (1981) (holding that U.S. Postal Service letterbox was not a “public forum” even if it was the most efficient means for communication); *Greer v. Spock*, 424 U.S. 828 (1976) (holding that bulletin board in a military base cafeteria not a public forum); *Lehman v. City of Shaker Heights*, 418 U.S. 298 (1974) (holding that advertising space made available in public transportation in a city was not a public forum despite being “specifically used for the communication of information and ideas”).

nonpublic forum provide the only means of contact with a particular audience.”²²³

A recent example of limited public forum analysis is *Christian Legal Society v. Martinez*,²²⁴ in which the Court rejected a First Amendment challenge to Hastings Law School (now UC Law San Francisco) conditioning official school registration on compliance with an “all-comers policy,” pursuant to which student groups must allow any student to participate regardless of that student’s beliefs.²²⁵ The Court found that the case “fit[] comfortably” within the limited public forum category.²²⁵ The Christian Legal Society, which brought the challenge, sought “what is effectively a state subsidy” and faced only indirect pressure to modify its membership policy; it was not prohibited from excluding individuals if it simply forwent the benefits of official school recognition, such as the ability to place announcements in a school newsletter and send emails using a Hastings official email.²²⁶ Further, limited public forum analysis better accounted for the fact that Hastings was “dangling the carrot of subsidy, not wielding the stick of prohibition.”²²⁷ The existence of alternative channels for the group was also significant in the Court’s reasonableness analysis—alongside the fact that Hastings’ interest in the all-comers policy, including its interest in minimizing the “daunting labor” of having to inquire into each club’s rejection of a student, was reasonable, the policy was made “all the more creditworthy” by the significant alternative channels available.²²⁸ Although an unregistered club could not take advantage of certain methods of communication, “the advent of electronic media and social-networking sites reduces the importance of those channels.”²²⁹

2. *Analogy to Trademark*

The trademark registration system can be most appropriately analogized to the limited public forum because it is a forum for private “speech” that offers limited access to benefits, that allows for significant doctrinal emphasis on access to alternative channels. The limited public forum, unlike commercial speech regulation, does not involve outright prohibition of speech. In fact,

223. *Cornelius*, 473 U.S. at 809 (emphasizing “access to alternative channels” beyond the federal contribution system in question, including “direct mail and in-person solicitation outside the workplace”).

224. 561 U.S. 671 (2010).

225. *Id.*

226. *Id.* at 682.

227. *Id.* at 683.

228. *Id.*

229. *Id.*

“[r]arely will a nonpublic forum provide the only means of contact with a particular audience.”²³⁰

As discussed in Section IV.1.a, in determining the existence of a limited public forum, the Court looks to (1) the policy and practice of the government, and (2) the nature of the property and its compatibility with expressive activity. In the first prong, the Court attempts to ascertain whether the government intended to designate a place not traditionally open to assembly and debate as a public forum. For trademark registration, it is clear that in promulgating various requirements for registration, Congress did not intend to designate such registration as a public forum open to all who wish to register trademarks. On the second prong, trademark registration was not created for the purpose of providing or incentivizing a forum for expressive activity.²³¹ That expressive activity may be compatible with trademark registration, “does not imply that the forum thereby becomes a public forum for First Amendment purposes.”²³²

As to “alternative channels,” trademark holders can “communicate” with their desired audience regardless of registration. In other words, the PTO is “dangling the carrot of subsidy” rather than compelling inclusion in a forum it controls.²³³ In *Christian Legal Society*, Hastings dangled the carrot of official club recognition with benefits such as access to school facilities, bulletin boards to advertise events, and use of Hastings’ name and logo. In the trademark registration context, the PTO dangles the carrot of registration with benefits such as a presumption of validity and nationwide notice of ownership as of the registration date.

The above is reaffirmed both by the fact that the Justices seemed to express preference for this model,²³⁴ and that the conclusion subjects trademark registration restrictions to a rational basis review without forcing it

230. *Cornelius*, 473 U.S. at 809 (“Here . . . the speakers have access to alternative channels, including direct mail and in person solicitation outside the workplace, to solicit contributions from federal employees.”).

231. *See supra* Section III.A (noting that federal trademark registration does not alter the rights retained in common law unregistered trademarks, which arise from use in commerce).

232. *Cornelius*, 473 U.S. at 802 (“[T]hat such [expressive] activity occurs in the context of the forum created does not imply that the forum thereby becomes a public forum for First Amendment purposes.”).

233. *See Christian Legal Soc’y*, 561 U.S. at 682.

234. Justice Alito, in his opinion in *Tam*, joined by three Justices, referred to the limited public forum as “potentially more analogous” to trademark registration than the frameworks of government subsidies or government programs. 137 S. Ct. at 1763. Justice Sotomayor, in her opinion in *Brunetti*, noted that in contrast to four Justices’ rejection of cash-subsidy programs as a model for understanding trademark registration, “[n]o Justice, meanwhile, rejected the limited-public-forum analogy.” 139 S. Ct. at 2316 n.10 (concurring in part and dissenting in part).

to become a vessel of false advertising law. In allowing the increased deference of the limited public forum framework, courts will create consistency and clarity for PTO examiners and trademark applicants—with the knowledge that certain provisions aren't subject to potential carveouts—and Congress—with the ability to make further decisions about how best to manage the federal trademark registration system on top of common law. Moreover, while Elster is free to build brand goodwill around the trademark TRUMP TOO SMALL without a registered trademark, preventing registration could potentially increase free expression. Were he to be granted the mark registration, he could more perhaps be more equipped to deter the creation of similar parodies of a national moment.

V. PARODIC USE CARVEOUT?

An illumination of the trademark registration system through First Amendment doctrinal analysis may not sufficiently quell a certain discomfort with the protection of former President Trump's rights to prevent trademark registration of a trademark making fun of him. Surely, for such extreme cases—a parody of a political figure and “the least private name in American life”²³⁵—one might hope there could be a parodic use carveout of § 2(c).²³⁶ Most presidents and political figures, in fact, tend to ignore the use of their names on products.²³⁷ Moreover, as discussed *supra* Section II.A, the purpose of § 2(c) is to protect rights of privacy and publicity of living persons; most states that recognize a right of publicity acknowledge a First Amendment defense.²³⁸ For example, California has adopted a test similar to copyright's fair use doctrine, which considers “whether the work in question adds significant creative elements so as to be transformed into something more than a mere celebrity likeness or imitation.”²³⁹

At the same time, there is a circuit split in the reasoning for such defenses, and courts generally have “failed to articulate a clear standard to resolve the

235. *In re Elster*, 2 F.4th 1328, 1335.

236. *Id.* at 1339 (noting that § 2(c) gave the PTO “no discretion to exempt trademarks that advance parody, criticism, commentary on matters of public importance, artistic transformation, or any other First Amendment interests”).

237. See Ernst, *supra* note 178, at 10 n.10 (citing Dave Gilson, *Most Presidents Ignore Products That Rip Off Their Names. Will Trump?*, MOTHER JONES (Feb. 13, 2017), <https://www.motherjones.com/politics/2017/02/trump-name-publicity-rights/> (“Despite his initial promise to keep a close watch on his image, Obama would eventually ignore thousands of products with no political message that likely infringed on his publicity rights.”).

238. See Rothman & Post, *supra* note 45, at 127.

239. *Id.* (citing *Comedy III Prods., Inc. v. Gary Saderup, Inc.*, 21 P.3d 797, 799 (Cal. 2001)).

conflict, resulting in a confusing morass of inconsistent, incomplete, or mutually exclusive approaches, tests, and standards.”²⁴⁰ Moreover, no proposed balancing test touches on the unique role of trademark registration—a process of content discrimination that does not actually trample on free speech in the way an injunction wholly restricting the sale of a product might. In *Comedy III Productions, Inc. v. Gary Saderup, Inc.*, the California Supreme Court stated that the “transformative” right of publicity test elaborated in the opinion was designed to “protect the right-of-publicity holder’s core interest in monopolizing the merchandising of celebrity images without unnecessarily impinging on the artists’ right of free expression.”²⁴¹ As emphasized in this Note, trademark registration refusal does not impinge on an artist’s free expression, nor does it chill speech in a way that warrants careful balancing with a famous person’s rights. It makes content-based decision based on rational legislative decisions about the kinds of marks to allow in the benefits of trademark registrations, in the way of a limited public forum.²⁴²

Moreover, no aspect of the limited public forum framework would prevent Congress from enacting legislation or amending § 2(c) to include something akin to copyright fair use. However, until the right of publicity is more fully understood—which it may be, as further support for a federal right to publicity is embraced²⁴³—it’s not practical for the PTO to select and apply one of many viable balancing tests to all future potentially parodic applications.²⁴⁴

240. *Id.* at 125 n.167 (citing Roberta Rosenthal Kwall, *A Perspective on Human Dignity, the First Amendment, and the Right of Publicity*, 50 B.C. L. REV. 1345, 1346 (2009) (identifying five different “balancing tests for determining how the right of publicity should be applied in cases presenting First Amendment challenges”).

241. 21 P.3d 797, 808 n.10 (2001).

242. As Rebecca Tushnet sums up her defense of the § 2 provisions, “if we are really that suspicious of government economic regulation picking winners and losers by way of speech, then we shouldn’t even have trademark registration.” Tushnet, *supra* note 41, at 424.

243. *See supra* note 181.

244. *See* Ernst, *supra* note 178 (“Who is to say whether ROYAL KATE jewelry is a brazen attempt to use Kate Middleton’s name to sell jewelry or is, instead (or in addition) a comment on the opulence and materialism of the British royal family?”). Professor Ernst also notes the absurdity in forcing the PTO to make a decision on which courts have been unable to come to agreement, but argues that such a lack of direction is reason not for § 2(c) to remain untouched, but to be struck down completely. He writes that, “[t]he other provisions to section 2 adequately allow the PTO to bar registration if there is deception, passing off, dilution, confusion or any of the other legitimate trademark concerns. Hence, it would appear to do no harm if section 2(c) were eliminated.” However, to do so would confine trademark registration to the realm of commercial speech, where a provision like § 2(c) is meant to do more than regulate deception. *See supra* Section II.B.1.

VI. CONCLUSION

Beyond the doctrinal fit with limited public forum, trademark registration as a limited public forum is a normatively preferable outcome because, in giving deference to the purpose of the forum (and therefore its registration provisions), it would create consistency for courts, applicants, examiners, and, in the case of § 2(c), third parties affected by applications. In addition, preventing registration of expressive language could actually increase free speech by limiting access to one outlet through which individuals could more easily attempt to control culturally and politically important speech. Finally, it would provide Congress with a foundation to enact further registration restrictions outside the bounds of policing deceptive and misleading speech.

In sum, trademark registration is not the regulation of expressive speech or commercial speech, nor is it a form of government speech through the “subsidy” of certain trademarks. In simply providing a benefit to registrants who comply with the requirements of the forum the government opened, registered trademarks look a lot like the registered organizations at UC Hastings. The school, in offering some benefits to “registered” organizations, was simply “dangling the carrot of subsidy,” to achieve compliance with its policy goal, but in no way restricting speech outside the forum for those benefits.²⁴⁵ It’s worth returning one last time to Chief Justice Roberts’ words about trademark registration decisions: “[n]o speech is being restricted; no one is being punished.”²⁴⁶ While it may seem counterintuitive to turn to a confusing corner of First Amendment doctrine to make this relatively simple point clear, it is ultimately the best way to prevent judicial overreach into valid congressional decisions about how to facilitate a federal trademark system in coexistence with common law.

245. *Christian Legal Soc’y*, 561 U.S. at 682.

246. *Brunetti*, 139 S. Ct. at 2302.

WHY IT'S TIME TO BAN GEOFENCE SEARCHES IN LIGHT OF *UNITED STATES V. CHATRIE*

Danny Drane[†]

TABLE OF CONTENTS

I.	INTRODUCTION	1308
II.	GEOFENCE SEARCHES THREATEN PRIVACY AND SPEECH	1310
A.	UNDERSTANDING GEOFENCE SEARCHES.....	1310
B.	GEOFENCE SEARCHES THREATEN PRIVACY.....	1311
C.	GEOFENCE SEARCHES THREATEN POLITICAL SPEECH.....	1313
III.	FOURTH AMENDMENT DOCTRINE ENCOURAGES TECH- SAVVY SURVEILLANCE	1316
A.	THE THIRD-PARTY DOCTRINE ENABLES MASS DIGITAL SURVEILLANCE.....	1316
B.	FOURTH AMENDMENT REMEDIES ARE POOR DETERRENTS.....	1319
1.	<i>The Exclusionary Rule</i>	1319
2.	<i>Civil Suits</i>	1323
C.	CHATRIE EPITOMIZES THE FOURTH AMENDMENT'S FAILURES....	1326
1.	<i>The Geofence Search Warrant in Chatrie</i>	1326
2.	<i>Chatrie is Not the Answer</i>	1327
IV.	THE IMPORTANCE OF A BLANKET BAN	1332
A.	WHY THE CONSTITUTION CANNOT REGULATE GEOFENCE SEARCHES.....	1332
B.	WHY PROPOSED LEGISLATION WILL NOT PROTECT SPEECH AND PRIVACY.....	1334
C.	WHY A BLANKET BAN IS THE ANSWER	1338
V.	CONCLUSION.....	1340

DOI: <https://doi.org/10.15779/Z38VM42Z6N>

© 2023 Danny Drane.

† J.D. Candidate, University of California, Berkeley, School of Law, Class of 2024.
Thank you to Professor Talha Syed for his mentorship, as well as the stellar editors at the
Berkeley Technology Law Journal.

I. INTRODUCTION

On December 13, 2018, Jorge Molina was arrested for a murder he did not commit.¹ At roughly 9 a.m., four police officers approached Molina at a Macy's department store and told him that they needed to speak with him.² The officers put Molina in handcuffs, drove him to the jailhouse, and interrogated him about a murder.³ In shock, Molina pleaded, "I didn't shoot anybody. I'm not that type of person."⁴ Yet the officers confidently retorted that they "knew, one hundred percent, without a doubt, that his phone was at the shooting scene."⁵ As it turned out, that was wrong.

The officers were confident Molina's phone was at the scene because they had issued a standard "geofence search warrant" to Google. In the week prior, police obtained surveillance footage of a car following the victim on the night he was killed.⁶ The officers then sent a geofence search warrant to Google, asking the company to identify "any wireless communication device that passed through the same geographical locations that the suspect vehicle did" on that night.⁷ Google complied with the request, sending back a list of four Google accounts that were in that area at the time.⁸ Then, when police asked for more details on each account, Google identified a device that was logged into Jorge Molina's Google account.⁹ Rather than pursue leads that would have uncovered the real culprit, police pinned this evidence on Molina, costing him his job, car, and reputation.¹⁰ Police were "blinded by data."¹¹

Since Molina's wrongful arrest, police use of geofence search warrants has skyrocketed nationwide. In 2020, the most recent year for which data is available, law enforcement issued over 11,000 geofence search warrants to

1. Meg O'Connor, *Avondale Man Sues After Google Data Leads to Wrongful Arrest for Murder*, PHX. NEW TIMES (Jan. 16, 2020), <https://www.phoenixnewtimes.com/news/google-geofence-location-data-avondale-wrongful-arrest-molina-gaeta-11426374>.

2. *Id.*

3. *Id.*

4. *Id.*

5. *Id.*

6. *Id.*

7. *Id.*

8. *Id.*

9. *Id.*

10. *See id.* Had police investigated further, it would have been "clear" that the culprit was Molina's stepfather. *Id.* Police learned that Molina owned the suspect vehicle, yet two months prior, police impounded the same car after Molina's stepfather was arrested for driving it without a license, which had occurred multiple times prior. *Id.* And if police had sought additional data on Molina's Google Account, they would have learned that Molina himself was in a different part of the city that night. *Id.*

11. *Id.*

Google—a 37% increase from 2019.¹² This trend concerns privacy advocates because a single geofence search can sweep up over a thousand people.¹³ Consequently, legal scholarship has dissected whether and when geofence search warrants violate the Fourth Amendment’s privacy protections.¹⁴

Troublingly, legal scholarship has largely ignored the on-the-ground impact of geofence searches on political speech. Although court records typically shield the details of search warrants, activists have discovered that police departments have used geofence searches to solve crimes committed at or near Black Lives Matter protests.¹⁵ This pattern suggests police are using geofence search warrants to target individuals who are expressing viewpoints with which police do not agree. But so far, legal scholarship on geofence search warrants is largely grounded in discussions on privacy, with very limited mentions of speech.¹⁶ This Note seeks to fill this gap in legal scholarship, in part because geofence search could become a potent tool against protestors. Protests have a high density of people concentrated in one area, and geofence searches offer police the unique ability to identify and track anyone present at a particular place, time, and location.

This Note proposes a simple legislative solution to the threats posed by geofence search warrants: a blanket ban on all geofence searches. Part II explains what geofence search warrants are, Google’s protocols for processing them, and how they threaten privacy and speech. Part III contends that, absent

12. Zack Whittaker, *Google Says Geofence Warrants Make Up One-Quarter of All US Demands*, TECHCRUNCH (Aug. 19, 2021), <https://techcrunch.com/2021/08/19/google-geofence-warrants/>.

13. Thomas Brewster, *Google Hands Fed 1,500 Phone Locations in Unprecedented ‘Geofence’ Search*, FORBES (Dec. 11, 2019), <https://www.forbes.com/sites/thomasbrewster/2019/12/11/google-gives-feds-1500-leads-to-arsonist-smartphones-in-unprecedented-geofence-search/?sh=3220433827dc>.

14. See, e.g., Haley Amster & Brett Diehl, *Against Geofences*, 74 STAN. L. REV. 385 (2022); Esteban De La Torre, *Digital Dragnets: How the Fourth Amendment Should Be Interpreted and Applied to Geofence Search Warrants*, 31 S. CAL. INTERDISC. L. J. 329 (2022); Cassandra Zietlow, *Reverse Location Search Warrants: Law Enforcement’s Transition to ‘Big Brother’*, 23 N.C. J.L. & TECH. 669 (2022); Note, *Geofence Warrants and the Fourth Amendment*, 134 HARV. L. REV. 2508 (2021).

15. See, e.g., Russell Brandom, *How Police Laid Down a Geofence Dagnet for Kenosha Protestors*, VERGE (Aug. 30, 2021), <https://www.theverge.com/22644965/kenosha-protests-geofence-warrants-atf-android-data-police-jacob-blake>; Zach Whittaker, *Minneapolis Police Tapped Google to Identify George Floyd Protestors*, TECHCRUNCH (Feb. 6, 2021), <https://techcrunch.com/2021/02/06/minneapolis-protests-geofence-warrant/>.

16. See, e.g., Amster & Diehl, *supra* note 14, at 396 (mentioning protests in only one sentence throughout the article); De La Torre, *supra* note 14, at 330 n.7, 330 n.8, 351 n.185 (citing three articles that mention protests in headlines but not stating “protest” or “speech” anywhere in the article); Zietlow, *supra* note 14, at 670–72, 678, 690 (mentioning the use of geofence searches against protestors several times without mentioning or contextualizing the accompanying threat to political speech).

Congressional legislation, courts will interpret the Fourth Amendment in ways that encourage, rather than limit, geofence searches' harms. To do so, Part III dissects the third-party doctrine, the Fourth Amendment's weak remedies, and how a recent case, *United States v. Chatrie*, epitomizes these doctrinal failures. Finally, Part IV criticizes alternative proposals to rely on courts and legislative reforms to showcase why a blanket ban is the most desirable solution.

II. GEOFENCE SEARCHES THREATEN PRIVACY AND SPEECH

A. UNDERSTANDING GEOFENCE SEARCHES

A geofence is a virtual perimeter that maps out a real-world geographic area during a specific timeframe. Geofences use GPS technology to identify digital devices that enter or exit the geofence boundaries. Law enforcement agencies conduct geofence searches to retroactively locate mobile devices that entered or exited the geofence.¹⁷ This entails submitting a geofence search warrant. Geofence search warrants are requests to a third-party company such as Google, for information on mobile devices that are associated with the accounts.¹⁸ Before requesting this information, law enforcement applies for a search warrant and describes the searches' terms to a magistrate judge.¹⁹

Take, for instance, the geofence search warrant ruled unconstitutional in *United States v. Chatrie*.²⁰ In response to a bank robbery, police in *Chatrie* issued a geofence search warrant to Google, compelling Google to identify every device that was within 17.5 acres of a bank between 4:20 p.m. and 5:20 p.m. on the day it was robbed.²¹ Put another way, police sought to identify every phone in an area equal to 3.5 blocks in New York City.²²

Geofence search warrants are primarily issued to Google, which processes warrants through a three-step protocol.²³ Google's specialists use data from Location History (LH), an opt-in feature on Google products and services.²⁴ In the first step, a specialist searches the entirety of Google's LH database and provides law enforcement with the requested information in an anonymized

17. Mark Harris, *A Peek Inside the FBI's Unprecedented January 6 Geofence Dragnet*, WIRED (Nov. 28, 2022), <https://www.wired.com/story/fbi-google-geofence-warrant-january-6/>.

18. *Id.*

19. See *Geofence Warrants and the Fourth Amendment*, *supra* note 14, at 2509, 2514.

20. 590 F. Supp. 3d 901 (E.D. Va. 2022).

21. *Id.* at 919.

22. *Id.* at 918 n.26.

23. *Id.*

24. *Id.* at 908–09.

format.²⁵ This data includes the time-stamped coordinates of, and Google accounts associated with, every device located within the geofence.²⁶ This data is ostensibly anonymous. However, at this point in the three-step protocol, without further information from Google, an officer can “observe each account’s reported location, track each account to his or her home, and pinpoint each account’s personal identity using publicly available resources.”²⁷ At step two, law enforcement reviews the list to identify devices they deem worth investigating and requests additional location data from Google.²⁸ During step three, Google provides information that identifies the users of these devices, including their full name, username, email addresses, birthdate, account type, account number, phone numbers, and the device’s make and model.²⁹

This three-step process is done at Google’s and law enforcement’s discretion, largely without the input of a judge.³⁰ Aside from approving the initial warrant, a neutral judge is not involved at any subsequent step of the process.³¹ There is also no requirement that officers narrow their request in step two.³² In fact, police often broaden the scope of their requests without seeking additional approval from a judge.³³ And, of course, Google is generally free to amend its three-step protocol at any point, which reduces the power of judges to limit geofence search warrants under the Fourth Amendment.³⁴

B. GEOFENCE SEARCHES THREATEN PRIVACY

Google’s Location History data is retroactive, precise, and comprehensive. The LH feature is automatically available on nearly every Android smartphone and on the Google Maps apps installed on any smartphone.³⁵ Considering 130 million Americans use an Android smartphone,³⁶ and one-third of active

25. *Id.* at 914–15.

26. *Id.* at 915–16.

27. *Id.* at 931 n.39.

28. *Id.* at 916–17.

29. *Id.* at 919 n.27.

30. *See Geofence Warrants and the Fourth Amendment*, *supra* note 14, at 2508, 2514–16.

31. *Id.*

32. *Chatrie*, 590 F. Supp. 3d at 923 (explaining that Google “typically require[s]” law enforcement to narrow the request but “has no firm policy as to precisely *when* a Step 2 request is sufficiently narrow”).

33. *Geofence Warrants and the Fourth Amendment*, *supra* note 14, at 2514–16.

34. *See Amster & Diehl*, *supra* note 14, at 437–44.

35. *Chatrie*, 590 F. Supp. 3d at 920 (quoting a law enforcement affidavit describing Google’s LH feature).

36. *Number of Android Smartphone Users in the United States from 2014 to 2022*, STATISTA, <https://www.statista.com/statistics/232786/> (last visited Nov. 1, 2023).

Google users have the LH feature enabled,³⁷ a back-of-the-napkin estimate of solely Android users suggests Google’s databases contain the minute-by-minute locations of, at the very least, 40 million Americans. Google can even pinpoint a phone’s location to within three meters.³⁸ Thus, with any geofence search aimed at finding a suspect, police have a good chance of finding detailed information, to say the least.

Geofence searches typically have wide margins of error and expansive geographic parameters. Google estimates that the data it provides to law enforcement fall within a 68% confidence interval, meaning there is only a 68% chance that the identified devices were within the given location.³⁹ And in step two of Google’s process, police often expand searches by requesting information on devices “outside the search parameters but within a ‘margin of error.’”⁴⁰ This means not only do police routinely identify people outside the scene of the relevant crime, but the information learned is often inaccurate.

These wide parameters raise distinct privacy concerns in urban areas. Urban police departments are more capable of deploying geofences than their rural counterparts due to superior staffing and resources.⁴¹ And urban police face more pressure to deploy geofences because urban areas also have higher rates of unsolved crimes—the exact situations where geofences are most valuable.⁴²

The urbanization of geofence searches is troubling for two reasons. First, the high population density of cities increases the number of innocent people swept up in searches.⁴³ Second, people of color are concentrated in urban

37. *Chatrie*, 590 F. Supp. 3d at 909.

38. *Id.*; see also *In re Search Warrant Application for Geofence Location Data Stored at Google Concerning an Arson Investigation*, 497 F. Supp. 3d 345, 360 (N.D. Ill. 2020) (“One only needs to look at one’s location in Google Maps to know that the location data is remarkably accurate.”).

39. *Chatrie*, 590 F. Supp. 3d at 909. A confidence interval is a statistical measure that, in simple terms, shows the probability that a given number falls within a certain range.

40. See, e.g., *In re Search of: Information Stored at Premises Controlled by Google*, 481 F. Supp. 3d 730, 745 (N.D. Ill. Aug. 24, 2020).

41. Lauren Weisner, H. Douglas Otto & Sharyn Adams, *Issues in Policing Rural Areas: A Review of the Literature*, ILL. CRIM. JUST. INFO. AUTHORITY (Mar. 18, 2020), <https://icjia.illinois.gov/researchhub/articles/issues-in-policing-rural-areas-a-review-of-the-literature>.

42. See *id.*; Maura Arnold, *Geofence Warrants: Useful Crime Solving Tool or Invasive Surveillance Tactic?*, J. HIGH TECH. L. BLOG. (Mar. 10, 2021), <https://sites.suffolk.edu/jhtl/2021/03/10/geofence-warrants-useful-crime-solving-tool-or-invasive-surveillance-tactic/>.

43. A. Reed McLeod, *Geofence Warrants: Geolocating the Fourth Amendment*, 30 WM. & MARY BILL RTS. J. 531, 557 (2021); see also *In re Search of: Info. Stored at Premises Controlled by Google*, No. 20 M 297, 2020 U.S. Dist. LEXIS 165185, at *1, *2 (N.D. Ill. July

areas.⁴⁴ There, racially disparate policing patterns are deeply rooted, well-documented, and typically reinforced when police acquire new tools and technology.⁴⁵ Given urban police departments are able and incentivized to use geofence searches, this tool will join a growing list of police technologies that perpetuate privacy invasions, structural racism, and mass incarceration.

C. GEOFENCE SEARCHES THREATEN POLITICAL SPEECH

Digital surveillance of protestors is not new. Since 2014, the FBI has used social media for long-term monitoring of Black Lives Matter activists.⁴⁶ Six federal agencies used facial recognition software to identify and criminally investigate people who protested the killing of George Floyd in 2020.⁴⁷ Customs and Border Patrol (CBP) used information collected from digital surveillance to curate dossiers of lawyers, activists, and journalists assisting migrants at the U.S.-Mexico border.⁴⁸ Federal, state, and local law enforcement agencies routinely share digital surveillance with each other through 80 federally funded “fusion centers.”⁴⁹

8, 2020) (highlighting that the requested geofence areas were within “a densely populated city” and captured individuals partaking in the “amenities associated with upscale urban living”).

44. Kim Parker, Juliana Menasce Horowitz, Anna Brown, Richard Fry, D’Vera Cohn & Ruth Igielnik, *Demographic and Economic Trends in Urban, Suburban, and Rural Communities*, PEW RES. CTR. (May 22, 2018), <https://www.pewresearch.org/social-trends/2018/05/22/demographic-and-economic-trends-in-urban-suburban-and-rural-communities/> (noting that 56% of the total population in urban counties are non-white).

45. *See, e.g.*, Michael Siegel, Rebecca Sherman, Cindy Li & Anita Knopov, *The Relationship Between Racial Residential Segregation and Black-White Disparities in Fatal Police Shootings at the City Level, 2013–2017*, 111 J. NAT’L MED. ASS’N. 580–87 (2019) (tracing racial disparities in policing and fatal shootings in cities to residential segregation); Will Douglas Heaven, *Predictive Policing Is Still Racist—Whatever Data It Uses*, MIT TECH. REV. (Feb. 5, 2021); <https://www.technologyreview.com/2021/02/05/1017560/predictive-policing-racist-algorithmic-bias-data-crime-predpol/> (discussing the racial bias encoded in modern, data-driven predictive policing tools).

46. George Joseph & Murtaza Hussain, *FBI Tracked An Activist Involved With Black Lives Matter As They Traveled Across the U.S., Documents Show*, INTERCEPT (Mar. 19, 2018, 8:29 AM), <https://theintercept.com/2018/03/19/black-lives-matter-fbi-surveillance/>.

47. Radhamely De Leon, *Six Federal Agencies Used Facial Recognition on George Floyd Protestors*, VICE (June 30, 2021), <https://www.vice.com/en/article/3aqpjm/six-federal-agencies-used-facial-recognition-on-george-floyd-protestors>.

48. Tom Jones, Mari Payton & Bill Feather, *Source: Leaked Documents Show the U.S. Government Tracking Journalists and Immigration Advocates Through a Secret Database*, NBC SAN DIEGO (Jan. 10, 2020), <https://www.nbcsandiego.com/news/local/source-leaked-documents-show-the-us-government-tracking-journalists-and-advocates-through-a-secret-database/3438/>.

49. SARAH BRAYNE, PREDICT AND SURVEIL: DATA, DISCRETION, AND THE FUTURE OF POLICING 9 (2020); *see* Rachel Levinson-Waldman & Ángel Díaz, *How to Reform Police Monitoring*

Make no mistake: law enforcement uses digital surveillance to retaliate against protestors, even those who do not commit crimes at protests.⁵⁰ Immigration and Customs Enforcement (ICE), for instance, recently arrested and initiated deportation proceedings against several of its critics shortly after they participated in protests.⁵¹ Sometimes police do not even wait for a protest to conclude. As Baltimore’s protestors mourned the death of Freddie Gray in 2015, police, in their words, “stay[ed] one step ahead” by using “real-time, location-based social media monitoring” to identify protestors with outstanding warrants and “arrest them directly from the crowd.”⁵²

Police used geofence search warrants during Black Lives Matter protests in recent years. During protests over George Floyd’s death at the hands of police, officers in Minneapolis asked Google to identify every device in an area with “dozens” of people to identify a person who broke the windows of an AutoZone store.⁵³ During protests in Kenosha, Wisconsin following the murder of Jacob Blake, federal agents issued six geofence search warrants that “stretch[ed] as long as two hours” and resembled a “dragnet[] spread over some of the [protests] busiest times and locations.”⁵⁴

Geofence searches will have chilling effects on political expression, particularly when they complement other forms of digital surveillance. In 2019, the Manhattan District Attorney, for instance, combined facial recognition, social media monitoring, and a geofence search to try to identify “members” of Antifa for a separate prosecution of right-wing Proud Boys.⁵⁵ This prosecution is particularly telling. Antifa has no real “membership.” It is an umbrella term that refers to small, loosely affiliated pockets of activists who are opposed to fascism.⁵⁶ Yet conservatives have warped “Antifa” into a catch-

of Social Media, BROOKINGS (July 9, 2020), <https://www.brookings.edu/techstream/how-to-reform-police-monitoring-of-social-media/>.

50. Levinson-Waldman & Díaz, *supra* note 49.

51. Alice Speri & Maryam Saleh, *An Immigrant Journalist Faces Deportation as ICE Cracks Down on its Critics*, INTERCEPT (Nov. 28, 2018), <https://theintercept.com/2018/11/28/ice-immigration-arrest-journalist-manuel-duran/>.

52. AM. CIVIL LIBERTIES UNION N. CAL., CASE STUDY: BALTIMORE COUNTY PD (2016), http://www.aclunc.org/docs/20161011_geofeedia_baltimore_case_study.pdf.

53. Whittaker, *supra* note 15.

54. Bandom, *supra* note 15.

55. Colin Moynihan, *How Police Used Antifa to Investigate Far-Right Proud Boys*, N.Y. TIMES (Aug. 8, 2019), <https://www.nytimes.com/2019/08/08/nyregion/proud-boys-antifa-trial.html>.

56. Mark Bray, *Five Myths About Antifa*, WASH. POST (Sept. 11, 2020), https://www.washingtonpost.com/outlook/five-myths/five-myths-about-antifa/2020/09/11/527071ac-f37b-11ea-bc45-e5d48ab44b9f_story.html; Michael Kenney & Colin Clarke, *What Antifa Is, What it Isn't, and Why it Matters*, WAR ON ROCKS (June 23, 2020), <https://warontherocks.com/2020/06/what-antifa-is-what-it-isnt-and-why-it-matters/>.

all term for left-leaning protestors, and this prosecution showcases law enforcement's willingness to exploit this narrative to the detriment of Black Lives Matter protestors.⁵⁷ Thus, one lesson rings clear: law enforcement will use geofence searches to disrupt protests, leading people to self-censor and expend additional resources to engage in political activities. And given empirical research confirms that the mere perception of online surveillance is sufficient to stifle the expression of political views, chilling effects will occur even if people's fears of geofence-based surveillance are misplaced.⁵⁸

Arguably, geofences pose greater risks than other forms of digital surveillance for two reasons. First, information revealed from a geofence search is more detailed than that of facial recognition and social media surveillance. After all, knowledge of a person's full name, usernames, birthdate, email address, and phone make and model is more likely to lead to arrests than, for example, a blurry photo put through facial recognition software.⁵⁹ Second, it is difficult to evade geofence surveillance. Through social media, for instance, police identify protestors largely because people voluntarily, and perhaps unwittingly, post photos and videos online. In response, activists have started warning protestors that "police can see your social media posts."⁶⁰ To evade surveillance and enable political speech, activists advise would-be protestors to communicate on encrypted platforms and refrain from posting another protestor's identifying information on social media.⁶¹

Whereas one can refrain from simply posting online, a protestor cannot as easily evade geofence-based location tracking. It would be counterproductive for a protestor to leave their phone at home because phones are invaluable for communication, coordination, and navigation to and from protests.⁶²

57. See Tina Nguyen, *How 'Antifa' Became a Trump Catch-All*, POLITICO (June 2, 2020), <https://www.politico.com/news/2020/06/02/how-antifa-became-a-trump-catch-all-297921>.

58. See Elizabeth Stoycheff, *Under Surveillance: Examining Facebook's Spiral of Silence Effects in the Wake of NSA Internet Monitoring*, 93 JOURNALISM & MASS COMM. Q. 296, 299–300 (2016).

59. See Clare Garvie, *Garbage In, Garbage Out: Face Recognition on Flawed Data*, GEO. L. CTR. ON PRIVACY & TECH. (2019), <https://www.law.georgetown.edu/privacy-technology-center/publications/garbage-in-garbage-out-face-recognition-on-flawed-data/> (noting that surveillance footage is often too low quality to identify suspects using facial recognition software).

60. See Corinne Purtill, *Before You Post That #Protest Selfie at the Inauguration Protests, Remember that Police Can See Your Social Media Posts*, QUARTZ (Jan. 20, 2017), <https://qz.com/889696/before-you-post-that-protest-selfie-at-the-inauguration-protests-remember-that-police-can-see-your-social-media-posts/>.

61. See *id.*

62. See, e.g., Christina Neumayer & Gitte Stald, *The Mobile Phone in Street Protest: Texting, Tweeting, Tracking, and Tracing*, 2 MOBILE MEDIA & COMM. 117, 118 (2014) (highlighting that cell phones allow street protestors to coordinate in real time, to send short and functional text

Alternatively, disabling location features on one's phone is confusing, burdensome, and requires some degree of technical knowhow. As a Google employee once described the process of deleting one's location history, "[it feels] like it is designed to make things possible, yet difficult enough that people won't figure . . . [it] out."⁶³ Finally, buying a burner phone is cost-prohibitive, and its inconvenience is incompatible with the spontaneity of many protests.⁶⁴ As a result, protests are full of phones pinging their minute-by-minute locations to Google's vast database that police can access. As the Supreme Court recently put it, "a phone goes wherever its owner goes, conveying to the wireless carrier . . . a detailed chronicle of a person's physical presence."⁶⁵

III. FOURTH AMENDMENT DOCTRINE ENCOURAGES TECH-SAVVY SURVEILLANCE

A. THE THIRD-PARTY DOCTRINE ENABLES MASS DIGITAL SURVEILLANCE

The Fourth Amendment protects people from "unreasonable searches" by requiring police obtain a warrant to search a person's "papers, houses, or effects."⁶⁶ A warrant is required only when the officer's conduct constitutes a "search," which occurs when police violate a person's "reasonable expectation of privacy"⁶⁷ or physically trespass on a person's property.⁶⁸

messages, and to document the actions of protestors and police); Allison Gordon, *Black Lives Matter Makes its Mark on Map Apps*, CNN (June 10, 2020), <https://www.cnn.com/2020/06/10/tech/map-protests-trnd/index.html> (highlighting the value of Snapchat in broadcasting and finding protests).

63. *Chatrie*, 590 F. Supp. 3d at 913 (quoting an Associated Press article that described the user interface as of August 13, 2018). The interfaces of Google's location products can be so convoluted that they confuse Google's own software engineers. Okello Chatrie's lawyers introduced evidence of emails from Google employees expressing confusion about Google's various location products. *Id.* at 914 n.17.

64. *See* Neumayer & Stald, *supra* note 62, at 118 ("The immediacy, mobility, and constant access afforded by mobile phones make them especially useful in ad hoc demonstrations.").

65. *Carpenter v. United States*, 138 S. Ct. 2206, 2211 (2018).

66. U.S. CONST. amend. IV.

67. *See* *Katz v. United States*, 389 U.S. 347, 360–61 (1967) (Harlan, J., concurring) (explaining that a reasonable expectation of privacy requires both a subjective expectation of privacy, and that the expectation is one that society is prepared to recognize as reasonable).

68. *United States v. Jones*, 565 U.S. 400, 407 (2012) (establishing that either the reasonable expectation of privacy test or a physical intrusion onto a persons' constitutionally protected area is sufficient to constitute a search within the meaning of the Fourth Amendment).

Technology has shaped the Fourth Amendment's privacy protections. *Kyllo v. United States*, for instance, held that police need a warrant to use thermal imaging devices that compile images of the inside of a person's home based on the home's distribution of heat.⁶⁹ Writing for the majority, Justice Scalia reasoned that, although police do not physically enter a person's home, the use of thermal imaging devices risks inadvertently revealing "intimate" information that traditionally could only be revealed by entering the home.⁷⁰

The third-party doctrine, however, has significantly undermined these protections. Under the third-party doctrine, police do not need a warrant to obtain information voluntarily given to a third party.⁷¹ A person forfeits their expectation of privacy because they "assume[] the risk" that the information will be disclosed to police.⁷² Since the person has no privacy expectation in disclosed information, an officer who obtains the information is not conducting a "search" for Fourth Amendment purposes and thus does not need a warrant. As such, the Supreme Court has held that police do not need a warrant to obtain a person's bank records,⁷³ or even the phone numbers of incoming and outgoing calls.⁷⁴ Effectively, the Fourth Amendment fails to protect Americans' digital information because virtually all digital information is shared with or stored by a third party.

The sole case where the Supreme Court declined to apply the third-party doctrine to digital information is *Carpenter v. United States*. There, police took advantage of the fact that cell phones send a signal to the nearest cell tower several times every minute.⁷⁵ Police obtained two sets of cell tower records without warrants: one retroactively traced the defendant's location over the course of 127 days, the other traced his location over two days.⁷⁶ The majority stressed that there was a significant privacy interest in "a person's physical presence compiled every day, every moment, over several years."⁷⁷

69. 533 U.S. 27, 40 (2001).

70. *Id.* at 38 ("The [device] might disclose, for example, at what hour each night the lady of the house takes her daily sauna and bath—a detail that many would consider 'intimate'; and a much more sophisticated system might detect nothing more intimate than the fact that someone left a closet light on . . . [And] no police officer would be able to know *in advance* whether his through-the-wall surveillance picks up 'intimate' details—and thus would be unable to know in advance whether it is constitutional.")

71. *Smith v. Maryland*, 442 U.S. 735, 744 (1979); *United States v. Miller*, 425 U.S. 435, 443 (1976).

72. *Smith*, 442 U.S. at 745.

73. *Miller*, 425 U.S. at 443.

74. *Smith*, 442 U.S. at 744.

75. *Carpenter*, 138 S. Ct. at 2211.

76. *Id.* at 2212.

77. *Id.* at 2220.

Additionally, since a phone shares its location “without any affirmative act” by the user, the person does not “voluntarily assume[] the risk” of disclosing their location “in [a] meaningful sense” for the purposes of the third-party doctrine.⁷⁸ Thus, the warrantless search of location information was unconstitutional.

In spite of *Carpenter*, however, lower courts still apply the third-party doctrine in countless scenarios where police collect intimate digital information about people. For instance, multiple circuit courts have held that police do not need a warrant to obtain basic subscriber information that customers must provide to use mobile applications, websites, and services like Google and Facebook.⁷⁹ This “basic” information usually includes IP addresses,⁸⁰ which are the identities of networks and devices on the internet, as well as a user’s first and last name, home address, email address, profile pictures, birthdate, location information, and device information.⁸¹

Under the logic of the third-party doctrine, a frightening amount of digital information is provided “voluntarily” and thus does not necessitate a warrant to search. Consider a universal experience: someone visits a website or downloads an app, then agrees to a privacy policy or a pop-up notice with the word “cookies.” Cookies are data that websites track, like a person’s web browsing history or online shopping carts.⁸² When a user agrees to or even ignores these terms, he or she consents to the website selling the user’s information to third parties, which are usually advertisers and data brokers that make profiles of your online activity.⁸³ Users agree to these terms 95–99% of the time, even when given an option to opt-out that is explicitly titled “Do Not

78. *Id.* (quoting *Smith*, 442 U.S. at 745) (emphasis added).

79. *See, e.g.*, *United States v. Rosenow*, 33 F.4th 529, 548 (9th Cir. 2022); *United States v. Ulbricht*, 858 F.3d 71, 97 (2d Cir. 2017); *United States v. Cairra*, 833 F.3d 803, 806 (7th Cir. 2016).

80. *See United States v. Forrester*, 512 F.3d 500, 510 (9th Cir. 2008).

81. LIZ WOOLERY, RYAN BUDISH & KEVIN BANKSTON, THE TRANSPARENCY REPORTING TOOLKIT: SURVEY & BEST PRACTICE MEMOS FOR REPORTING ON U.S. GOVERNMENT REQUESTS FOR USER INFORMATION, THE BERKMAN CTR. FOR INTERNET & SOC’Y AT HARV. UNIV. (Mar. 2016), https://cyber.harvard.edu/sites/cyber.harvard.edu/files/Final_Transparency.pdf. The messaging app Kik has a FAQ website for law enforcement that includes this information in its definition of “basic subscriber information.” *See Kik FAQ*, LAW ENFORCEMENT HELP CTR., <https://medialablawenforcementhelp.zendesk.com/hc/en-us/articles/4404983340187-What-s-included-in-Basic-Subscriber-Information-> (last visited Nov. 1, 2023).

82. Jon Healey, *What Are Those Annoying Website Popups About Cookies? And What Should You Do About Them?*, L.A. TIMES (Sept. 1, 2021), <https://www.latimes.com/business/technology/story/2021-09-01/what-are-website-cookies-how-do-they-impact-internet-data>.

83. *Id.*

Sell My Personal Information.”⁸⁴ In one experiment, 74% of people agreed to a website’s privacy policy without reading it, and 93% agreed to a condition to give up their first-born child.⁸⁵

This habit of doling out digital information has gifted police with endless opportunities for warrantless digital surveillance. Police are free to commercially purchase troves of data from online advertisers, and do not need a warrant or subpoena.⁸⁶ Thanks to this “constitutional loophole,” law enforcement can simply “us[e] its checkbook to get around *Carpenter*.”⁸⁷ Though advertisers’ data is anonymized, its precision makes it easy to identify people.⁸⁸ For example, when given access to one digital advertising dataset, New York Times staffers were “quickly able to match more than 2,000 supposedly anonymous devices . . . with email addresses, birthdays, ethnicities, ages, and more.”⁸⁹ It is no exaggeration to say the third-party doctrine “threatens to nullify the Fourth Amendment.”⁹⁰

B. FOURTH AMENDMENT REMEDIES ARE POOR DETERRENTS

1. *The Exclusionary Rule*

The primary remedy for Fourth Amendment violations is the Exclusionary Rule: evidence uncovered from an unconstitutional search cannot be admitted

84. INTERACTIVE ADVERTISING BUREAU, IAB CCPA BENCHMARK SURVEY SUMMARY 6 (Nov. 12, 2020), <https://www.iab.com/insights/iab-ccpa-benchmark-survey/> (finding that opt-out rates were only 1–5%). This study examined, among other things, the rate at which website users opted out of websites selling third-party cookies after the passage of the California Consumer Privacy Act (CCPA). CAL. CIV. CODE §§ 1798.100–1798.198 (2018). The CCPA requires websites that sell data to provide a “clear and conspicuous link on the business’s internet homepages, titled ‘Do Not Sell or Share My Personal Information,’ to an internet web page that enables a consumer . . . to opt out of the sale of the consumer’s personal information.” CAL. CIV. CODE § 1798.135(a)(1).

85. Jonathan A. Obar & Anne Oeldorf-Hirsch, *The Biggest Lie on the Internet: Ignoring the Privacy Policies and Terms of Service Policies of Social Networking Services*, 23 INFO., COMM., & SOC’Y 128 (2020).

86. Tim O’Brien, *Suspicionless Search: Geofence Warrants and the Fourth Amendment* 28 (Feb. 13, 2023) (unpublished manuscript), <https://ssrn.com/abstract=3834623>.

87. See Charles Levinson, *Through Apps, Not Warrants, ‘Locate X’ Allows Federal Law Enforcement to Track Phones*, PROTOCOL (Mar. 5, 2020), <https://www.protocol.com/government-buying-location-data>; Isabelle Canaan, *A Fourth Amendment Loophole?: An Exploration of Privacy and Protection through the Muslim Pro Case*, 6 HUM. RTS. L. REV. 95, 104 (2021).

88. Canaan, *supra* note 87, at 104.

89. Charlie Warzel & Stuart A. Thompson, *They Stormed the Capitol. Their Apps Tracked Them*, N.Y. TIMES (Feb. 5, 2021), <https://www.nytimes.com/2021/02/05/opinion/capitol-attack-cellphone-data.html>.

90. Gabriel Broshteyn, *If These Walls Could Talk: The Smart Home and the Fourth Amendment Limits of the Third Party Doctrine*, 130 HARV. L. REV. 1924, 1931 (2017).

against the defendant.⁹¹ To those who believe that an effective remedy should make a person whole, excluding evidence hardly ameliorates the harms of pretrial detention, including disruptions in wages and employment,⁹² housing stability,⁹³ familial relationships,⁹⁴ and mental⁹⁵ and physical health.⁹⁶ The Supreme Court has openly acknowledged the Exclusionary Rule's inability to

91. *Weeks v. United States*, 232 U.S. 383, 393 (1914).

92. See Will Dobbie & Crystal Yang, *The Economic Costs of Pretrial Detention*, BROOKINGS PAPERS ON ECON. ACTIVITY, Spring 2021, at 251, 260, https://www.brookings.edu/wp-content/uploads/2021/03/15872-BPEA-SP21_WEB_DobbieYang.pdf (“Even a short period of pretrial detention can be destabilizing . . . resulting in immediate job loss . . .”). On average, pretrial detention reduces a person’s earnings by \$948 per year over the 3–4 years following detention. *Id.* at 13.

93. See GINA CLAYTON, ENDRIA RICHARDSON, LILY MANDLIN & BRITTANY FARR, ESSIE JUSTICE GRP., *BECAUSE SHE’S POWERFUL: THE POLITICAL ISOLATION AND RESISTANCE OF WOMEN WITH INCARCERATED LOVED ONES* 62 (2018), https://www.becauseshespowerful.org/wp-content/uploads/2018/05/Essie-Justice-Group_Because-Shes-Powerful-Report.pdf (“[Fifty percent] of women who have owed money to a bail bonds agency faced housing insecurity as a result.”).

94. See Sara Wakefield & Lars Højsgaard Andersen, *Pretrial Detention and the Costs of System Overreach for Employment and Family Life*, 7 SOCIO. SCI. 342 (2020) (finding that people detained pretrial but not convicted have a statistically higher risk of no longer living with their partner or child after release); see also CREASIE FINNEY HAIRSTON, ANNIE E. CASEY FOUND., *KINSHIP CARE WHEN PARENTS ARE INCARCERATED: WHAT WE KNOW, WHAT WE CAN DO. A REVIEW OF THE RESEARCH AND RECOMMENDATIONS FOR ACTION* (2009), <https://eric.ed.gov/?id=ED507722> (documenting how incarcerated mothers seek care for their children by relying on their children’s grandparents, extended family, and foster care).

95. See, e.g., JENNIFER BRONSON, JESSICA STROOP, STEPHANIE ZIMMER & MARCUS BERZOFSKY, U.S. DEP’T OF JUSTICE, *DRUG USE, DEPENDENCE, AND ABUSE AMONG STATE PRISONERS AND JAIL INMATES*, 2007-2009, at 3 (2017), <https://www.bjs.gov/index.cfm?ty=pbdetail&iid=5966> (finding that two-thirds of jail inmate have a substance use disorder); Andrew P. Wilper, Steffie Woolhandler, J. Wesley Boyd, Karen E. Lasser, Danny McCormick, David H. Bor, & David U. Himmelstein, *The Health and Health Care of US Prisoners: Results of a Nationwide Survey*, 99 AM. J. PUB. HEALTH 666 (2009) (finding that half of all inmates who have previously been treated for a psychiatric condition receive no medical treatment while in jail).

96. See, e.g., Amy Katzen, *African American Men’s Health and Incarceration: Access to Care upon Reentry and Eliminating Invisible Punishments*, 26 BERKELEY J. GENDER, L. & JUST. 221, 228 (2011) (noting that poor ventilation and overcrowding in jails cause higher rates of tuberculosis); Shabbar I. Ranapurwala, Meghan E. Shanahan, Apostolos A. Alexandridis, Scott K. Proescholdbell, Rebecca B. Naumann, Daniel Edwards, Jr., & Stephen W. Marshall, *Opioid Overdose Mortality Among Former North Carolina Inmates: 2000–2015*, 108 AM. J. PUB. HEALTH 1207, 1208 (2018) (attributing an increase in mortality rates not to higher incarceration rates among substance users, but to the fact that one’s tolerance for drugs decreases while behind bars, thereby increasing the risk of overdose upon release).

compensate for harms, having characterized the rule as simply a “deterrent” against violations of the Fourth Amendment.⁹⁷

Yet the Court has carved out numerous exceptions that swallow the Exclusionary Rule’s deterrent effect. For instance, evidence can only be excluded if the defendant proves that the officer intentionally or recklessly violated the Fourth Amendment.⁹⁸ Even if a defendant overcomes that hurdle, prosecutors can still use unlawfully gained evidence to impeach any witness, including the defendant,⁹⁹ against a different defendant whose Fourth Amendment rights were not violated,¹⁰⁰ and when the officer had a “good-faith” reason for not knowing that their search was illegal.¹⁰¹ The inevitable discovery doctrine, too, is a “colossal loophole” that allows police to use illegally gained evidence if other practices would have otherwise yielded the evidence.¹⁰² Additionally, unconstitutional searches typically yield topics for further investigation, including physical evidence and witness identifications. Under yet another exception, prosecutors can admit anything police learn from follow-up actions to an unconstitutional search so long as intervening circumstances render the evidence gained to be sufficiently “attenuated” from the initial constitutional violation.¹⁰³

97. *See, e.g.*, *United States v. Leon*, 468 U.S. 897, 906 (1984) (quoting *Stone v. Powell*, 428 U.S. 465, 540 (White, J., dissenting)).

98. *Herring v. United States*, 555 U.S. 135, 144 (2009).

99. *United States v. Havens*, 446 U.S. 620, 628 (1980).

100. *Rakas v. Illinois*, 439 U.S. 128, 130 (1978).

101. *Leon*, 468 U.S. at 906–08.

102. *Nix v. Williams*, 467 U.S. 431, 432, 444 (1984) (announcing the inevitable discovery doctrine); *see generally* Tonja Jacobi & Elliot Louthen, *The Corrosive Effect of Inevitable Discovery on the Fourth Amendment*, 171 U. PA. L. REV. 1, 2 (2022) (charting the application of the inevitable discovery doctrine and arguing that tests adopted by many lower courts have devolved into a more relaxed standard than the one set out by the Supreme Court in *Nix v. Williams*).

103. Initially intended to allow for the admission of evidence gained as a result of unforeseeable intervening circumstances, the attenuation doctrine is a three-part test announced in *Brown v. Illinois*. 422 U.S. 590 (1975). To determine whether the admitted evidence is sufficiently attenuation from the officer’s unconstitutional conduct, courts analyze the 1) temporal proximity between the officer’s actions and the seizure of the evidence, 2) whether there are intervening circumstances, and 3) the flagrancy of the officer’s conduct. *Id.* at 603–605. The Supreme Court vastly expanded the application of this test in *Utah v. Strieff*, concluding that evidence obtained by police during an unlawful stop is not subject to the exclusionary rule if the police discover that the person stopped has a warrant out for their arrest. 136 S. Ct. 2056, 2059–63 (2016). The majority reasoned that the discovery of this warrant, though merely minutes after the stop began, was sufficiently attenuated from the unlawful stop. *Id.* For a discussion of this flawed holding’s likely impacts on the Fourth Amendment and officer misconduct, see Matthew E. Sweet, *Stretching the Attenuation Doctrine to Its Limits: How the Supreme Court Erred in Utah v. Strieff and What Can Be Done to Preserve the Doctrine*, 25 GEO. MASON. L. REV. 861, 871–880 (2018).

To illustrate how these broad, categorical exceptions have made the Exclusionary Rule “Swiss cheese,”¹⁰⁴ consider a hypothetical. Police illegally raid a person’s home without a warrant and find nothing except a box of drugs owned by the homeowner’s friend. Although the homeowner’s Fourth Amendment rights were violated, this box is admissible evidence in a criminal prosecution against the friend because his home was not raided.¹⁰⁵ Let’s add to this hypothetical. A label on the box contains the friend’s phone number, so police find an unconstitutional way to intercept his outgoing texts that say, “I am currently at a large meeting of drug dealers.” Police then contact each drug dealer at this meeting, and each dealer snitches on each other. At his trial, the friend testifies that he was not at this meeting, so the prosecutor reads these texts out loud for impeachment purposes.¹⁰⁶ In reality, the jury just heard a smoking gun confession disguised as an impeachment.¹⁰⁷ And in subsequent prosecutions, every drug dealer is out of luck because their confessions were “attenuated” from these Fourth Amendment violations. As this hypothetical showcases, “[w]hat ultimately matters to defendants is not where their constitutional rights begin and end, but rather the more pragmatic question of whether or not evidence is actually admitted.”¹⁰⁸

Thus, for the Court to call the Exclusionary Rule a “deterrent” ignores the obvious. For a deterrent to work, it must impose sufficient costs on bad actors. To borrow from economics literature on deterrence, an officer will violate a person’s Fourth Amendment rights “if the expected benefits to the police officer exceed the expected costs.”¹⁰⁹ The expected benefit to an officer would be a criminal conviction, or merely pretrial detention itself, which would allow police to confiscate contraband, interrogate the suspect, and perhaps temporarily prevent a crime.¹¹⁰ An officer weighs these benefits against the

104. Christopher Slobogin, *Why Liberals Should Chuck the Exclusionary Rule*, 1999 U. ILL. L. REV. 363, 375.

105. *See Rakas v. Illinois*, 439 U.S. 128, 130 (1978).

106. *See generally* *United States v. Havens*, 446 U.S. 620, 628 (1980).

107. The exclusionary rule’s exception for impeachment evidence has been rightfully criticized for its prejudicial impact on criminal defendants and the fact that it deters defendants from testifying at trial. *See, e.g.*, Richard D. Friedman, *Minimizing the Jury Over-Valuation Concern*, 2003 MICH. ST. L. REV. 967, 981 (“It is unlikely that jurors use [impeachment evidence] to assess the credibility of the accused Inevitably though, they are tempted to use the evidence for a purpose for which they are *not* supposed to consider it—in this case, determining that the accused is a bad person. This means that the impeachment evidence has a serious biasing effect—and because of that, the threat of such evidence often intimidates a defendant from exercising his fundamental right to testify in his own defense”)

108. Jacobi & Louthen, *supra* note 102, at 1–2.

109. Michael Cicchini, *An Economics Perspective on the Exclusionary Rule and Deterrence*, 75 MO. L. REV. 459, 469 (2010).

110. *Id.*

probability that the Exclusionary Rule frees the criminal suspect.¹¹¹ Unfortunately, this probability is “near zero” given the frequency with which police lie at hearings, the pressure defendants face during plea negotiations, and the Exclusionary Rule’s many exceptions.¹¹²

The Exclusionary Rule also imposes zero personal costs on officers. Deterrence generally requires that the targeted person perceive a sufficiently high probability and severity of punishment.¹¹³ In economics terms, effective punishments cause officers to “internalize the harm” that they cause which incentivizes them to refrain from future misconduct.¹¹⁴ Merely excluding evidence, however, does not affect officers personally because the outcomes of evidentiary hearings only affect defendants, and defendants lose 99% of the time.¹¹⁵ In fact, court surveys of police demonstrate that officers twist the facts at evidentiary hearings so often that police coined a term for it: “testilying.”¹¹⁶ When officers “so widely, willingly, and cavalierly lie[] to courts about their Fourth Amendment actions,” the Supreme Court is wrong to suggest that the Exclusionary Rule sufficiently deters officers from violating the Fourth Amendment.¹¹⁷

2. *Civil Suits*

There is also little deterrent value in lawsuits brought under 42 U.S.C. § 1983, which authorizes civil suits against state government officials who violate a person’s constitutional rights.¹¹⁸ To prevail and earn money damages

111. *See id.* at 470–81 (theorizing that the expected costs are the probability of evidence suppression, the cost of a lost conviction, and secondary sanctions against the officer such as “civil lawsuits, job-related sanctions, and public condemnation.”).

112. *See id.* at 470–71 (quoting commentary on the Mollen Report, a survey of New York City police officers that documented a common practice of testifying untruthfully during suppression hearings); Jamie Fellner, *An Offer You Can’t Refuse: How U.S. Federal Prosecutors Force Drug Defendants to Plead Guilty*, 26 FED. SENT’G REP. 276, 277–80 (2013) (discussing the effects of mandatory minimum sentencing provisions on plea negotiations of criminal defendants facing federal drug charges).

113. *Cf.* Nuno Garoupa, *The Theory of Optimal Law Enforcement*, 11 J. ECON. SURVEYS 267, 268 (1997) (referencing Gary Becker’s seminal papers on the deterrence theory of criminal punishment); *see also* Cicchini, *supra* note 109, at 470–81 (theorizing that officers may face “secondary sanctions” such as job-related sanctions, civil lawsuits, and public condemnation).

114. Robert Cooter, *Three Effects of Social Norms on Law: Expression, Deterrence, and Internalization*, 79 OR. L. REV. 1, 16 (2000).

115. *See* Albert Alschuler, *Studying the Exclusionary Rule: An Empirical Classic*, 75 U. CHI. L. REV. 1365, 1375 (2008) (citing empirical studies finding that courts exclude evidence in, at most, 1.3 percent of criminal cases).

116. *Id.* at 1376–77.

117. David Harris, *How Accountability-Based Policing Can Reinforce—or Replace—the Fourth Amendment Exclusionary Rule*, 7 OHIO ST. J. CRIM. L. 149, 162, 162 n.53 (2009).

118. 42 U.S.C. § 1983.

under § 1983, plaintiffs must overcome the affirmative defense of qualified immunity.¹¹⁹ This requires that the officer violated a “clearly established” right of which “every reasonable officer” under the circumstances would be aware.¹²⁰

Rather than interpret constitutional rights at a high level of generality, courts distinguish the facts of qualified immunity cases at a granular level, leaving very few rights “clearly established” in the eyes of individual police officers.¹²¹ In an article describing the extreme degree to which courts distinguish facts to favor police, Professor Mark Brown highlighted troubling precedent from the Eleventh Circuit: “[f]or qualified immunity to be surrendered, pre-existing law must dictate, that is, truly compel (not just suggest or allow to raise a question about), the conclusion for every like-situated, reasonable government agent that what [they are] doing violates federal law in the circumstances.”¹²² Indeed, “*minor* variations in some facts” including “an *arguably significant* fact . . . might be very important” from the perspective of an officer and therefore make a right not “clearly established.”¹²³

This tendency to overly distinguish cases is particularly harmful in Fourth Amendment doctrine, which entails highly fact-specific tests.¹²⁴ The Supreme Court requires parsing existing law “with [such] a high degree of specificity” that a search’s constitutionality under the Fourth Amendment is “beyond debate.”¹²⁵ Although technologies like geofence searches are too new to have many cases surrounding their use in general, qualified immunity is only overcome with “controlling authority” or “a robust consensus of cases of persuasive authority” that directly bear on the facts of a particular search.¹²⁶

119. *Harlow v. Fitzgerald*, 457 U.S. 800, 815 (1982).

120. *Ashcroft v. al-Kidd*, 563 U.S. 731, 741 (2011) (internal citations omitted). *Al-Kidd* heightened the standard of qualified immunity to emphasize what “every” reasonable officer would know, as opposed to *Harlow*’s original phrasing: “a reasonable officer.” 457 U.S. at 815.

121. John C. Jeffries Jr., *What’s Wrong with Qualified Immunity*, 62 FLA. L. REV. 851, 854–65 (2010).

122. Mark R. Brown, *The Fall and Rise of Qualified Immunity: From Hope to Harris*, 9 NEV. L.J. 185, 198 (2008) (quoting *Rowe v. City of Ford Lauderdale*, 279 F.3d 1271, 1280 (11th Cir. 2002) (emphasis omitted)).

123. *Id.* at 198–99 (quoting *Marsh v. Butler Cty.*, 268 F.3d 1014, 1032 (11th Cir. 2001) (emphasis added)).

124. Jeffries, *supra* note 121, at 859–60. The primary test for evaluating whether the Fourth Amendment has been violated is “totality of the circumstances.” *Illinois v. Gates*, 462 U.S. 213, 230 (1983); *Lange v. California*, 141 S. Ct. 2011, 2018 (2021) (applying the totality of the circumstances test through the officer’s perspective to evaluate whether the “exigent circumstances” exception to the warrant requirement applied).

125. *District of Columbia v. Wesby*, 138 S. Ct. 577, 589 (2018).

126. *Id.* at 589–90 (quoting *Wilson v. Layne*, 526 U.S. 603, 617 (1999)); see also *Chatrue*, 590 F. Supp. 3d at 936 (holding that a police officer conducted an unconstitutional geofence search

And ultimately, even if an innocent criminal suspect somehow overcomes qualified immunity, money damages are hard to square with the one-time violation of privacy from an unconstitutional search. One Justice Department study found that, out of 12,000 lawsuits against federal officers for alleged constitutional violations, plaintiffs were paid damages in only five cases.¹²⁷ It was unknown whether any of those cases involved Fourth Amendment searches.¹²⁸ The study attributed this, in part, to the fact that illegal searches generally “do[] not cause the kind of actual damages that our tort system compensates.”¹²⁹ All of this is to say, the low prospect of money damages provides virtually no deterrent for police use of novel technology because qualified immunity offers police tremendous freedom to experiment with our Fourth Amendment rights.

Injunctive relief, the other § 1983 remedy to stop or deter unconstitutional police practices, is difficult to pursue due to *City of Los Angeles v. Lyons*.¹³⁰ In *Lyons*, the Supreme Court held that, when seeking injunctive relief, a plaintiff’s case is moot¹³¹ unless they demonstrate that they are likely to be injured *again*

in good-faith belief of its constitutionality given “rapidly advancing technology” and the lack of “judicial guidance” on employing geofences).

127. See Donald Dripps, *Beyond the Warren Court and Its Conservative Critics: Toward a Unified Theory of Constitutional Criminal Procedure*, 23 U. MICH. J.L. REFORM 591, 629 (1990) (internal citation omitted).

128. *Id.*

129. *Id.*; see also Sherry F. Colb, *Innocence, Privacy, and Targeting in Fourth Amendment Jurisprudence*, 96 COLUM. L. REV. 1456, 1521 n.173 (1996) (“Damages may be minimal in the ordinary case because there is little injury to a person or to property.”). Another convincing explanation is that juries are unwilling to award damages to suspected criminals. See Tracey Maclin, *When the Cure for the Fourth Amendment is Worse than the Disease*, 68 S. CAL. L. REV. 1, 31 (1994) (“If the majority of the public is willing to sacrifice the Fourth Amendment to stop illegal drug use, why should anyone believe that jurors in civil damages cases will protect the Fourth Amendment rights of guilty drug couriers?”).

130. 461 U.S. 95 (1983).

131. For background on mootness, federal courts may only hear cases that are “ripe,” as opposed to “moot.” That is, the injury must actively persist at the time of litigation such that resolution of the case would affect the plaintiff’s rights. See, e.g., *DeFunis v. Odegaard*, 416 U.S. 312, 316–19 (1974) (describing the Court’s mootness doctrine, then explaining that the plaintiff, a law school applicant seeking an injunction to be admitted into a law school, had a moot case because they were ultimately admitted into the school after the litigation commenced). The relevant exception to mootness in *Lyons* was that federal courts will hear a moot case when the injury is capable of repetition, yet evading review. The “classic” example of this exception is that courts will hear cases where the injury is related to pregnancy. See *Roe v. Wade*, 410 U.S. 113, 125 (1973). To require a litigant be pregnant throughout a lawsuit would be unrealistic because litigation can be a lengthy endeavor. Thus, to dismiss pregnancy-related injuries as moot would allow defendants to repeatedly evade identical lawsuits solely due to the temporary nature of injuries they cause. Hence, courts created the “capable of repetition, yet evading review” exception to mootness doctrine.

by the practice alleged to be unconstitutional.¹³² Thus, Lyons was unable to challenge the L.A. Police Department's use of chokeholds on Fourth Amendment grounds because he failed to show that he specifically would be choked again by L.A. police.¹³³ In the context of a novel technology, how could a person credibly predict that in the near future they will be captured in, say, a geofence search?¹³⁴ And given this new search method has largely evaded judicial scrutiny, how could courts craft injunctions to accommodate the Fourth Amendment's many exceptions? After all, the court would need to "answer a seemingly limitless set of hypothetical situations addressing a seemingly limitless set of possible exceptions[.]"¹³⁵

Altogether, the Fourth Amendment's weak remedies provide overly broad discretion to police over people's privacy.¹³⁶ With new technology, privacy infringements are becoming even cheaper and more convenient. If the Fourth Amendment exists only as a subject on which police experiment, then our privacy protections "might as well be stricken from the Constitution."¹³⁷

C. CHATRIE EPITOMIZES THE FOURTH AMENDMENT'S FAILURES

1. *The Geofence Search Warrant in Chatrie*

The story of Okello Chatrie's arrest and conviction in the Eastern District of Virginia is as follows. After a bank robbery in May of 2019, police in Midlothian, Virginia issued a geofence search warrant to Google, seeking to identify every cell phone within 17.5 acres of the bank between 4:20 p.m. and 5:20 p.m. on the day it was robbed.¹³⁸ The geofence initially had a diameter of 300 meters, which was "longer than three football fields" and included the bank, a church, and a nearby wooded area.¹³⁹

In his application for the search warrant, the officer told a magistrate judge what information he planned to request from Google. In sum, this is what the

132. 461 U.S. at 110.

133. *Id.* at 111–12.

134. *See, e.g., Clapper v. Amnesty Int'l USA*, 568 U.S. 398 (2013) (holding that a plaintiff failed to establish an injury-in-fact for standing purposes by means of a probabilistic theory that the National Security Agency's foreign surveillance program was reasonably likely to intercept the plaintiff's communications).

135. Orin Kerr, *The Limits of Fourth Amendment Injunctions*, 7 J. ON TELECOMM. & HIGH TECH. L. 127, 134–35 (2009) (charting, by means of example, the inherent difficulties in crafting an injunction against a warrantless search of a home).

136. Slobogin, *supra* note 104, at 364.

137. *Weeks v. United States*, 232 U.S. 383, 393 (1914) (announcing the exclusionary rule, then characterizing it as a means of deterring unconstitutional conduct, thereby securing Fourth Amendment protections).

138. *Chatrie*, 590 F. Supp. 3d at 914–15.

139. *Id.* at 922–23.

officer said would occur during his search: at step one of Google’s process, the officer planned to ask for anonymized information on devices within the geofence;¹⁴⁰ then, at step two, law enforcement promised to “attempt[] to narrow” this list and request additional “contextual data points” that illustrate each person’s travel,¹⁴¹ where these data points would expand the geofence’s radius to 387 meters—“more than twice as large as the original geofence”—and add thirty minutes to the beginning and to the end of the initial timeframe;¹⁴² and finally, at step three, Google would provide account-identifying information.¹⁴³

The magistrate judge reviewed this information for, at most, fifteen-to-thirty minutes.¹⁴⁴ He had completed his magistrate training program only three months prior and did not have a law degree, which is allowed under Virginia law.¹⁴⁵ Predictably, the magistrate judge signed off on this search warrant’s “sweeping and powerfully intrusive” terms.¹⁴⁶

Then, the officer contradicted the terms approved by the magistrate judge as he executed the geofence search. In step one, the officer requested anonymized information on 19 individuals detected within the geofence.¹⁴⁷ In step two, however, the officer “did not ‘attempt to narrow down’” his request despite making that exact promise to the magistrate judge days before.¹⁴⁸ Rather, “in contravention to Google’s policy, and without consulting [the judge],” the officer repeatedly asked Google for the full names, usernames, email addresses, and other identifying information on all 19 people.¹⁴⁹ Not only that, the also officer doubled the geofence’s time and location parameters in these subsequent requests without narrowing the initial list of suspects.¹⁵⁰ It was only after Google’s specialist personally called the officer did the latter finally narrow his request.¹⁵¹

2. *Chatric is Not the Answer*

On a motion to suppress evidence, the district court held that the geofence search warrant was invalid for two reasons, but nevertheless denied the

140. *Id.* at 919–20.

141. *Id.* at 919.

142. *Id.* at 922–23.

143. *Id.* at 919.

144. *Id.* at 939.

145. *Id.* (citing VA. CODE §§ 19.2-37).

146. *Id.*

147. *Id.* at 920.

148. *Id.* at 921.

149. *Id.*

150. *Id.*

151. *Id.* at 922–23.

motion. First, the third-party doctrine did not apply because the government could not point out when Chatrie enabled the feature that disclosed his location.¹⁵² This, coupled with Google’s confusing interfaces, showcased the government’s failure to prove that Chatrie voluntarily shared his location to a third party.¹⁵³ Because the third-party doctrine did not apply under these facts, police needed a warrant to conduct the geofence search that identified Chatrie. Hence, the court proceeded to its second line of reasoning: the warrant was not “sufficiently particular” in outlining probable cause for the individuals to be searched or the information sought.¹⁵⁴

First, the third-party doctrine did not apply. The district court held multiple evidentiary hearings on Google’s various products and services. These hearings revealed that, at the time the geofence search was conducted in summer 2018, Google’s interfaces made it difficult for users to learn the extent of Google’s location tracking, let alone delete their location history data.¹⁵⁵ Due to the “messiness of the current record as to when Chatrie ‘gave consent’” for a third party to track his location, the trial court did not find that Chatrie voluntarily forfeited his expectation of privacy under the third-party doctrine.¹⁵⁶

Then, in what appears to be dicta, the district court cited *Carpenter* to argue that, more broadly, the third-party doctrine does not apply to geofences searches and thus a warrant is required. Prosecutors urged the opposite, distinguishing the two rationales given in *Carpenter* regarding cell-site location information. They argued that the geofence captured “just two hours” of Chatrie’s location, which raises a smaller privacy interest than the days’ worth of information revealed in *Carpenter*.¹⁵⁷ Second, they argued that a geofence search can only track those who enable Google’s Location History feature, which is a voluntary, “affirmative step” to disclosing one’s location, unlike the automatic pings to cell towers in *Carpenter*.¹⁵⁸

The district court rejected the prosecutors’ arguments by citing powerful language in *Carpenter*. First, the court stated that Chatrie did have a privacy interest in “just two hours” of location data because, “perhaps even more so than” the information in *Carpenter*, Chatrie’s location was “detailed,

152. *Id.* at 935.

153. *Id.* at 935–36.

154. *Id.* at 927–33.

155. *See id.* at 913, 914 n.17 (quoting Google employees and engineers who called the process of deleting one’s location history confusing).

156. *Id.* at 935.

157. *Id.*

158. *See id.* at 935–36.

encyclopedic, and effortlessly compiled.”¹⁵⁹ Second, the trial court concluded that Chatrie did not voluntarily provide his location because Google provided users with “limited and partially hidden warnings” regarding the frequency and precision of its location tracking.¹⁶⁰ Consequently, whatever “affirmative steps” Chatrie took in enabling the Location History feature would not “constitute a full assumption of the attendant risk of permanently disclosing one’s whereabouts during almost every minute of every hour of every day.”¹⁶¹ Indeed, “a user simply cannot forfeit the protections of the Fourth Amendment for years of precise location information by selecting ‘YES, I’M IN’ at midnight while setting up [an app].”¹⁶²

The *Chatrie* district court’s reasoning, though persuasive, will not broadly question the application of the third-party doctrine to geofence search warrants challenged in future cases. The court’s criticism rested primarily on the government’s failure to prove consent under these particular facts, which hinged on Google’s inaccessible interfaces. But both of these hurdles are fixable in future prosecutions. The court repeatedly noted that Google’s interfaces were confusing and incomplete as of summer 2018.¹⁶³ Yet in the months following the search in *Chatrie*, Google introduced several “controls that made it easier for users to manage their data.”¹⁶⁴ And today, Google automatically deletes location data after 18 months, gives the option of automatically deleting data every three months, and offers a “Privacy Checkup tool” that allows users to see and control all information that Google collects.¹⁶⁵ With Google’s since-updated privacy policies, in the future

159. *Id.* (quoting *Carpenter*, 138 S. Ct. at 2216).

160. *Id.* at 936 (“In the Google Assistant set-up process, the device likely provided Chatrie a single pop-up screen informing him that ‘[t]his data may be saved and used in any Google service where [he was] signed in to give [him] more personalized experiences,’ and that he ‘can see [his] data, delete it and change [his] settings at account.google.com.’ . . . However, the consent flow did not detail, for example, how frequently Google would record Chatrie’s location (every two to six minutes); the amount of data Location History collects (essentially *all* location information); that even if he ‘stopped’ location tracking it was only ‘paused,’ meaning Google retained in its Sensorvault all his past movements; or, how precise Location History can be (*i.e.*, down to twenty or so meters).”).

161. *Id.*

162. *Id.*

163. *See id.* at 911, 914 n.17, 936.

164. *Id.* at 913–14.

165. Jessica Bursztynsky, *Google Just Announced It Will Automatically Delete Your Location History by Default*, CNBC (June 24, 2020), <https://www.cnbc.com/2020/06/24/google-will-automatically-delete-location-history-by-default.html>; Todd Haselton, *Google Collects Information About Many Things You Do Online—Here’s How to Stop It*, CNBC (May 1, 2019), <https://www.cnbc.com/2019/05/01/how-to-stop-google-from-collecting-your-private-information.html>.

prosecutors can more credibly argue defendants “assume the risk” of disclosing their whereabouts to police. It bears repeating that if a court concludes that the third-party doctrine applies, whatever search law enforcement conducted is not a search within the meaning of the Fourth Amendment and no warrant is required.

The *Chatrrie* court’s conclusion that geofence search warrants lack sufficient particularity likewise cannot be used to broadly question future geofence searches. Search warrants must have probable cause, a “fair probability” that the search will reveal evidence of a crime based on the “totality of the circumstances.”¹⁶⁶ The court in *Chatrrie* emphasized that the requirement of particularity in warrants limits the officers’ discretion while they conduct searches. To limit infringements on privacy to only what is necessary for law enforcement, “discretion must be confined to the signing magistrate, not to the executing officers or a third party.”¹⁶⁷ Thus, the geofence search warrant in *Chatrrie* was invalid because steps two and three of Google’s protocol did not require police to narrow the list of identified devices. Accordingly, the warrant failed to meet the particularity requirement because it did not provide the officer with “clear standards from which he or she could reasonably . . . ascertain and identify . . . the place to be searched [or] the items to be seized.”¹⁶⁸

Crucially, the court in *Chatrrie* emphasized that it was not ruling that all geofence search warrants would lack particularity. The court referenced a case from the Northern District of Illinois that upheld a search warrant with six geofences that contained smaller timeframes and locations where few bystanders were present.¹⁶⁹ The court then suggested it would be constitutional for police to begin with an initial search for anonymized information, then broaden the search over the course of several successive approvals from magistrate judges.¹⁷⁰ Yet in the same breath, the court acknowledged that

166. *Gates*, 462 U.S. at 233, 238.

167. *Chatrrie*, 590 F. Supp. 3d at 935 (citing *United States v. Chadwick*, 433 U.S. 1, 9 (1977)).

168. *Id.* (quoting *In re Search of: Information Stored at Premises Controlled by Google*, 481 F. Supp. 3d 730, 754 (2020)); *United States v. Blakeney*, 949 F.3d 851, 861 (2022)).

169. *Id.* (citing *In re Search Warrant Application for Geofence Location Data Stored at Google Concerning an Arson Investigation*, 497 F. Supp. 3d 345, 361–62 (N.D. Ill. 2020)). In this case, six interrelated geofences covered areas related to two strings of suspected arsons. The geofences pinged devices present at multiple timeframes near multiple commercial and residential parking lots and roadways that connect these locations, then conducted searches of the same areas months later. 497 F. Supp. 3d at 351–53.

170. *See Chatrrie*, 590 F. Supp. 3d at 933 (“In certain situations, then, law enforcement likely *could* develop initial probable cause to acquire from Google *only* anonymous data from devices within a narrowly circumscribed geofence at Step 1 From there, officers likely could use

anonymized data can reveal shocking amounts of intimate information.¹⁷¹ Thus, if the alternative approach laid out in *Chatrie* were adopted, police would not need to obtain additional judge approval to conduct follow-up searches because all the information they need—and much more—would already be at their fingertips.

In its final illustration of the Fourth Amendment’s limited ability to restrict geofence search warrants, the court in *Chatrie* still admitted evidence gathered from the unconstitutional search. Unlawfully gained evidence can be admitted when an officer conducted the search in “good faith.”¹⁷² In *Chatrie*, the court reasoned that the officer had a good-faith belief that the geofence search warrant was constitutional, in part, because of “rapidly advancing technology and lack of judicial guidance on this novel investigatory technique.”¹⁷³ Another way to prove an officer’s good faith is to show that they “reasonably” relied on the fact that a magistrate judge approved the search warrant, even when the approval itself was a “sweeping and powerfully intrusive” constitutional error.¹⁷⁴ Because the officer in *Chatrie* “reasonably” relied on the fact that magistrates had previously approved three similarly broad geofence search warrants, the court held that the officer acted in good faith.¹⁷⁵

Chatrie leaves unanswered an important question: why is it “reasonable” or in “good faith” for an officer to not follow protocol? Okello Chatrie spent time behind bars because this officer “reasonably” believed it was legal to use Google’s inadequate protocol on three prior occasions. The Fourth Amendment did not allow the court in *Chatrie* to question the good faith of an officer who “inexplicably” told the judge he had already found nineteen suspects before he even spoke with Google.¹⁷⁶ Perhaps the officer did not narrow his “sweeping and powerfully intrusive” request because he knew, as

that narrow, anonymous information to develop probable cause particularized to specific users. Importantly, officers likely could then present that particularized information to a magistrate or magistrate judge to acquire successively broader and more invasive information.”).

171. *See id.* at 931 n.39 (“The fact that data points obtained during Steps 1 and 2 are anonymized when Google reports them does not completely quell this Court’s concerns about the invasiveness of this warrant. Even ‘anonymized’ location data—from innocent people—can reveal astonishing glimpses into individuals’ private lives when the Government collects data across even a one or two hour period.”).

172. *Leon*, 468 U.S. at 923.

173. *Chatrie*, 590 F. Supp. 3d at 936.

174. *Id.* at 939; *Leon*, 468 U.S. at 922–23.

175. *Chatrie*, 590 F. Supp. 3d at 937–38.

176. *Id.* at 920.

Chatrle showcases, that the judiciary fails to hold police accountable when they violate an individual's Fourth Amendment rights.¹⁷⁷

Rather than provide a cause for celebration, *Chatrle* epitomizes the Fourth Amendment's failure to deter and remedy infringements of privacy and free speech when police use novel technology. Thus, privacy and speech advocates should not rely on courts to restrict geofence searches.

IV. THE IMPORTANCE OF A BLANKET BAN

A. WHY THE CONSTITUTION CANNOT REGULATE GEOFENCE SEARCHES

Existing literature on geofences primarily discusses the constitutionality of geofence searches. Just as the court in *Chatrle* did, many articles argue or assume that geofence searches require a warrant because of the holding in *Carpenter* that people have a privacy interest in their location data and because geofences often capture information on people in their homes.¹⁷⁸ But courts likely will not adopt a “bright-line rule” that warrantless geofence searches are unconstitutional.¹⁷⁹ That is because, as *Chatrle* and other cases exhibit, the constitutionality of any search warrant turns on its degree of particularity and the “totality of the circumstances.”¹⁸⁰ And every hole in our “Swiss cheese” Fourth Amendment weakens the promise of that already deprived test.

In fact, there are at least five justices on the Supreme Court who could rule that geofence searches categorically do not require a warrant. Justices Alito and Thomas are obvious candidates. Both justices dissented in *Carpenter*, that *Carpenter* had no privacy interest in any amount of location data—even data with “GPS-level precision”—because customers have no property rights over cell phone records.¹⁸¹ A third candidate is Justice Gorsuch, who separately dissented in *Carpenter* on originalist, property-based grounds, under which one scholar has argued geofence searches would not require a warrant.¹⁸² Fourth, Chief Justice Roberts, who authored *Carpenter*, could plausibly distinguish the

177. *See id.* at 921.

178. *See De La Torre, supra* note 14, at 329–30.

179. *See id.*

180. *Id.*; *Gates*, 462 U.S. at 230; *Chatrle*, 590 F. Supp. 3d at 927; *In re Search of: Information Stored at Premises Controlled by Google*, 481 F. Supp. 3d 730, 740–41 (2020).

181. *Carpenter*, 138 S. Ct. at 2224.

182. *See id.* at 2261–72 (Gorsuch, J., dissenting); Reed Sawyers, *For Geofences: An Originalist Approach to the Fourth Amendment*, 29 GEO. MASON L. REV. 787, 796–809 (2021) (appraising Justice Gorsuch's originalist framework and, in part, analogizing geofence search warrants to compelled subpoenas from early American history).

privacy interest and voluntariness of geofence search data from that of cell towers like the prosecutors in *Chatrie* did.

Either Justice Kavanaugh or Justice Barrett could be a fifth vote. Although Justices Kavanaugh and Barrett have not yet ruled on a Fourth Amendment case, neither one is a reliable vote for privacy. Professor Orin Kerr pegged Kavanaugh's likely Fourth Amendment jurisprudence as "somewhere in the ballpark" of Justice Kennedy, who wrote the primary *Carpenter* dissent, or Chief Justice Rehnquist, who voted to weaken the Fourth Amendment dozens of times.¹⁸³ Neither hypothesis is promising if the Supreme Court hears a geofence challenge. On the D.C. Circuit, then-Judge Kavanaugh wrote that the National Security Agency's bulk collection of metadata was "entirely consistent with the Fourth Amendment"—a position that has troubled digital privacy advocates.¹⁸⁴ On the Seventh Circuit, then-Judge Barrett twice ruled to exclude evidence, but neither case involved the search of a cell phone.¹⁸⁵ In her sole case that involved digital privacy interests, she ruled to admit evidence obtained from a warrantless border search of a traveler's cell phone.¹⁸⁶ Again, this holding is not promising if the Court decides to hear a challenge to geofence searches, particularly when Justice Barrett is a self-avowed originalist like Justices Gorsuch and Thomas.

Even when geofence searches require a warrant, this requirement itself does not adequately protect speech. Setting aside the numerous relevant exceptions to the warrant requirement,¹⁸⁷ the Fourth Amendment is not

183. Orin Kerr, *Judge Kavanaugh on the Fourth Amendment*, SCOTUSBLOG (July 20, 2018), <https://www.scotusblog.com/2018/07/judge-kavanaugh-on-the-fourth-amendment/> (analyzing five of Justice Kavanaugh's rulings on the D.C. Circuit Court of Appeals); Craig M. Bradley, *Rehnquist's Fourth Amendment: Be Reasonable*, 82 MISS. L.J. 259, 260, 268 (2013) (noting that in over thirty years on the bench, there was only one non-unanimous Fourth Amendment case where Chief Justice Rehnquist voted for the defendant).

184. *Klayman v. Obama*, 805 F.3d 1148, 1148–49 (D.C. Cir. 2015) (Kavanaugh J., concurring). One expert wrote that Kavanaugh's opinion foreshadowed his "unwillingness to consider how technological changes have affected rights afforded by the Fourth Amendment." Susan Landau, *Brett Kavanaugh's Failure to Acknowledge the Changes in Communications Technology: The Implications for Privacy*, LAWFARE (Aug. 3, 2018), <https://www.lawfareblog.com/brett-kavanaughs-failure-acknowledge-changes-communications-technology-implications-privacy>.

185. *Amy Coney Barrett and Privacy*, ELECTRONIC PRIVACY INFO. CTR., <https://archive.epic.org/privacy/barrett/> (last visited Nov. 8, 2023) (dissecting then-Judge Barrett's opinions on the Seventh Circuit).

186. *Id.*

187. One exception is when the facts facing the officer present "exigent circumstances." See, e.g., *Brigham City v. Stuart*, 547 U.S. 398, 400–01 (2006) (holding that the safety of an occupant inside a home creates exigent circumstances, then finding this safety threatened when police overheard a fist fight inside a person's home); *Kentucky v. King*, 563 U.S. 452, 470 (2011) (holding that it does not violate the Fourth Amendment for police to deliberately create exigent circumstances).

equipped to address speech concerns. The Supreme Court effectively ruled as much when it held that police may raid a newsroom as long as there is a fair probability that that doing so will reveal evidence of a crime.¹⁸⁸ The fact that raiding a newsroom would have harmed the free flow of information, a core tenet of the First Amendment, did not change the Court's analysis because the warrant requirement's *raison d'être* is to limit invasions of privacy, not speech.¹⁸⁹

Frankly, merely requiring a warrant is not a panacea for privacy concerns, either. The heart of the Fourth Amendment's warrant requirement is the notion that police simply need to ask permission before they violate your privacy or enter your home. Thus, if a cop and a judge suspect you have committed a crime, you no longer have an expectation of privacy over your personal information when that information is relevant to a crime. If one's priority is effective law enforcement, that makes sense. But the scope of criminal law has become so broad that it extends to the act of protesting itself,¹⁹⁰ activities that occur near or during protests,¹⁹¹ and even people who attend protests with outstanding arrest warrants.¹⁹² So long as our overly broad criminal law remains the filter through which the Fourth Amendment's warrant requirement operates, courts will allow geofence searches to the detriment of people, privacy, and speech.

That is, unless Congress acts. If courts will not offer meaningful, much-needed restrictions on geofence searches in the coming years, then privacy advocates must seek a different avenue. Legislative action is thus necessary.

B. WHY PROPOSED LEGISLATION WILL NOT PROTECT SPEECH AND PRIVACY

Generally, legislative proposals argue that geofence search warrants should have greater detail than a typical search warrant. For instance, one scholar

188. *Zurcher v. Stanford Daily*, 436 U.S. 547, 553 (1978).

189. *See generally id.*

190. *See, e.g., US Protest Law Tracker*, INT'L CTR. FOR NOT-FOR-PROFIT L., <https://www.icnl.org/usprotestlawtracker/> (last visited Nov. 8, 2023) (documenting the 18 states that have criminalized protests against oil and gas infrastructure since 2017); Kaylana Mueller-Hsia, *Anti-Protest Laws Threaten Indigenous and Climate Movements*, BRENNAN CTR. FOR JUST. (Mar. 17, 2021), <https://www.brennancenter.org/our-work/analysis-opinion/anti-protest-laws-threaten-indigenous-and-climate-movements> (“The combination of overly broad language and steep penalties in critical infrastructure laws make it likely that future activists and supporting organizations will be discouraged from exercising their First Amendment-protected protest rights.”).

191. *See, e.g., Whittaker, supra* note 12; Brandom, *supra* note 15.

192. *See* AM. CIVIL LIBERTIES UNION N. CAL., *supra* note 52 (noting that police used real-time social media monitoring to identify protesters and “directly” arrest them from the crowd).

proposes that all geofence search warrants should have a printed map that illustrates the geofence's parameters.¹⁹³ This scholar contends that this would educate magistrate judges on privacy concerns before they approve warrants.¹⁹⁴

This proposal touches on, yet does not fully grasp the implications of, a wealth of evidence that magistrate judges are weak checks against police. One landmark study revealed that judges, on average, takes less than three minutes to review and approve a warrant.¹⁹⁵ It is common for police to go “shopping” for magistrates who tend to favor police.¹⁹⁶ Highly technical information in search warrants is systematically reduced to boilerplate explanations and surface-level descriptions like “cellular phone analysis.”¹⁹⁷ Yet as police request tens of thousands of geofence search warrants per year, continuing education programs for magistrate judges did not have a single class with the word “geofence” in 2021.¹⁹⁸ Worse, as *Chatrle* showcased, magistrate judges do not need a law degree to authorize geofence search warrants.¹⁹⁹ To the extent that magistrates understand the Fourth Amendment, police deference is practically hardwired into its doctrine. All of this, coupled with informational asymmetries between police and magistrates, causes the latter to routinely defer to the former.²⁰⁰

In the face of these enormous structural problems, it is improbable to think that reforms like the inclusion of a printed map in a geofence search warrant application would sway a magistrate judge. Consider the photo below in Figure 1, which was contained in the geofence search warrant application in *Chatrle*.²⁰¹ To put it mildly, nothing about the photo illustrates the privacy and speech interests at play because the photo is blurry, black and white, and wholly non-descriptive as to what the captured buildings are and who may be inside them.

193. Mohit Rathi, *Rethinking Reverse Location Search Warrants*, 111 J. CRIM. L. & CRIMINOLOGY 805, 832 (2021).

194. *Id.* at 832–33.

195. RICHARD VAN DUIZEND, L. PAUL SUTTON & CHARLOTTE A. CARTER-YAMAUCHI, *THE SEARCH WARRANT PROCESS: PRECONCEPTIONS, PERCEPTIONS, AND PRACTICES* 31(1985).

196. *Id.* at 23–26.

197. *Id.*; O'Brien, *supra* note 86, at 24.

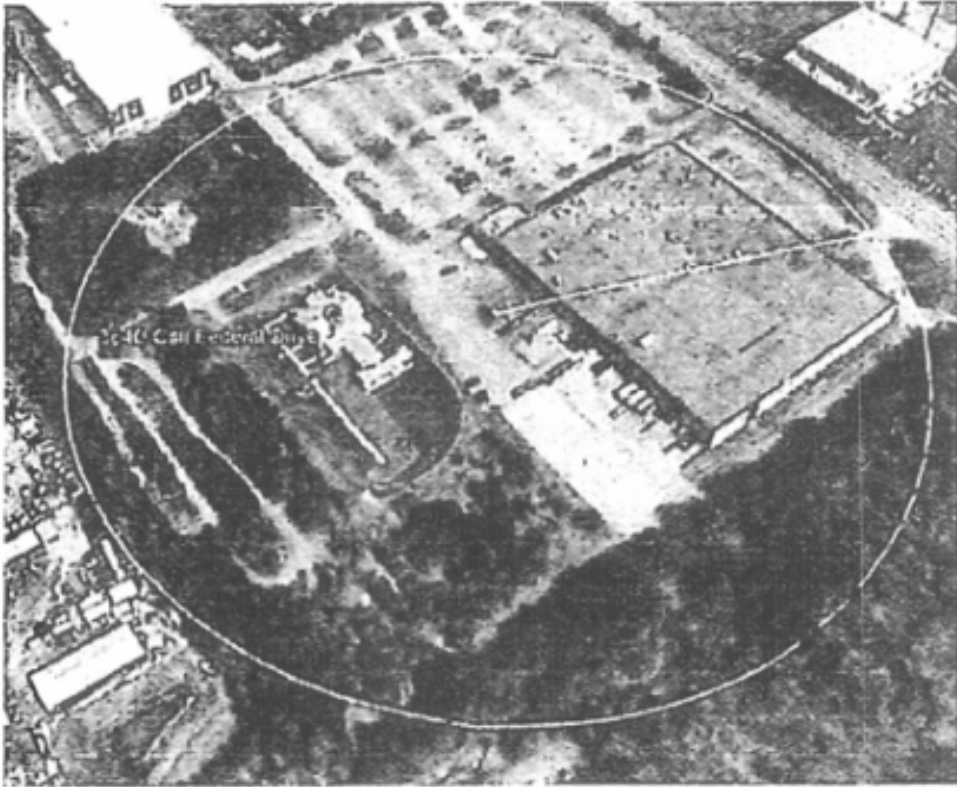
198. O'Brien, *supra* note 86, at 24–25.

199. *Chatrle*, 590 F. Supp. 3d at 939.

200. *See generally* O'Brien, *supra* note 86.

201. *Chatrle*, 590 F. Supp. 3d at 919.

Figure 1. The Parameters of *Chattie*'s Geofence Search²⁰²



Another set of legislative reforms calls on Congress to regulate Google's process for evaluating warrants. Among these proposals is the requirement that Google provide step-one information in a de-identified format.²⁰³ Echoing the *Chattie* court's suggestion, another idea proposes that law enforcement must seek further judge approval after some—or all—steps in Google's protocol.²⁰⁴ And police must narrow requests at step two or three of Google's protocol, rather than give police discretion over this decision.²⁰⁵

But legislation that focuses solely on the process of seeking search warrants will be ineffective. Recall what the court in *Chattie* concluded: using Google's anonymous data, police could “observe each account's reported location, track each account to his or her home, and pinpoint each account's personal identity

202. *Id.*

203. Rathi, *supra* note 193, at 833.

204. *Chattie*, 590 F. Supp. 3d at 933.

205. Rathi, *supra* note 193, at 834–35.

using publicly available resources.”²⁰⁶ These resources include the Data Broker Loophole.²⁰⁷ Using anonymized data, one can identify the names, addresses, consumption habits, ethnicities, and ages of hundreds of people.²⁰⁸ When Americans have exposed so much of their private lives in the digital era, simply tinkering with the warrant-seeking process will not protect privacy. The mere existence of geofence searches weaponizes our ubiquitous internet footprints.

A similar category of legislative proposals are restrictions on the circumstances under which geofence searches can be sought. One scholar suggested they only be approved in “exigent circumstances.”²⁰⁹ Under this proposal, Congress would require police to demonstrate geofence searches are “a last resort.”²¹⁰ Judges would engage in an explicit balancing inquiry, approving warrants only when “the public safety threat would significantly outweigh the privacy [risks].”²¹¹

There are two primary problems with this proposal. First, as a practical matter, police could easily manipulate the statutory language. Second, exceptions will do little to quell the perception of surveillance, which inhibits speech.

First, case law shows just how easily police and courts would manipulate the language in the proposed restriction. Consider the phrase “public safety threat.” In a criminal procedure ruling, the Supreme Court referred to a suspect who was disarmed, already in handcuffs, and in an empty supermarket in the middle of the night as a “threat to the public safety” that “outweigh[ed]” his Fifth Amendment rights.²¹² In an evidence case, a drug-deal shooting that occurred twenty-five minutes prior with no follow-up activity was an “ongoing emergency.”²¹³ The Court also interprets phrases “last resort” and “exigent circumstances” broadly. In Fourth Amendment cases, the term “exigent circumstances” describes situations where there is “no time to secure a

206. *Chatrre*, 590 F. Supp. 3d at 931 n.39.

207. *See id.* (quoting an American Bar Association report that discussed the power of anonymized data); *supra* Section III.A.

208. Warzel & Thompson, *supra* note 89.

209. Cassandra Zietlow, *Reverse Location Search Warrants: Law Enforcement’s Transition to ‘Big Brother,’* 23 N.C. J.L. & TECH. 669, 698 (2022).

210. *Id.* at 697.

211. *Id.* at 700.

212. *New York v. Quarles*, 467 U.S. 649, 651–52, 58 (1984).

213. *Michigan v. Bryant*; 562 U.S. 344, 351–52 (2011); *see also Quarles*, 467 U.S. at 879–85 (Scalia, J., dissenting).

warrant,²¹⁴ which lower courts construed to include testing a person's urine²¹⁵ and smelling marijuana then hearing people moving inside an apartment.²¹⁶ With this in mind, legislators must ask whether they can trust courts and police to interpret even strongly worded limitations in a way that protects people, privacy, and speech.

Second, carving out piecemeal exceptions for geofence searches cannot ameliorate the harms to political speech. Empirical evidence shows that disruptions in political speech flow from the mere perception of surveillance.²¹⁷ In response to perceived surveillance, people self-censor, expend additional resources to organize political activities, and refrain from protests.²¹⁸ A proposal that does not ban all geofence searches will prove ineffective because the public will continue to correctly perceive that police can exploit the law's vagueness to use geofence searches as a surveillance tactic.

C. WHY A BLANKET BAN IS THE ANSWER

A blanket ban is the most effective way to address the impending harm of geofence searches. Regulated or not, geofence searches will inevitably lead to harassment of peaceful activists, intrusions on privacy, and unwarranted incarceration. The best avenue would be an act of Congress because federal legislation affects not just state and local police, but federal officers as well. Preventing federal officers from using geofence search warrants is crucial because the vast reach and resources of federal agencies like ICE make them uniquely able to maintain the worst harms of our surveillance state.²¹⁹

In September 2021, New York introduced legislation to ban geofence searches. Congress should follow suit. Under Assembly Bill A84A, “no court shall issue a reverse location court order” and “no government entity shall seek, from any court, a reverse location court order.”²²⁰ “Reverse location court order” is the bill’s term for a court-issued geofence search warrant.²²¹ If

214. See, e.g., *Lange*, 141 S. Ct. at 2018; *Missouri v. McNeely*, 133 S. Ct. 1552, 1559 (2013); *Michigan v. Tyler*, 436 U.S. 499, 509 (1978).

215. *State v. Hanson*, 588 N.W.2d 885, 889 (S.D. 1999); see also Emily J. Sovell, *State v. Hanson: Has the Exigent Circumstances Exception to the Warrant Requirement Swallowed the Rule?*, 45 S.D. L. REV. 163, 179–185 (2000) (criticizing the *Hanson* decision).

216. *Kentucky v. King*, 563 U.S. 452, 470 (2011).

217. See generally Stoycheff, *supra* note 58, at 299–300.

218. *Id.*

219. See Speri & Saleh, *supra* note 51.

220. Assemb. B. A84A, 2021–2022 Leg., Reg. Sess. §§ 695.10, 695.20(2) (N.Y. 2021).

221. *Id.* § 695.00(3) (“‘Reverse location court order’ means any court order, including a search warrant, compelling the disclosure of records or information pertaining to electronic devices or their users or owners, whose scope extends to an unknown number of electronic devices present in a given geographic area at a given time as measured via global positioning

the bill is passed, criminal defendants may make a motion to exclude evidence gained from geofence searches.²²² To ensure deterrence and compliance with the law, the bill authorizes civil suits by “any individual whose records were obtained by any government entity” in violation of its terms.²²³ Google, Microsoft, and Yahoo all support the bill.²²⁴

A common counterargument to a blanket ban is one that reifies our carceral state: banning geofence searches is “too extreme” because it would hurt law enforcement.²²⁵ Yet the value that geofence searches add is, at best, indeterminate. Recent statistics suggest 11,000 geofence search warrants were executed in 2020.²²⁶ There is no data on how many convictions these 11,000 searches led to, or even a breakdown of the crimes that were investigated.²²⁷ It is also unclear how often geofences prove necessary; police have plenty of other cheap, effective, and less racially disparate investigative tools at their disposal.²²⁸

And broadly speaking, whatever benefit of solving crimes occurs is linked to over-policing and mass incarceration, largely against Black and Brown communities.²²⁹ This has immense human and social costs.²³⁰ Not only is

system coordinates, cell tower connectivity, Wi-Fi data, and/or any other form of location detention.”).

222. *Id.* § 695.30.

223. *Id.* § 695.40.

224. Zack Whittaker, *Google, Microsoft, and Yahoo Back New York Ban on Controversial Search Warrants*, TECHCRUNCH (May 10, 2022), <https://techcrunch.com/2022/05/10/google-new-york-geofence-keyword-warrant/>; Matthew Guariglia, *Geofence Warrants and Reverse Keyword Warrants are So Invasive, Even Big Tech Wants to Ban Them*, ELECTRONIC FRONTIER FOUND. (May 13, 2022), <https://www.eff.org/deeplinks/2022/05/geofence-warrants-and-reverse-keyword-warrants-are-so-invasive-even-big-tech-wants>.

225. Zietlow, *supra* note 14, at 695; cf. A. Spencer Davies, *A Californian Algorithm: Amendment Assembly Bill 2261 to Regulate Law Enforcement’s Use of Facial Recognition Technology in Post Hoc Criminal Investigations*, 26 BERKELEY J. CRIM. L. 27, 68 (2021) (noting this argument against a ban on law enforcement’s use of facial recognition technology).

226. Whittaker, *supra* note 12.

227. *See id.*

228. *See* Nadine Deslauriers-Varin & Francis Fortin, *Improving Efficiency and Understanding of Criminal Investigations: Toward an Evidence-Based Approach*, 36 J. OF POLICE & CRIM. PSYCH. 635, 635 (2021) (“In recent years, we are, however, witnessing a growth of empirical studies that aim at providing support to police forces and specialized investigation units, and improving the efficiency of their practices using a proactive and evidence-based approach. This [is] particularly true for sexual crimes and homicides[.]”); *see also generally id.* at 636 (previewing a special issue of the *Journal of Police and Criminal Psychology* that contains 11 articles related to “innovative” investigative techniques, processes, and decision-making strategies).

229. *See supra* Section II.B (discussing how the urbanization of geofence searches will disproportionately affect racial minorities).

230. *See, e.g.*, Michael McLaughlin, Carrie Pettus-Davis, Derek Brown, Chris Veeh & Tanya Reen, *The Economic Burden of Incarceration in the United States* 4–5, (Inst. for Justice Research

incarceration's human toll important in its own right, but it is counterproductive because it aggravates the root causes of crime, thereby creating a revolving door of release, recidivism, and reincarceration.²³¹ For every person that a geofence search puts behind bars, there is a family and a community made less whole.

The premise of a blanket ban is that a small number of crimes may go unsolved if doing so safeguards people, privacy, and speech consistent with the values enshrined in the Constitution and Bill of Rights. Although this argument is not fully reflected in Fourth Amendment doctrine, Congress can and should enact a law with this principle in mind.

V. CONCLUSION

Geofence searches pose tremendous privacy and speech risks that neither Fourth Amendment nor legislative reforms will meaningfully mitigate. And *United States v. Chatrue*, despite its celebrated reasoning, showcased the failures of the Fourth Amendment to deal with these pending risks. Accordingly, Congress must enact a blanket ban on their use.

The harms of geofence searches are similar, and will add to, those of other forms of digital surveillance. The failures of the third-party doctrine and the Fourth Amendment's remedies to address these types of surveillance should give us pause as well. Going forward, legislators should consider whether the arguments fleshed out above justify blanket prohibitions on police use of commercial data, social media surveillance, facial recognition technology, and so much more. Without further action by Congress, the First and Fourth Amendments' promises will remain just that, promises.

& Development, Working Paper No. IJRD-072016, 2016) (finding that the aggregate economic impact of incarceration is \$1 trillion in losses to income, health, and other measures); *see also supra* notes 92–96 and accompanying text (documenting, in great detail, the individual harms of pretrial detention, including disruptions in “wages and employment, housing stability, familial relationships, and mental and physical health”).

231. Criminology literature offers several theories for why incarceration may reduce crime, including deterrence of crime and incapacitating people from committing crimes. For a thorough critique of this literature on theoretical, methodological, and empirical grounds, see David Roodman, *The Impacts of Incarceration on Crime* (July 9, 2020) (unpublished manuscript), <https://ssrn.com/abstract=3635864>; see Alexi Jones, *Reforms Without Results: Why States Should Stop Excluding Violent Offenses From Criminal Justice Reforms*, PRISON POLICY INITIATIVE (Apr. 2020), <https://www.prisonpolicy.org/reports/violence.html> (summarizing Roodman's findings aptly: “incarceration can be counterproductive: While a prison sentence can incapacitate people in the short term, it actually increases the risk that someone will commit a crime after their release.”).

DATA SURVEILLANCE AND ABORTION BANS AFTER *DOBBS*

Leila Nasrolahi[†]

TABLE OF CONTENTS

I.	INTRODUCTION	1342
II.	ABORTION JURISPRUDENCE AND STATE LAWS POST-<i>DOBBS</i>	1343
III.	DATA SURVEILLANCE WILL BE THE PRIMARY MODE OF ENFORCING ABORTION BANS.....	1345
A.	PRE-ROE ENFORCEMENT.....	1346
B.	POST-ROE ENFORCEMENT.....	1348
1.	<i>Modern Digital Data Technologies Reveal the User's Thoughts Before They</i> <i>Act on Them.....</i>	<i>1349</i>
2.	<i>Moving Data Trails.....</i>	<i>1350</i>
3.	<i>Data Trails Specific to Reproductive Health.....</i>	<i>1351</i>
4.	<i>Data is Easy for Law Enforcement to Obtain.....</i>	<i>1354</i>
C.	DIGITAL DATA ANSWERS QUESTIONS THAT EVEN MEDICINE CANNOT.....	1355
IV.	CHILLING EFFECTS	1356
A.	DATA SURVEILLANCE WILL CHILL ACCESS TO LEGAL ABORTION CARE.....	1357
B.	DATA SURVEILLANCE WILL CHILL THE PROVISION OF LEGAL HEALTH CARE.....	1360
C.	DATA SURVEILLANCE WILL CHILL LEGAL INFORMATION SHARING	1362
V.	POSSIBLE SOLUTIONS	1364
A.	TECH COMPANIES TO THE RESCUE?	1365
1.	<i>Evidence of Broken Privacy Promises.....</i>	<i>1367</i>
2.	<i>Placing the Responsibility on Users.....</i>	<i>1368</i>
3.	<i>Clear Conflict of Interest.....</i>	<i>1368</i>
B.	FEDERAL PRIVACY LEGISLATION.....	1369
1.	<i>Overview of Proposed Federal Legislation.....</i>	<i>1369</i>

DOI: <https://doi.org/10.15779/Z38QV3C51V>

© 2023 Leila Nasrolahi.

[†] J.D., 2024, University of California, Berkeley, School of Law.

VI. CONCLUSION 1372

I. INTRODUCTION

Is it acceptable if enforcing criminal law requires us to give up digital privacy? How much of ourselves are we willing to sacrifice for the perfect enforcement of crimes? State laws banning abortions following *Dobbs v. Jackson Women's Health Organization* give rise to these unanswered questions.

In that case, the Supreme Court overturned *Roe v. Wade* and held that the Constitution does not confer a right to abortion.¹ Since then, fourteen states have enacted laws banning almost all abortions, and even more enacted laws placing gestational limits on abortions.² In the most hostile states, abortion providers can face up to ninety-nine years in prison—even when the pregnancy was the result of rape or incest.³

Much has changed since abortion was last illegal. Most notably, digital technology now pervades reproductive healthcare.⁴ Members of the public use the internet to obtain health-related information, period-tracking apps to record their menstrual cycles, GPS to navigate to doctor's appointments, and social media to engage with others on reproductive health topics.⁵

While digital technology provides users with efficient tools for information access, it also provides law enforcement with efficient tools for criminal investigations. In fact, surveillance is the “dominant philosophy for how police enforce laws in 2022.”⁶ Post-*Dobbs*, there is increasing concern about how pregnant people's digital data will be used against them.⁷

1. *Dobbs v. Jackson Women's Health Org.*, 142 S. Ct. 2228, 2284 (2022).

2. *Tracking Abortion Bans Across the Country*, N.Y. TIMES (Nov. 7, 2023), <https://www.nytimes.com/interactive/2022/us/abortion-laws-roe-v-wade.html> [hereinafter *Tracking Abortion Bans*].

3. *See, e.g.*, ALA. CODE §§ 26-23H-4, 13A-5-6 (2022).

4. *See* Cynthia Conti-Cook, *Surveilling the Digital Abortion Diary*, 50 UNIV. BALT. L. REV. 1, 24 (2020).

5. *See id.* at 13.

6. Alfred Ng, *'A Uniquely Dangerous Tool': How Google's Data Can Help States Track Abortions*, POLITICO (July 18, 2022), <https://www.politico.com/news/2022/07/18/google-data-states-track-abortions-00045906> (explaining how Google's location data can help states track abortions).

7. *See, e.g.*, Jay Edelson, *Post-Dobbs, Your Private Data Will Be Used Against You*, BLOOMBERG L. (Sept. 22, 2022), <https://news.bloomberglaw.com/us-law-week/post-dobbs-your-private-data-will-be-used-against-you>.

Setting aside the substantive issue of abortion criminalization as an attack on bodily autonomy,⁸ this Note describes the “chilling effects”—the over-deterrence of legal activity—that will result from digital data surveillance used in abortion prosecutions. Because there is no viable way to enforce abortion bans via data surveillance without chilling legal activity, this Note argues that abortion bans should not be enforced this way, even if that means settling for lesser enforcement.

This Note proceeds in four Parts. Part II first reviews the Supreme Court’s abortion jurisprudence relevant to this piece, namely the trimester framework under *Roe v. Wade*, its modification by the undue burden test in *Planned Parenthood v. Casey*, and finally its subsequent reversal in *Dobbs*. Part II then summarizes the status of state abortion laws to date, paying particular attention to states with the most restrictive bans.

Part III posits that digital data surveillance will be the primary mode of enforcement in abortion actions. In contrast to how abortion laws were enforced pre-*Roe*, Part III describes how data surveillance allows for maximal enforcement. Part IV argues that data surveillance in abortion actions will result in dangerous chilling effects on legal activities. Part V offers solutions and concludes.

II. ABORTION JURISPRUDENCE AND STATE LAWS POST-*DOBBS*

In 1973, the Court held 7-2 that the Due Process Clause of the Fourteenth Amendment protects a fundamental right to privacy, which encompasses the right to an abortion.⁹ In *Roe v. Wade*, the Court distinguished between the different stages of pregnancy to delineate how a state could regulate abortion.¹⁰ Before fetal viability, a state could not regulate a person’s decision to seek an abortion.¹¹ Once the fetus reached viability, the point at which it could potentially survive outside the mother’s womb, a state could regulate or prohibit abortions except when necessary to save the life of the mother.¹²

8. See, e.g., *The Constitutional Right to Reproductive Autonomy: Realizing the Promise of the 14th Amendment*, CTR. FOR REPRODUCTIVE RTS. (July 2022), <https://reproductiverights.org/wp-content/uploads/2022/07/Final-14th-Amendment-Report-7.26.22.pdf> (discussing the constitutional rights and guarantees in U.S. law underlying the right to and importance of reproductive autonomy).

9. *Roe v. Wade*, 410 U.S. 113, 164 (1973).

10. *Id.* at 163–64.

11. *Id.*

12. *Id.*

In 1992, the Court reluctantly reaffirmed *Roe* in *Planned Parenthood of Southeastern Pennsylvania v. Casey*.¹³ However, *Casey* discarded the stages-of-pregnancy distinctions from *Roe* and instead imposed the “undue burden” standard, which asked whether a state regulation had the purpose or effect of placing a substantial obstacle in the way of a woman seeking an abortion before viability.¹⁴

In *Dobbs v. Jackson Women’s Health Organization*, the Supreme Court overturned *Roe* and *Casey* and held that the Constitution does not confer a right to abortion.¹⁵ For the first time since 1973, states are empowered to place total, unrestricted bans on abortion.¹⁶ *Dobbs* involved a Mississippi law that generally prohibited abortion after the fifteenth week of pregnancy, well before the viability line announced in *Roe*.¹⁷ In a 6-3 decision, the Court overruled *Roe* and *Casey*, reasoning that the Constitution makes “no reference to abortion, and no such right is implicitly protected by any constitutional provision.”¹⁸ Thus, after *Dobbs*, abortion legality is determined by states.

Fourteen states anticipated the reversal of *Roe* and wrote trigger laws banning abortion that immediately took effect after *Dobbs*.¹⁹ In Alabama, Arkansas, Idaho, Indiana, Kentucky, Louisiana, Mississippi, Missouri, North Dakota, Oklahoma, South Dakota, Tennessee, Texas, and West Virginia, abortion is banned with no exceptions for rape or incest.²⁰ Arizona, Florida, Georgia, Nebraska, North Carolina, and South Carolina, and Utah have gestational limit abortion bans, prohibiting abortion as early as six weeks from the last missed period.²¹ Separately, over 100 bills restricting access to abortion were introduced in 2022.²²

13. *Planned Parenthood of Se. Pa. v. Casey*, 505 U.S. 833, 846, 853 (1992) (“While we appreciate the weight of the arguments made on behalf of the State in the cases before us, arguments which in their ultimate formulation conclude that *Roe* should be overruled, the reservations any of us may have in reaffirming the central holding of *Roe* are outweighed by the explication of individual liberty we have given combined with the force of *stare decisis*.”).

14. *Id.* at 879.

15. *Dobbs*, 142 S. Ct. at 2284.

16. *See id.*

17. *Id.* at 2242.

18. *Id.* at 2284.

19. Larissa Jimenez, *60 Days After Dobbs: State Legal Developments on Abortion*, BRENNAN CTR. FOR JUST. (Aug. 24, 2022), <https://www.brennancenter.org/our-work/research-reports/60-days-after-dobbs-state-legal-developments-abortion>.

20. *Tracking Abortion Bans*, *supra* note 2.

21. Amy Schoenfeld Walker, *Most Abortion Bans Include Exceptions. In Practice, Few Are Granted*, N.Y. TIMES (Jan. 21, 2023), <https://www.nytimes.com/interactive/2023/01/21/us/abortion-ban-exceptions.html>.

22. Jimenez, *supra* note 19.

On the other side, several progressive states introduced legislation to expand abortion coverage following the *Dobbs* decision. In 2023, at least sixteen states passed legislation protecting abortion access.²³ New Jersey passed a bill to codify a constitutional right to freedom of reproductive choice.²⁴

III. DATA SURVEILLANCE WILL BE THE PRIMARY MODE OF ENFORCING ABORTION BANS

Data surveillance will be the primary mode of enforcement of abortion bans because: (1) it captures the widest possible range of potential criminal activity; (2) it answers questions that even medicine cannot; and (3) there is already evidence of it being used.

Dobbs must be considered against the backdrop of unprecedented technological advances in data surveillance²⁵ that have developed since *Roe*—technologies that allow law enforcement to achieve the most capacious mode of enforcement. That is, modern data surveillance captures as much potential criminal activity as possible—what I refer to as “maximal enforcement.” Data surveillance offers law enforcement an efficient and effective way to track criminal activity.²⁶ This is especially relevant in the abortion context since the activity at issue is inherently intimate and private. Moreover, digital data can answer a question about abortions that even medicine cannot: the difference between a medical abortion and a miscarriage.²⁷ That is, since the abortion pill

23. *Id.*

24. N.J. STAT. ANN. § 10:7-1(a).

25. Since *Roe*, Google was founded in 1998, portable GPS devices became available in 1999, Facebook was founded in 2004, and the first iPhone was sold in 2007. With that, increasing connections between technology and policing have developed. See *From the Garage to the Googleplex*, ABOUT GOOGLE, https://about.google/intl/ALL_us/our-story/ (last visited Nov. 11, 2023); Geotab Team, *History of GPS Satellites and Commercial GPS Tracking*, GEOTAB (June 23, 2020), <https://www.geotab.com/blog/gps-satellites/>; Nicholas Carlson, *At Last—The Full Story of How Facebook was Founded*, BUS. INSIDER (Mar. 5, 2010), <https://www.businessinsider.com/how-facebook-was-founded-2010-3>; Ben Gilbert & Sarah Jackson, *Steve Jobs Unveiled the First iPhone 16 Years Ago—Look How Primitive It Seems Today*, BUS. INSIDER (Jan. 9, 2023), <https://www.businessinsider.com/first-phone-anniversary-2016-12>. For research suggesting that technological improvements have increased police capabilities, see CHRISTOPHER KOPER, CYNTHIA LUM, JAMES WILLIS, DAN WOODS & JULIE HIBDON, REALIZING THE POTENTIAL OF TECHNOLOGY IN POLICING: A MULTISIDE STUDY OF THE SOCIAL, ORGANIZATIONAL, AND BEHAVIORAL ASPECTS OF IMPLEMENTING POLICING TECHNOLOGIES (2015), <https://nij.ojp.gov/library/publications/realizing-potential-technology-policing-multisite-study-social-organizational>.

26. See *infra* Section III.B.

27. The abortion pill works by stimulating the same process as a naturally occurring miscarriage. See Jessica Beaman, Christine Prifti, Eleanor Bimla Schwarz & Mindy Sobota,

stimulates the same process as a naturally occurring miscarriage, a doctor cannot readily discern whether a patient who is purporting to have a miscarriage in fact took an abortion pill. However, that patient's search history and location data may provide an answer.

Today's surveillance technology is what will separate pre-*Roe* abortion bans from post-*Dobbs* bans. Whereas abortion bans pre-*Roe* depended on physical evidence to prosecute lawbreakers, today digital data surveillance will be the primary mode of enforcing abortion bans.²⁸

A. PRE-ROE ENFORCEMENT

Dobbs must be considered in light of the unprecedented technological advances in data surveillance that have taken place since *Roe*. It is helpful to first understand enforcement mechanisms pre-*Roe* as a contrast to the pervasive possibilities that data surveillance now offers.

In the early 1900s, before data surveillance was available as an enforcement mechanism, abortion laws were enforced primarily through obtaining dying declarations of women who received abortions and through police raids.²⁹ When a woman in the early twentieth century died from an illegal abortion, the state prosecuted the "abortionist" by using dying declarations as a crucial piece of evidence.³⁰ In fact, some thought that without the dying declaration, it was "almost 'impossible' to obtain evidence of criminal abortion any other way."³¹ Since early abortion practices were often unsafe and performed illegally by non-physicians, women often called their physicians when they experienced post-abortion complications.³² Prosecutors primarily focused on cases where women died and were considered "victims" of a crime.³³ When this happened to Carolina Petrovitis, her doctor asked, "Who did it for you[?] If you won[']t tell me what was done to you I can't handle your case."³⁴ Petrovitis eventually revealed that a midwife performed her abortion and her doctor informed police officers.³⁵ As Petrovitis realized she would soon die,

Medication to Manage Abortion and Miscarriage, 35 J. GEN. INTERNAL MED. 2398 (2020), <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7403257/> (noting that "for both medication abortion and medical management of early miscarriage, the standard of care is to provide oral mifepristone followed by misoprostol tablets").

28. *Infra* Section III.B.

29. LESLIE J. REAGAN, WHEN ABORTION WAS A CRIME: WOMEN, MEDICINE, AND LAW IN THE UNITED STATES, 1867–1973, at 114, 161 (1997).

30. *Id.* at 114.

31. *Id.* at 118.

32. *See id.* at 119.

33. *Id.* at 116.

34. *Id.* at 113.

35. *Id.*

the police collected a statement that implicated the midwife who performed her abortion.³⁶ The police brought the midwife to the hospital and Petrovitis identified her as the person who performed her illegal abortion.³⁷

To gather evidence to prosecute abortionists in the early 1900s, the state needed to have physicians reporting abortions and collecting dying declarations from their patients, which many doctors were reluctant to do.³⁸ But doctors were convinced to side with the state because they feared the investigative process would be “turned against them.”³⁹ This fear was not irrational; records from medical society meetings describe doctors’ experiences being indicted as an accessory to murder for failing to call the coroner or obtain a dying declaration from a patient.⁴⁰ Even when doctors were acquitted of abortion charges, they were excommunicated by their medical communities.⁴¹ To protect themselves, physicians were advised to “deny medical care to a woman who had had an abortion until she made a statement.”⁴² As a result, “doctors found themselves caught in the middle between their responsibilities to their patients and the demands of government officials.”⁴³

In addition to dying declarations, by the 1940s the state relied on aggressive raids to enforce abortion laws. Rather than only focusing on women’s deaths by unsafe abortionists, prosecutors “worked to shut down the trusted and skilled abortionists, many of them physicians, who had operated clinics for years with little or no police interference.”⁴⁴ Consider the story of an underground abortion clinic in Pennsylvania. After receiving a tip from a suspicious neighbor,

police officers . . . hid in the nearby fields . . . waiting and watching . . . [T]he officers unlocked the front door . . . [T]hey found one woman wearing only a slip in one room, two lying in bed in another, and two more who, having removed their skirts and underwear, sat waiting for their abortions in a third.⁴⁵

36. *Id.*

37. *Id.*

38. *Id.* at 120.

39. *Id.*

40. *Id.* at 120–21.

41. *Id.*

42. *Id.* at 122.

43. *Id.* at 116.

44. *Id.* at 161.

45. Leslie J. Reagan, *Caught in the Net*, SLATE (Sept. 10, 2021), <https://slate.com/news-and-politics/2021/09/enforcement-of-abortion-laws-before-roe-v-wade.html>.

These raids were the primary mode of enforcement in the 1950s and 1960s.⁴⁶ Police officers raided offices and apartments where abortion providers worked and escorted women to male doctors who would determine whether a surgical procedure had been performed.⁴⁷ Doctors would then testify in court as to their findings.⁴⁸ Meanwhile, the women who received abortions were forced to testify in court against their abortion provider.⁴⁹

Pre-*Roe* enforcement tools relied on physical confrontations that took place after abortion care was administered. Much has changed since then. Whereas in the early 1900s it may have been impossible to imagine abortion prosecutions without dying declarations,⁵⁰ today's digital age allows law enforcement to obtain a wealth of information without relying on physical confrontation and well before an abortion occurs.

B. POST-*ROE* ENFORCEMENT

Data surveillance is a promising way to determine whether someone had or is planning to have an abortion because of how pervasive and informative the data is. Search history data provides information about a person's thoughts and considerations before any actions have necessarily been taken. Location data provides information connected to one's movements—where they go and when they go.⁵¹ Data from reproductive health applications, websites, and social media pages provides information specific to abortion care.⁵² This data about a person is produced “as an unintended byproduct of access to internet search tools, social-media platforms and other communication apps, and web-based services to make purchases or access services via a smartphone or other wired device.”⁵³ Data surveillance gives information about “individuals' physical states, movements, interests, and moods on a minute-by-minute basis.”⁵⁴

This Section, III.B, discusses three categories of data surveillance that are relevant to abortion criminal law enforcement. First, the data from search history that reveals the user's thoughts; second, location data that follows users' physical movements; and lastly, medical data that offers concrete

46. REAGAN, *supra* note 29, at 160–62.

47. *Id.*

48. *Id.*

49. *Id.* at 165.

50. *Id.* at 118.

51. *Infra* notes 66–76.

52. *Infra* notes 77–86.

53. Aziz Z. Huq & Rebecca Wexler, *Digital Privacy for Reproductive Choice in the Post-Roe Era*, 98 N.Y.U. L. REV. 555, 569–70 (2023).

54. *Id.* at 570.

information about pregnancy, menstruation, and other markers of reproductive health.

1. *Modern Digital Data Technologies Reveal the User's Thoughts Before They Act on Them.*

Search history data allows an evidence trail to begin much earlier than ever before—Google might be the first to find out someone is pregnant. Our search history is an extension of our thoughts.⁵⁵ *What do my symptoms mean? How much does an abortion cost?* Our online data follows our most intimate wonderings, blurring the lines between our physical and digital selves.⁵⁶ Pregnant people are likely to search for health-related information online, especially during the early stages of pregnancy.⁵⁷ Pregnant people “prefer the online experience because of . . . the ability to manage their health in what feels like a private manner.”⁵⁸ Search history sheds light on the questions people may be too afraid to ask in-person.

Law enforcement can require Google to turn over search history data by using a “keyword warrant.” A keyword warrant is when police request data in “reverse” by asking Google to disclose everyone who searched a keyword, without necessarily having a specific suspect in mind.⁵⁹ For example, in a 2020 arson-murder investigation, police sent a search warrant requesting information on users who searched the address of the residence around the time of the arson.⁶⁰ Google complied with the data request, and three teenagers who searched the address were charged with murder.⁶¹ In a fraud investigation, police requested “any/all user or subscriber information related to the Google searches of ‘Douglas [REDACTED]’ for the timeframe of December 1st, 2016 thru January 7th, 2017.”⁶² The warrant specified that the

55. SETH STEPHENS-DAVIDOWITZ, *EVERYBODY LIES: BIG DATA, NEW DATA, AND WHAT THE INTERNET CAN TELL US ABOUT WHO WE REALLY ARE* (2017).

56. *Id.*

57. See generally Padaphet Sayakhot & Mary Carolan-Olah, *Internet Use by Pregnant Women Seeking Pregnancy-Related Information: A Systematic Review*, *BMC PREGNANCY CHILDBIRTH* (Mar. 28, 2016), <https://pubmed.ncbi.nlm.nih.gov/27021727/>.

58. Conti-Cook, *supra* note 4, at 24.

59. Alfred Ng, *Google Is Giving Data to Police Based on Search Keywords, Court Docs Show*, *CNET* (Oct. 8, 2020), <https://www.cnet.com/news/privacy/google-is-giving-data-to-police-based-on-search-keywords-court-docs-show/>.

60. Julia Love, *Google Keyword-Search Warrants Questioned by Colorado Lawyers*, *BLOOMBERG* (Jan. 12, 2023), <https://www.bloomberg.com/news/articles/2023-01-12/google-keyword-search-warrants-questioned-by-colorado-lawyers>.

61. *Id.*

62. Application for Search Warrant, No. 27-CR-CV-17-1 (Feb. 1, 2017), <https://www.documentcloud.org/documents/3519211-Edina-Police-Google-Search-Warrant-Redacted.html>.

information should include names, addresses, phone numbers, dates of birth, social security numbers, email addresses, payment information, account information, and IP addresses of all persons who made the Google search.⁶³ In each of these scenarios, law enforcement used keyword warrants to obtain critical search history data.

Search history evidence is not new, but post-*Dobbs* abortion bans give it new power. When Latice Fisher was prosecuted for second-degree murder for the death of her newborn after stillbirth, her online search, “buy Misopristol Abortion Pill Online,” was key evidence.⁶⁴ In future abortion investigations, law enforcement can utilize a reverse keyword search to locate individuals who searched “Planned Parenthood address” or “abortion pills”—without having any specific suspect in mind. Albert Fox Cahn, the executive director of the Surveillance Technology Oversight Project, likened keyword warrants to “going to a library and then trying to search every person who checked out a specific book,” arguably something we “would never allow . . . in the analog world.”⁶⁵

2. *Moving Data Trails*

In addition to our intimate thoughts, our digital data also follows our physical movements. Many cellphone applications enable “location services,” which provide information about the geographic position of the device, even when the app is not actively being used.⁶⁶ Google tracks location data from the IP address of a device’s internet connection, a web search that includes a location in it, and Google Maps usage. Location-based data and analytics can identify where users are traveling from, how often they are visiting a location, and traveler demographics.⁶⁷ Location History logs a user’s location on average every two minutes.⁶⁸ By using geofencing technology, companies can direct advertisements at smartphone users located in a designated area through browsers and applications on their devices.

63. *Id.*

64. See Conti-Cook, *supra* note 4, at 3 n.3.

65. Bobby Allyn, *Privacy Advocates Fear Google Will be Used to Prosecute Abortion Seekers*, NPR (July 11, 2022), <https://www.npr.org/2022/07/11/1110391316/google-data-abortion-prosecutions>.

66. *Location, Location, Location: Tips on Controlling Mobile Tracking*, ST. CAL. DEP’T JUST.: OFF. ATT’Y GEN. (Oct. 2015), <https://oag.ca.gov/privacy/facts/online-privacy/location>.

67. Emily Carroll, *What is Location-Based Data?*, DRIVERRESEARCH (July 8, 2019), <https://www.driverresearch.com/market-research-company-blog/what-is-location-based-data-market-research-company/>.

68. Cullen Seltzer, *Google Knows Where You’ve Been. Should It Tell the Police?*, SLATE (May 16, 2022), <https://slate.com/technology/2022/05/google-geofence-warrants-chatrie-location-tracking.html>.

Even before *Dobbs*, digital data was used to monitor and target individuals who sought abortions. For example, in 2017, Copley Advertising created mobile geofences at reproductive health centers that read “Pregnancy Help,” “You Have Choices,” and “You’re Not Alone.”⁶⁹ Copley was hired by pro-life religious groups to target “abortion-minded” women.⁷⁰

By obtaining geofence warrants, police can make requests to Google for data on devices logged in at a specific area and time.⁷¹ Google received 982 geofence warrants in 2018, 8,396 in 2019, and 11,554 in 2020.⁷² Google does not publish information about how often it complies with geofence warrants or whether it rejects overly broad requests.⁷³ Geofence warrants, like keyword warrants, are “reverse” warrants because they identify people—anyone—who was near a certain area in a specified time frame. A geofence warrant “doesn’t start with a suspect or even an account; instead police request data on every device in a given geographic area during a designated time period, regardless of whether the device owner has any link at all to the crime under investigation.”⁷⁴ Police have used geofence warrants to determine the suspects in a burglary⁷⁵ and attendees at a protest.⁷⁶ Rather than conducting a physical raid to prove someone received an abortion, today police can draw a 200-foot boundary around an abortion clinic and use Google location data to determine the identity of everyone who entered the area at any given moment.

3. *Data Trails Specific to Reproductive Health*

In addition to the general information offered by search history and location tracking, there is an amalgam of digital data specific to reproductive

69. *AG Reaches Settlement with Advertising Company Prohibiting ‘Geofencing’ Around Massachusetts Healthcare Facilities*, MASS.GOV (Apr. 4, 2017), <https://www.mass.gov/news/ag-reaches-settlement-with-advertising-company-prohibiting-geofencing-around-massachusetts-healthcare-facilities>.

70. *Id.*

71. *Id.*

72. Zack Whittaker, *Google Says Geofence Warrants Make Up One-quarter of All US Demands*, TECHCRUNCH (Aug. 19, 2021), <https://techcrunch.com/2021/08/19/google-geofence-warrants/>.

73. *Id.*

74. Jennifer Lynch, *First Court in California Suppresses Evidence from Overbroad Geofence Warrant*, ELEC. FRONTIER FOUND. (Oct. 11, 2022), <https://www.eff.org/deeplinks/2022/10/california-court-suppresses-evidence-overbroad-geofence-warrant>.

75. Note, *Geofence Warrants and the Fourth Amendment*, 134 HARV. L. REV. 2508, 2508 (2021).

76. Matthew Guariglia, Mukund Rathi, Houston Davidson & Jennifer Lynch, *Geofence Warrants Threaten Civil Liberties and Free Speech Rights in Kenosha and Nationwide*, ELEC. FRONTIER FOUND. (Sept. 10, 2021), <https://www.eff.org/deeplinks/2021/09/geofence-warrants-threaten-civil-liberties-and-free-speech-rights-kenosha-and>.

health. As of the last decade, there has been an explosion of *femtech*⁷⁷ tools, products, services, wearable technology, and software that “use technology to address women’s health issues, including menstrual health, reproductive health, sexual health, maternal health and menopause.”⁷⁸ Femtech apps like Flo (a menstrual tracking app), Glow (a fertility tracking app), and Ava (a fertility tracking bracelet) store data about users that is specific to their reproductive health, including menstruation data and sexual activity. Period tracking apps are a common tool for people to anticipate their cycle symptoms, log menstruation dates, and family-plan. Some apps can predict pregnancy more than a week before at-home pregnancy tests can.⁷⁹ One study found that nearly a third of women in the United States use a period-tracking app.⁸⁰ Flo, a popular app with millions of users, includes articles, quizzes, and even a community for discussing sexual and reproductive health issues.⁸¹ Post-*Dobbs*, experts say period-tracking data may become a target for investigators.⁸² Used in combination with search history and location data, a period tracking app may give law enforcement evidence that someone received an illegal abortion.

Another critical change since the *Roe* era is the way digital communications are captured on social media. Social media is increasingly used as a source of political news and discussion.⁸³ Countless Instagram accounts specifically offer abortion-related content, offering anything from mutual aid funds, political opinions, personal stories, and information to obtain abortions in states where it is illegal.⁸⁴ Law enforcement is already using social media data in abortion

77. “Femtech” was coined by Ida Tin, co-founder of Clue, a menstrual health app. Ida Tin, *The Rise of a New Category: Femtech*, CLUE (Sept. 14, 2016), <https://hellocue.com/articles/culture/rise-new-category-femtech>.

78. Linda Rosencrance, *What Is Femtech?*, TECHTARGET (Apr. 2022), <https://www.techtarget.com/whatis/definition/femtech>.

79. Huq & Wexler, *supra* note 53, at 573.

80. Carly Page, *Supreme Court Overturns Roe v. Wade: Should You Delete Your Period-Tracking App?*, TECHCRUNCH (May 5, 2022), <https://techcrunch.com/2022/05/05/roe-wade-privacy-period-tracking/>.

81. See FLO HEALTH, <https://flo.health/> (last visited Nov. 22, 2023).

82. See, e.g., Leah Fowler & Michael Ulrich, *Femtechdystopia*, 75 STAN. L. REV. 1233, 1313 (2023) (“Period- and fertility-tracking apps are the most obvious consumer technologies but by no means the only ones that could be instrumentalized to criminalize abortion and other behaviors during pregnancy.”).

83. Dam Hee Kim, Brian E. Weeks, Daniel S. Lane, Lauren B. Hahn & Nojin Kwak, *Sharing and Commenting Facilitate Political Learning on Facebook: Evidence From a Two-Wave Panel Study*, 7 SOC. MEDIA + SOC’Y (Sept. 27, 2021), <https://journals.sagepub.com/doi/full/10.1177/20563051211047876>.

84. See, e.g., Nat’l Network of Abortion Funds (@abortionfunds), INSTAGRAM, <https://www.instagram.com/abortionfunds/> (last visited Nov. 22, 2023); Liberate Abortions (@liberateabortion), INSTAGRAM, <https://www.instagram.com/liberateabortion/> (last visited

investigations. For example, a Nebraska mother was sentenced to two years in prison for giving abortion pills to her pregnant daughter after 20 weeks of pregnancy.⁸⁵ Law enforcement obtained a warrant for their Facebook messages which allegedly discussed their plans to terminate the pregnancy at home.⁸⁶

In just the first half of 2021, Google received approximately 150,000 government requests for disclosure of users' account information pursuant to a subpoena in all cases and a search warrant in criminal cases.⁸⁷ Google complied with almost 80% of those requests.⁸⁸ Apple received 12,589 government requests and complied in 90% of cases.⁸⁹ Facebook received 237,414 requests and provided data in 76.1% of cases.⁹⁰ Data-driven law enforcement "lets police become aggressively more proactive."⁹¹ A supervising police detective said, "tech providers, especially social media platforms, offer a trove of information that can help solve [crimes]. Everything happens on Facebook. The amount of information you can get from people's conversations online—it's insane."⁹²

Combined, all the data that companies collect from their users make up what has been coined as "surveillance capitalism": "the unilateral claiming of private human experience as free raw material for translation into behavioral data."⁹³ So long as these surveillance mechanisms exist, law enforcement and

Nov. 22, 2023); Abortion Photograph (@theabortionproject), INSTAGRAM, <https://www.instagram.com/theabortionproject/> (last visited Nov. 22, 2023).

85. Margery A. Beck, *Nebraska Mother Sentenced to 2 years in Prison for Giving Abortion Pills to Pregnant Daughter* AP NEWS (Sept. 22, 2023, 2:31 PM), <https://apnews.com/article/abortion-charges-nebraska-sentence-36b3dcaadd6b705ca2315bc95b99bdc1>.

86. *Id.*

87. *Global Requests for User Information*, GOOGLE: TRANSPARENCY REP., <https://transparencyreport.google.com/user-data/overview> (last visited Nov. 22, 2023).

88. *Id.*

89. APPLE, APPLE TRANSPARENCY REPORT: GOVERNMENT AND PRIVATE PARTY REQUESTS 1 (2021), <https://www.apple.com/legal/transparency/pdf/requests-2021-H1-en.pdf> ("Types of legal requests Apple receives from the United States can be: subpoenas, court orders, search warrants, pen register/trap and trace orders, or wiretap orders.").

90. *Facebook Transparency Report*, FACEBOOK, <https://transparency.fb.com/data/government-data-requests/> (last visited Nov. 22, 2023).

91. *How Data-driven Policing Threatens Human Freedom*, ECONOMIST (June 4, 2018), <https://www.economist.com/open-future/2018/06/04/how-data-driven-policing-threatens-human-freedom>.

92. Matt O'Brien & Michael Liedtke, *How Big Tech Created a Data 'Treasure Trove' for Police*, COURTHOUSE NEWS SERV. (June 22, 2021), <https://www.courthousenews.com/how-big-tech-created-a-data-treasure-trove-for-police/>.

93. Shoshana Zuboff, a professor at Harvard Business School, coined the term "surveillance capitalism" in 2014. Zuboff notes it was "Google that first learned how to capture surplus behavioral data, more than what they needed for services, and used it to

private enforcers will continue to take full advantage of available data. As the saying goes, “if you build it, they will come.”

4. *Data is Easy for Law Enforcement to Obtain*

It will not be difficult for law enforcement to access the troves of data created by our digital devices. To obtain a warrant for users’ data, police must satisfy a probable cause showing. However, “warrants will offer only very limited protection against restrictionist law enforcement demands” because probable cause is such a low bar.⁹⁴ Police who seek a keyword warrant for users who searched “abortion” will likely be able to articulate probable cause just by “point[ing] to criminal statutes in seeking evidence about abortion.”⁹⁵

Police can also circumvent warrant requirements by purchasing data directly from data brokers. Widespread data surveillance supports what is known as a data economy, a “digital ecosystem in which the producers and consumers of data—business and individuals—and government and municipal agencies gather, organize, and share accumulated data from a wide variety of sources.”⁹⁶ Users’ data is pervasively shared and sold to third party data brokers who compile it and resell it to whoever seeks to buy it—including individuals, advertisers, marketing firms, and law enforcement.⁹⁷ In August 2022, the Federal Trade Commission sued Kochava Inc., a data broker allegedly selling non-anonymized mobile geolocation data that could be used to track consumers’ visits to sensitive locations including abortion providers.⁹⁸ To prove how easy it is to obtain location data of people who visit abortion clinics, a reporter bought a week’s worth of data on where people who visited

compute prediction products that they could sell to their business customers, in this case advertisers.” John Laidler, *High Tech is Watching You*, HARV. GAZETTE (Mar. 4, 2019), <https://news.harvard.edu/gazette/story/2019/03/harvard-professor-says-surveillance-capitalism-is-undermining-democracy/>.

94. Huq & Wexler, *supra* note 53, at 578.

95. *Id.*

96. *Capitalizing on the Data Economy*, MIT TECH. REV. (Nov. 16, 2021), <https://www.technologyreview.com/2021/11/16/1040036/capitalizing-on-the-data-economy/>.

97. *See, e.g.*, Bennett Cyphers, *Inside Fog Data Science, the Secretive Company Selling Mass Surveillance to Local Police*, ELEC. FRONTIER FOUND. (Aug. 31, 2022), <https://www.eff.org/deeplinks/2022/08/inside-fog-data-science-secretive-company-selling-mass-surveillance-local-police>.

98. *FTC Sues Kochava for Selling Data that Tracks People at Reproductive Health Clinics, Places of Worship, and Other Sensitive Locations*, FED. TRADE COMMISSION (Aug. 29, 2022), <https://www.ftc.gov/news-events/news/press-releases/2022/08/ftc-sues-kochava-selling-data-tracks-people-reproductive-health-clinics-places-worship-other>.

Planned Parenthood came from and went afterward for just \$160 from a data broker called SafeGraph.⁹⁹

Law enforcement—and anyone else—can purchase data directly from data brokers without any judicial oversight. One data broker, Fog Data Science, contracts with police to provide “easy and often warrantless access to the precise and continuous geolocation of hundreds of millions of unsuspecting Americans.”¹⁰⁰ Fog purchases billions of data points across thousands of mobile apps from millions of devices, which it then sells to law enforcement agencies for a cheap subscription fee.¹⁰¹

Post-*Dobbs* abortion law enforcement will look drastically different from the rudimentary pre-*Roe* methods. An overwhelming amount of information about individuals’ thoughts, ideas, preferences, and movements is collected by Big Tech companies. Law enforcement will capitalize on this data to identify as much abortion-related activity as possible.

C. DIGITAL DATA ANSWERS QUESTIONS THAT EVEN MEDICINE CANNOT

Data surveillance is a feasible way to determine whether someone is planning to have an abortion. Unless a pregnant person specifically goes out of their way to avoid a digital trace completely, their location data and search history will implicate them. Data surveillance offers law enforcement the tools to achieve as close to perfect enforcement as possible. Moreover, digital data answers a question that medicine often cannot: the difference between a miscarriage and a medical abortion. From a medical perspective, “there is no physically significant difference between a medication abortion and a spontaneously occurring miscarriage. For example, the medicines used in medication abortion are used to help safely manage an incomplete miscarriage.”¹⁰² Digital data has the power to fill in the gaps. In states where abortion is banned, consider the following scenario: a pregnant person takes an abortion pill and experiences excessive bleeding. She goes to her doctor but does not want to disclose that she took abortion pills. Her doctor provides

99. Joseph Cox, *Data Broker Is Selling Location Data of People Who Visit Abortion Clinics*, VICE (May 3, 2022), <https://www.vice.com/en/article/m7vzjb/location-data-abortion-clinics-safegraph-planned-parenthood>.

100. Matthew Guariglia, *What Is Fog Data Science? Why Is the Surveillance Company So Dangerous?*, ELEC. FRONTIER FOUND. (Aug. 31, 2022), <https://www.eff.org/deeplinks/2022/06/what-fog-data-science-why-surveillance-company-so-dangerous>.

101. *Id.*

102. *Consumer Health Info: Medication Abortion and Miscarriage*, NAT’L WOMEN’S HEALTH NETWORK (Aug. 15, 2019), <https://nwhn.org/abortion-pills-vs-miscarriage-demystifying-experience/>.

treatment—the same treatment used for both abortions and miscarriage. Perhaps the doctor is suspicious that it was a self-managed abortion but cannot diagnose because there is no way to distinguish from a spontaneous miscarriage. But her search history shows searches for abortion pills. Geolocation data places her at a clinic that was known to provide abortions before *Dobbs*. Suddenly, her digital data enables a medical diagnosis.

IV. CHILLING EFFECTS

Part III described the pervasive data surveillance that will be used in abortion-related criminal investigations. This Part considers the repercussions of that surveillance, which I argue are chilling effects on various legal activities. I use “chilling effects” to mean that a rule will involve some ambiguity or error in application, causing people to avoid beneficial conduct that society would otherwise like them to engage in.

The “chilling effect” is a phenomenon in which people refrain from engaging in legal expression for fear of breaking a law and the subsequent retaliation, prosecution, or punitive governmental action.¹⁰³ In states where abortion bans are in place, people will be deterred from breaking criminal abortion laws, but they will also refrain from participating in legal activities like providing life-saving abortions and sharing information about reproductive health.¹⁰⁴ Law enforcement’s use of data surveillance—the enforcement mechanism—will be the primary cause of this deterrence, rather than the severity of punishment itself. Criminal deterrence scholars have posited that the *certainty of punishment* has a greater impact on deterrence than the *severity of punishment*:

Certainty refers to the likelihood of being caught and punished for the commission of a crime. Research underscores the more significant role that certainty plays in deterrence than severity—certainty of being caught deters a person from committing crime, not the fear of being punished or the severity of the punishment. Effective policing that leads to swift and certain (but not necessarily severe) sanctions is a better deterrent than the threat of incarceration.¹⁰⁵

103. David L. Hudson, Jr., *Chilling Effect Overview*, FOUND. FOR INDIVIDUAL RTS. & EXPRESSION, <https://www.thefire.org/research-learn/chilling-effect-overview>.

104. See, e.g., *Further Restricting Abortions in NC Will Have ‘Chilling’ Effect, Doctors Say*, DUKE TODAY (Feb. 17, 2023), <https://today.duke.edu/2023/02/further-restricting-abortion-nc-will-have-chilling-effect-doctors-say>.

105. *Five Things About Deterrence*, NAT’L INST. JUST. (June 5, 2016), <https://nij.ojp.gov/topics/articles/five-things-about-deterrence>.

Applying this logic, the probability of whether someone will be caught performing or receiving an abortion has a greater impact on behavior than the length of the sentence imposed. Since the probability of punishment is determined by the pervasiveness of data surveillance, it follows that the more surveillance there is, the more behavior—both legal and illegal—will be deterred. The likelihood of enforcement in the abortion context is dependent on the invasiveness of digital surveillance.¹⁰⁶ Without it, enforcement mechanisms will look like they did pre-*Roe* and will be inefficient and largely ineffective. Therefore, without a fine net of data, the concerns of the chilling effects described below would be much less. Conversely, the more data surveilled, the greater the chilling effects will become.

Data surveillance as an enforcement mechanism for abortion bans gives rise to three major chilling effects. First, there will be a chilling effect on legal abortion access. Second, there will be a chilling effect on legal non-abortion reproductive care. Third, there will be a chilling effect on legal information sharing about reproductive health. Each is discussed in turn.

A. DATA SURVEILLANCE WILL CHILL ACCESS TO LEGAL ABORTION CARE

Data surveillance will have a chilling effect on legal abortions because increasing the certainty of enforcement will make doctors more risk averse to perform abortions in gray areas. As they stand, abortion laws target providers and others who assist in performing an abortion.¹⁰⁷ But even the strictest states have exceptions when abortion is necessary to save the life of the mother.¹⁰⁸ Other less restrictive states also include exceptions when the pregnancy was the result of rape or incest.¹⁰⁹ As abortion laws are more intensely enforced via data surveillance, these important exceptions will be undermined because doctors will be fearful of being wrongfully accused of performing an illegal abortion.¹¹⁰

In 2021, Alabama made it a Class A felony to perform an abortion except in cases where it is necessary to “prevent a serious health risk to the unborn child’s mother,” which the legislature defined as death or serious risk of substantial physical impairment of a major bodily function.¹¹¹ Class A felonies are punishable by up to ninety-nine years in prison.¹¹² Therefore, there will be

106. See *supra* Part III.

107. See, e.g., ALA. CODE § 26-23H-4 (2021); IDAHO CODE § 18-622 (2020).

108. See statutes cited *supra* note 107.

109. See Walker, *supra* note 21.

110. See *id.*

111. ALA. CODE §§ 26-23H-4–8 (1975).

112. ALA. CODE §§ 13A-5–6 (2019).

instances in which doctors must ask and answer questions like: *Is this patient's condition close enough to death? How much blood loss must occur before an ectopic pregnancy is considered life-threatening under Alabama's law? How serious is a "serious risk"? How should "substantial" impairment be quantified?*¹¹³ These are all questions that remain unanswered and will inevitably unfold as cases are litigated. What if, in investigating whether the mother's life was truly endangered, law enforcement obtains search history data that indicates the woman was seeking an abortion?

Several doctors have articulated their fears. One Indiana doctor described a patient whose ultrasound showed a miscarriage was inevitable and the mother's life was potentially in danger, but Kentucky doctors refused to terminate the pregnancy.¹¹⁴ In Kentucky, abortion is completely banned except for when necessary to save the mother's life.¹¹⁵ The patient was able to travel to Indiana, where doctors were able to "provide that pregnancy termination for her, save her uterus, and potentially save her life."¹¹⁶ Even though the patient's pregnancy could not continue, and her life was potentially in danger, Kentucky doctors "did not feel that they were legally able to [terminate the pregnancy]. So they sent her away."¹¹⁷

In Ohio, Tara George's ultrasound showed there was no amniotic fluid around the fetus, indicating that the fetus was in kidney failure and had multiple heart defects.¹¹⁸ Before Ohio's recent amendment to its constitution,¹¹⁹ it banned abortions after six weeks, except to prevent the death of the mother or the serious risk of substantial and irreversible impairment of a major bodily function.¹²⁰ If Tara carried the fetus to term, it would survive for no more than a few hours. Doing so would also put Tara's life at risk, since she had various medical conditions that put her "at high risk for hemorrhaging, clotting and preeclampsia—all potentially deadly complications."¹²¹ Tara's best

113. J. David Goodman & Azeen Ghorayashi, *Women Face Risks as Doctors Struggle With Medical Exceptions on Abortion*, N.Y. TIMES (July 20, 2022), <https://www.nytimes.com/2022/07/20/us/abortion-save-mothers-life.html>.

114. *Doctors Refusing Potentially Life-saving Abortion Treatment Over Legal Fears, Indiana Doctor Says*, ABC NEWS (Aug. 24, 2022), <https://www.radioalabama.net/news/national/doctors-refusing-potentially-life-saving-abortion-treatment-over-legal-fears-indiana-doctor-says>.

115. KY. REV. STAT. ANN. § 311.723 (West 2019).

116. *Doctors Refusing Potentially Life-saving Abortion Treatment Over Legal Fears*, *supra* note 114.

117. *Id.*

118. Elizabeth Cohen & Danielle Herman, *Ohio's New Abortion Law Forces Doctor to Fight to Protect Her Patient's Life*, CNN (Sept. 22, 2022), <https://www.cnn.com/2022/09/22/health/ohio-abortion-patient-doctor/index.html>.

119. Julie Carr Smyth, *Ohio Voters Just Passed Abortion Protections, When and How They Take Effect is Before the Courts*, AP NEWS (Nov. 24, 2023), <https://apnews.com/article/abortion-ohio-constitutional-amendment-republicans-courts-fb1762537585350caeee589d68fe5a0d>.

120. S.B. 23, 133rd Gen. Assem. (Ohio 2019).

121. Cohen & Herman, *supra* note 118.

option was to terminate the pregnancy, but Ohio hospital lawyers advised her doctor not to do so because there was uncertainty as to “how sick is sick enough.”¹²² Since doctors could lose their medical license, face fines, and be incarcerated for performing an illegal abortion, “doctors and hospitals are reluctant to get even close to violating it.”¹²³ Life-saving abortions are legal and desirable, but the risk of it being miscategorized as an illegal abortion deters doctors who are reasonably fearful of the criminal liability.

In addition to life-saving exceptions, some state laws allow abortions in cases of rape or incest. Although these abortions are legal, doctors must decide whether their patients’ claims are valid. Abortion clinics across these states have noted, “while the law may allow people to terminate their pregnancy in those instances, it will likely be easier to get patients across state lines for an abortion than try to clear the hurdles associated with obtaining one legally in their home state.”¹²⁴ One provider in Wyoming’s only clinic said, “I don’t want to go to jail. I don’t want to break the law, but I also can’t imagine a patient who has been raped or assaulted and is pregnant and calling for help and, as a gynecologist, to say to her, ‘Sorry, you’re on your own.’ It’s just horrific.”¹²⁵ The same experience has occurred in Texas, where some physicians with training in abortion procedures have been unable to offer even abortions allowed by SB8 because nurses and anesthesiologists, concerned about being seen as “aiding and abetting,” have declined to participate.¹²⁶

The better data surveillance is at capturing abortion, the more likely it is that doctors will be chilled from engaging in legal, desirable behavior. A pregnant person’s digital search for abortion-inducing medication, location data revealing presence at a reproductive health clinic, and information from a period tracking app can all be deployed in criminal proceedings. Since doctors are the primary target of these criminal laws, knowing that law enforcement has the capacity to track their patients’ locations, desires, and plans via their digital data will cause doctors to feel hyperaware that their decision-making process can be readily scrutinized.

122. *Id.*

123. *Id.*

124. Megan Messerly, *In States That Allow Abortion for Rape and Incest, Finding a Doctor May Prove Impossible*, POLITICO (June 27, 2022), <https://www.politico.com/news/2022/06/27/abortion-exceptions-doctor-shortage-00042373>.

125. *Id.*

126. Whitney Arey, Klaira Lerma, Anitra Beasley, Lorie Harper, Ghazaleh Moayedí & Kari White, *A Preview of the Dangerous Future of Abortion Bans—Texas Senate Bill 8*, 387 NEW ENG. J. MED. 388, 388–89 (2022).

B. DATA SURVEILLANCE WILL CHILL THE PROVISION OF LEGAL HEALTH CARE

Second, data surveillance for abortion ban enforcement will have a chilling effect on the provision of legal health care because many medications that treat a variety of non-abortion-related conditions have side-effects related to pregnancy. Rheumatoid arthritis patients use methotrexate, which can cause miscarriage or serious birth defects, for pain relief.¹²⁷ Mifepristone—the pill given for medication abortions—is also used to manage miscarriages, treat cancer, and control hyperglycemia in patients with Type 2 diabetes.¹²⁸ Isotretinoin treats severe acne, but causes severe birth defects.¹²⁹ Of course, treating arthritis, miscarriages, cancer, and skin conditions is completely legal and desirable activity. Nevertheless, increasing the certainty of criminal punishment for abortions makes providers more risk averse.

While no state laws impose restrictions on birth control, the prospect of criminal liability under abortion bans adds a new uncertainty. For example, in Louisiana, one doctor prescribed Cytotec to make IUD insertion less painful. Despite birth control being completely legal, a Walgreens pharmacy refused to fill the prescription because “they could not be sure [they] weren’t prescribing this for an abortion.”¹³⁰ At the University of Idaho, the school’s general counsel sent a memo to staff stating that employees cannot “dispens[e] drugs classified as emergency contraception by the FDA, except in the case of rape.”¹³¹ Even though contraceptives remain legal in Idaho—and protected under the Constitution—the university intended the memo to “help

127. Maria Angeles Lopez-Olivo, Harish R. Siddhanamatha, Beverley Shea, Peter Tugwell, George A. Wells & Maria E. Suarez-Almazor, *Methotrexate for Treating Rheumatoid Arthritis*, COCHRANE DATABASE SYS. REV., no. 6, 2014, <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7047041/>.

128. Margaret Beal & Kathy Simmonds, *Clinical Uses of Mifepristone: An Update for Women’s Health Practitioners*, 47 J. MIDWIFERY & WOMEN’S HEALTH 451 (2014), <https://pubmed.ncbi.nlm.nih.gov/12484667/>.

129. June Seek Choi, Gideon Koren & Irena Nulman, *Pregnancy and Isotretinoin Therapy*, 185 CANADIAN MED. ASS’N J. 411 (2013), <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3602257/>.

130. Emily Woodruff, *As Abortion Ban Is Reinstated, Doctors Describe ‘Chilling Effect’ on Women’s Care*, NOLA (July 10, 2022), https://www.nola.com/news/healthcare_hospitals/article_238af184-ff02-11ec-9bce-dfd660a21ce1.html.

131. Kelcie Moseley-Morris, *University of Idaho Releases Memo Warning Employees That Promoting Abortion Is Against State Law*, IDAHO CAP. SUN (Sept. 26, 2022), <https://idahocapitalsun.com/2022/09/26/university-of-idaho-releases-memo-warning-employees-that-promoting-abortion-is-against-state-law/>.

employees understand the legal significance and possible ramifications of the law, which includes individual criminal prosecution.”¹³²

Even in circumstances further removed from the pregnancy context, patients have experienced the spillover effects of abortion criminalization. In Tennessee, where it is illegal to have an abortion after six weeks of pregnancy, Becky Hubbard “decided to get sterilized so that she can go back on the only medication that has relieved her disabling pain from rheumatoid arthritis for the last eight years.”¹³³ Her Tennessee doctor gave her an ultimatum: “if she wanted to stay on . . . methotrexate, she was told she had to go on birth control despite her age and history of infertility.”¹³⁴ Because methotrexate can also end a pregnancy, doctors and pharmacists could be held criminally liable for prescribing to pregnant people.¹³⁵ Increasingly, pharmacies are changing policies to require diagnosis codes to ensure the prescription will not be used to end a pregnancy.¹³⁶ One rheumatologist described how dangerous this can be: “It becomes a huge problem if we see [a] patient on Thursday or Friday and we don’t get the pharmacy to call back The patient can’t get treatment for three or four days, which can be agonizing.”¹³⁷

Treatment for miscarriages post-*Dobbs* may be especially controversial since patients with miscarriage complications are often given the same medication that is used for abortions. In Washington D.C., which has among the least restrictive abortion laws in the country, Christina Zielke’s ultrasound showed her fetus had no heartbeat.¹³⁸ Her doctors confirmed that she miscarried and told her the pregnancy tissue would eventually come out on its own.¹³⁹ Soon after, due to miscarriage complications she experienced excessive, life-threatening bleeding.¹⁴⁰ At the time, she happened to be on a trip in Ohio, where abortion was banned after six weeks of pregnancy except

132. Kelcie Moseley-Morris, *White House Calls Idaho Abortion Laws ‘Extreme and Backwards’ in Response to University Memo*, IDAHO CAP. SUN (Sept. 27, 2022), <https://idahocapitalsun.com/2022/09/27/white-house-calls-idaho-abortion-laws-extreme-and-backwards-in-response-to-university-memo/>.

133. Katie Shepherd & Frances Stead Sellers, *Abortion Bans Complicate Access to Drugs for Cancer, Arthritis, Even Ulcers*, WASH. POST (Aug. 8, 2022), <https://www.washingtonpost.com/health/2022/08/08/abortion-bans-methotrexate-mifepristone-rheumatoid-arthritis/>.

134. *Id.*

135. *Id.*

136. *Id.*

137. *Id.*

138. Selena Simmons-Duffin, *Her Miscarriage Left Her Bleeding Profusely. An Ohio ER Sent Her Home to Wait*, NPR (Nov. 15, 2022), <https://www.npr.org/sections/health-shots/2022/11/15/1135882310/miscarriage-hemorrhage-abortion-law-ohio>.

139. *Id.*

140. *Id.*

where there is a medical emergency.¹⁴¹ She was bleeding profusely for hours, but Ohio doctors discharged her, saying “they needed to prove there was no fetal development.”¹⁴² Despite D.C. doctors already having diagnosed a miscarriage, Ohio doctors told her “the pregnancy could still be viable.”¹⁴³ To ensure compliance with the state’s abortion ban and avoid liability, Ohio doctors delayed treatment and may have endangered a patient’s life.

Doctors delaying treatments and turning patients away is reminiscent of their behaviors before *Roe*, when they prioritized securing dying declarations from patients that would clear them of liability.¹⁴⁴ The difference now is that doctors face the added pressure of knowing every patient they see is being digitally surveilled. Doctors know that the chances of getting caught, even wrongfully, are high.

C. DATA SURVEILLANCE WILL CHILL LEGAL INFORMATION SHARING

Perhaps the most devastating chilling effect will be overdeterrence of legally seeking, sharing, and accessing information. There is evidence that censorship of abortion-related speech is already occurring, and data surveillance only exacerbates the issue.

At the University of Idaho, the same memo that cautioned staff against giving emergency contraceptives also directed staff to “avoid language that could be seen as counseling in favor of, referring for, or promoting abortion.”¹⁴⁵ The memo was in response to Idaho’s No Public Funds for Abortion Act. Since the university is public, its legal team “highly recommend[ed] employees in charge of the classroom remain neutral or risk violating this law.” Even though abortion-related speech may be protected by the First Amendment,¹⁴⁶ professors are erring on the side of caution. One faculty member said the guidance could “cause individual faculty members, frankly, particularly those who don’t have job protection like tenure, to be very, very careful. To refrain from saying things they might otherwise say[.]”¹⁴⁷

141. *Id.*

142. *Id.*

143. *Id.*

144. *See supra* Section III.A.

145. Rachel Sun, *UI Employees Say Memo on Abortion, Contraception Creating Chilling Effect in Classroom*, NW PUB. BROAD. (Oct. 3, 2022), <https://www.nwpb.org/2022/10/03/ui-employees-say-memo-on-abortion-contraception-creating-chilling-effect-in-classroom/>.

146. Jeremy W. Peters, *First Amendment Confrontation May Loom in Post-Roe Fight*, N.Y. TIMES (June 30, 2022), <https://www.nytimes.com/2022/06/29/business/media/first-amendment-roe-abortion-rights.html> (presenting commentary that people have “the right, ostensibly, to talk about abortion”).

147. Sun, *supra* note 145.

When *The New York Times* asked to interview a Texas doctor about patients' experiences with abortion, her hospital's public relations office asked the doctor to decline to comment. The doctor told CNN, "They're censoring me."¹⁴⁸ The doctor was not allowed to tell media where she works and could not communicate with journalists on her work email or using her work computer. At a different hospital, residents who posted an Instagram photo stating "Abortion is healthcare" were forced by university lawyers to take it down.¹⁴⁹ Perfectly legal communication about abortion—especially when housed online where law enforcement has unbridled access to it—poses too high of a risk for hospitals who fear liability.

Even though learning about abortion is completely legal, medical students and residency programs in restrictive states are discontinuing abortion training. Pamela Merritt, the executive director of Medical Students for Choice, said some medical schools are "so risk averse, they're shutting down all access. They're in a political pickle."¹⁵⁰ OB-GYN residency programs, which are required to provide clinical abortion experience, are facing difficulties sending residents out-of-state to get trained.¹⁵¹ Since clinical capacity is limited, out-of-state programs cannot accommodate every program in an abortion-restrictive state.¹⁵²

In addition to providers being deterred from legally sharing abortion-related information, pregnant people will also be deterred from seeking information to learn their options. Moments after *Dobbs* came down, Instagram and Facebook removed posts that offered women information about how to obtain abortion pills.¹⁵³ Nikolas Guggenberger, the executive director at the Yale Information Society Project, said that "[j]ust the possibility of using phone surveillance to enforce abortion bans will hang over the heads of people seeking abortions or helping others get them."¹⁵⁴ Following *Dobbs*,

148. Elizabeth Cohen, Justin Lape & Danielle Herman, 'Heartbreaking' Stories Go Untold, Doctors Say, As Employers 'Muzzle' Them in Wake of Abortion Ruling, CNN (Oct. 12, 2022), <https://www.cnn.com/2022/10/12/health/abortion-doctors-talking/index.html>.

149. *Id.*

150. Olivia Goldhill, *After Dobbs, U.S. Medical Students Head Abroad for Abortion Training No Longer Provided by Their Schools*, STAT (Oct. 18, 2022), <https://www.statnews.com/2022/10/18/medical-students-heading-abroad-for-abortion-training/>.

151. *Id.*

152. *Id.*

153. *Instagram and Facebook Begin Removing Posts Offering Abortion Pills*, NPR (June 28, 2022), <https://www.npr.org/2022/06/28/1108107718/instagram-and-facebook-begin-removing-posts-offering-abortion-pills>.

154. Geoffrey A. Fowler & Tatum Hunter, *For People Seeking Abortions, Digital Privacy is Suddenly Critical*, WASH. POST (June 24, 2022), <https://www.washingtonpost.com/technology/2022/05/04/abortion-digital-privacy/>.

hundreds of online posts urged women to delete their period tracking apps.¹⁵⁵ Even if the abortion ban does not apply to their activities, people are nonetheless deterred because they fear the mere possibility of being surveilled.

Maximal enforcement by way of maximal surveillance will result in chilling effects on legal and desirable activities such as performing a life-saving abortion, promptly treating miscarriage complications, and discussing information online. People are afraid of being placed in a situation of potential criminal liability. Their fear is reasonable—with all the possibilities data surveillance has to offer, the certainty of punishment can be extremely high.

When the enforcement mechanism of a criminal law requires us to give up digital privacy, should the law be enforced that way? How much of our legal, desirable activity are we willing to sacrifice for the enforcement of crimes? The central tension here involves the tools for administrability in one field—data surveillance in criminal law—directly threatening the values in another—control over one’s information in privacy law. I argue that chilling legal abortions, legal non-abortion healthcare, and legal information sharing is too great an externality. Data surveillance must be curtailed even if that means capturing less effective enforcement of abortion bans.

V. POSSIBLE SOLUTIONS

Using data surveillance to enforce abortion bans creates too high of a privacy cost. The question becomes, who is responsible for protecting individuals’ privacy? Some look to Big Tech, whose business practices create the troves of data that law enforcement exploits. But others point out that tech companies’ data practices are perfectly legal, and instead argue that it is the federal government’s responsibility to protect data privacy.

While tech companies do have the capability to alleviate abortion-related privacy concerns, it would be naïve to rely on their goodwill. Federal privacy legislation is necessary, but largely ineffective if it continues to allow exceptions for law enforcement’s requests. Thus, I conclude that the solution is to limit law enforcement’s ability to request sensitive data from Big Tech companies.

155. See, e.g., Gennie Gebhart & Daly Barnett, *Should You Really Delete Your Period Tracking App?* ELEC. FRONTIER FOUND. (June 30, 2022), <https://www.eff.org/deeplinks/2022/06/should-you-really-delete-your-period-tracking-app>; @ECMcLaughlin, X (May 3, 2022, 10:36 AM), <https://web.archive.org/web/20220504013052/https://twitter.com/ECMcLaughlin/status/1521467912162226176> (“If you are using an online period tracker or tracking your cycles through your phone, get off it and delete your data. Now.”).

A. TECH COMPANIES TO THE RESCUE?

Post-*Dobbs*, tech companies have faced pressure to respond to growing concerns about data privacy.¹⁵⁶ Privacy experts and the general public have called on Big Tech to help women seeking abortions and have suggested a variety of rationales as to why they should do so. Privacy advocates have urged tech firms to provide better encryption, delete abortion-related data on users, and educate users about their data privacy.¹⁵⁷ Since tech companies hold what will be the critical evidence in abortion ban enforcement, many rightfully believe that the onus is on tech companies to stop collecting and storing this sensitive data in the first place.

Many tech companies have entered the dialogue by supporting their own employees who receive abortions, but are quieter when it comes to their data privacy practices. For example, an Apple spokesperson stated that, “[Apple] supports employees’ right to make their own decision regarding their reproductive health. For more than a decade, Apple’s comprehensive benefits have allowed our employees to travel out-of-state for medical care if it is unavailable in their home state.”¹⁵⁸ Microsoft released a statement saying it “will provide travel expense reimbursement for employees seeking abortions and gender-affirming care anywhere in the country.”¹⁵⁹ Amazon added a \$4,000 employee benefit to cover out-of-state travel for reproductive healthcare or other medical issues.¹⁶⁰ Lyft’s statement explicitly mentioned *Dobbs*: “In the wake of the Supreme Court decision on *Dobbs v. Jackson Women’s Health Organization*, we’re committed to providing team members with uninterrupted access to safe and critical healthcare services.”¹⁶¹ A Meta spokesperson told ABC News that the company “plans to offer coverage of

156. Kimberly Adams & Jesus Alvarado, *With Roe Overturned, Tech Companies Will Have to Weigh Big Data Questions*, MARKETPLACE TECH (June 27, 2022), <https://www.marketplace.org/shows/marketplace-tech/with-roe-overturned-tech-companies-will-have-to-weigh-big-data-questions/>.

157. Aziz Huq & Rebecca Wexler, *Big Tech Can Help Women in a Post-Roe World. Will it?*, WASH. POST (June 1, 2022), <https://www.washingtonpost.com/outlook/2022/06/01/roe-dobbs-big-tech/>.

158. *Companies Respond to Abortion Ruling That Overturns Roe v. Wade*, B.C. CTR. FOR CORP. CITIZENSHIP (June 30, 2022), <https://ccc.bc.edu/content/ccc/blog-home/2022/06/companies-respond-to-abortion-ruling.html>.

159. *Id.*

160. *Id.*

161. *Id.*

travel expenses for some employees seeking an abortion.”¹⁶² A Google memo told employees they may relocate from states banning abortion.¹⁶³

While Big Tech companies have shown a commitment to employees’ reproductive health, their commitment to users’ reproductive health remains largely opaque.¹⁶⁴ Many companies that released statements regarding new employee policies have declined to respond to media inquiries into their post-*Dobbs* policies and requests for data from law enforcement.¹⁶⁵ Huq and Wexler note that the distinction between users and employees is ultimately untenable because employees are also users whose privacy is compromised.¹⁶⁶

There have been some exceptions to the general silence about abortion-related data privacy. Most notably, Google released a statement in July 2022 vowing to delete location history data from abortion clinics:

Some of the places people visit—including medical facilities like counseling centers, domestic violence shelters, abortion clinics, fertility centers, addiction treatment facilities, weight loss clinics, cosmetic surgery clinics, and others—can be particularly personal. Today, we’re announcing that if our systems identify that someone has visited one of these places, we will delete these entries from Location History soon after they visit. This change will take effect in the coming weeks.¹⁶⁷

Google’s decision came after Alphabet Workers Union, a minority labor union, demanded that Google delete any personal data that law enforcement could use to prosecute people who receive abortions.¹⁶⁸ The announcement did not make any commitments as to how Google will handle data requests from law enforcement, nor did it commit to automatically deleting search records about abortions. Instead, “[u]sers must individually opt to delete their search history.”¹⁶⁹

162. *Id.*

163. Jennifer Elias, *Google Memo on End of Roe v. Wade Says Employees May Apply to Relocate Without Justification*, CNBC (June 27, 2022), <https://www.cnbc.com/2022/06/24/google-memo-to-employees-on-roe-v-wade-overturn.html>.

164. Huq & Wexler, *supra* note 53, at 590–91.

165. *Id.*

166. *Id.* at 592.

167. Jen Fitzpatrick, *Protecting People’s Privacy on Health Topics*, GOOGLE: KEYWORD (July 1, 2022), <https://blog.google/technology/safety-security/protecting-peoples-privacy-on-health-topics/>.

168. Nico Grant, *Google Says It Will Delete Location Data When Users Visit Abortion Clinics*, N.Y. TIMES (July 1, 2022), <https://www.nytimes.com/2022/07/01/technology/google-abortion-location-data.html>.

169. *Id.*

Some privacy experts believe that the onus is on tech companies to stop collecting and storing this sensitive data in the first place.¹⁷⁰ However, while it is true that tech companies have the primary power to stop collecting or distributing sensitive data, I argue that we cannot rely on Big Tech to protect abortion access. First, evidence of tech companies' broken privacy promises diminishes confidence that they will live up to their policies. Second, tech companies often place the responsibility on the user to opt out of sensitive data collection, making it unlikely that unsophisticated users will do so. Finally, tech companies whose primary revenue comes from data collection cannot be left to self-regulate.

1. *Evidence of Broken Privacy Promises*

In 2021, the aforementioned period and ovulation tracker Flo, shared users' sensitive fertility data with third parties, in violation of its express privacy claims. Flo's privacy policy misleadingly represented that third parties could not use consumers' personal information "for any other purpose except to provide services in connection with the App."¹⁷¹ However, for five years the app included tools from a variety of third-party marketing and analytics firms that gathered records of users' interactions on the app.¹⁷² When a user entered pregnancy-related information on the app, third parties received analytics records with the word "pregnancy" attached.¹⁷³ Flo settled with the FTC over the allegations.¹⁷⁴ Flo agreed to notify users about how their data was shared and receive an audit of its privacy practices, but did not admit any wrongdoing.¹⁷⁵

In May 2022, Twitter was fined \$150 million for allegedly breaking its privacy promises. It asked users to provide their contact information to "safeguard your account," but it failed to mention that it was also used to deliver targeted ads.¹⁷⁶ In November 2022, Apple, who has a reputation for

170. Jordan Famularo & Richmond Wong, *How the Tech Sector Can Protect Personal Data Post-Roe*, BROOKINGS INST. (Oct. 27, 2022), <https://www.brookings.edu/techstream/how-tech-firms-can-protect-personal-data-after-roe-us-privacy-abortion-surveillance/>.

171. Lesley Fair, *Health App Broke Its Privacy Promises by Disclosing Intimate Details About Users*, FED. TRADE COMMISSION (Jan. 13, 2021), <https://www.ftc.gov/business-guidance/blog/2021/01/health-app-broke-its-privacy-promises-disclosing-intimate-details-about-users>.

172. *Id.*

173. *Id.*

174. *Id.*

175. *Id.*

176. Lesley Fair, *Twitter to Pay \$150 Million Penalty for Allegedly Breaking Its Privacy Promises—Again*, FED. TRADE COMMISSION (May 25, 2022), <https://www.ftc.gov/business-guidance/blog/2022/05131/twitter-pay-150-million-penalty-allegedly-breaking-its-privacy-promises-again>.

strong consumer privacy protections, was sued in a class action over tracking of users' activity in violation of the California Invasion of Privacy Act.¹⁷⁷

2. *Placing the Responsibility on Users*

The typical privacy framework for digital data processing in the United States is a “strict opt-out” option, allowing consumers to request that the company does not sell or share their personal information.¹⁷⁸ Tech companies place the burden on consumers to “exercise their rights and take action to prevent an organization from processing their data.”¹⁷⁹ The opposite approach, an “opt-in” system, requires the company to affirmatively obtain consumer consent, rather than assuming it exists to begin with.¹⁸⁰ Notably, opt-in systems are far less common in the United States.¹⁸¹

In the current privacy framework, tech companies can “shift[] the work onto the user to figure out how to delete their data.”¹⁸² Unfortunately, just like users likely do not read terms and conditions policies, they do not typically change default data collection settings.¹⁸³ Shoshana Zuboff, a surveillance capitalism scholar, describes the power asymmetry under this framework: “Take a minute and just feel how intolerable it is for us to essentially be supplicants toward a massively wealthy, massively powerful data company, saying, ‘Please, please, please stop collecting sensitive data.’”¹⁸⁴

3. *Clear Conflict of Interest*

It is unrealistic to rely on tech companies to safeguard privacy to the necessary extent because minimizing data collection is contrary to their profit models. To ask Big Tech to solve a problem it created is to ask it to dismantle surveillance capitalism and its economic imperatives. Google is a \$150 billion

177. Sarah Perez, *Apple Faces New Lawsuit Over Its Data Collection Practices in First-Party Apps, Like the App Store*, TECHCRUNCH (Nov. 14, 2022), <https://techcrunch.com/2022/11/14/apple-faces-new-lawsuit-over-its-data-collection-practices-in-first-party-apps-like-the-app-store/>.

178. Sarah Rippey, *Opt-in vs. Opt-out Approaches to Personal Information Processing*, INT'L ASS'N PRIVACY PROFESSIONALS (May 10, 2021), <https://iapp.org/news/a/opt-in-vs-opt-out-approaches-to-personal-information-processing/>.

179. *Id.*

180. *Id.*

181. *Id.*

182. Geoffrey A. Fowler, *Okay, Google: To Protect Women, Collect Less Data About Everyone*, WASH. POST (July 1, 2022), <https://www.washingtonpost.com/technology/2022/07/01/google-privacy-abortion/>.

183. Editorial Board, *America, Your Privacy Settings Are All Wrong*, N.Y. TIMES (Mar. 6, 2021), <https://www.nytimes.com/2021/03/06/opinion/data-tech-privacy-opt-in.html>.

184. Casey Newton, *Why Abortion is Tech's Next Big Reputational Risk*, KAIROS FELLOWSHIP (July 13, 2022), <https://www.kairosfellows.org/news/tag/Data+Privacy>.

advertising business. It was the first to create “lucrative markets to trade in human futures, what we now know as online targeted advertising, based on their predictions of which ads users would click.”¹⁸⁵ It relies on access to users’ data to develop its services and products. Sundar Pichai, Google’s chief executive officer, wrote an editorial in *The New York Times* titled “Privacy Should Not Be A Luxury Good.”¹⁸⁶ Just months later, the *Daily News* reported that unhoused people were lined up to get a \$5 gift card in exchange for uploading their face scan to Google.¹⁸⁷ Facebook has acted similarly. In 2019, Mark Zuckerberg announced at a conference that “the future is private.”¹⁸⁸ Just weeks later, a lawyer for Facebook argued in a user privacy case that the “very act of using Facebook negates any reasonable expectation of privacy as a matter of law.”¹⁸⁹ Rather than relying on Big Tech’s goodwill, we need strong federal privacy legislation.

B. FEDERAL PRIVACY LEGISLATION

Post-*Dobbs*, the case for federal privacy legislation is stronger than ever. As it currently stands, there are two abortion-specific data privacy bills that have recently been introduced, the My Body, My Data Act and the Health and Location Data Protection Act.

1. Overview of Proposed Federal Legislation

In June 2022, Representative Sara Jacobs introduced the My Body, My Data Act in the House. The proposed bill establishes that “commercial entities, including individuals, nonprofits, and common carriers, may not collect, retain, use, or disclose personal reproductive or sexual health information except (1) with the express written consent of the individual to whom such information relates, or (2) as is strictly necessary to provide a requested product or service.”¹⁹⁰ The Act would also give users the right to access or delete their personal data by requiring commercial entities to “provide individuals with access to, and a reasonable mechanism to delete, any of their reproductive or

185. Shoshana Zuboff, *You Are Now Remotely Controlled*, N.Y. TIMES (Jan. 24, 2020), <https://www.nytimes.com/2020/01/24/opinion/sunday/surveillance-capitalism.html>.

186. Sundar Pichai, *Google’s Sundar Pichai: Privacy Should Not Be a Luxury Good*, N.Y. TIMES (May 7, 2019), <https://www.nytimes.com/2019/05/07/opinion/google-sundar-pichai-privacy.html>.

187. Ginger Adams Otis & Nancy Dillon, *City Worker Saw Homeless People Lined Up to Get \$5 Gift Card for Face Scan Uploaded to Google*, N.Y. DAILY NEWS (Jan. 31, 2020), <https://www.nydailynews.com/news/national/ny-witness-saw-homeless-people-selling-face-scans-google-five-dollars-20191004-j6z2vonllnerpiuakt6wrp6l44-story.html>.

188. Zuboff, *supra* note 185.

189. *Id.*

190. H.R. 8111, 116th Cong. (2022).

sexual health information upon request.”¹⁹¹ The bill is endorsed by Planned Parenthood, NARAL Pro-Choice America, National Abortion Federation, United for Reproductive & Gender Equity, National Partnership for Women & Families, Feminist Majority, and the Electronic Frontier Foundation.¹⁹²

My Body, My Data is a step in the right direction to limit health-related data collection, but it likely does not do enough to prevent or mitigate law enforcement’s access to and use of abortion-related data. Representative Jacobs recognized that “it’s unconscionable that information could be turned over to the government or sold to the highest bidder and weaponized against us.”¹⁹³ However, scholars pointed out that the Act “does not block, or indeed even mention, warrants, subpoenas, or other court orders.”¹⁹⁴ Based on the bill’s language, only collection of voluntarily shared data would be disallowed.¹⁹⁵ While limiting data collection in any way possible is a positive step, Representative Jacobs’ bill likely does not do enough to prevent abortion criminalization via data surveillance.

Additionally, Senators Warren, Wyden, Murray, Whitehouse, and Sanders introduced the Health and Location Data Protection Act in June 2022.¹⁹⁶ The proposed bill bans data brokers from selling or transferring health and location data, but makes exceptions for HIPAA-compliant activities, protected First Amendment speech, and validly authorized disclosures.¹⁹⁷ Again, the bill falls short in specifically addressing how law enforcement can obtain abortion related data to surveil potentially pregnant people.

The Fourth Amendment Is Not For Sale Act, although not specifically about sensitive health data, does specifically address law enforcement’s ability to obtain data. The bipartisan Act, introduced in 2021 by Senators Wyden, Paul, and eighteen other senators, seeks to “close the legal loophole that allows

191. *Id.*

192. Hayley Tsukayama & India McKinney, *Pass the “My Body, My Data” Act*, ELEC. FRONTIER FOUND. (June 21, 2022), <https://www.eff.org/deeplinks/2022/06/pass-my-body-my-data-act>.

193. SARA JACOBS, MY BODY, MY DATA ACT OF 2022, <https://sarajacobs.house.gov/uploadedfiles/mybodymydataactonepager.pdf> (last visited Nov. 24, 2023).

194. Huq & Wexler, *supra* note 53, at 634–35. Notably, Huq and Wexler are the first to propose creating an evidentiary privilege for abortion-relevant data. While I endorse this as an *ex-post* solution, *ex-ante* legislation is also necessary.

195. *Id.*

196. S. 4408, 117th Cong. (2022).

197. *Warren, Wyden, Murray, Whitehouse, Sanders Introduce Legislation to Ban Data Brokers from Selling Americans’ Location and Health Data*, ELIZABETH WARREN (June 15, 2022), <https://www.warren.senate.gov/newsroom/press-releases/warren-wyden-murray-whitehouse-sanders-introduce-legislation-to-ban-data-brokers-from-selling-americans-location-and-health-data>.

data brokers to sell Americans' personal information to law enforcement and intelligence agencies without any court oversight."¹⁹⁸ While this would prevent the government from getting around the Fourth Amendment by simply paying for the data, police are still allowed to get a court order to compel that data.¹⁹⁹ This solution does not go far enough in protecting privacy, especially considering the ease with which warrants for health-related data can be obtained.

Ultimately, federal privacy legislation has much work to do. On the tech companies' side, legislation like My Body, My Data is needed to limit the information companies are allowed to collect and use. Doing so will at least limit the voluntary information collected, even if it still requires companies to disclose data to law enforcement. Data brokers selling sensitive data to law enforcement is perhaps the most obviously problematic—the Fourth Amendment Is Not For Sale Act can help reduce the amount of data law enforcement receives that is completely unregulated. Finally, even when law enforcement does have a warrant, there is a question of whether the warrant should have been granted in the first place. For data as sensitive as health information, it may be appropriate to outlaw reverse-search warrants entirely.²⁰⁰

These privacy reforms go beyond opinions on abortion constitutionality. Across party lines, Americans support federal data privacy legislation.²⁰¹ Even Republican Senator Josh Hawley, who openly rejects a constitutional right to abortions, considers data surveillance “a separate question altogether.”²⁰² Regardless of whether abortion is a crime, there should be rights to data privacy that apply even if it makes things harder for prosecutors.

198. Wyden, *Paul and Bipartisan Members of Congress Introduce The Fourth Amendment Is Not For Sale Act*, RON WYDEN (Apr. 21, 2021), <https://www.wyden.senate.gov/news/press-releases/wyden-paul-and-bipartisan-members-of-congress-introduce-the-fourth-amendment-is-not-for-sale-act->.

199. Huq & Wexler, *supra* note 53, at 635 n.355 (“the Act provides no protection against warrants or indeed against any other form of legal process applied to the majority of abortion-relevant data that does not fall within existing Fourth Amendment doctrine.”).

200. Indeed, California introduced a bill to prohibit any government entity from seeking a reverse-keyword or reverse-location demand. *See* A.B. 793, 2023–2024 Reg. Sess. (Cal. 2023).

201. Chris Teale, *More Than Half of Voters Back a National Data Privacy Law*, MORNING CONSULT (Jan. 12, 2022), <https://morningconsult.com/2022/01/12/federal-data-privacy-legislation-polling/>.

202. Matt Laslo, *The Shaky Future of a Post-Roe Federal Privacy Law*, WIRED (Sept. 15, 2022), <https://www.wired.com/story/adppa-roe-democrats-congress/>.

VI. CONCLUSION

Enforcement of abortion bans post-*Dobbs* will look vastly different than they did pre-*Roe*. *Dobbs* must be considered against a backdrop of unprecedented technological advances in data surveillance that have developed since *Roe*. Modern technology allows law enforcement to achieve increasingly expansive enforcement of abortion laws. Digital data contains an enormous amount of information much about users. Search history data, location data, and even data specific to reproductive health provide a mechanism to achieve maximal enforcement of abortion laws. Our thoughts, movements, habits, and preferences are constantly tracked and sold to third parties, including law enforcement. But giving up this privacy is too high a cost. Even if it means letting some criminal abortion activity go undetected, choosing less invasive enforcement mechanisms is worth avoiding the chilling effects on legal activity.

CLEARLY REPUGNANT: CORRECTING THE COURT’S FAILED APPROACH TO ANTITRUST ENFORCEMENT

M. A. Katz[†]

TABLE OF CONTENTS

I.	INTRODUCTION	1373
II.	ANTITRUST AND REGULATION AT ODDS	1377
	A. THE CLEAR REPUGNANCY DOCTRINE.....	1377
	B. THE DISTORTED FRAMEWORK OF IMPLIED IMMUNITY	1381
	1. Trinko.....	1382
	2. Credit Suisse.....	1386
III.	PARANOIA, UNDERENFORCEMENT, AND THE PSN MARKET	1389
	A. THE CONSEQUENCES OF UNDERENFORCEMENT.....	1390
	B. A CASE STUDY: <i>FTC V. META PLATFORMS, INC.</i>	1394
IV.	AN OPPORTUNITY FOR RECONCILIATION	1401
V.	CONCLUSION	1403

Where . . . ‘[t]here is nothing built into the regulatory scheme which performs the antitrust function,’ . . . the benefits of antitrust are worth its sometimes considerable disadvantages. Just as regulatory context may in other cases serve as a basis for implied immunity, . . . it may also be a consideration in deciding whether to recognize an expansion of the contours of § 2.¹

I. INTRODUCTION

For the past fifty years, regulation and antitrust have maintained a dysfunctional relationship in the United States. Although they effectively

DOI: <https://doi.org/10.15779/Z38GH9BB13>

© 2023 M. A. Katz.

[†] J.D. Candidate, University of California, Berkeley, School of Law, Class of 2024. Sincere thanks to Professor Talha Syed, my fellow students in the 2022 Law & Technology Writing Workshop, and the Berkeley Technology Law Journal editors.

1. *Commc’ns, Inc. v. Law Offices of Curtis V. Trinko, LLP*, 540 U.S. 398, 412 (2004) (citing *Silver v. N.Y. Stock Exch.*, 373 U.S. 341, 358 (1963)).

operated in tandem throughout the 1950s and 60s,² the Supreme Court has increasingly aligned itself with big business interests³ and engaged in overt antitrust antitextualism⁴ since the 1970s. This has resulted in a simultaneous reduction in both antitrust enforcement⁵ and regulatory power.⁶ The disastrous effects of this laissez-faire model in the United States are apparent from the unprecedented consolidation of market power across sectors⁷ and its accompanying effects on both consumers⁸ and labor conditions.⁹ The effects of this dysfunction are uniquely obvious in the realm of “Big Tech”—a moniker often applied to companies like Google, Microsoft, Amazon, and Meta. In an increasingly digital economy, a handful of companies wield an outsized influence over our daily lives and “[t]here is bipartisan agreement that

2. While antitrust enforcement receded somewhat in the wake of New Deal reforms and through World War II, a wave of market consolidation in the 1950s prompted amendments to the Clayton Act and renewed investment in antitrust enforcement. *See* Gene M. Gressley, *Thurman Arnold, Antitrust, and the New Deal*, 38 BUS. HIST. REV. 214, 227 (1964) (describing U.S. antitrust enforcement in the wake of the National Recovery Act and the beginnings of World War II); Debra A. Valentine, *The Evolution of U.S. Merger Law: Prepared Remarks before INDECOPI Conference*, FED. TRADE COMM. (Aug. 13, 1996), <https://www.ftc.gov/news-events/news/speeches/evolution-us-merger-law> (“A dominant theme driving the 1950 amendments was a fear of what was considered to be a rising tide of economic concentration in the American economy. In 1909, the 200 largest non-banking corporations owned about one-third of all corporate assets; in 1928 they owned 48%; in the early thirties they owned 54%; by 1940 they held 55%.”).

3. *See generally* Filippo Lancieri, Eric A. Posner & Luigi Zingales, *The Political Economy of the Decline of Antitrust Enforcement in the United States*, NAT’L BUREAU OF ECON. RESEARCH, WORKING PAPER NO. 30326 (2022) (empirically connecting the decline in U.S. antitrust enforcement with the advancement of big business interests).

4. *See generally* Daniel A. Crane, *Antitrust Antitextualism*, 96 NOTRE DAME L. REV. 1205 (2021).

5. *See, e.g.*, Jon Dubrow, Noah Feldman Greene & Gregory Heltzer, *DOJ to Merging Parties: The Time of “Underenforcement” is Over; Fix-It-First or Risk Being Challenged*, JDSUPRA (Sept. 21, 2022), <https://www.jdsupra.com/legalnews/doj-to-merging-parties-the-time-of-8445648/>.

6. *See, e.g.*, Charlie Savage, *E.P.A. Ruling Is Milestone in Long Pushback to Regulation of Business*, N.Y. TIMES (June 30, 2022), <https://www.nytimes.com/2022/06/30/us/supreme-court-epa-administrative-state.html>.

7. REBECCA GIBLIN & CORY DOCTOROW, CHOKEPOINT CAPITALISM: HOW BIG TECH AND BIG CONTENT CAPTURED CREATIVE LABOR MARKETS AND HOW WE’LL WIN THEM BACK 4–5 (2022).

8. *See, e.g.*, Janet Nguyen, *Money and Millennials: The Cost of Living in 2022 vs. 1972*, MARKETPLACE (Aug. 17, 2022), <https://www.marketplace.org/2022/08/17/money-and-millennials-the-cost-of-living-in-2022-vs-1972/>.

9. *See, e.g.*, Greg Iacurci, *U.S. Is Worst Among Developed Nations for Worker Benefits*, CNBC (Feb. 4, 2021), <https://www.cnn.com/2021/02/04/us-is-worst-among-rich-nations-for-worker-benefits.html>.

the status quo is just not working.”¹⁰ However, there is little agreement over how and where to begin repairing antitrust enforcement’s role in regulation.

This Note adopts the argument that the appropriate relationship between antitrust and regulation is neither adversarial nor cyclical, but symbiotic.¹¹ Using the Federal Trade Commission’s (FTC) antitrust suit against Meta over its WhatsApp and Instagram acquisitions as a case study,¹² this Note then goes on to argue that the market of Personal Social Networks (PSNs) is the perfect place to begin restoring the balance between antitrust and regulation. PSNs are uniquely underregulated because they arose and grew in the midst of regulatory and antitrust decay—companies like Meta began in the early 2000s and exploded in the mobile device era of internet access.¹³ As a result of this unchecked growth, PSNs must first be broken down to a manageable size by antitrust enforcement before regulation can be crafted to effectively protect consumers from harms like hate speech and privacy invasions.

To make this argument, this Note first addresses the history of the Sherman Antitrust Act¹⁴ and of the Supreme Court’s anti-textualist approach to its interpretation since the 1970s. To do so, this Note focuses on the Court’s inversion of the “clear repugnancy” doctrine¹⁵ into a doctrine of implied antitrust immunity—while the Court once staunchly maintained that antitrust should rarely be precluded by the existence of regulation, that standard has

10. Shannon Bond, *Facebook, Twitter, Google CEOs Testify Before Congress: 4 Things to Know*, NPR (Mar. 25, 2021), <https://www.npr.org/2021/03/25/980510388/facebook-twitter-google-ceos-testify-before-congress-4-things-to-know>.

11. See Robert A. Jablon, Anjali G. Patel & Latif M. Nurani, *Trinko and Credit Suisse Revisited: The Need for Effective Administrative Agency Review and Shared Antitrust Responsibility*, 34 ENERGY L.J. 627, 627 (2013) (“[A]gencies must continue to have significant antitrust roles but . . . judicial antitrust enforcement must also be fully available.”).

12. Fed. Trade Comm’n v. Facebook, Inc., 581 F. Supp. 3d 34 (D.D.C. 2022).

13. Substitute Amended Complaint for Injunctive and Other Equitable Relief at 3, Fed. Trade Comm’n v. Meta Platforms, Inc., No. 1:20-cv-03590 (D.D.C. Sept. 8, 2021), ECF No. 82 [hereinafter *Meta Complaint*].

14. The impetus behind the Sherman Act grew out of a moment of severe state deregulation and a correlated growth in new forms of national corporate structure in the late 1800s. See Daniel A. Crane, *Lochnerian Antitrust*, 1 NYU J.L. & Liberty 496, 506–08 (2005) (attributing the rise of monopolies in the late 19th century to a liberalization of state corporate law).

15. See *Tex. & Pac. Ry. Co. v. Abilene Cotton Oil Co.*, 204 U.S. 426, 437 (1907) (“[A] statute will not be construed as taking away a common-law right existing at the date of its enactment, . . . unless it be found that the pre-existing right is so repugnant to the statute that the survival of such right would in effect deprive the subsequent statute of its efficacy[.]”).

now been almost fully inverted.¹⁶ The Court has done so even where regulation contains specific antitrust saving clauses.¹⁷

This Note then explores the consequences of this doctrinal inversion and its resulting underenforcement of anticompetitive conduct. When antitrust and regulation are treated as adversarial, the intended symbiotic system of checks and balances between them breaks down.¹⁸ The gaps between enforcement grow longer, the enforcement itself gets weaker, and new corporate structures become so big that no one knows where to start—especially in an increasingly global and technology-centric economy. Antitrust and regulation have different but complementary roles in protecting consumers. A rise in antitrust lawsuits should not be dismissed as inhibitive of regulation—it should be treated as a call to arms for regulation to step up and assess what has gone wrong and how it can do better to protect consumers.

Finally, this Note explains the FTC's case against Meta and why the proposed divestment remedy is necessary to make prospective regulation more feasible and to protect consumers. The Note concludes with suggestions on how to craft regulation around PSNs that explicitly accounts for the ongoing role of antitrust, specifically in the form of a well-crafted saving clause that the Court cannot ignore.¹⁹ Regulation must change to keep up with the markets, and antitrust is both the alarm bell and the fire extinguisher that buys regulation the time it needs to catch up when it falls behind. A symbiotic approach between antitrust and regulation is both necessary *and* more authentic to the original intended purpose of the Sherman Antitrust Act.

16. *Trinko*, 540 U.S. at 406.

17. *Id.*

18. See generally Howard Shelanski, *Antitrust and Deregulation*, 127 YALE L.J. 1922, 1922 (2018) (discussing the countercyclical role of antitrust enforcement during periods of deregulation). Antitrust scholars like Howard Shelanski argue that this cyclical ebb and flow has provided valuable information on the respective roles of antitrust and regulation, but now is the time to act on that knowledge to “restore antitrust as a complement, rather than substitute, for rules in regulated markets.” *Id.* at 1959.

19. This is necessary to avoid what occurred in the telecommunications market. After the Bell System was broken up in 1982, Congress made sure to explicitly include an antitrust saving clause in the Telecommunications Act of 1996 in recognition that regulation could never replace the bluntness of antitrust enforcement where necessary. See Andrew Pollack, *Bell System Breakup Opens Era of Great Expectations and Great Concern*, N.Y. TIMES (Jan. 1, 1984), <https://www.nytimes.com/1984/01/01/us/bell-system-breakup-opens-era-of-great-expectations-and-great-concern.html>; Saving the Savings Clause: Congressional Intent, the *Trinko* Case, and the Role of the Antitrust Laws in Promoting Competition in the Telecom Sector: Hearing Before the H. Comm. on the Judiciary, 108th Cong. 1–2 (2003) (statement of Hon. F. James Sensenbrenner, Jr., Chairman, Committee on the Judiciary).

II. ANTITRUST AND REGULATION AT ODDS

Regulation and antitrust enforcement have not always been so exaggeratedly at odds. Until 2004, the Supreme Court somewhat consistently applied a “clear” or “plain” repugnancy standard in the balancing of statutes with common-law rights,²⁰ so that “[r]epeals of the antitrust laws by implication from a regulatory statute are strongly disfavored.”²¹ This standard allowed regulation and antitrust enforcement to work in tandem, each tool available as the facts of a case demanded.

However, the Court has distorted this standard since the 2000s. As Shelanski summarizes in his argument for rebalancing antitrust and regulation, “[t]he Supreme Court’s trend in adopting blunt forms of claim preclusion in regulated industries throws out good cases along with the bad, treats private cases identically to those brought by public enforcement agencies, and makes no provision for the comparative advantages of antitrust and regulation in different settings.”²² The Court’s new interpretation of the relationship between antitrust laws and regulatory schemes has worsened an already dire underenforcement of anticompetitive behaviors.

To provide context for this Note’s subsequent arguments, this Part covers background on the origins of the plain repugnancy standard and its application throughout the 20th century. It then provides more detail on *Trinko* and *Credit Suisse*, two cases which narrowed the scope of antitrust enforcement in regulated markets, and their reframing of that standard. Finally, this Part further details the consequences of these decisions and the effects they have had on antitrust enforcement.

A. THE CLEAR REPUGNANCY DOCTRINE

The Sherman Antitrust Act of 1890 was born at the tail end of the Gilded Age, a period characterized by steep wealth inequality, across-the-board

20. See *Abilene Cotton*, 204 U.S. at 437 (“[A] statute will not be construed as taking away a common-law right existing at the date of its enactment . . . unless it be found that the pre-existing right is so repugnant to the statute that the survival of such right would in effect deprive the subsequent statute of its efficacy[.]”); see also *Keogh v. Chicago & N.W. Ry. Co.*, 260 U.S. 156, 162 (1922) (holding that regulation of rates does not bar government actions, but does bar private antitrust actions under the facts described, by stating that “[t]he fact that these rates had been approved by the Commission would not, it seems, bar proceedings by the government. It does not, however, follow that Keogh, a private shipper, may recover damages under section 7 because he lost the benefit of rates still lower, which, but for the conspiracy, he would have enjoyed.”).

21. *United States v. Phila. Nat’l Bank*, 374 U.S. 321, 350–51 (1963).

22. Howard Shelanski, *The Case for Rebalancing Antitrust and Regulation*, 109 MICH. L. REV. 683, 731 (2001).

market consolidation under the so-called Robber Barons, and the growth of an increasingly active and agitated labor rights movement.²³ In advocating for the bill to his colleagues, Senator John Sherman made it clear that the Act was intended to address the newly national scale of a common law problem already regulated by the states.²⁴ While Chicago School academics since Robert Bork have argued that the sole goal of the Act was the protection of consumer welfare,²⁵ that interpretation of the legislative history does not account for the legal reality out of which the Act emerged.

As indicated by Sherman's repeated references to the inhuman nature of corporations and the "corporate rights open to all," consumer welfare was ancillary to his general concern over monopolies and anticompetitive behavior.²⁶ His federalist framing indicates that the promotion of competition in and of itself is the goal because the existence of free competition across the states is a check on the "undue influence" that a corporation can otherwise accrue in a single state.²⁷ As Sherman noted, his intent was to combat "the law of selfishness, uncontrolled by competition," and not to single out "a particular trust, but the system" writ large.²⁸ To focus on the secondary effect of consumer welfare is to convolute and subjugate the actual goal of the Act: promoting competition to prevent the concentration of unchecked market power.

Keeping this context in mind, the plain repugnancy standard traces back to 1907, and the *Lochner* Court's general hostility toward regulation during that period.²⁹ Despite the fact that antitrust may seem at odds with *Lochnerian* conceptions of freedom to contract, "the freedom of the consumer, individual producer, artisan, or trader from the coercion of government-sanctioned monopolies . . . is reflected amply in the pre-Sherman Act common law and in the antitrust ideology of the *Lochner* era."³⁰ From this perspective, antitrust enforcement is actually a restraint on government overreach, both in the form of regulation *and* calculated de-regulation. When one state de-regulates to collude with corporate power, the citizens of all states are threatened by the

23. See Lina M. Khan, *The Ideological Roots of America's Market Power Problem*, 127 Yale L.J.F. 960, 965 (2018).

24. 21 CONG. REC. 2456 (1890).

25. See generally Robert H. Bork, *Legislative Intent and the Policy of the Sherman Act*, 9 J.L. & ECON. 7 (1966).

26. 21 CONG. REC. 2457 (1890).

27. *Id.*

28. *Id.*

29. *Abilene Cotton*, 204 U.S. at 437.

30. Crane, *supra* note 14, at 497 (emphasis omitted).

wrongly state-sanctioned and un-checked growth of that chosen corporation and its effects on interstate commerce.

This populist framing of antitrust carried through the first third of the 20th Century. As the Court wrote in a now-overturned 1933 case,

As a charter of freedom, the [Sherman Anti-Trust Act] . . . call[s] for vigilance in the detection and frustration of all efforts unduly to restrain the free course of interstate commerce, but [does] not seek to establish a mere delusive liberty either by making impossible the normal and fair expansion of that commerce or the adoption of reasonable measures to protect it from injurious and destructive practices and to promote competition upon a sound basis.³¹

Antitrust intervention was deemed necessary to ensure that private entities did not exploit the market and that regulatory entities could not facilitate such behavior by intentionally or unintentionally cementing certain monopolies. Antitrust was thus not only a check on anticompetitive behavior, but on regulatory frameworks themselves.

After the New Deal's expansion of the administrative state, and as World War II distracted from enforcement, however, antitrust took a backseat to regulation.³² In a 1948 opinion upholding a vertical merger in the steel industry, Justice Reed wrote, “[i]t is not for courts to determine the course of the Nation’s economic development If businesses are to be forbidden from entering into different stages of production that order must come from Congress, not the courts.”³³ This was a massive departure from the original intent of the Sherman Act, given Sherman’s express intention for the statute to “be construed liberally, with a view to promote its object.”³⁴ Nor was this departure missed by Justice Douglas in his dissent focused on the problem of “bigness.”³⁵

Despite this retraction of antitrust law, the Court consistently displayed a hesitance to disregard repugnancy standards until the 1960s.³⁶ By that time, this was clearly articulated as the “plain” or “clear” repugnancy doctrine:

31. *Appalachian Coals v. United States*, 288 U.S. 344, 359–60 (1933), *overruled by* *Copperweld Corp. v. Indep. Tube Corp.*, 467 U.S. 752 (1984).

32. *See* Gressley, *supra* note 2, at 227 (“By the fall of 1940 . . . [i]t became increasingly clear that attack on monopoly was being given a holiday.”).

33. *United States v. Columbia Steel Co.*, 334 U.S. 495, 526 (1948).

34. 21 CONG. REC. 2456 (1890).

35. *Columbia Steel*, 334 U.S. at 535 (Douglas, J., dissenting).

36. *Pan Am. World Airways, Inc. v. United States*, 371 U.S. 296, 304–05 (1963) (“[W]e hesitate here, as in comparable situations, to hold that [a] new regulatory scheme . . . was designed completely to displace antitrust laws—absent an unequivocally declared congressional purpose so to do.”).

“Repeals of the antitrust laws by implication from a regulatory statute are strongly disfavored, and have only been found in cases of plain repugnancy between the antitrust and regulatory provisions.”³⁷ In the absence of a clause explicitly ruling out antitrust enforcement, the Court sought to maintain both statutory regulation and antitrust actions to the fullest extent possible. But this standard began to weaken in the 1970s. This is illustrated by the stark difference between antitrust decisions from the ‘60s, like *Silver v. New York Stock Exchange*,³⁸ and those from the ‘70s, like *Gordon v. New York Stock Exchange*.³⁹

In each case, the Court considered the interaction between antitrust enforcement and regulatory oversight by the Securities and Exchange Commission (SEC). In *Silver*, the Court emphasized reconciliation of antitrust and regulation,⁴⁰ but just twelve years later in *Gordon*, mere potential interference was enough for the Court to find implied antitrust immunity.⁴¹ *Gordon* thus marked the beginning of an era of antitrust anti-textualism and over-enforcement paranoia that is only now beginning to change.

Throughout this period, antitrust laws were gradually weakened. After *Gordon*, the Court clarified that implied immunity can apply “even absent active regulatory supervision of the specific conduct at issue . . . if the challenged conduct could be allowed under the statute and if the agency generally exercised ‘the kind of administrative oversight of private practices that Congress contemplated.’”⁴² Essentially, this means that an antitrust action can be precluded even when brought under a statute with an antitrust saving clause if that clause is not explicitly written to prevent such a result. It also means that an agency does not even need to be actively monitoring certain conduct for an antitrust action to be precluded—all that matters is if the agency is capable of cobbling together some type of oversight and remedy *ex post*. This perspective reduces the “strongly disfavored” preclusion of antitrust to a standard that requires mere “oversight” by a regulatory body.

Such a perspective also grossly misinterprets the concept of what function regulation must actually perform to preclude antitrust. An accurate example of repugnance would be a price manipulation claim in an industry where prices are set by a regulatory body—not a price manipulation claim in an industry

37. *Philadelphia Nat. Bank*, 374 U.S. at 350–51.

38. *Silver*, 373 U.S. at 341.

39. *Gordon v. New York Stock Exch.*, 422 U.S. 659 (1975).

40. Shelanski, *supra* note 22, at 687.

41. *Id.*

42. *Id.* at 688 (quoting *United States v. Nat’l Ass’n of Sec. Dealers*, 422 U.S. 694, 728 (1975)).

where a regulatory body merely has some kind of its own regulatory mechanism in place to punish such misconduct. An entity may be subject to duplicative punishment in the form of both antitrust monetary damages and regulatory penalties, but such hefty punishment may in fact be desirable given the current excesses of market concentration across industries.

While the cases cited here concern securities law, the Court made clear that the repugnancy standard applies in other industries as well, particularly telecoms and energy. In *Otter Tail Power Co. v. United States*, the Court “declined to find that the Federal Power Act provided immunity from the government’s claim that the defendant had violated the antitrust laws by refusing to supply either interconnection to distribution facilities or power to competing municipal utilities.”⁴³ The Court distinguished between “duplicative” claims and repugnant claims in a way that bolstered the repugnancy standard yet again, but this 1973 case was the last to do so.

The clear repugnancy standard is evidence that antitrust and regulation not only can but should operate in connection with each other. More importantly, however, it is proof that antitrust enforcement is necessary as a check to regulation—it should be treated as a canary in the coal mine of regulatory capture, and responded to as a call to update how regulation operates. As the following Section illustrates, there are serious consequences when antitrust actions are precluded and regulatory bodies are left to stagnate.

B. THE DISTORTED FRAMEWORK OF IMPLIED IMMUNITY

Since the 1970s, legal precedent has reflected a certain hostility towards antitrust. This hostility is largely attributable to the influence of big business interests on the Court.⁴⁴ Such an attitude is reminiscent of antitrust enforcement reluctance in the 1930s and 1940s, but without a similarly robust strengthening of the administrative state. Today’s wealth inequality rivals that of the 19th century,⁴⁵ and a historically weakened labor movement is unable to fight back.⁴⁶ Yet, neither Congress nor the judiciary has responded to reign in the markets and address the unprecedented concentration of market power.

43. *Id.*

44. *See* Lancieri, Posner & Zingales, *supra* note 3, at 57 (“Large business interests have always been opposed to strong enforcement of antitrust law. If we want to attribute the decline of antitrust enforcement to the pressure exerted by big business, we need to explain why starting in the mid-1970s these interest groups succeeded where they had failed before.”).

45. *See* CONG. BUDGET OFF., TRENDS IN THE DISTRIBUTION OF FAMILY WEALTH, 1989 TO 2019 (Sept. 27, 2022), <https://www.cbo.gov/publication/57598> (tracking wealth distribution in the U.S. from 1989 to present).

46. While there were historically at least 200 work stoppages per year from 1947 to 1979, those numbers have steadily dropped to the point that there are fewer than 50 each year since

This national hostility toward antitrust enforcement is crystalized in two cases from 2004 and 2011 in which the Court suggested that antitrust immunity can exist in regulated markets. Until the first, the Supreme Court had never held that an antitrust action could be precluded by a regulatory scheme with an explicit antitrust saving clause.

1. Trinko

In 2004, the Court held in *Verizon Communications Inc. v. Law Offices of Curtis V. Trinko, LLP* (“*Trinko*”) that while the Telecommunications Act of 1996 “preserves claims that satisfy existing antitrust standards,” the creation of new non-traditional antitrust claims depends heavily on “the existence of a regulatory structure designed to deter and remedy anticompetitive harm.”⁴⁷ Justice Scalia’s application of the plain repugnancy standard in the majority opinion diverged significantly from past precedent and set the stage for further erosion.

There, a customer of AT&T sued Verizon for denying competitors “access to interconnection support services, making it difficult for those competitors to fill their customers’ orders.”⁴⁸ The plaintiff argued that Verizon’s conduct violated § 2 of the Sherman Act by impeding downstream local telephone service offerings. Aside from an obvious issue of standing, which the concurrence would have declined to go beyond,⁴⁹ the issue at the core of the case was whether a breach of duties imposed by the Telecommunications Act of 1996 could form the basis of a claim under § 2 of the Sherman Act.

In considering this issue, the opinion emphasized the regulatory response that occurred prior to the filing. As Justice Scalia noted, to take “advantage of the opportunity provided by the 1996 Act for incumbent LECs to enter the long-distance market . . . required Verizon to satisfy, among other things . . . compliance with the Act’s network-sharing duties.”⁵⁰ As a result of this and other interconnection agreements with rivals, Verizon was subject to oversight from both state regulators and the Federal Communications Commission (FCC).

1990. *Annual Work Stoppages Involving 1,000 or More Workers, 1947 - Present*, U.S. BUREAU LAB. STAT. (Aug. 11, 2023), <https://www.bls.gov/web/wkstp/annual-listing.htm>.

47. *Trinko*, 540 U.S. at 407, 412 (2004).

48. *Pac. Bell Tel. Co. v. linkLine Commc’ns, Inc.*, 555 U.S. 438, 449 (2009) (quoting *Trinko*, 540 U.S. at 404–05).

49. *See Trinko*, 540 U.S. at 416–17 (Stevens, J., concurring) (“In complex cases it is usually wise to begin by deciding whether the plaintiff has standing to maintain the action . . . I would not go beyond the first step in this case.”).

50. *Id.* at 402–03.

Upon reports from rivals that Verizon was not properly fulfilling service orders from other local exchange carriers, simultaneous investigations were opened by the FCC and New York's Public Service Commission (PSC).⁵¹ While the PSC issued a series of orders including heightened reporting requirements and a \$10 million fine, the FCC compelled Verizon to pay \$3 million to the U.S. Treasury and to enter a consent decree with additional requirements and penalties.⁵² The law offices of Curtis V. Trinko, a customer of Verizon's affected rival AT&T, filed suit "[t]he day after Verizon entered its consent decree with the FCC."⁵³

As Justice Stevens's concurrence in the judgment noted, the most obvious issue in this case was standing. Justice Stevens wrote that the threshold question is "whether, assuming the truth of its allegations, respondent is a 'person' within the meaning of § 4 [of the Clayton Act]."⁵⁴ According to precedent, § 4 is not read literally, "particularly in cases in which there is only an indirect relationship between the defendant's alleged misconduct and the plaintiff's asserted injury."⁵⁵ The rationale behind this is to avoid "either the risk of duplicate recoveries on the one hand, or the danger of complex apportionment of damages on the other,"⁵⁶ which harkens back to a duplicative standard espoused in various 1960s and 1970s antitrust cases.⁵⁷ Justice Stevens would have declined to go beyond this issue of standing because the claim was "purely derivative of the injury that AT&T suffered."⁵⁸

However, instead of halting the inquiry there, Justice Scalia's opinion went much further and convoluted the repugnancy standard by reframing the issue of duplication. First, he posed the question of "what effect (if any) the 1996 Act has upon the application of traditional antitrust principles."⁵⁹ The effect of this framing is clear from the holding that results: while the 1996 Act's antitrust-specific saving clause "preserves claims that satisfy existing antitrust standards, it does not create new claims that go beyond existing antitrust standards."⁶⁰ While the concurrence would have clearly maintained the

51. *Id.* at 403–04.

52. *Id.*

53. *Id.* at 404.

54. *Id.* at 416 (Stevens, J., concurring).

55. *Id.* (Stevens, J., concurring).

56. *Id.* at 416–17 (Stevens, J., concurring) (quoting *Associated Gen. Contractors of Cal., Inc. v. Carpenters*, 459 U.S. 519, 529–535 (1983)).

57. *See Illinois Brick Co. v. Illinois*, 431 U.S. 720, 731 (1977) (quoting *Hawaii v. Standard Oil Co. of Cal.*, 405 U.S. 251, 264 (1972)) ("[W]e are unwilling to 'open the door to duplicative recoveries' under s 4.").

58. *Trinko*, 540 U.S. at 417 (Stevens, J., concurring).

59. *Id.* at 405.

60. *Id.* at 407.

repugnancy standard by refusing to allow a duplicative and administratively difficult claim to proceed, Justice Scalia's holding created a new inquiry entirely: does the claim arise out of "traditional antitrust principles," or is it something "new" created by the regulation itself?

Justice Scalia's reframing had two simultaneous functions. First, it narrowed possible government antitrust actions in a regulated market to only those that are "traditional" without precisely defining what that means. Second, it diluted the rationale for why "new" claims should not proceed. As Justice Stevens explained, the issues of duplicative punishment and impossible administrability are arguably reasonable justifications for why a particular anticompetitive claim cannot proceed in a regulated market.⁶¹ This has nothing to do with the "newness" of the claims, however. Theoretically, if a "new" type of anticompetitive behavior arose as a result of the 1996 Act that was not enforced by a regulatory body and for which there were reduced administrability concerns, there is nothing in the Court's antitrust precedent that should prevent the claim from proceeding simply because it is not "traditional." The emergence of such a claim should instead inspire an audit of the regulatory body to determine if it needs new or additional resources to address new problems.

Instead, Justice Scalia further distorted the standard of clear repugnancy by focusing on a derivative doctrine of implied immunity. In his discussion of the additional requirements imposed by the 1996 Act, Justice Scalia wrote:

That Congress created these duties, however, does not automatically lead to the conclusion that they can be enforced by means of an antitrust claim. Indeed, a detailed regulatory scheme such as that created by the 1996 Act ordinarily raises the question whether the regulated entities are not shielded from antitrust scrutiny altogether by the doctrine of implied immunity.⁶²

While the idea of implied immunity has floated around in various contexts, it had never existed in antitrust until *Trinko*, except as the strongly disfavored result of clear repugnancy analysis.

For example, the first case Justice Scalia cited in referring to the doctrine, *United States v. National Association of Securities Dealers, Inc. (NASD)*, considered "whether certain sales and distribution practices employed in marketing securities of open-end management companies, popularly referred to as 'mutual funds,' are immune from antitrust liability."⁶³ The Court held that

61. *Id.* at 416–17 (Stevens, J., concurring).

62. *Id.* at 406.

63. 422 U.S. 694, 697 (1975).

mutual funds are immune, but only because antitrust enforcement of those particular claims would have been too duplicative and could result in inconsistent standards. Despite the application of this duplicative standard, the *NASD* majority still cited to the clear repugnancy standard as their guiding principle.

As the dissent in *NASD* indicates, this line of cases ostensibly preserved clear repugnancy but still set the stage for the erosion into implied immunity that occurred in the 2000s. As Justice White wrote in the *NASD* dissent:

Under that holding, in light of the context of this case, implied antitrust immunity becomes the rule where a regulatory agency has authority to approve business conduct whether or not the agency is directed to consider antitrust factors in making its regularity decisions and whether or not there is other evidence that Congress intended to displace judicial with administrative antitrust enforcement.⁶⁴

In other words, the clear repugnancy standard has been watered down to the point that any type of regulatory oversight is enough to justify the dismissal of antitrust claims. This holds true regardless of whether the framework can address anticompetitive harms and regardless of the fact that duplicative punishment can be beneficial where market consolidation has run amok.

Justice Scalia's holding in *Trinko* built on the weakening of the repugnancy standard by focusing the analysis of "traditional antitrust principles" on their enforcement in a regulated market. While the rule-of-reason doctrine has long required an analysis of the totality of the circumstances in cases lacking a *per se* violation of the Sherman Act, Scalia's analysis brought the existence of a regulatory framework to the foreground. In fact, he completely set aside the 1996 Act and its specific enforced infrastructure sharing by arguing that absent regulation,

[To compel] such firms to share the source of their advantage is in some tension with the underlying purpose of antitrust law, since it may lessen the incentive for the monopolist, the rival, or both to invest in those economically beneficial facilities. Enforced sharing also requires antitrust courts to act as central planners, identifying the proper price, quantity, and other terms of dealing—a role for which they are ill suited. Moreover, compelling negotiation between competitors may facilitate the supreme evil of antitrust: collusion.⁶⁵

64. *Id.* at 736 (White, J., dissenting).

65. *Trinko*, 540 U.S. at 407–08.

In other words, Justice Scalia went beyond the already weakened duplicative standard to argue that traditional antitrust law can rarely sustain such a claim at all, absent very specific circumstances of prior dealing. To do so, Justice Scalia says, would actually be antithetical to antitrust doctrine.

This focus on the difference between regulated and unregulated markets misappropriates the doctrine of implied immunity and gives it undue significance in the analysis. Scalia acknowledged that Congress “precluded that interpretation,” but he also argued that the clause only preserves “traditional” antitrust claims and that creating a “new claim[] . . . would be equally inconsistent with the saving clause’s mandate that nothing in the Act ‘modify, impair, or supersede the applicability’ of the antitrust laws.”⁶⁶ This interpretation reveals a blatant disregard for actual antitrust doctrine, which has long included rule-of-reason analysis as a means of recognizing so-called “new” claims.

Justice Scalia even acknowledged this precedent of contextually grounded antitrust claims in the final part of the opinion. He wrote that, “[a]ntitrust analysis must always be attuned to the particular structure and circumstances of the industry at issue,”⁶⁷ but he did so to emphasize the weighty role he believes regulatory frameworks should play in that analysis. According to Justice Scalia:

One factor of particular importance is the existence of a regulatory structure designed to deter and remedy anticompetitive harm. Where such a structure exists, the additional benefit to competition provided by antitrust enforcement will tend to be small, and it will be less plausible that the antitrust laws contemplate such additional scrutiny.”⁶⁸

This framework expands the issue beyond duplicative claims and the practical justifications for why certain antitrust suits cannot proceed in light of regulation. It further appears to imply that a more lenient balancing test should be used rather than strict clear repugnancy analysis. Justice Scalia did not outright say this, but why else would he fail to cite to clear repugnancy *at all*?

2. Credit Suisse

In 2007, the Court in *Credit Suisse Sec. (USA) LLC v. Billing* (“*Credit Suisse*”) further eroded enforcement in the Court’s most recent ruling on the application of antitrust laws to regulated markets.⁶⁹ In *Credit Suisse*, the Court

66. *Id.* at 406, 407.

67. *Id.* at 411.

68. *Id.* at 412.

69. 551 U.S. 264 (2007).

held that “the threat of antitrust mistakes” was too great to allow an antitrust suit to proceed despite the saving clauses of the securities acts under consideration.⁷⁰ This holding placed an outsized emphasis on the potential havoc antitrust intervention could wreak, but that fear was largely unfounded.⁷¹

While *Credit Suisse* cites to the clear repugnancy standard, it does so only to further undercut its disfavored status. In an opinion authored by Justice Breyer, the Court defined the clear repugnancy standard according to an interpretation distilled from *Silver*, *Gordon*, and *NASD* as “clear incompatibility” between “securities law and the antitrust complaint.”⁷² But Justice Breyer first listed a number of specific factors to consider in that approach:

(1) the existence of regulatory authority under the securities law to supervise the activities in question; (2) evidence that the responsible regulatory entities exercise that authority; and (3) a resulting risk that the securities and antitrust laws, if both applicable, would produce conflicting guidance, requirements, duties, privileges, or standards of conduct. We also note (4) that in *Gordon* and *NASD* the possible conflict affected practices that lie squarely within an area of financial market activity that securities law seeks to regulate.⁷³

While it is possible to interpret this definition as maintaining the status quo, these additional factors actually water down the clear repugnancy analysis in several ways. Specifically, the third factor merely requires a “risk” that there is some kind of conflict, which is a serious departure from the historically disfavored status of antitrust preclusion.

Justice Breyer went even further by quickly dismissing the possibility that §§ 77p(a) and 78bb(a) of the Securities Act and the Securities Exchange Act could be interpreted “as saving clauses so broad as to preserve all antitrust actions.”⁷⁴ According to the framing of the Court, if a saving clause does not explicitly mention antitrust, courts must determine if and how antitrust law might be precluded: “Those determinations may vary from statute to statute, depending upon the relation between the antitrust laws and the regulatory program set forth in the particular statute, and the relation of the specific conduct at issue to both sets of laws.”⁷⁵ By posing the issue in this manner,

70. *Id.* at 282.

71. See Lancieri, Posner & Zingales, *supra* note 3 (empirically connecting the decline in U.S. antitrust enforcement with the advancement of big business interests).

72. *Credit Suisse*, 551 U.S. at 275.

73. *Id.* at 275–76.

74. *Id.* at 275.

75. *Id.* at 271.

Breyer essentially lost track of the purpose of the plain repugnancy standard and reduced the analysis to a specific interrogation of securities markets.

Breyer wrote that he fully accepted petitioners' argument that despite having full control over the matter and actively disapproved of the behavior, the SEC only regulates some of the conduct in question. But he argued that this lack of intervention is an intentional exercise of discretion intended by Congress, and that "there is no practical way to confine antitrust suits so that they challenge only activity of the kind the investors seek to target, activity that is presently unlawful and will likely remain unlawful under the securities law."⁷⁶ Not only does this interpretation grossly underestimate courts' ability to discern between approved and unapproved conduct, but it also completely erases the original purpose of antitrust as a check on poorly functioning regulation. If the SEC is not addressing clearly harmful behavior that it itself disapproves of, and if the antitrust action would be in line with the SEC's goals, then why should the case be precluded?

Justice Thomas wrote a compelling dissent in *Credit Suisse* pointing to just this issue that harkens back to the origins of the clear repugnancy standard and antitrust as a common law right. As he noted, the texts of both §§ 77p(a) and 78bb(a) preserve "any and all other rights and remedies that may exist at law or in equity," but make no specific reference to antitrust.⁷⁷ He reasoned,

[T]he mere existence of targeted saving clauses does not demonstrate—or even suggest—that antitrust remedies are not included within the "any and all" other remedies to which the securities saving clauses refer. Although Congress may have singled out antitrust remedies for special treatment in some statutes, it is not precluded from using more general saving provisions that encompass antitrust and other remedies. Surely Congress is not required to enumerate every cause of action—state and federal—that may be brought. When Congress wants to preserve all other remedies, using the word "all" is sufficient.⁷⁸

This analysis reflected Senator Sherman's assertion that, "[t]he purpose of [the Sherman Antitrust Act] is to enable the courts of the United States to apply the same remedies against combinations which injuriously affect the interests of the United States that have been applied in the several States to protect local interests."⁷⁹ By requiring that a saving clause explicitly mention antitrust, the majority limited the reach of enforcement of nearly any regulated industry.

76. *Id.* at 282.

77. *Id.* at 287 (Thomas, J., dissenting).

78. *Id.* at 289 (Thomas, J., dissenting).

79. 21 CONG. REC. 2456 (1890).

In a similar vein, Justice Stevens's concurrence also rejected the majority's determination that antitrust action is precluded and would instead have argued that the plaintiffs simply failed to state a cognizable claim. He wrote that,

Surely I would not suggest, as the Court did in *Twombly*, and as it does again today, that either the burdens of antitrust litigation or the risk "that antitrust courts are likely to make unusually serious mistakes," . . . should play any role in the analysis of the question of law presented in a case such as this.⁸⁰

Justice Stevens explicitly pointed to the unjustified paranoia that the courts are incapable of properly addressing antitrust claims in regulated markets.

In sum, the Court has muddied the waters of antitrust in regulated markets based on unfounded fears. As a result, they have led a misguided attempt to uphold administrative agencies regardless of whether they are effectively protecting and promoting competition. And they have driven a wedge between antitrust and regulation that is difficult to dislodge.

III. PARANOIA, UNDERENFORCEMENT, AND THE PSN MARKET

The Court's unfounded fear of antitrust overenforcement comes from a misunderstanding of the goals of antitrust and regulation and how they relate. This concern about overenforcement is palpable in the language used repeatedly by the Court in its latest applications of the clear repugnancy standard. As Justice Breyer gravely opines in *Credit Suisse* with respect to securities markets, not only is "any enforcement-related need for an antitrust lawsuit . . . unusually small," but "to allow an antitrust lawsuit would threaten serious harm to the efficient functioning of the securities markets."⁸¹ As far as the majority is concerned, antitrust is merely a burden on regulation which itself is already a burden on markets—to allow both at once would stifle efficiency and expose market participants to potentially duplicative or even conflicting obligations. This fear is symptomatic of the past fifty years of Court opinions subverting the clear repugnancy doctrine into one of implied immunity.

More importantly, this fear is also reflective of a larger cultural subservience to the cult of business. As Lancieri, Posner, and Singales illustrated and argued based on a unique empirical study, there has been a serious decline in antitrust enforcement since the 1950s as a result of both

80. *Credit Suisse*, 551 U.S. at 287 (Stevens, J., concurring).

81. *Id.* at 283.

regulatory underenforcement and judicial antitrust anti-textualism.⁸² They write,

Since the 1970s, no president advocated for a reduction in antitrust enforcement, no Congress voted for reduced enforcement except indirectly in obscure budget bills, and no Senate knowingly confirmed nominees to the FTC or DOJ, or to the Supreme Court, who openly promised to reduce antitrust enforcement (again, with some limited exceptions). The decline of antitrust enforcement took place at the hands of regulators and judges with little to no open political support[.]⁸³

Despite ostensible support for antitrust enforcement across the political spectrum, their research shows that big business interests have successfully swayed the Court towards implied immunity and reliance on under-resourced regulation. The warping of antitrust into antibusiness in the eyes of regulators and the Court has led the United States to an unprecedented point of market concentration and deregulation.

Was any of this fear warranted? Or has it actually manifested even greater harms than the Court's imaginary antitrust bogeyman?

A. THE CONSEQUENCES OF UNDERENFORCEMENT

Deregulation and reduced enforcement have failed to generate greater efficiency and more robust competition. Even worse, they have resulted in a concentration of market power across industries large enough to rival the era of Robber Barons that inspired the Sherman Act.

The failures of deregulation are apparent from Lancieri, Posner, and Singales's research results. Contrary to Chicago School promises of increased efficiency, they found that unlike otherwise-similar nations, annual growth in output per hour worked in the United States has actually decreased significantly since the '70s,⁸⁴ and “[w]hile median earnings of male full-time workers in the United States grew 36% in real terms between 1960 and 1980, they did not change at all between 1980 and 2016.”⁸⁵ Moreover, as a result of profits concentrating around larger firms, “during the 1980–2020 period, the share of income earned by the top 1% of the income distribution grew from 10% to 19% in the United States, versus an increase from 8% to 13% in the United Kingdom and from 7% to 10% in France.”⁸⁶ In other words, the U.S.

82. See Lancieri, Posner & Zingales, *supra* note 3, at 41.

83. *Id.*

84. *Id.* at 54.

85. *Id.* at 55.

86. *Id.* at 56.

economy has not benefitted from decreased antitrust enforcement. Rather, a handful of corporations have benefitted by building monopoly power across industries.

The failures of underenforcement are market consolidation and its accompanying harms to consumers in the long run—higher prices, lower quality products, and worse overall quality of life for the average American both as a worker and as a customer. These failures are evident from looking at nearly any market in the country. As Rebecca Giblin and Cory Doctorow write,

Just a handful of firms—or sometimes only one—now control everything from the arts (publishing, movies, music, streaming, comics, bookselling, movie theaters, talent agencies, games, wrestling, radio stations) . . . to agribusiness (seeds, livestock, tractors, fertilizer, pesticides, precision agriculture, and the production of meat, eggs, grain, and produce) and everything in between (cruise lines, cheerleader uniforms, groceries, pharmaceuticals, glass bottles, medical devices, airlines, eyeglasses, athletic shoes, fast food, food delivery, and pet food).⁸⁷

They attribute much of this concentration to a “tsunami of mergers: the number of [U.S.] publicly traded companies dropped by half even as they increased by 50% in other developed nations.”⁸⁸ But they also attribute it to new forms of anticompetitive behavior that have only become possible through technology, like “data moats” and “network effect moats” that are creating “chokepoints that separate producers from consumers so [corporations] can capture a disproportionate share of the value of other people’s work.”⁸⁹ Monopoly may temporarily lower prices for consumers, but when monopsony power kicks in, those same consumers lose out on more value from their labor as workers and eventually are still subjected to higher prices through gimmicks like shrinkflation⁹⁰ and because firms have realized they *can* just raise prices in the wake of disasters like the COVID-19 pandemic.⁹¹

87. GIBLIN & DOCTOROW, *supra* note 7, at 4–5.

88. *Id.* at 5.

89. *Id.* at 6.

90. See, e.g., ‘Shrinkflation’ Accelerates Globally as Manufacturers Quietly Shrink Package Sizes, NAT’L PUB. RADIO (June 8, 2022), <https://www.npr.org/2022/06/08/1103766334/shrinkflation-globally-manufacturers-shrink-package-sizes>.

91. Even though pandemic-induced shortages have ended, prices have remained unusually high. See, e.g., Rachel Layne, *Why Are Prices So High Right Now—and Will They Ever Return to Normal?*, HARV. BUS. SCH. (Feb. 10, 2022), <https://hbswk.hbs.edu/item/why-are-prices-so-high-right-now-inflation>.

For PSNs, this concentration is readily apparent from the handful of technology companies that dominate global internet service provider (ISP) markets, including Meta. Of the largest corporations in the world by market capitalization, Apple, Microsoft, Alphabet, Amazon, and Meta are all in the top ten.⁹² In 2023, three of the top four social media interfaces were owned by Meta with over a billion monthly active users on each: Facebook, WhatsApp, and Instagram.⁹³ PSNs are defined by the number of users they provide access to for advertisers and app developers, and by the amount of time they can capture users' attention and have them engage with those ads and apps. These companies arose in the current cycle of drawn-out underenforcement and have never been adequately regulated. Their unprecedented rate and scale of growth is indicative of those origins. The PSN model lays bare the failure of the "consumer welfare" ethos that continues to dominate antitrust discourse.

Proponents of a "consumer welfare" antitrust ethos like Herbert Hovenkamp argue that antitrust is *meant* to be limited and narrowly focused, and its standards have simply been misapplied by its dissenters on either side of the political spectrum. According to Hovenkamp, "bigness" is not a problem under the enlightened consumer welfare standard because,

While small competitors of a large low cost and high output firm can be injured, many other small firms benefit, including suppliers and retailers. A good illustration is Amazon, which is a very large firm that generally sells at low prices and has maintained high consumer satisfaction. Amazon has undoubtedly injured many small firms forced to compete with its prices and distribution. At the same time, however, Amazon acts as broker for millions of small firms who use its retail distribution services. When a very large firm produces more, it creates opportunities for other firms that sell complements, that distribute the products that a large firm produces, or that supply it with inputs.⁹⁴

As far as Hovenkamp is concerned, antitrust has no place interfering with a firm like Amazon because it hasn't hurt consumers and it props up as many small businesses as it crushes. But he only tells half the story.

92. *The 100 Largest Companies in the World by Market Capitalization in 2022*, STATISTA (Aug. 5, 2022), <https://www.statista.com/statistics/263264/top-companies-in-the-world-by-market-capitalization/>.

93. *Most Popular Social Networks Worldwide as of January 2023, Ranked by Number of Monthly Active Users*, STATISTA (Feb. 14, 2023), <https://www.statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users/#:~:text=Meta%20Platforms%20owns%20four%20of,monthly%20core%20Family%20product%20users.>

94. Herbert J. Hovenkamp, *Antitrust: What Counts as Consumer Welfare?*, at 5, UNIV. OF PENNSYLVANIA CAREY LAW SCHOOL ALL FACULTY SCHOLARSHIP, PAPER NO. 2194 (2020), https://scholarship.law.upenn.edu/faculty_scholarship/2194/.

Hovenkamp leaves out everything he has decided does not have a place in antitrust law. As a result of market power concentration like Amazon's domination of online shopping and shipping, consumers have experienced a significant decrease in the quality of their experience as the company becomes too unwieldy to manage. For Amazon, this has meant a proliferation of fraudulent and shoddy products on the platform.⁹⁵ This benefits Amazon by making their own branded knock-off products—the same product ideas they have blatantly stolen from small businesses—more appealing.⁹⁶ Moreover, market power like Amazon's allows a firm to wield power over producers and suppliers that ultimately can reduce and control the choices to which consumers have access. When only a handful of firms control what is available to consume, creative expression can be stifled, innovation can be deterred, and consumers have nowhere to turn for alternatives that better align with their needs, like privacy rights or promoting their local community and economy. The extrinsic costs of oversized behemoths are unaccounted for yet notable, like the environmental impact of unchecked consumerism and the perpetuation of imperialist harms through global extractionism.⁹⁷

From a Neo-Brandeisian perspective, the very bigness that Hovenkamp derides as irrelevant is the real problem. As Lina Khan and Sandeep Vaheesan argue, market concentration results in regressive wealth redistribution, enables the accumulation of political clout, and threatens the sanctity of democracy itself.⁹⁸ And nowhere is this more easily observable than in the market of PSNs that the FTC is now addressing under Khan's leadership.

95. See Paul Conley, *Nearly a Third of Amazon Shoppers Are Disappointed by Quality or Timeliness*, DIGITAL COM. 360 (Jan. 19, 2022), <https://www.digitalcommerce360.com/2022/01/19/nearly-a-third-of-amazon-shoppers-are-disappointed-by-quality-or-timeliness/> (“The challenge is that Amazon spent quite a bit of time pre-pandemic, saying, ‘we are taking control of our own logistics, we have our Amazon trucks, we have our Amazon Prime shipping,’ Ng says. ‘Yes, there are things out of everyone’s control, but they’ve spent so much time talking about efficiency and scale that it is actually hurting them.’”).

96. Not only are physical goods an environmental problem, the storage of data and cloud computing are substantial as well. See Ashleigh Hollowell, *Why Data Has a Sustainability Problem*, VENTURE BEAT (July 7, 2022), <https://venturebeat.com/data-infrastructure/why-data-has-a-sustainability-problem/>.

97. See MICHAEL HARDT & ANTONIO NEGRI, ASSEMBLY 167 (2017) (“Capitalist industry and commodification have long had destructive effects, but in some respects extractivism today brings that process to a head and a point of no return. *Capital against the earth*—one or the other may survive, but not both.”).

98. Lina M. Khan & Sandeep Vaheesan, *Market Power and Inequality: The Antitrust Counterrevolution and Its Discontents*, 11 HARV. L. & POL'Y REV. 235 (2017).

The following Sections apply this perspective on bigness to the FTC's amended complaint against Meta regarding its acquisitions of WhatsApp and Instagram.

B. A CASE STUDY: *FTC V. META PLATFORMS, INC.*

In its amended complaint, the FTC defines personal social networking as a unique way of maintaining personal connections that encompasses a multitude of modes of interaction.⁹⁹ PSNs are similar to natural monopolies, like energy and telephone networks, “characterized by strong network effects: the value of the service to individual consumers increases with the number of other consumers that use the service.”¹⁰⁰ Also like energy and telecommunications networks, PSNs have come to occupy an indisputably important place in contemporary society.

Although it began as a juvenile tool for ranking the appearances of college classmates,¹⁰¹ Facebook (“FB”) has since become one of a handful of sites hosting virtually all online speech in the United States. It is not difficult to illustrate the place FB has secured in American culture: there are “over 300 million [users] in the United States alone.”¹⁰² In 2020, every single member of Congress posted on Facebook—and they posted to Facebook and Twitter over 2.2 million times just that year.¹⁰³ In 2021 alone, Facebook’s Law Enforcement Response Team (LERT) received nearly 120,000 “legal process requests” according to their own recordkeeping.¹⁰⁴ And Meta even offers a popular marketplace for goods and has increasingly branched out into other financial services.¹⁰⁵ In other words, Facebook is so integrated with social,

99. *Meta Complaint*, *supra* note 13, at 7.

100. *Id.* at 8.

101. See Julia Reinstein, *Mark Zuckerberg Tells Congress: No, Facebook Wasn't Invented to Rank Hot Girls, That Was My Other Website*, BUZZFEED NEWS (Apr. 11, 2018), <https://www.buzzfeednews.com/article/juliareinstein/facemash>.

102. Connor M. Correll, *Facebook, Crime Prevention, and the Scope of the Private Search Post-Carpenter*, 56 GA. L. REV. 787, 787 (2022).

103. Connor Perrett, *Members of the 116th Congress Rail Against Social-media Companies But Posted to Twitter and Facebook a Record 2.2 Million Times*, INSIDER (Jan. 26, 2021), <https://www.businessinsider.com/congress-members-social-media-records-analysis-2021-1>.

104. *Government Requests for User Data - United States*, META, <https://transparency.fb.com/data/government-data-requests/country/US/> (last visited Oct. 11, 2023).

105. See, e.g., Hannah Murphy & Kiran Stacey, *Facebook's Libra: the Inside Story of How the Company's Cryptocurrency Dream Died*, FIN. TIMES (Mar. 10, 2022), <https://www.ft.com/content/a88fb591-72d5-4b6b-bb5d-223adfb893f3>.

economic, and political life in America that some have argued it is a state actor¹⁰⁶ or even operating as a government in and of itself.¹⁰⁷

Despite the outsized societal importance of social media, the market for PSNs is uniquely underregulated and underenforced. Unlike utilities and telecoms that have been federally regulated since the early 20th century, PSNs originated in the mid-2000s at a low point in antitrust enforcement and grew exponentially with the development and widespread adoption of mobile devices¹⁰⁸ throughout the 2010s. Concurrently, regulatory agencies contended with a growing antagonism to their role in a federalist system.¹⁰⁹ Further complicating the issue, the underlying technology behind PSNs is constantly developing and poorly understood by both legislators and the judiciary, while legal barriers such as the First Amendment and § 230 of the Communications Decency Act have made regulation difficult to formulate or enact.¹¹⁰

The need for regulation, however, is clear from the similarities between PSNs and utilities and telecoms.¹¹¹ First, the technology underlying PSNs, and their data-driven business models (DDBMs) make anticompetitive harms difficult to assess and remedies a challenge to administer. Second, the high barriers to entry, including infrastructural demands, give early market entrants an inordinate advantage over new competitors. And third, the social

106. See, e.g., Jed Rubenfeld, *Are Facebook and Google State Actors?*, LAWFARE (Nov. 4, 2019), <https://www.lawfareblog.com/are-facebook-and-google-state-actors>.

107. See, e.g., Editorial Board, *Facebook is Looking a Lot Like a Government*, WASH. POST (Feb. 23, 2020), https://www.washingtonpost.com/opinions/facebook-is-looking-a-lot-like-a-government/2020/02/23/2977a204-53f1-11ea-929a-64efa7482a77_story.html.

108. See *Meta Complaint*, *supra* note 13, at 3.

109. See, e.g., Noah Feldman, *Gorsuch v. the Administrative State Is Really Heating Up*, BLOOMBERG (Jan. 15, 2022), <https://www.bloomberg.com/opinion/articles/2022-01-15/supreme-court-conservative-fight-against-regulatory-state-is-really-heating-up#xj4y7vzkg>.

110. See VALERIE C. BRANNON, CONG. RES. SERV., R45650, FREE SPEECH AND THE REGULATION OF SOCIAL MEDIA CONTENT 15–16 (Mar. 27, 2019), https://www.everycrsreport.com/files/20190327_R45650_9f272501744325782e5a706e2aa76781307abb64.pdf (“[C]ourts have often dismissed lawsuits attempting to hold social media providers liable for regulating users’ content, whether because the court concludes that the First Amendment does not apply to the actions of these private actors or because the court holds that Section 230(c)(2) of the CDA bars the lawsuit Particularly because of Section 230, there are few, if any, federal or state laws that expressly govern social media sites’ decisions about whether and how to present users’ content.”).

111. For a general understanding of the debate regarding PSNs as infrastructure and thus as comparable to other natural monopolies like electricity and telecommunications, compare Luigi Zingales, *The Silent Coup*, PROMARKET (Jan. 11, 2021), <https://www.promarket.org/2021/01/11/facebook-twitter-ban-trump-parler-concentration-power/>, with Carlo Amenta, Michele Boldrin & Carlo Stagnaro, *Digital Platforms May Be Monopolistic Providers, But They Are Not Infrastructure*, PROMARKET (Jan. 26, 2021), <https://www.promarket.org/2021/01/26/digital-platforms-monopolistic-infrastructure-free-speech/>.

essentiality of PSNs' speech-hosting functions creates a hefty public interest in favor of government oversight.

From an antitrust perspective, the PSN market is complicated because it is two-sided. Two-sided markets, also known as multi-sided platforms (MSPs), are organizations that “have two key features beyond any other requirements (such as indirect network effects or non-neutrality of fees): [t]hey enable direct interactions between two or more distinct sides and [e]ach side is affiliated with the platform.”¹¹² While the classic example of an MSP is a credit card company, this type of business model is increasingly common as a result of the internet and targeted advertising. PSNs like Meta put the majority of their resources into increasing and tracking user engagement so that they can sell their attention to advertisers and app developers. PSNs are designed to maximize the interactions between these three groups and thus facilitate third-party transactions. Users theoretically benefit because the networking services remain “free,” while advertisers and developers benefit through the volume of consumers they can reach in a maximally efficient way.

However, as the FTC articulates in their complaint, a lack of market oversight has enabled Meta to acquire monopolistic control through its anticompetitive purchases of WhatsApp and Instagram.¹¹³ And these harms are serious precisely because Meta has come to occupy such an important role for consumers, advertisers, and developers. “Data is the new oil,” as mathematician Clive Humby quipped in 2006.¹¹⁴ A PSN like Meta generates, stores, and analyzes more personal data than almost any other business—data provided for free by users who pay in attention instead. And that attention, refined from the raw masses of data, is invaluable to anyone trying to sell a product, service, or idea. According to the FTC's narrative, Meta struggled to maintain their hold on this data and keep up with competitors as mobile devices proliferated during the 2010s and changed the ways and degree to which people interact online.¹¹⁵ Old data is practically useless for sales, so Meta must maintain constant streams of fresh data to attract advertisers, better target user attention, and further refine their insight-generating algorithms.¹¹⁶

112. Andrei Hagiu & Julian Wright, *Multi-Sided Platforms*, 43 INT'L J. INDUS. ORG. 162, 163 (2015).

113. *Meta Complaint*, *supra* note 13, at 25–26.

114. Jo Ann Barefoot, *The Case for Placing AI at the Heart of Digitally Robust Financial Regulation*, BROOKINGS (May 24, 2022), <https://www.brookings.edu/research/the-case-for-placing-ai-at-the-heart-of-digitally-robust-financial-regulation/>.

115. *Meta Complaint*, *supra* note 13, at 3.

116. See, e.g., Robert Springer, *Data Is Useless Without Meaning: The Importance of Insight*, TILT (Sept. 9, 2019), <https://www.thetilt.com/content/data-meaning-insight> (“Digital marketers are not regularly cleaning their customer data, which the InfoGroup report says should be

To overcome these struggles in the market, the FTC alleges that Meta engaged in two specific types of anticompetitive conduct. First, they leveraged the size of their network and access to the interoperability of their Application Programming Interface (API)¹¹⁷ to attract third party developers.¹¹⁸ But once those developers were reliant on the system and Meta had derived extensive benefits from the user base they attracted and locked in, “Facebook imposed several other policies restricting app developers’ use of Facebook Platform, including Facebook APIs.”¹¹⁹ As a result, “[w]ith the implementation of these anti-competition policies, developers who had relied on Facebook’s expressions of openness suddenly found themselves targeted by Facebook.”¹²⁰ The court dismissed this part of the complaint because Meta ended these anticompetitive policies, and § 2 of the Sherman Act cannot be applied retroactively. However, it survived as factual support for the other claim.¹²¹

Second, Meta strategically identified threats to their market dominance and then acquired those threats in order to build a digital moat around itself.¹²² To do so, Meta first acquired Onavo, a firm that “marketed itself to users as providing secure virtual private networking services, but—unknown to many users—it also tracked users’ activity online.”¹²³ Using the intelligence gathered from this service, Meta would identify targets and then “acquire a potential rival and keep the rival’s mechanics deployed to frustrate others’ efforts to gain scale using similar mechanics.”¹²⁴ This is how Meta came to own both Instagram and WhatsApp, two apps that threatened Meta’s growth by innovating in the areas of photo-sharing and mobile-messaging respectively. At the same time, Meta also acquired Octazen (a contact importing service)

done weekly or at least monthly. Despite that, ‘it is a pervasive problem in the industry to see large companies sitting on years of inactive files,’ the report states.”).

117. An API is software that facilitates interaction between two or more computer programs.

118. *Meta Complaint, supra* note 13, at 14.

119. *Id.* at 45.

120. *Id.*

121. *Fed. Trade Comm’n v. Facebook, Inc.*, 581 F. Supp. 3d 34, 60–61 (D.D.C. 2022) (“The question therefore is what to do with Count II: should the Court dismiss the portion that encompasses challenges to the Platform policies, or must it allow the count to remain given its incorporation of the acquisitions? The Court concludes that the latter is the better course, with an important caveat In the meantime, the Court will not award the FTC a discovery windfall for using Count II as a Trojan horse to smuggle in the Platform policies. Instead, it will not permit what would certainly be time-consuming and costly discovery on such policies.”).

122. *Meta Complaint, supra* note 13, at 34.

123. *Id.* at 23.

124. *Id.* at 24.

and Glancee (a geolocation service).¹²⁵ These two acquisitions allowed Meta to further cut off competitors from vital growth services and dig an even deeper moat around itself.

Meta's control of the market through these acquisitions can be concretely assessed in a few ways. To measure Meta's dominance in the market, the FTC analyzed three main metrics: time spent, daily active users (DAUs), and monthly active users (MAUs).¹²⁶ Not only are these similar to the metrics by which Meta judges its own performance and those used in antitrust analysis abroad, they are also indicative of the core mechanism of the PSN business model: captured attention. The value of a PSN comes from both the volume of users it can attract through the network effect and from the ability to keep each of those users engaged with the platform for as long as possible each day. As Tim Wu writes, "Zuckerberg . . . understood advertising's potential to degrade his product . . . the Holy Grail was advertising that people actually wanted to see; Facebook figured that nanotargeting could make that happen."¹²⁷ By monopolizing users and their attention—and keeping competitors from doing the same—Meta was able to dominate the targeted advertising space.

According to the FTC, the reason why this behavior is problematic is threefold: Meta's "better to buy than compete" strategy deprived consumers and advertisers of innovation, quality improvements, and choice.¹²⁸ By acquiring existing companies to deter competition rather than to improve the user experience, Meta has slowly stripped WhatsApp and Instagram of the qualities that once made them appealing to users, transforming their features into whatever will help quash the latest new competitor.¹²⁹ In consolidating all these services into one company, Meta has also subjected the users of Instagram and WhatsApp to the same infrastructure as Meta, making them all more vulnerable to simultaneous service outages and privacy breaches.¹³⁰ And by dominating the market in this unethical manner, Meta has also denied consumers and advertisers greater choice of services and privacy protection.¹³¹

While the FTC's complaint stops there, Meta's domination of attention is also a threat to democracy. As illustrated by the proliferation of mass manipulation and disinformation campaigns run during elections around the

125. *Id.* at 24–25.

126. *Id.* at 61.

127. TIM WU, THE ATTENTION MERCHANTS: THE EPIC SCRAMBLE TO GET INSIDE OUR HEADS 296–97 (2016).

128. *Meta Complaint*, *supra* note 13, at 1–2.

129. *Id.* at 42.

130. *Id.* at 42–43.

131. *Id.* at 73.

world over the past decade,¹³² Meta’s massive user base enables private interest groups and political actors to easily target vulnerable audiences. This is a risk involved with any social media or social network, but the sheer scale of Meta combined with the deeply personal data it collects make it particularly appealing to malicious actors.

As noted in the 2021 Annual Threat Assessment of the U.S. Intelligence Community, “[f]oreign states use cyber operations to steal information, influence populations, and damage industry, including physical and digital critical infrastructure.”¹³³ Because FB, WhatsApp, and Instagram all share the same infrastructure and provide access to over a billion people around the world, it is an appealing one-stop-shop for anyone buying or selling influence. It is also a major stress point for malicious actors to halt global communications, as evidenced by the global outcry every time Meta products (and all the third-party apps that rely on it for log-in functionality) crash simultaneously.¹³⁴ Meta even settled a class action lawsuit in 2022 as a result of their role in Cambridge Analytica’s mass manipulation of voters in 2016.¹³⁵ Based on all this, it is clear that the sheer size and scale of Meta is the biggest part of the problem.

But how does the government remedy a problem of this scale? Antitrust is often criticized as being difficult to administer, especially when it comes to natural monopolies. As Richard Posner wrote in 1968, a natural monopoly is defined by “the relationship between demand and the technology of supply.”¹³⁶ For utilities and telecoms, the technologies in question are the vast infrastructural networks and machinery required to operate at scale—much of which requires specialized expertise to understand. PSNs are not so different, and there is little overlap between those with internet savvy and members of the judiciary or Congress. As of 2020, the average age of a federal judge was

132. See Report: *Digital Election Interference Widespread in Countries Across the Democratic Spectrum*, FREEDOM HOUSE (Dec. 7, 2020), <https://freedomhouse.org/article/report-digital-election-interference-widespread-countries-across-democratic-spectrum>.

133. OFF. DIRECTOR NAT’L INTELLIGENCE, ANNUAL THREAT ASSESSMENT OF THE U.S. INTELLIGENCE COMMUNITY 20 (Apr. 9, 2021), <https://www.dni.gov/files/ODNI/documents/assessments/ATA-2021-Unclassified-Report.pdf>.

134. See Rhona Ascierio, *Too Big to Fail? Facebook’s Global Outage*, DATA CTR. DYNAMICS (Oct. 15, 2021), <https://www.datacenterdynamics.com/en/opinions/too-big-to-fail-facebooks-global-outage/> (“Facebook’s hours long outage on October 4th snarled completely unrelated applications globally, underscoring the criticality—and fragility—of publicly shared digital infrastructure.”).

135. James Vincent, *Meta Agrees to Pay \$725 Million to Settle Cambridge Analytica Class Action Lawsuit*, VERGE (Dec. 23, 2022), <https://www.theverge.com/2022/12/23/23523862/meta-cambridge-analytica-class-action-lawsuit-settlement-725-million>.

136. Richard Posner, *Natural Monopoly and Its Regulation*, 21 STAN. L. REV. 548, 548 (1968).

about sixty-nine.¹³⁷ As of 2021, the average age of a congressman was fifty-eight in the House and sixty-four in the Senate.¹³⁸ And in the same year, 2021, only 50% of U.S. adults over the age of sixty-five used Facebook and YouTube, and no more than 20% used any of the other major social media sites.¹³⁹

This general lack of understanding of PSNs and their business models exacerbates a long-standing problem of antitrust law: intervention should account for the specificities of the market in question. Under traditional principles of antitrust analysis, courts consider the particularities of the industry and adjust application of the law to the circumstances including the existence of regulation.¹⁴⁰ An unproblematic behavior in one market may be anticompetitive in another, depending on factors like market share, price regulation, and entry barriers. One answer to this problem has been to shift enforcement away from antitrust by enacting regulation with some antitrust-like functions such as the 1996 Telecommunications Act.

But how can Congress even begin to regulate something as global and unmanageable as a network used by over two billion people each month?¹⁴¹ When questioned by Congress about the difficulties of moderation in 2018, Zuckerberg acknowledged that, “this is an arms race, right? . . . which is why one of the things I mentioned before is we’re going to have more than 20,000 people, by the end of this year, working on security and content review across the company.”¹⁴² With only 20,000 people moderating a network of nearly two billion, it’s no surprise that problems slip through the cracks—problems that could be avoided by shrinking Meta’s scale while simultaneously growing its human moderation.

Fortunately, the remedy suggested by the FTC is simple and easily administered: divestiture of WhatsApp and Instagram.¹⁴³ While FB will still be a substantial and unwieldy PSN in and of itself, the divestiture of both WhatsApp and Instagram will substantially reduce Meta’s size and allow the

137. Francis Shen, *Aging Judges*, 81 OHIO ST. L.J. 235, 235 (2020).

138. *117th United States Congress: A Survey of Books Written by Members*, LIBR. CONG. (June 15, 2022), <https://guides.loc.gov/117th-congress-book-list> (discussing the average age of Congress members).

139. Brooke Auxier & Monica Anderson, *Social Media Use in 2021*, PEW RES. CTR. (Apr. 7, 2021), <https://www.pewresearch.org/internet/2021/04/07/social-media-use-in-2021/>.

140. See generally Posner, *supra* note 136.

141. See *Transcript of Mark Zuckerberg’s Senate Hearing*, WASH. POST (Apr. 10, 2018), <https://www.washingtonpost.com/news/the-switch/wp/2018/04/10/transcript-of-mark-zuckerbergs-senate-hearing/> (questioning Zuckerberg on his ability to manage a platform with two billion users).

142. *Id.*

143. *Meta Complaint*, *supra* note 13, at 79.

two platforms to move away from Meta’s infrastructure. Moreover, severing them from Meta’s ownership will end the anticompetitive “digital moat” conduct by which Meta has unfairly prevented innovation in mobile photo sharing and messaging. The most complex part of the divestiture will be the infrastructural issues, but that is not a problem for the courts—it is a prompt to establish or designate an appropriate regulatory body to administer the remedy. And even if WhatsApp and Instagram don’t survive the separation, at least the PSN market will have space for new firms to compete and offer users better privacy protections, better APIs, and unique ways to share content and connect with others.

IV. AN OPPORTUNITY FOR RECONCILIATION

The FTC’s case against Meta provides a unique opportunity to assess the proper relationship between antitrust and regulation in an industry emblematic of new, technology-driven anticompetitive behavior. The difference between how the court treated each of the FTC’s claims indicates the boundaries of where antitrust’s reach ends and the need for regulation begins. By assessing these differences and the appropriate boundaries between antitrust and regulation, it is possible to reconcile the two and provide a concrete example of how a regulatory framework can not only survive antitrust intervention but become stronger as a result.

First, it is necessary to address consumer harms that antitrust cannot reach with some form of regulation. This is because interoperability issues, like shutting developers out of an API on which they have become reliant, are not typically redressable via antitrust enforcement. According to Shelanski and William Rogerson, there are three main reasons why regulation can effectively supplement antitrust for digital platforms: (1) antitrust enforcement has been targeted by “well-founded criticism”; (2) regulation offers different and potentially more effective tools than antitrust; and lastly, (3) “because of network effects, conduct that courts ordinarily judge under antitrust law’s general rule of reason might have different presumptive effects, and therefore be better governed by a more specific set of standards, in digital platform industries.”¹⁴⁴ While the first point is questionable given the unfounded paranoia discussed above, the next two points are important to explore further.

Because PSNs have strong network effects and high infrastructural barriers,¹⁴⁵ there is a certain amount of scale that will always be necessary for

144. William P. Rogerson & Howard Shelanski, *Antitrust Enforcement, Regulation, and Digital Platforms*, 168 U. PA. L. REV. 1911, 1914–15 (2020).

145. *See id.*

one to function. In fact, it can be difficult to imagine a social networking market composed of smaller, more localized firms because the internet is inherently global and detached from locality. But that does not mean such an exercise is not worthwhile. Shelanski and Rogerson extol the potential benefits of what they call “light-handed pro-competitive” (LHPC) regulation that “could include interconnection/interoperability requirements (such as access to application programming interfaces (APIs)), limits on discrimination, both user-side and third-party-side data portability rules, and perhaps additional restrictions on certain business practices subject to rule of reason analysis under general antitrust statutes.”¹⁴⁶ While these are all reasonable suggestions that could be included in any potential future regulatory framework, they are not enough without antitrust intervention first because none of these solutions target bigness.

This is where it becomes clear that antitrust and regulation must be redefined in relation to each other. When antitrust is viewed from a “consumer welfare” perspective, regulation seems sufficient to handle any problems that could arise and antitrust feels inappropriate because it is a blunt tool in comparison. As stated at the beginning of this Note, antitrust isn’t meant to protect consumers like regulation does, it is meant to protect competition.

Regulation is perfectly adequate at protecting consumers in some ways, such as mandated sharing of infrastructure, but it can’t break up a behemoth into regulatable size. Antitrust arose in the context of federalism and was intended to prevent state monopolies from becoming national monopolies,¹⁴⁷ but underenforcement prevented that goal from being achieved. Now, national monopolies have become international monopolies that not only harm U.S. competitors, but that stall the development of other nations’ internet innovations. Renewed antitrust enforcement is therefore critical because antitrust must come first for regulation to be effective and then exist concurrently to keep regulation effective.

Moreover, not only is the need for antitrust enforcement clear—the possibility is quite feasible as well. The infrastructure of PSNs is distinct from that of energy or telecommunications because it is virtual, and it is more divisible in some ways. If API interoperability is regulated, the storage of data becomes the biggest issue. And even though the internet is not localized, data is. Largely as a result of jurisdictional evidence collection issues, it has become

146. *Id.* at 1915.

147. *See* Crane, *supra* note 14 (discussing antitrust’s role as a check on state-sanctioned monopolies).

common practice for internet firms to store data locally.¹⁴⁸ This means that a Baby Bell-esque break down of a company like Meta would actually be easier to administer than one might think, so long as a regulatory framework is in place to maintain interoperability.

To summarize, the current state of PSNs makes the proper relationship between antitrust and regulation simple. If antitrust can break up a market to prevent power concentration and facilitate more effective oversight, then regulation can take over to protect consumers from other harms. When regulation begins to falter, as it did in the area of telecommunications, antitrust can step in again to address concentration and urge regulation changes to prevent that concentration from building *in that manner*. For this to work, both antitrust and regulation must be constantly vigilant and simultaneously engaged in monitoring a given market. Antitrust is a check on poorly functioning regulation and on regulatory capture, and it bluntly rebalances markets so regulation can be more effective. A renewed investment in antitrust enforcement is also necessary to address the now-global scale of monopolies like Meta.

V. CONCLUSION

The current state of the PSN market clarifies this symbiotic relationship and its importance—and provides an opportunity to establish that relationship as the norm. It is unlikely that a comprehensive regulatory framework for PSNs will exist any time soon because there are larger political splits over how and what content should be regulated. But if any framework is ever proposed, it must include an antitrust saving clause that goes further than the Telecommunications Act of 1996. Instead of just preserving antitrust actions, it should explicitly proscribe the appropriate and continued role of antitrust enforcement as a check on regulation and the ways in which such regulation can entrench certain firms into monopolistic power.

Moreover, although the courts have been hostile to antitrust enforcement over the past fifty years, the Supreme Court has shifted so far towards being equally hostile towards Big Tech that a rehabilitation of the saving clause may now be possible. Justice Breyer and Justice Scalia, two of the biggest influences on the withdrawal of antitrust from regulated markets, are no longer on the Court. And Justice Thomas, a staunch advocate for antitrust saving clauses

148. See Erol Yayboke, Carolina G. Ramos & Lindsey R. Sheppard, *The Real National Security Concerns Over Data Localization*, CTR. FOR STRATEGIC & INT'L STUD. (July 23, 2021), <https://www.csis.org/analysis/real-national-security-concerns-over-data-localization> (“[G]overnments are increasingly seeking to maintain ‘digital sovereignty’ and control through protectionist data localization mandates”).

despite his general hostility to antitrust, has more power than ever before as the most senior member. When reframed in an originalist light as a *Lochnerian* check on state-sanctioned monopolies, antitrust can be made appealing even to the most anti-government of libertarians.¹⁴⁹

Deregulation and poorly functioning regulation must be recognized as the means through which state-sanctioned monopoly power operates. The solution is increased antitrust enforcement. As unchecked monopoly power now reaches a global scale, support from across the political spectrum has never been more necessary to return antitrust enforcement to its rightful place in both regulated and unregulated markets as a check on the regulatory state.

149. Crane, *supra* note 14, at 513.

PROTECTING THE COMPETITIVE PROCESS IN VERTICAL MERGER

Zhudi Huang[†]

TABLE OF CONTENTS

I.	INTRODUCTION	1406
II.	VERTICAL MERGERS ENFORCEMENT IN THE UNITED STATES.....	1409
III.	THE UNSATISFYING HISTORICAL APPROACHES TO VERTICAL MERGER ENFORCEMENT	1412
A.	THE STRUCTURALIST APPROACH.....	1413
B.	THE CHICAGO SCHOOL’S WELFARE-BASED APPROACH	1414
1.	<i>The Early Chicago School, Efficiencies, and the 1982 and 1984 Merger Guidelines</i>	1415
2.	<i>The Post-Chicago School, Consumer Welfare Standard and the 2020 Vertical Merger Guidelines</i>	1418
C.	CRITIQUE OF CURRENT VERTICAL MERGER ENFORCEMENT.....	1421
IV.	VERTICAL MERGER REFORM.....	1423
A.	CRITIQUES OF A RETURN TO THE STRUCTURALIST APPROACH	1424
B.	THE TWO STANDARDS OF PROTECTING COMPETITION: TRADING PARTNER WELFARE STANDARD AND PROTECTION OF THE COMPETITIVE PROCESS	1426
C.	PROTECTION OF COMPETITIVE PROCESS.....	1429
1.	<i>Framework for Assessing the Competitive Process</i>	1430
2.	<i>The Competitive Process Standard is Better Suited to Protect Competition</i>	1431
a)	The Competitive Process Serves Both Consumers and Competitors.....	1431
b)	Welfare Cannot Capture Dynamic Harms, at Least Not Without Sacrificing Administrability	1433
V.	CONCLUSION.....	1435

DOI: <https://doi.org/10.15779/Z380C4SM2W>

© 2023 Zhudi Huang.

† J.D. Candidate, 2024, University of California, Berkeley, School of Law. Many thanks to Professor Talha Syed and my fellow students in the Law & Technology Writing Workshop.

I. INTRODUCTION

*“Antitrust is sexy again.”*¹

With the rise of tech giants, public discourse over their concentrated economic power is more vibrant than ever. Many have claimed that America has a competition problem.² However, the evidence of rising industrial concentration and whether economic concentration in fact indicates a decline in competition is still inconclusive.³ Nonetheless, antitrust enforcers, commentators, and politicians alike are taking a closer look at antitrust enforcement, particularly in the context of mergers. Reformers of antitrust law have argued that merger enforcement has been “overly lax” and needs to be invigorated.⁴

The contests over merger enforcement are attributable to the antitrust statutes’ open-ended articulation of competition and the predictive nature of merger enforcement. While § 7 of the Clayton Act expressly condemns mergers that may substantially lessen competition,⁵ none of the statutes define what “competition” means.⁶ Thus, since the enactment of the Sherman Act in 1890,⁷ misplaced debates over the proper goal of antitrust—rather than

1. Carl Shapiro, *Antitrust in a Time of Populism*, 61 INT’L J. INDUS. ORG. 714, 714 (2018).

2. See, e.g., John Mauldin, *America Has a Monopoly Problem*, FORBES (Apr. 11, 2019), <https://www.forbes.com/sites/johnmauldin/2019/04/11/america-has-a-monopoly-problem/?sh=4b9877e42972>; Lina Khan, *The New Brandeis Movement: America’s Antimonopoly Debate*, 9 J. EUR. COMPETITION L. & PRAC. 131 (2018); Carl Shapiro, *Protecting Competition in the American Economy: Merger Control, Tech Titans, Labor Markets*, 33 J. ECON. PERSPECTIVES No. 3, 2019, at 69.

3. *Compare Does America Have a Monopoly Problem?: Examining Concentration and Competition in the US Economy: Hearing Before the S. Comm. on Antitrust, Competition Pol’y, & Consumer Rts.*, 116th Cong. (2019) (statement of Robert B. Reich, Carmel P. Friesen Professor of Public Policy, University of California, Berkeley), <https://www.judiciary.senate.gov/imo/media/doc/Reich%20Testimony.pdf> (arguing that the consolidation in American economy is contributing to problems beyond the narrow concepts of consumer welfare), *with Does America Have a Monopoly Problem?: Examining Concentration and Competition in the US Economy: Hearing Before the Subcomm. on Antitrust, Competition Pol’y, & Consumer Rts.*, 116th Cong. (2019) (statement of A. Douglas Melamed, Professor, Stanford Law School), <https://www.judiciary.senate.gov/imo/media/doc/Melamed%20Testimony.pdf> (arguing that there is no concrete evidence that America has a market power issue).

4. See, e.g., Shapiro, *supra* note 2, at 70 (“The clearest area where antitrust enforcement has been overly lax is the treatment of mergers.”); TIM WU, *THE CURSE OF BIGNESS: ANTITRUST IN THE NEW GILDED AGE* 127 (“The priority for Neo-Brandeisian antitrust is the reform of merger review.”).

5. 15 U.S.C. § 18.

6. Herbert Hovenkamp & Carl Shapiro, *Horizontal Mergers, Market Structure, and Burdens of Proof*, 127 YALE L.J. 1996, 2030 (2018).

7. 15 U.S.C. §§ 1–38.

healthy conversations about how to achieve the goal—have dominated the public discourse. While economists and legal scholars have clustered into different ideologies, with each claiming a different goal for antitrust law, these debates are simply a red herring.⁸ The statutory mandate of antitrust law is clear: the goal is to preserve and promote competition. A misplaced focus on the goal of antitrust consequently results in commentators relying on distorted legislative history, limited economic theory, and various political agendas to purport their self-reinforcing interpretations. But perhaps most dangerously, it prevents a discussion of the deeper normative values underpinning antitrust law that balance the “need for protecting individualism and community . . . in the private economic sphere.”⁹

The language of the Sherman Act and Clayton Act is intentionally broad to assert competition as the “preferred governor of markets”¹⁰ while allowing for debates as to the means to measure and achieve the goal. The legislative history of the Sherman Act reveals various concerns regarding the statute, some economic while others social and political.¹¹ But the goal of antitrust laws is neither to promote market efficiencies, nor to promote wealth equality, nor to tackle private political power. Instead, antitrust law reflects a careful balancing between competing concerns through “the preservation of free and fair competition or trade.”¹² Granted, this framing does not answer the precise questions of what conduct constitutes competition on the merits. But defining “competition” is difficult precisely because “competition” refers to a process rather than a result.¹³ The goal of antitrust is to safeguard the dynamic, robust

8. See, e.g., Barak Orbach, *Antitrust's Pursuit of Purpose, Foreword to Symposium: The Goals of Antitrust*, 81 FORDHAM L. REV. 2151, 2153 (2013) (introducing the debate on the goals of antitrust and noting the “unproductive nature of the debate”); Eleanor M. Fox, *Against Goals*, 81 FORDHAM L. REV. 2157, 2159 (“The typical framing of the debate on the goals of U.S. antitrust law is misleading.”).

9. John J. Flynn, *Antitrust Policy and the Concept of a Competitive Process*, 35 N.Y. L. SCH. L. REV. 893, 898 (1990) (“Antitrust policy plays a fundamental part in defining the scope of property rights in our society by balancing the rights of individuals and communities in the private economic sphere.”).

10. Eleanor M. Fox, *The Modernization of Antitrust: A New Equilibrium*, 66 CORNELL L. REV. 1140, 1153 (1981).

11. See generally Christopher Grandy, *Original Intent and the Sherman Antitrust Act: A Re-examination of the Consumer-Welfare Hypothesis*, 53 J. ECON. HIST. 359 (1993) (rejecting a consumer welfare goal of antitrust).

12. *Id.* at 363.

13. Fox, *supra* note 10, at 1154 (arguing that the competition process is the “preferred governor of markets,” and that competition as a process has unified three major concerns in antitrust law, them being distrust of power, concern for consumers, and commitment to opportunity of entrepreneurs); see also Flynn, *supra* note 9, at 896 (noting the importance of

competitive process itself, as opposed to ordaining certain market results.¹⁴ This is far from a novel idea. Indeed, most contemporary antitrust scholars can agree that the principal function of antitrust enforcement is to preserve the competitive process.¹⁵ Where they differ are the correct policies and standards to achieve that goal.

This Note begins by clarifying and reasserting that the sole goal of antitrust, as mandated by statute and interpreted by the Supreme Court, is to regulate between business decisions that are part of competition and those that “suppress or even destroy competition.”¹⁶ If antitrust reform is to proceed, the question should be framed as: what is the best method to identify conduct that does not compete on the merits and thereby harms the competitive process. This framing is necessary because competition is not static. The role of the “market” and regulations in our political economy are reflections of our deep, complex societal values and should be informed by progress in economics, social sciences, and technologies.¹⁷ There is no short shrift to these substantive and normative questions. This Note does not aim to resolve the underlying normative debates. Rather, it evaluates different approaches to antitrust enforcement as different proxies to competition. The best approach to

understanding the concept of competition as “competition as a process,” and deriving a multi-disciplinary meaning of it).

14. Fox, *supra* note 8, at 2160.

15. See, e.g., Carl Shapiro, *Antitrust: What Went Wrong and How to Fix It*, 35 ANTITRUST, no. 3, 2021, at 33, 33 (“Part of my thesis today is that the goal of antitrust law should be to protect and promote competition. Period.”); A. Douglas Melamed & Nicholas Petit, *Before “After Consumer Welfare” – A Response to Professor Wu*, COMPETITION POL’Y INT’L (July 1, 2018), https://www.pymnts.com/cpi_posts/before-after-consumer-welfare-a-response-to-professor-wu/ (“As will be seen, both elements of the antitrust offense in the [consumer welfare] paradigm are about ‘protecting a process.’”); A. Douglas Melamed, *Antitrust Law Is Not That Complicated*, 130 HARV. L. REV. F. 163, 166 (2016) (“With a couple of refinements, U.S. antitrust law makes it illegal to cause an increase in market power by conduct that is not competition on the merits.”); Herbert Hovenkamp, *The Slogans and Goals of Antitrust Law*, 25 N.Y.U. J. LEGIS. & PUB. POL’Y, at 89 (forthcoming 2023), <https://ssrn.com/abstract=4121866> (“Antitrust is properly focused on competition.”); Frank H. Easterbrook, *Limits of Antitrust*, 63 TEX. L. REV. 1, 1 (1984) (“The goal of antitrust is to perfect the operation of competitive markets.”); Einer Elhauge, *Should the Competitive Process Test Replace the Consumer Welfare Standard?*, PROMARKET (May 24, 2022), <https://www.promarket.org/2022/05/24/should-the-competitive-process-test-replace-the-consumer-welfare-standard/> (“Kanter is right that antitrust law protects ‘competition and the competitive process.’”); Fox, *supra* note 8 (explaining the substantial consensus of the goal of antitrust to be a robust market).

16. *Bd. of Trade v. United States*, 246 U.S. 231, 238 (1918).

17. See Flynn, *supra* note 9, at 898 (noting that antitrust law requires “a deeper, more sophisticated understanding of the normative values underlying antitrust policy, contract law, property law, and various schools of economic and political thought.”).

antitrust enforcement should balance the administrability concerns against the need for normative discussions.

This Note focuses the debates of antitrust enforcement on vertical mergers. Vertical mergers, those “that combine firms or assets at different stages of the same supply chain,”¹⁸ do not directly eliminate competitors and, therefore, present unique and hotly contested considerations in analyzing their competitive effects. Part II of this Note examines the unique role of agencies in merger enforcement in the United States, and then the unique considerations for vertical merger analysis. Part III traces back the history of vertical merger enforcement, and then outlines the drastic changes in merger policy—with their implications for contemporary ideologies—and offers critiques to the *status quo*. Part IV analyzes potential approaches to process-based antitrust reform. It first rejects a return to the structuralist approach, and then discusses the differences between two purported standards that each claim to protect the competitive process. The final Part, Part V, asserts that antitrust law must protect competition as a process to align the law and regulations with antitrust’s goal of promoting robust competition.

II. VERTICAL MERGERS ENFORCEMENT IN THE UNITED STATES

Mergers can be divided into horizontal mergers and non-horizontal mergers.¹⁹ Horizontal mergers involve mergers between actual or potential direct competitors.²⁰ Therefore, the potential anticompetitive harm of horizontal mergers arises from the direct elimination of competitors because an increase in market share post-merger can directly influence firms’ competitive incentives.²¹ In contrast, non-horizontal mergers, which include vertical, diagonal, and conglomerate mergers, have indirect impacts on competition.²²

Vertical mergers are particularly tricky due to their efficiency-enhancing nature. On one hand, vertical mergers can have inherent efficiency gains from

18. *Id.*

19. 2010 *Horizontal Merger Guidelines*, U.S. DEP’T JUSTICE & FED. TRADE COMM’N, <https://www.justice.gov/atr/horizontal-merger-guidelines-08192010> (Aug. 19, 2010) [hereinafter 2010 Horizontal Merger Guidelines]; 2020 *Vertical Merger Guidelines*, U.S. DEP’T JUSTICE & FED. TRADE COMM’N, https://www.ftc.gov/system/files/documents/reports/us-department-justice-federal-trade-commission-vertical-merger-guidelines/vertical_merger_guidelines_6-30-20.pdf (2020) [hereinafter 2020 Vertical Merger Guidelines].

20. 2010 Horizontal Merger Guidelines, *supra* note 19, § 1.

21. *Id.* § 5.

22. 2020 Vertical Merger Guidelines, *supra* note 19, § 1.

the elimination of double marginalization (EDM).²³ EDM occurs when the upstream firm transfers input at marginal cost instead of a marked-up price premerger.²⁴ Therefore, when the input supplier and the output producer are merged, the integrated firm can efficiently supply input to itself and thereby eliminate one of two markups.²⁵ On the other hand, vertically integrated downstream firms also have “an inherent exclusionary incentive” against unintegrated downstream competitors to preclude supplies.²⁶ Thus, despite the inherent efficiencies gained from vertical mergers, not all approaches consider them as part of the competitiveness analysis. Even for those who agree on efficiency as a competitive benefit, there is no clear consensus as to how to factor for these potential efficiencies in merger analysis.

Merger analysis considers efficiency claims in two ways.²⁷ First, efficiencies may be considered in the prima facie case. That is, efficiencies may be part of the inquiry of whether a given merger would have an anticompetitive effect in a given market.²⁸ Second, it has been argued that out-of-market efficiencies should be credited as merger benefits even after the plaintiff has established their prima facie case.²⁹ That is, efficiencies can be viewed as an affirmative defense to anticompetitive harm if the efficiencies are substantial enough.³⁰ At the heart of these debates are three fundamental questions: First, how would a vertical merger harm the competitive process? Second, how would a vertical merger benefit or strengthen the competitive process? And third, how should the agencies and courts balance the potential harms and benefits of the merger, if both exist? The courts have yet to give satisfying answers to these questions, partially due to their lack of expertise and the piecemeal nature of common law merger jurisprudence.³¹ While the Supreme Court has rejected efficiencies

23. Steven C. Salop, *Invigorating Merger Enforcement*, 127 YALE L.J. 1962, 1970 (2018).

24. *Id.*

25. Michael H. Riordan & Steven C. Salop, *Evaluating Vertical Merger: A Post-Chicago Approach*, 63 ANTITRUST L.J. 513, 526 (1995).

26. Jonathan B. Baker, Nancy L. Rose, Steven C. Salop & Fiona Scott Morton, *Five Principles for Vertical Merger Enforcement Policy*, 33 ANTITRUST, no. 3, 2019, at 12–13.

27. *See* Robert D. Willig, Steven C. Salop & F.M. Scherer, *Merger Analysis, Industrial Organization Theory, and Merger Guidelines*, 1991 BROOKINGS PAPERS ON ECON. ACTIVITY MICROECON. 281, 290 (1991), for a different framing of the two ways efficiencies enter into the analytic process.

28. Herbert Hovenkamp, *Appraising Merger Efficiencies*, 24 GEO. MASON L. REV. 703, 706 (2017).

29. PHILLIP E. AREEDA & HERBERT HOVENKAMP, ANTITRUST LAW: AN ANALYSIS OF ANTITRUST PRINCIPLES AND THEIR APPLICATION ¶ 1040d (5th ed., 2022 Cum. Supp. 2015–2021).

30. *Id.*

31. *See* Hillary Greene, *Guideline Institutionalization: The Role of Merger Guidelines in Antitrust Discourse*, 48 WM. & MARY L. REV. 771, 775 (2006) (noting that the common law piecemeal

as an affirmative defense,³² merger enforcement has deviated from the early courts' skeptical views on efficiencies claims.³³

The antitrust enforcement agencies play a crucial role in shaping the standard for antitrust enforcement. In the United States, § 7 of the Clayton Act expands on the Sherman Act of 1890³⁴ and prohibits mergers whose effect “may be substantially to lessen competition, or tend to create a monopoly.”³⁵ With the addition of the Federal Trade Commission Act of 1914,³⁶ the Antitrust Division of the Department of Justice (DOJ) and the Federal Trade Commission (FTC) have joint authority to arrest anticompetitive mergers in their incipiency.³⁷ If the agency decides that the merger raises competition concerns, it may work with the parties to resolve the issues by entering into a negotiated consent agreement with provisions that will cure the competition concerns.³⁸ Alternatively, the agency may seek to stop the transaction by filing for a preliminary injunction in federal court pending a full examination of the proposed deal in an administrative proceeding.³⁹ Most mergers and acquisitions are able to proceed without much intervention from the agencies, and only a few mergers are litigated in court.⁴⁰

Since 1968, the DOJ, later joined by the FTC, began to issue “Merger Guidelines” that outlined the agencies’ analytical techniques and enforcement policies to determine whether to challenge a merger.⁴¹ Though not binding,

nature of the antitrust jurisprudence and reliance on agencies’ action have resulted in increasing reliance on agency guidelines).

32. *F.T.C. v. Procter & Gamble Co.*, 386 U.S. 568, 580 (1967) (stating that possible economies cannot defend against illegality, and that Congress struck the balance in favor of protecting competition over some competition-harming mergers creating economies).

33. Hovenkamp, *supra* note 28, at 706.

34. 15 U.S.C. §§ 1–38.

35. 15 U.S.C. § 18.

36. 15 U.S.C. §§ 41–58.

37. Hovenkamp, *supra* note 28, at 703 (“While private plaintiffs are also empowered to enforce Section 7 through both damages and equity actions, their impact on merger law has been relatively small.”).

38. *Premerger Notification and the Merger Review Process*, FED. TRADE COMM’N, <https://www.ftc.gov/advice-guidance/competition-guidance/guide-antitrust-laws/mergers/premerger-notification-merger-review-process> (last visited Apr. 17, 2023).

39. *Id.*

40. In the fiscal year of 2020, 1,637 transactions were reported under the HSR Act but only 43 of them were challenged in court. FED. TRADE COMM’N & U.S. DEP’T. JUST. ANTITRUST DIV., HART-SCOTT-RODINO ANNUAL REPORT: FISCAL YEAR 2020, at 1, https://www.ftc.gov/system/files/documents/reports/hart-scott-rodino-annual-report-fiscal-year-2020/fy2020_-_hsr_annual_report_-_final.pdf.

41. *1968 Merger Guidelines* § 1, U.S. DEP’T OF JUSTICE, <https://www.justice.gov/archives/atr/1968-merger-guidelines> (last visited July 31, 2023) [hereinafter 1968 Merger Guidelines].

courts have generally referred to the Guidelines as a persuasive framework for merger analysis.⁴² The Merger Guidelines, therefore, create a critical channel in framing the legal debates.⁴³ Recently, new leadership at the executive branch has reinvigorated this debate by urging the agencies to strengthen vertical merger enforcement. On June 15, 2021, Lina Khan, a key scholar of the New Brandeis movement, was sworn in as the Chair of the FTC.⁴⁴ Shortly after, the FTC majority voted 3-2 to rescind its approval of the Vertical Merger Guidelines (VMG) issued in 2020.⁴⁵ Less than a month later, President Biden issued an Executive Order, encouraging the Attorney General and the FTC Chair to review and consider whether to revise the horizontal and vertical Merger Guidelines.⁴⁶ More recently, on July 19, 2023, the FTC and DOJ issued a draft update of the Merger Guidelines and requested public comments.⁴⁷ The next Part evaluates the unsatisfying historical approaches of vertical merger enforcement as reflected in the various revisions of the Merger Guidelines.

III. THE UNSATISFYING HISTORICAL APPROACHES TO VERTICAL MERGER ENFORCEMENT

Two approaches emerged throughout the evolution of merger enforcement—a structuralist approach and a welfare-based approach. The structuralist approach was manifested in the 1968 Merger Guidelines, and the welfare-based approach appeared in the 1984 Merger Guidelines as well as the newly rescinded 2020 Vertical Merger Guidelines. This Part examines the structuralist and welfare-based approaches to separate anticompetitive and procompetitive vertical mergers. First, this Part discusses the structuralist approach's populist roots as evident in the 1968 Merger Guidelines. Second, this Part examines the Chicago School's welfare-based consumer welfare standard and its lasting impact on antitrust enforcement and jurisprudence. Lastly, this Part elaborates on the issues behind the current application of the

42. Greene, *supra* note 31, at 817.

43. *Id.* at 821.

44. *Lina M. Khan Sworn in as Chair of the FTC*, FED. TRADE COMM'N (June 15, 2021), <https://www.ftc.gov/news-events/news/press-releases/2021/06/lina-m-khan-sworn-chair-ftc>.

45. *Fed. Trade Commission Withdraws Vertical Merger Guidelines and Commentary*, FED. TRADE COMM'N (Sept. 15, 2021), <https://www.ftc.gov/news-events/news/press-releases/2021/09/federal-trade-commission-withdraws-vertical-merger-guidelines-commentary>.

46. Exec. Order No. 14036 on Promoting Competition in the American Economy, 86 Fed. Reg. 36,987, § 5(c) (July 9, 2021).

47. *FTC and DOJ Seek Comment on Draft Merger Guidelines*, FED. TRADE COMM'N (July 19, 2023), <https://www.ftc.gov/news-events/news/press-releases/2023/07/ftc-doj-seek-comment-draft-merger-guidelines>.

consumer welfare standard that has caused gradually diminished vertical merger enforcement.

A. THE STRUCTURALIST APPROACH

The 1968 Merger Guidelines took the structuralist position that mergers are anticompetitive if they result in a highly concentrated market structure.⁴⁸ This position is hardly surprising, considering the political economy at the time. From the 1940s to 1960s, the prevailing industrial organization economics doctrine was dominated by the “Structure-Conduct-Performance” framework purported by Harvard economists such as Donald Turner, Edward Chamberlain, and Joe Bain.⁴⁹ In their view, high market concentration tends to result in anticompetitive behavior.⁵⁰ Accordingly, although the 1968 Guidelines addressed horizontal and vertical mergers separately and identified different theories of anticompetitive harm, the enforcement policies for each were almost exclusively based on market share.⁵¹ Especially for vertical mergers, the DOJ identified foreclosure and barriers to entry as potential anticompetitive effects, but noted that vertical merger enforcement “can be satisfactorily stated by . . . [framing] primarily in terms of the market shares of the merging firms and the conditions of entry which already exist in the relevant markets.”⁵² The 1968 Guidelines also expressly rejected efficiencies as justification for all mergers except under exceptional circumstances.⁵³

In addition to the prevailing economic theory at the time, the 1968 Guidelines’ embracement of a structuralist approach was motivated by socio-political considerations. Beginning in the 1940s, commentators and legislators became increasingly concerned over the “rising tide of economic concentration in the American economy.”⁵⁴ In *Alcoa*, Judge Learned Hand famously rejected pure economic considerations and enunciated the socio-

48. 1968 Mergers Guidelines, § 2.

49. Thomas A. Piraino, *Reconciling the Harvard and Chicago Schools: A New Antitrust Approach for the 21st Century*, 82 IND. L.J. 346, 348 (2007); see also Herbert Hovenkamp, *Robert Bork and Vertical Integration: Leverage, Foreclosure, and Efficiency*, 79 ANTITRUST L.J. 983, 990 (2014) (“Although [the Harvard economists] did not recommend a per se rule, they did find a strong link between integration and monopoly control. These views were reflected in the 1968 Merger Guidelines[.]”).

50. Hovenkamp, *supra* note 49, at 990.

51. 1968 Mergers Guidelines, § 4 (addressing horizontal mergers), § 11 (addressing vertical mergers).

52. *Id.* § 11.

53. *Id.* § 16.

54. *Brown Shoe Co. v. United States*, 370 U.S. 294, 315 (1962).

political goals of antitrust law.⁵⁵ He argued that Congress had intended to preserve “an organization of industry in small units” in spite of the possible cost.⁵⁶ Consistent with the 1968 Guidelines, the Supreme Court famously came very close to entirely ruling out efficiencies as a consideration from the merger analysis in *Brown Shoe Co. v. United States*.⁵⁷ Indeed, the Supreme Court treated protection of competition and the pursuit of efficiencies as directly conflicting objectives.⁵⁸ When balancing between competing considerations of integrated efficiencies and market concentration, the Court concluded that Congress resolved them in favor of decentralization.⁵⁹ The Court’s analysis reflected the prevailing mid-century idea that achieving efficiencies through merger is not a part of the competitive process, and merger that would result in a concentrated market structure is anticompetitive.

B. THE CHICAGO SCHOOL’S WELFARE-BASED APPROACH

Beginning in the late 1960s, a group of legal scholars and economists associated with the University of Chicago began to challenge this interventionist approach underlying the 1968 Guidelines.⁶⁰ The Chicago School aimed to provide a scientific tool for antitrust analysis, which led to a focus on the outcome of the mergers. While not the first to introduce economic analysis in antitrust, the Chicago School explicitly recognized economics in judicial and administrative literature.⁶¹ Indeed, the Chicago School’s widely influential consumer welfare standard dominates the mainstream antitrust analysis to this day. Yet, it is a diverse school of thought and has progressed drastically over the years. This Section first discusses the early Chicago School’s laissez-faire approach to vertical mergers, the establishment of the consumer welfare standard, and the standard’s impact on vertical merger enforcement. This Section then examines both the so-called

55. *United States v. Aluminum Co. of America*, 148 F.2d 416, 427 (2d Cir. 1945) (noting that Congress forbade all trusts, regardless of good or bad, not only because of economic motives, but also because of the indirect social or moral effect to prefer a system of small producers).

56. *Id.* at 429.

57. 370 U.S. 294, 294 (condemning mergers between two firms with small market shares, in part because the integrated firms can achieve cost-savings).

58. William J. Kolasky & Andrew R. Dick, *Merger Guidelines and the Integration of Efficiencies in Antitrust Review of Horizontal Mergers*, 71 ANTITRUST L.J. 207, 209 (2003).

59. *Brown Shoe*, 370 U.S. at 344 (signaling adherence to Congress’s decision to favor decentralization, in the face of competing concerns that maintaining fragmented industries and markets might occasionally create higher costs and prices).

60. Scholars in the early Chicago School that purported this view include but are not limited to: Robert Bork, John McGee, Lester Telser, Richard Posner, and Ward Bowman.

61. Herbert Hovenkamp, *Post-Chicago Antitrust: A Review and Critique*, 2001 COLUM. BUS. L. REV. 257, 265 (2001).

Post-Chicago School's critics of the Chicago School and the welfare-based analytical framework for vertical merger under the more complex and sophisticated Post-Chicago School.

1. *The Early Chicago School, Efficiencies, and the 1982 and 1984 Merger Guidelines*

Relying on the belief that in the long run markets tend to self-correct, the early Chicago School rejected the structuralist approach of the 1960s and advocated for a laissez-faire approach to antitrust enforcement.⁶² Overenforcement was considered an evil and agencies were directed to intervene only when it was clear that certain anticompetitive conduct was threatening consumer welfare.⁶³

The Chicago School offered three insights into the definition of the competitive process. First, it offered a “coherent and elegant ideology” that shifts the focus of antitrust from market structure to a purely economic calculation.⁶⁴ In that sense, the use of economic models in administrative and judicial decision-making promised a rigorous, value-neutral approach to market regulation.⁶⁵ Second, and relatedly, the Chicago School prescribed welfare as the sole determination of whether conduct is procompetitive or anticompetitive.⁶⁶ Jurist and scholar Robert Bork coined the term “consumer welfare standard” as the *only* value to be considered by a court.⁶⁷ Yet, Bork, a lawyer by training, departed from the traditional economic textbook definition and interpreted consumer welfare as the “the maximization of wealth” increased through market efficiency.⁶⁸ In classic economics, what Bork referred to is the total welfare in the market, irrespective of the distribution of surplus between consumers and producers.⁶⁹ The biggest difference between total welfare and a true consumer welfare approach is that under a true consumer welfare approach, only welfare gained by the consumer would be

62. Piraino, *supra* note 49, at 350.

63. *Id.*

64. Hovenkamp, *supra* note 61, at 258, 265.

65. *Id.* at 265.

66. See A. Douglas Melamed & Nicholas Petit, *The Misguided Assault on the Consumer Welfare Standard in the Age of Platform Markets*, 54 REV. INDUS. ORG. 741, 746 (2019) (“[T]he CW paradigm makes clear that antitrust laws are about conduct that reduces or is likely to reduce economic welfare and is not intended to prevent noneconomic harms such as harm to the political process or to serve other social objectives.”).

67. Robert H. Bork, *Legislative Intent and the Policy of the Sherman Act*, 9 J.L. & ECON. 7, 10–11 (1966) (emphasis added).

68. *Id.* at 7.

69. Barak Orbach, *The Antitrust Consumer Welfare Paradox*, 7 J. COMPETITION L. & ECON. 133, 162 (2010).

credited as competitively beneficial. But under the Borkean view, mergers that increase efficiencies are procompetitive, regardless of their harm to competitors and even consumers. As a result, efficiency is not only an affirmative defense, but also the benchmark for competitiveness.

This embrace of total welfare naturally led to the Chicago School's third influence—a view that vertical mergers are generally competitively neutral or procompetitive, and therefore should be presumed to be procompetitive to avoid false positive errors and overdeterrence.⁷⁰ Although the Chicago School did identify some competitive concerns, its early proponents “placed little credence in the harm from foreclosure” and collusion.⁷¹ First, the Chicago School rejected the theory of foreclosure, on the ground that the unintegrated rival may gain access to input elsewhere by realigning purchasing patterns.⁷² Second, it rejected the theory of leverage, based on an oversimplified “single monopoly profit” model that claimed that the integrated firm cannot enjoy more than one monopoly profit.⁷³ Lastly, the Chicago School viewed vertical mergers as “invariably highly efficient,” in large part because of the elimination of double marginalization.⁷⁴

Bork pushes the presumption of procompetitive effect further by famously rejecting calculation for individual efficiencies.⁷⁵ Bork relied on a “beguilingly simple” theory: to the extent that vertical integration creates efficiencies, it may deter entry, but only as a result of increased competition through cost-savings; to the extent that a vertical merger is not efficient, it would not impede entry.⁷⁶ Unlike Oliver Williamson's welfare tradeoff model, which would balance the productive efficiencies gain against consumer welfare loss to determine the total welfare impact of a merger, Bork argued that an individualized calculation of net welfare gain is neither necessary nor possible.⁷⁷ Instead, Bork believed that efficiencies would be presumed to exist in all vertical mergers.⁷⁸ Although Bork's extreme views on vertical integration have subsequently been doubted by other Chicago scholars, they have important and lasting impacts in courts'

70. Salop, *supra* note 23, at 1972; *see also* Orbach, *supra* note 69, at 162 n.38 (explaining false positive and false negative errors in antitrust enforcement).

71. Riordan & Salop, *supra* note 25, at 518.

72. *Id.* at 516.

73. *Id.* at 517.

74. Salop, *supra* note 23, at 1970.

75. Hovenkamp, *supra* note 49, at 983.

76. *Id.* at 994.

77. Nancy L. Rose & Jonathan Sallet, *The Dichotomous Treatment of Efficiencies in Horizontal Mergers: Too Much? Too Little? Getting it Right*, 168 U. PA. L. REV. 1941, 1952 (2021).

78. *Id.*

considerations of efficiencies benefits in vertical mergers, particularly with regard to the gradually pro-defendant burden of proof in proving efficiencies.

In response to the adoption of the consumer welfare standard and the Chicago School's economics-centered analysis, efficiencies analysis began to enter the merger review framework in the 1982 Merger Guidelines.⁷⁹ Unsurprisingly, the Guidelines marked a radical change from a market-structure-based approach to a market-power-based approach for both horizontal and vertical mergers.⁸⁰ For vertical mergers, the DOJ would no longer rely on a structure-based presumption of anticompetitive harm. Instead, the agency would evaluate the competitive effects of a merger based on specific theories of harm. The 1982 Guidelines emphasized that vertical mergers lack direct impact on market concentration.⁸¹ Moreover, just two years later, the 1984 Guidelines marked a "more dramatic departure from earlier positions."⁸² Influenced by the Chicago School's endorsement of market efficiency, the 1984 Guidelines began by expressly claiming that "[t]he primary benefits of mergers to the economy is their efficiency-enhancing potential."⁸³ Most notably, the Guidelines noted that "the Department will give relatively more weight to expected efficiencies" for vertical mergers than horizontal mergers.⁸⁴ Under the 1984 Guidelines, the DOJ would allow mergers that it otherwise would challenge if the parties could establish by clear and convincing evidence that the merger will achieve net efficiencies.⁸⁵

However, despite Bork's misnomer, the agencies and courts mostly interpreted consumer welfare as consumer surplus, not total surplus, and, accordingly, rejected efficiencies as an affirmative defense. The 1984 Guidelines took this view by presenting efficiencies as a factor to consider, not as a defense.⁸⁶ The then-Assistant Attorney General Paul McGrath clarified that under this approach the DOJ "would not balance expected efficiencies against expected anticompetitive consequences."⁸⁷ In doing so, the agencies

79. *1982 Merger Guidelines*, U.S. DEP'T OF JUSTICE, <https://www.justice.gov/archives/atr/1982-merger-guidelines> (last visited July 31, 2023).

80. *Id.* § 1.0.

81. *Id.* § 4.1A.

82. Rose & Sallet, *supra* note 77, at 1953.

83. *1984 Merger Guidelines* § 3.5, U.S. DEP'T OF JUSTICE, <https://www.justice.gov/archives/atr/1984-merger-guidelines> (last visited July 31, 2023).

84. *Id.*

85. *Id.*

86. *Id.* §§ 3.5, 4.135.

87. Richard A. Pogue, Harry M. Reasoner, John H. Shenefield & Richard A. Whiting, *60 Minutes with J. Paul McGrath*, 54 ANTITRUST L.J. 131, 134–35, 141 (1985).

reaffirmed that an efficiencies gain must result in an increase in consumer surplus to be credited as procompetitive.

2. *The Post-Chicago School, Consumer Welfare Standard and the 2020 Vertical Merger Guidelines*

Beginning in the 1990s, armed with a more sophisticated understanding of microeconomics, commentators concluded that the Chicago School's economic models were overly simplistic.⁸⁸ The Post-Chicago School debunked the "single monopoly profit theories" and the assumption of a perfectly competitive market with unhindered free flow of information and low entry barriers.⁸⁹ Rather, the Post-Chicago School observed that in imperfectly competitive markets, evaluating the net competitive effect of a merger is a question of fact, not theory, which often requires sophisticated econometric modeling.⁹⁰ Informed by game theory and industrial organization economics, the Post-Chicago School offered a newer, more realistic methodology to market structure in which vertical mergers can have anticompetitive effects.⁹¹

The Post-Chicago School made three major contributions to the vertical merger evaluation. First, it incorporated market imperfection into economic analysis and offered tools for analyzing both unilateral and coordinated harms in vertical mergers.⁹² In terms of unilateral harms, the Post-Chicago School provided "a metered alternative" to the largely binary concept of foreclosure.⁹³ The idea is that an integrated firm may reduce sales or increase prices to downstream unintegrated rivals and thereby make it more costly for downstream rivals to do business.⁹⁴ The Post-Chicago School measured harms under foreclosure not by a competitor's exit, but instead by the increase in equilibrium prices.⁹⁵ Additionally, the Post-Chicago School argued that vertical mergers can facilitate exclusionary conduct based on competitively sensitive information obtained through the mergers.⁹⁶ The agencies adopted the Post-Chicago School's view on competitive harms and incorporated the theories of raising rivals' costs and access to competitively sensitive information in the 2020 Vertical Merger Guidelines—the first revision of vertical merger review

88. Piraino, *supra* note 49, at 364.

89. *See generally* Salop, *supra* note 23 (rejecting the Chicago School's assumption and providing analytical framework for foreclosure and leverage theories).

90. *Id.* at 1974.

91. *Id.*

92. Hovenkamp, *supra* note 61, at 324.

93. Herbert Hovenkamp, *Competitive Harm from Vertical Merger*, 59 REV. INDUS. ORG. 139, 144 (2021).

94. *Id.*

95. *Id.*

96. Riordan & Salop, *supra* note 25, at 520.

since 1984.⁹⁷ Under the 2020 VMG, the agencies consider the ability and incentive for the merged firm to raise rivals' costs (RRC) or foreclose sale(s).⁹⁸ The central question for a firm's ability is "whether the downstream rivals have good substitutes for the input in question."⁹⁹ If the downstream firms have no good substitute, their ability to compete is weakened if the merged firm denies access to or charges higher prices for the input. In that case, the merged firm has the ability to RRC or foreclose inputs. The key question for determining incentive is whether weakening "downstream rivals would enhance profit of the merged firm due to diverted downstream sales."¹⁰⁰ If a merged firm has both the ability and the incentive to RRC or foreclose input, the merger harms the downstream competitors. But a final balancing of the RRC and efficiencies claims is still required.¹⁰¹ Generally, this last step requires complex econometrics modeling and simulation.

Second, the Post-Chicago School reaffirmed the welfare-based consumer welfare standard but clarified that consumer welfare, rather than total welfare, should be the benchmark to determine whether business conduct is procompetitive or anticompetitive. While the Post-Chicago School has acknowledged market imperfection and the likelihood of foreclosure in vertical mergers, it has shifted the focus of competitive injury away from the destruction of rivals purported by the 1960s structuralists.¹⁰² For the Post-Chicago School, harm to rivals was simply part of the competitive process if and only if the merger would not make consumers worse off.¹⁰³ While the Post-Chicago School acknowledged the intrinsic EDM effect for most vertical mergers in imperfectly competitive markets,¹⁰⁴ it asserted that EDM and other efficiencies gains do not always pass down to consumers.¹⁰⁵ Accordingly, under a consumer welfare standard, any efficiency gains that do not pass on to consumers theoretically should not be credited.

The agencies adopted this approach in the 2020 VMG, under which efficiency claims must be merger-specific, cognizable, and verifiable to be

97. 2020 Vertical Merger Guidelines, *supra* note 19.

98. *Id.* § 4(a).

99. Carl Shapiro, *Vertical Mergers and Input Foreclosure: Lessons from the AT&T/Time Warner Case*, 59 REV. INDUS. ORG. 303, 306 (2021).

100. *Id.*

101. *Id.*

102. Hovenkamp, *supra* note 61, at 318.

103. Hovenkamp, *supra* note 93, at 174.

104. Salop, *supra* note 23, at 1972.

105. *See id.* at 1974 ("[E]ven if EDM or other efficiencies do create downward pricing pressure, that downward pressure does not necessarily dominate the upward pricing pressure from the incentive of the upstream merging firm to raise its input price to rivals.").

credited as competitive benefits.¹⁰⁶ Efficiencies are merger-specific when they are likely to be achieved through the merger and unlikely to be accomplished through any *practical alternatives*.¹⁰⁷ For example, in assessing the merger-specificity of EDM, the agencies “examine whether it would likely be less costly for the merged firm to self-supply inputs [post-merger] . . . than for the downstream firm to purchase them from one or more independent firms absent the merger.”¹⁰⁸ Efficiency claims are cognizable if they do not arise from anticompetitive reduction in output or service.¹⁰⁹ Efficiency claims are verifiable if the merging parties are able to meet the burden to substantiate their claims by showing that they are not merely speculative.¹¹⁰ However, while the Post-Chicago School explicitly rejected a total welfare standard, the 2020 VMG retained a footnote indicating that agencies may consider out-of-market efficiencies.¹¹¹

Additionally, the 2020 VMG provided that harms to downstream unintegrated rivals are not sufficient to constitute harm to competition. Rather, consistent with most of the Post-Chicago commentators,¹¹² the 2020 VMG required evaluating the competitive effect of a merger on the actual or potential buyers of the downstream firms.¹¹³ The VMG expressly acknowledged that while the merged firm may have the ability and incentive to foreclose its rival or raise their costs, the merger can also create procompetitive effects that offset or even outweigh the incentive to harm customers.¹¹⁴ A merger that harms downstream unintegrated competitors may nonetheless be benign if it does not harm downstream consumers. Under the Guidelines, the agencies would take an additional step to evaluate “the likely net effect on the competition.”¹¹⁵

Third, to balance the potential efficiency benefits from the merger against the potential foreclosure or coordinated effect of a vertical merger, the Post-

106. 2020 Vertical Merger Guidelines, *supra* note 19, § 6.

107. *Id.* (emphasis added).

108. *Id.*

109. *Id.*

110. *Id.*

111. *See* 2020 Vertical Merger Guidelines, *supra* note 19, § 6 n.6 (“The Agencies in their prosecutorial discretion may also consider efficiencies not strictly in the relevant market[.]”).

112. *See* Shapiro, *supra* note 99, at 320 (following the 2020 Vertical Merger Guidelines’ approach of evaluating input foreclosure concerns based on their impact on downstream customers); Riordan & Salop, *supra* note 25, at 561 (“In evaluating input foreclosure, we concluded that proof that input prices would rise is insufficient. It also is necessary to show injury to consumers.”).

113. 2020 Vertical Merger Guidelines, *supra* note 19, § 1.

114. *Id.*

115. *Id.* § 4(a).

Chicago School proposed a welfare tradeoff model to evaluate the net competitive effect of the merger.¹¹⁶ Where vertical mergers create significant efficiency benefits and raise significant competitive concerns, “those conflicting effects must be weighed and balanced.”¹¹⁷ The Post-Chicago School’s sophisticated welfare-balancing approach dominates the contemporary antitrust analysis. But the increasing reliance on econometrics and expert testimony imposes a great challenge for lawyers and judges to evaluate the accuracy and presumptions behind these complex economic models. The next section addresses the shortcomings of the modern Post-Chicago approach.

C. CRITIQUE OF CURRENT VERTICAL MERGER ENFORCEMENT

Commentators who are discontent over the *status quo* of vertical merger enforcement have argued for a reform in merger enforcement and offered three main critiques. First, the misleading phrasing of “consumer” welfare has led to neglect in identifying merger harms to input markets such as the labor market.¹¹⁸ Second, although the merger guidelines acknowledge non-price harms, the consumer welfare standard, as currently applied, focuses almost exclusively on economic factors such as price, output, or efficiencies, and rarely considered less-quantifiable theories based on reduced product quality, variety, and diminished innovation.¹¹⁹

Third and relatedly, under the Chicago School’s continuing influence within the agencies and the judiciary, merger enforcement has been indoctrinated with pro-defendant assumptions that vertical mergers are mostly procompetitive. Therefore, in practice, contrary to the incipiency standard mandated by the Clayton Act, a plaintiff challenging a vertical merger faces a heavy burden to show competitive harm under the three-step burden-shifting framework outlined in *United States v. AT&T, Inc.*¹²⁰ For example, under a

116. Hovenkamp, *supra* note 28, at 715.

117. Riordan & Salop, *supra* note 25, at 523.

118. *See, e.g.*, Fed. Trade Comm’n, Statement of Chair Lina M. Khan, Commissioner Rohit Chopra, and Commissioner Rebecca Kelly Slaughter on the Withdrawal of the Vertical Merger Guidelines, at 7–8 (Sept. 15, 2021), https://www.ftc.gov/system/files/documents/public-statements/1596396/statement_of_chair_lina_m_khan_commissioner_rohit_chopra_and_commissioner_rebecca_kelly_slaughter_on.pdf (advocating for new vertical merger guidelines that include framework for evaluating non-price harms); Hiba Hafiz, *Labor Antitrust Paradox*, 86 CHI. L. REV. 381 (2020) (arguing that the consumer welfare standard is ill-equipped to address the labor market); Eric A. Posner, Glen Weyl & Suresh Naidu, *Antitrust Remedies for Labor Market Power*, 132 HARV. L. REV. 536 (2018) (noting the neglect of labor market harm in merger review and developing a variety of analytic tools to evaluate labor markets).

119. Lina Khan, Note, *Amazon’s Antitrust Paradox*, 126 YALE L.J. 710, 721–22 (2017).

120. 916 F.3d 1029, 1032 (D.C. Cir. 2019).

bargaining theory based on pricing pressure, a plaintiff must quantify the net effect on competition and persuade the skeptical generalist judge, who is often ill-equipped to evaluate complex economic models and tends to err on the side of the defendant.¹²¹

The burden on plaintiffs seems to be even heavier for those who allege *prima facie* theories of harms other than price and output. Even after the revision of the 2020 VMG and the 2010 HMG, which explicitly outline various non-price theories of harms, courts are reluctant to embrace theories of harms other than price and output.¹²² Plaintiffs are frequently required to show econometric proof that relies on data access.¹²³ In addition, it is not sufficient for a plaintiff to merely show that the merged firm has the ability and incentive to harm competition as a profit maximizing entity. The plaintiff must also consider any historical business practices that would prevent the merged firm from behaving anticompetitively *and* show that the merged firm would harm downstream consumers.

For example, in a recent vertical merger case, Judge Carl Nichols of the District Court of Columbia rejected the government's vertical data misuse theory for a merger between UnitedHealth Group and Change Healthcare, a health care technology company that operated the largest electronic data interchange (EDI)¹²⁴ clearinghouse in the United States.¹²⁵ Under the vertical data misuse theory, the government claimed that UnitedHealthcare, the nation's biggest commercial health insurer, would have access and use rights to the claims data of its rivals and would thereby deter its rivals from innovating out of the fear that UnitedHealthcare will free ride off their innovation.¹²⁶ Judge Nichols found that the government failed to establish fact-specific showings that United would "uproot its entire business strategy and corporate culture," intentionally violate firewall policies and existing contractual commitments, and sacrifice significant financial reputational interests.¹²⁷ And, perhaps most alarmingly, Judge Nichols further reasoned that even if the government had shown that the merged firm has an incentive

121. See Shapiro, *supra* note 99, for a detailed account of the heavy evidentiary burden to show harms in *United States v. AT&T, Inc.*

122. See 2010 Horizontal Merger Guidelines, *supra* note 19, § 6.4; 2020 Vertical Merger Guidelines, *supra* note 19, § 4.b.

123. Elhauge, *supra* note 15.

124. IBM, *What is electronic data interchange (EDI)?*, <https://www.ibm.com/topics/edi-electronic-data-interchange> (last visited Feb. 3, 2024).

125. *United States v. UnitedHealth Grp., Inc.*, 1:22-CV-0481, 2022 WL 4365867, at *15–*26 (D.D.C. Sept. 21, 2022).

126. *Id.* at *15.

127. *Id.* at *16.

to misuse claims data obtained through Change’s EDI clearing house, the government failed to demonstrate that rival payers would innovate less post-merger.¹²⁸ And, even if rival payers would scale back on innovation, the government must also have proved that the reduction in innovation would *substantially* lessen competition.¹²⁹ As this example shows, the immense burden on plaintiffs to establish *prima facie* harms has a serious chilling effect on potential private plaintiffs and agencies who have limited resources. The result is that, for the past forty years, only a handful of cases were litigated where the focus was mainly on the vertical aspects of the merger—and the agencies lost each of them.¹³⁰

IV. VERTICAL MERGER REFORM

Reformers wishing to reinvigorate vertical merger enforcement generally fall under one of the three camps: (1) a return to the structuralist approach; (2) a trading partner welfare approach (rebranded as “protecting competition”); or (3) a protection of competitive process standard. Reformers under the second and third camps share the same explicit acknowledgement of antitrust’s competition goal and urge focus on the merger’s impact on the competitive process. Yet, the two groups differ as to how to evaluate the competitive impact. The trading partner welfare standard expands the consumer welfare standard’s narrow focus of a merger’s impact on direct consumers onto trading partners on the other side of the market. Meanwhile, the protection of competitive process standard rejects the use of welfare as a proxy and argues that antitrust law should directly separate “fair and foul.”¹³¹

This Part asserts that a protection of competitive process standard—the third approach—is needed to truly capture the concept of competition and safeguard the long-term interests of consumers, producers, and workers. The first Section, IV.A, argues that a return to the first, structuralist approach is undesirable. The second Section, IV.B, comparatively analyzes the frameworks under the trading partner welfare standard and the protection of competition approach. The last Section, IV.C, asserts that the rebranded trading partner welfare standard is inadequate to safeguard competition, and analyzes

128. *Id.* at *24–*25.

129. *Id.* at *25.

130. *See* United States v. AT&T, Inc., 916 F.3d 1029 (D.C. Cir. 2019); Fed. Trade Comm’n v. Illumina, Inc., No. CV 21-873, 2021 WL 1546542 (D.D.C. Apr. 20, 2021); *UnitedHealth*, 2022 WL 4365867.

131. Tim Wu, *After Consumer Welfare, Now What? The “Protection of Competition” Standard in Practice* 8, SSRN (Apr. 5, 2018), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3249173.

application of each the second and third approaches in vertical merger enforcement.

A. CRITIQUES OF A RETURN TO THE STRUCTURALIST APPROACH

At the outset, this Note rejects a return to the structuralist approach. Relying on the original legislative intent and the populist root of the Sherman Act, some reformers advocate that concentrated private power is an evil in and of itself and should be prohibited or regulated.¹³² Some even propose for a return to the 1968 Merger Guidelines.¹³³ However, a focus on market structure risks the same pitfalls as the Chicago School's consumer welfare standard, as they both focus on results rather than process. The 1960's structuralist approach emerged in an era when there were no satisfactory tools for case-by-case assessment for mergers.¹³⁴ The "Structure Conduct Performance" framework prevalent at that time was supplanted long ago within industrial organization economics.¹³⁵ Additionally, the New Brandeis' argument based on the murky legislative intent of the Sherman Act is as unpersuasive today as when it was raised by Judge Bork for his prescription of the consumer welfare standard. Although the Sherman Act was inspired by various social, political, and economic concerns of monopolies, antitrust law is not designed to solve all of these concerns.¹³⁶ What's more, regardless of Congress's intent over a hundred years ago, antitrust statutes' broad mandates have been generally viewed as a common law-like process evolving overtime.¹³⁷ A fixation over the original intent of the law is neither meaningful nor productive to the discussion.

It is not to say that antitrust law is solely for promoting economic goals. It does not. A narrow view that antitrust should only look at economic welfare of the society, however defined, is misguided.¹³⁸ Indeed, unlike many commentators who reject a return to a structuralist approach, this Note

132. See Khan, *supra* note 119, at 797.

133. Open Markets Institute, American Economic Liberties Project, Frank Pasquale & Maurice Stucke, Comment on Draft Vertical Merger Guidelines (Feb. 2020), https://www.ftc.gov/system/files/attachments/798-draft-vertical-merger-guidelines/comment_to_ftc-doj_re_vertical_merger_guidelines.pdf.

134. Gregory J. Werden, *Back to School: What the Chicago School and New Brandeis School Get Right* 10, SSRN (Oct. 10, 2018), <https://ssrn.com/abstract=3247116>.

135. Shapiro, *supra* note 15, at 34.

136. *Id.* at 42.

137. *State Oil Co. v. Khan*, 522 U.S. 3, 20 (1997) (noting that when Congress promulgated the Sherman Act it "expected the courts to give shape to the statute's broad mandate by drawing on common-law tradition"); Melamed & Petit, *supra* note 66, at 746 ("Antitrust has long been understood to evolve over time through a common-law like process.").

138. See Fox, *supra* note 10; *supra* note 13 and accompanying text.

recognizes that the New Brandeis movement and earlier progressive commentators are correct to identify that antitrust laws promote competition to serve a variety of interests.¹³⁹ The structuralist's biggest pitfall, however, is that it views socio-political benefits of a decentralized market a *prevailing and dispositive* consideration. Its favoritism of decentralized markets would bypass the process of vigorous competition and ignore the benefits of increased efficiencies.

Some New Brandeis commentators enunciate several principles of antitrust in hope to replace the Chicago School's efficiency obsession. Some of these objectives are within the purview of antitrust law, while others are not. For example, reformers aim for "the preservation of open markets, the protection of producers and consumers from monopoly abuse and the dispersion of political and economic control."¹⁴⁰ The first two aims fall squarely within the idea of protecting competition, while the connection between protecting competition and the "dispersion of political and economic control" is less direct. It is important to note that antitrust law is not and should not be the only body of law that addresses the political control of private entities, the inequitable distribution of wealth, and many other social issues.¹⁴¹ It is antithetical to the basic idea of competition to punish firms for being big and successful if they achieve their size lawfully.¹⁴² The line between size and power is a thin one, and commentators may disagree vigorously about what strategies are or are not lawful. But a shortcut based on the size of the firm alone contradicts the long-held distinction between the mere possession of market power and abuse of market power.¹⁴³ Perhaps more importantly, a structuralist approach, like the welfare approach, predetermines the role of antitrust law in our democratic republic and forecloses normative discussions about fairness, justice, and market competition. Therefore, a presumption or

139. Khan, *supra* note 119, at 739.

140. *Id.* at 743.

141. See Shapiro, *supra* note 15, at 42 (arguing that lax antitrust enforcement is not the central cause of social and economic problems in America); see also Herbert Hovenkamp, *Whatever Did Happen to the Antitrust Movement?*, 93 NOTRE DAME L. REV. 583, 594 (2018) (noting that antitrust is only one of many legal policies that address the concern of what citizens are entitled to expect from business and their economy).

142. *Id.*

143. See, e.g., *Verizon Commc'n Inc. v. L. Offices of Curtis V. Trinko, LLP*, 540 U.S. 398, 407 (2004) (clarifying that merely possessing monopoly power and charging monopoly prices, without anticompetitive conduct, is not only not unlawful, but critical to the free market in attracting "business acumen," inducing risk-taking, and incentivizing innovation).

even per se illegality based on a firm's size and market power is not a desirable policy.¹⁴⁴

B. THE TWO STANDARDS OF PROTECTING COMPETITION: TRADING PARTNER WELFARE STANDARD AND PROTECTION OF THE COMPETITIVE PROCESS

The idea of protecting competition means little without defining what types of business conduct are deemed proper and which are not. Therefore, while commentators generally agree on the goal of antitrust law as protecting competition or the competitive process, how to administer that goal is particularly divisive. This Section discusses two existing standards and their key differences.

Under the guise of various names, a number of mainstream progressives argue for a "trading partner" welfare approach to protect the competitive process.¹⁴⁵ In its essence, compared to the Post-Chicago School's consumer welfare standard, the trading partner welfare standard expands the recognition of harm from consumers to trading partners on the other side of the market.¹⁴⁶ Trading partners thus include product and labor suppliers, and welfare is defined broadly to include product variety, product quality, and innovation.¹⁴⁷ While commentators supporting the trading partner welfare standard acknowledge that promoting competition is the goal of antitrust, they also emphasize that economic measurements are necessary to make this goal operational.¹⁴⁸ The issue with underenforcement in vertical mergers, they

144. See Fox, *supra* note 10, at 1182 (declining to include the preservation of small size for its own sake as a possible goal of antitrust because of the potential conflict between that objective and consumers' interests); see also Shapiro, *supra* note 1, at 745 ("Economic growth will be undermined if firms are discouraged from competing vigorously for fear that they will be found to have violated the antitrust laws, or for fear they will be broken up if they are too successful.").

145. See, e.g., Shapiro, *supra* note 15, at 38 ("A business practice is judged to be anticompetitive if it harm trading parties on the other side of the market as a result of disrupting the competitive process."); C. Scott Hemphill & Nancy L. Rose, *Mergers that Harm Sellers*, 127 YALE L.J. 2078, 2080 (2018) (arguing that reduced competition between buyers is unlawful even where there is no harm to downstream purchasers).

146. Shapiro, *supra* note 15, at 38; see also *The Consumer Welfare Standard in Antitrust: Outdated, or a Harbor in a Sea of Doubt: Hearing Before the Subcomm. on Antitrust, Consumer Protection and Consumer Rights of the S. Comm. on the Judiciary*, 115th Cong. (2017) (statement of Carl Shapiro, Professor of Business Strategy, University of California, Berkeley), <http://faculty.haas.berkeley.edu/shapiro/consumerwelfarestandard.pdf>.

147. Shapiro, *supra* note 15, at 38.

148. *The Consumer Welfare Standard in Antitrust: Outdated, or a Harbor in a Sea of Doubt: Hearing Before the Subcomm. on Antitrust, Consumer Protection and Consumer Rights of the S. Comm. On the Judiciary*, 115th Cong. (2017) (statement of Carl Shapiro Professor of Business Strategy,

argue, is not that the welfare approach is wrong, but that it has not been applied properly under the influence of the Chicago School.¹⁴⁹ Under this view, the issues of false negative errors and the plaintiff's high evidentiary burden to establish prima facie case can be fixed by establishing rebuttable presumptions.¹⁵⁰ Similarly, the critique over the welfare standard's price fixation is not inherent to the consumer welfare paradigm.¹⁵¹ In sum, the trading partner welfare standard does not signify a fundamental change in a welfare approach.¹⁵² Rather, it is a rebranding of the consumer welfare standard to escape the inconsistent history of the term since Bork and the Chicago School.¹⁵³ Under this approach, mergers that harm trading partners on the other side of the market are anticompetitive and should be condemned.

In contrast, some commentators associated with the New Brandeis school reject the use of welfare as a proxy and argue that antitrust law should directly separate "fair and foul" under a protection of competitive process standard.¹⁵⁴ In an influential paper prior to her appointment to the FTC, Chairwoman Lina Khan criticized the use of consumer welfare as "inadequate to promote real competition."¹⁵⁵ She identified that the issue with the welfare standard is that it focuses on an outcome, as opposed to process.¹⁵⁶ The right inquiry is about a business conduct's impact on the neutrality of the competitive process and the openness of the market, which must be viewed in relation to the market structure.¹⁵⁷ Many critics of the New Brandeis movement have characterized the school as advocating for a structuralist return under the "big is bad" motto.¹⁵⁸ But the New Brandeis School is much more diverse and nuanced than that. In *Amazon's Antitrust Paradox*, Khan clarified that she was not advocating for "a strict return to the structure-conduct-performance

University of California, Berkeley), <http://faculty.haas.berkeley.edu/shapiro/consumerwelfarestandard.pdf>.

149. Salop, *supra* note 70, at 1963.

150. Shapiro, *supra* note 15, at 39.

151. See Melamed & Petit, *supra* note 15, at 5 ("But that inhospitality to pricing cases can hardly be called a problem of price fixation, and its correction does not required abandonment of the [Consumer Welfare] paradigm.").

152. Shapiro, *supra* note 15, at 38 ("I have seen no evidence whatsoever that the "consumer welfare" standard is somehow outdated, so long as one accepts that the goal of antitrust is to promote competition.").

153. *Id.*

154. Wu, *supra* note 131, at 8.

155. Khan, *supra* note 119, at 744.

156. *Id.*

157. *Id.* at 745–46.

158. See, e.g., Aurelien Portuese & Joshua Wright, *Antitrust Populism: Towards a Taxonomy*, 21 STAN. J.L. BUS. FIN. 1, 18 (2020); cf. Khan, *supra* note 2, at *3 (distinguishing antimonopoly from "big is bad").

paradigm.”¹⁵⁹ Instead, she argued that market structure provides insights on how power is distributed in a given market, which is crucial to determine whether a business decision would prevent competition on the merits.¹⁶⁰ In other words, the New Brandeis’ view of antitrust does not blanketly prohibit big firms from engaging in certain business conducts just because of their size. Rather, it simply recognizes that dominant firms may be prohibited from engaging in certain conduct where their smaller rivals would not because dominant firms have the power to distort the competitive outcome.¹⁶¹ Khan defined “distorting” as “a single player [having] enough control to dictate outcomes.”¹⁶² Under this definition, large firms that engage in conduct or agreements that do not give them the power to dictate a competitive result are free to do so without scrutiny. In the horizontal merger context, agencies, economists, and courts have long recognized that concentration in a given market is a good indicator of whether a horizontal merger would raise substantial competitive concerns. Under the 2010 HMG, a firm with the largest market share would not be allowed to merge with the second largest competitor in a concentrated market, while a merger between two small competitors may not raise similarly competitive concerns. Market structure has always mattered and should continue to matter in antitrust analysis.

The disconnect between these two competition-focused approaches stems from their different definitions of “competition on the merits.” Whereas mainstream progressives view conduct that does not harm trading partners as competition on the merits, the protection of competitive process standard recognizes that conduct that does not harm consumers or suppliers may nonetheless harm the competitive process.¹⁶³ Thus, their key disagreement is the role of regulation in shaping the competitive process.

The proponents of the welfare approach view markets as strictly driven by economics. That is, market regulations and policies (including antitrust policy) exist to facilitate the best allocation of resources. New Brandeis proponents, on the other hand, view markets as defined by economic justice, fairness, and opportunities.¹⁶⁴ Therefore, while economic learnings may guide our understanding of the economic effects of certain business conduct, a determination of legality requires additional examination of the equitable

159. Khan, *supra* note 119, at 745.

160. *Id.*

161. *Id.*

162. *Id.* at 746 n.189.

163. Tim Wu, *The “Protection of the Competitive Process” Standard* (Columbia Public Law Research Paper, No. 14-612, 2018), https://scholarship.law.columbia.edu/faculty_scholarship/2290.

164. Fox, *supra* note 10, at 1178.

effects of that conduct. The New Brandeis movement presented several factors to consider in determining the neutrality of the competitive process, including entry barriers, conflict of interest, the emergence of gatekeepers or bottlenecks, the use of and control over data, and the dynamics of bargaining power.¹⁶⁵

To resolve this dispute of whether the rebranded welfare standard is adequate to protect competition, we must answer the underlying question of what the competitive process seeks to protect. The trading partner welfare standard offers compelling reasons to limit actionable harms to those suffered by trading partners and to use welfare, broadly defined, as a measurement of harm. But the welfare-based approach deviates from the principle of protecting competition in significant ways when applied to vertical mergers.¹⁶⁶ The next Sections explain why a process-based approach is needed to effectively protect competition.

C. PROTECTION OF COMPETITIVE PROCESS

This Section begins (in Section IV.C.1) by establishing a framework for analyzing vertical mergers under the protection of competitive process standard.¹⁶⁷ The Section then (in Section IV.C.2) evaluates the two competition-based standards by looking at their substantive abilities to capture and protect the essence of competition, as well as their administrability. Section IV.C.2.a first argues that the trading partner welfare standard fails to recognize that “competition” is not limited to the relationships and interactions between sellers and buyers, but also includes dynamics between sellers who compete in the market. Section IV.C.2.b then notes that balancing the various harms and benefits among trading partners is no more administrable than asking courts

165. *Id.* at 746.

166. The same is true when applied to vertical restraints cases. For criticism on the use of neoclassic efficiencies standard on vertical restraints cases, see John J. Flynn & James F. Ponsoldt, *Legal Reasoning and the Jurisprudence of Vertical Restraints: The Limitations of Neoclassical Economic Analysis in the Resolution of Antitrust Disputes*, 62 N.Y.U. L. REV. 1125 (1987).

167. After this note is drafted, on July 19, 2023, the Federal Trade Commission and the Department of Justice released a draft update of the Merger Guidelines for public comment. *Merger Guidelines for Public Comment*, U.S. DEP'T JUSTICE & FED. TRADE COMM'N (July 19, 2023), https://www.ftc.gov/system/files/ftc_gov/pdf/p859910draftmergerguidelines2023.pdf. In many ways, the proposed guidelines are consistent with the process-based approach, by directing the focus to the competition among rivals and market structure. *See id.* at 15. (“Mergers should not substantially lessen competition by creating a firm that controls products or services that its rivals may use to compete.”); *see also id.* at 17, n.52 (“(“In addition to this structural analysis, many vertical mergers can also be analyzed under the ability and incentive analysis in Guideline 5. Either can be a sufficient basis to warrant concern.”).

to make equitable judgments of business conduct under the protection of competition standard.

1. *Framework for Assessing the Competitive Process*

A true process-based standard should look at (1) the incentives and abilities of the merged firm to prevent downstream rivals from competing on the merits, and (2) whether the entry barriers in the downstream market are sufficiently high to raise competitive concerns. Under the first prong, a focus on the competitive process standard should debunk the popular but often misquoted slogan that antitrust protects competition, not competitors.¹⁶⁸ Protecting competitors against conduct that impedes competing on the merits *is* protecting competition. A showing of the merged firm's incentives and abilities to foreclose or RRC through sophisticated econometric modeling is sufficient but not necessary. Downstream competitors' abilities to compete on the merits can be further defined as offering goods or services at cheaper prices, better quality, or in any other way that attracts consumers.¹⁶⁹

Under the second prong, regarding entry barriers, harms to downstream rivals by themselves are not sufficient to render a merger anticompetitive. A competitive process protects the robustness of the market as a whole, not any particular unintegrated downstream rival.¹⁷⁰ The requirement that plaintiffs must bear the burden of proving high barriers to entry in the downstream market would safeguard this principle. If the downstream market has low entry barriers and the loss of competition from the foreclosure effect can practically be replenished, the competitive process of the downstream market would not be harmed. Conversely, if the downstream market has high entry barriers or the if vertical merger is likely to result in high entry barriers, such as by creating the need for two-tier entry, the downstream competition would be harmed and the merger is anticompetitive.

Moreover, efficiency claims should not be credited as a defense when the two prongs are met.¹⁷¹ It is true that any efficiencies gained through vertical integration may give the merged firm an incentive to pass those efficiencies down for the benefit of consumers. Accordingly, some may argue that

168. *Atlantic Richfield Co. v. USA Petroleum Co.*, 495 U.S. 328, 338 (1990) (quoting *Brown Shoe Co. v. United States*, 82 S. Ct. 1502, 1521(1962)) (“The antitrust laws were enacted for ‘the protection of competition, not competitors.’”).

169. *See* Wu, *supra* note 131, at 9 (arguing that enforcers should consider whether the complained-of conduct is “competition on the merits,” namely a better or cheaper product).

170. *See* *United States v. Aetna Inc.*, 240 F. Supp. 3d 1, 18 (D.D.C. 2017) (“[T]he Clayton Act protects ‘competition,’ rather than any particular competitor.”).

171. Of course, defendants can still rebut the *prima facie* harm by showing that the plaintiffs fail to meet the two prongs test.

efficiency should be credited as a part of “competition on the merits.” However, conducts that harm the competitive process may increase efficiency in the short run. For example, efficiency gained from vertical integration may simultaneously create incentives to foreclose downstream unintegrated rivals.¹⁷² For mergers that do not impede on the competitive process, which are most mergers, firms are free to achieve efficiencies through vertical integration. But for vertical mergers that generate substantial efficiency benefits, the potential harm to competition is also more likely.¹⁷³ Crediting efficiency gains based on overall consumer welfare increases¹⁷⁴ would put consumers’ benefits before harms to a fair competitive process. In light of the incipency standard enunciated by the Clayton Act, harms to the competitive process, once established, cannot be cured through efficiency claims.

2. *The Competitive Process Standard is Better Suited to Protect Competition*

This Section addresses two reasons why the protection of competitive process standard is better than the trading partners welfare standard. First, competition serves to safeguard both consumers and competitors. Thus, the trading partner welfare standard is not sufficient to capture the essence of competition when it ignores harms to competitors. Second, the trading partner welfare standard cannot capture dynamic competitive harms and thus is no more administrable than the competitive process standard.

a) *The Competitive Process Serves Both Consumers and Competitors*

Mainstream progressives argue that competition is fundamentally intended to serve consumers.¹⁷⁵ Mere harms to competitors are not actionable harms because “many forms of legitimate competition harm rivals but benefit customers.”¹⁷⁶ Thus, the trading partners welfare standard becomes a useful tool to separate legitimate competitive conduct from illegitimate conduct. However, the trading partner welfare standard presents an interesting issue when applied to the vertical merger context. Since vertical mergers necessarily involve two stages of a supply chain, trading partners in a vertical merger can arguably include downstream unintegrated rivals who rely on the upstream

172. See Yongmin Chen, *On Vertical Mergers and Their Competitive Effects*, 32 RAND J. ECON. 667, 681 (2001) (“[A] firm can raise rivals’ cost through vertical integration if and only if its own cost is reduced through the integration.”).

173. *Id.*

174. See 2020 Vertical Merger Guidelines, *supra* note 19, § 6 (identifying efficiencies gains that lead to lower prices to consumers as potential procompetitive benefits that would counterbalance incentive to foreclose or raise rivals’ costs).

175. Shapiro, *supra* note 15, at 38.

176. *Id.*

supplier. Under a consumer welfare standard, as outlined in the 2010 HMG and 2020 VMG, a “consumer” was considered to be the direct consumer of the downstream firm.¹⁷⁷ Therefore, harms to downstream rivals were not considered in the analysis, despite the fact that they are consumers of the upstream firm prior to the merger.

Under a trading partner welfare standard, a similar question arises: are downstream unintegrated rivals considered as trading partners whose harms are recognized under the expanded standard of trading partner welfare?¹⁷⁸ The mainstream progressives fail to give a consistent answer to this question. The general consensus is that harms to downstream unintegrated rivals are not by themselves sufficient to render a merger anticompetitive.¹⁷⁹

But tensions arise when a vertically integrated firm has the incentive and ability to foreclose downstream rivals, yet at the same time generates cognizable, merger-specific efficiencies that benefit downstream customers.¹⁸⁰ Some commentators acknowledge that when the merged firm has the ability and incentive to raise costs for the unintegrated downstream rivals, the impact “could be said to disrupt competition on the merits.”¹⁸¹ On the other hand, the welfare-based reformers nonetheless suggest a final balancing of the welfare effect on the consumers. Some commentators proposed using a burden-shifting rule that allows a plaintiff to shift the burden to the merging parties once the plaintiff establishes harms to downstream rivals.¹⁸² Then, the merging parties must bear the burden to produce evidence of merger-specific benefits, including accounting for the elimination of double marginalization and other efficiency claims.¹⁸³ Next, if the merging parties are able to rebut the prima facie case, the plaintiff bears the ultimate burden of persuasion to show the net effect on the downstream customers.¹⁸⁴

Under a foreclosure or RRC theory, for example, prima facie harm is established if the plaintiff can show that the merged firm has the ability and

177. *See supra* notes 112–113 and accompanying text.

178. *See* Salop, *supra* note 23, at 1985 (“One key legal and policy issue raised here is whether it should be sufficient for the government just to prove likely higher prices or other injury to the customers of the upstream firms (i.e., the unintegrated downstream competitors) or whether it is also necessary to show harm to the customers of the downstream competitors.”).

179. Shapiro, *supra* note 99, at 320 (“The 2020 VMGs evaluate input foreclosure concerns based on their impact on downstream customers . . . I believe there is a consensus that this is the proper way to evaluate vertical mergers.”).

180. Salop, *supra* note 23, at 1985.

181. *Id.* at 1985.

182. *Id.* at 1986.

183. *Id.*

184. *Id.*

incentive to foreclose or significantly raise rivals' costs.¹⁸⁵ Then, the merging parties can rebut the prima facie case by proving that the efficiencies are cognizable and merger-specific.¹⁸⁶ Ultimately, the case comes down to modeling the welfare tradeoffs on downstream firms' consumers when both foreclosure effect and efficiencies are present.¹⁸⁷

The welfare-based reformers claim that a change in the burden of proof would remove the undue burden on the plaintiff and encourage the parties to "seriously balanc[e]" the pricing effect when necessary.¹⁸⁸ But if the goal is to protect competition on the merits and impede harms to downstream rivals on their ability to compete with the merged firm on the merits, why engage in the final balancing at all? This gap between the harms to the process and harms to consumers demonstrate that mergers that harm the competitive process may not always result in harms to consumers. Alternatively, firms that compete vigorously can produce sub-optimal allocations of resources and may not directly benefit consumers economically. Most vertical mergers that harm downstream rivals are likely to result in harm to consumers.¹⁸⁹ But equating consumers' economic welfare to the vigorousness of competition is both under- and over-inclusive.

b) Welfare Cannot Capture Dynamic Harms, at Least Not Without Sacrificing Administrability

Even if the ultimate goal of competition is to serve consumers, a trading welfare standard can easily fall into the same fraught fixation over qualifiable evidence as the consumer welfare standard. To the extent that the mainstream progressives' rebranding is successful, a trading partner welfare standard is likely to be extremely hard to administer. To begin, "welfare" in a technical sense does not necessarily cover the general notion of consumer interest or supplier interest.¹⁹⁰ Granted, in theory, a welfare standard can be defined broadly enough to encompass a broad range of long-term interests, such as innovation, consumer satisfaction, etc. But in reality, courts often require quantifiable economic analysis as evidence, starting from market definition to the defendant's abilities and incentives to engage in anticompetitive conduct

185. Shapiro, *supra* note 99, at 332 (delineating how agencies may seek to prove a prima facie case of harm to competition).

186. *Id.*

187. *Id.*

188. *Id.*

189. *Id.*

190. Fox, *supra* note 10, at 1161.

post-merger.¹⁹¹ A simple rebranding of consumer welfare would not resolve the issues of technocracy and inhospitality towards less quantifiable interests, particularly when proof of additional harm to consumers is required.

Second and relatedly, broadening the concept of welfare to encompass a broad range of interests, such as innovation and product quality, would greatly undermine the administrability of the welfare standard. To illustrate, consider a vertical merger that would allow the merged firm to gain access to competitively sensitive information about its downstream rivals. Like the government in *UnitedHealth Group*,¹⁹² the plaintiff would allege a harm to innovation based on the theory of data misuse. Assuming the plaintiff can successfully establish a prima facie harm, the defendant would aim to rebut the case by arguing that the merger-specific efficiency benefits of the merger lower prices for consumers. How should a court balance alleged long-term consumer harm stemming from the potential loss of innovation against efficiency gains by the defendant and the alleged short-term consumer welfare gain? While the welfare approach provides a helpful model to trading off conflicting welfare effects, it gives little instruction on how to tradeoff between different types of welfare harms.

Critics of the protection of competitive process standard have frequently attacked the New Brandeis school's process-based approach for its indeterminacy, administrability, and unsophistication.¹⁹³ But it is no less indeterminant or un-administrable than the trading partner welfare standard. Protecting competition and confronting novel business practices is no easy task. It is particularly true if the goal is to avoid false negative error,¹⁹⁴ in light of weakened merger enforcement under the Chicago School's dooming influence. In that sense, the protection of competitive process standard offers a clean slate to define the role of markets and unfair business conduct.

As aforementioned, a process-based standard need not deviate from sound economic learning. Industrial economic theories are and will continue to be helpful in identifying changes to firms' incentives and abilities to prevent rivals from competing on the merits. The protection of competitive process standard

191. See *United States v. AT&T, Inc.*, 310 F. Supp. 3d 161, 190 (D.D.C. 2018), *aff'd*, *United States v. AT&T, Inc.*, 916 F.3d 1029 (D.C. Cir. 2019) (holding that the government must base its case on evidence and facts, instead of antitrust theory and speculation).

192. *UnitedHealth*, 2022 WL 4365867.

193. See, e.g., Hovenkamp, *supra* note 15, at 89 (arguing that "protection for competitive process" operates as a slogan, not as a goal because it "lacks sufficient definition and does not create a meaningful target for measurement").

194. A false negative error in this context means finding no anticompetitive effect when the merger in fact has. See *supra* note 70 (defining false negative and false positive errors in merger analysis).

emphasizes that economic theories and models are tools that help us understand the relationships between business entities, rather than limiting principles. The protection of competition is itself a protection of economic liberty, which inevitably has the indirect effect of protecting other social and political values. It should be driven by a determination of right and wrong conduct in the market. The inquiry is purely about the economy, but not purely economic. Defining desirable market conduct thus requires deeper discussions to draw the line between fairness and efficiencies, individuals and communities.

V. CONCLUSION

Vertical merger enforcement has become the front and center of antitrust debates. It offers a great opportunity to reevaluate the role of antitrust law in our society. Despite the decades of debates over the proper goal of antitrust, antitrust law is about protecting a competitive process. Congress's decision in entrusting "competition" as the governor of the market reflects a careful balancing between the benefits of integrated efficiencies and deconcentrated economic powers; between private contractual and property rights and the broader sense of fairness embedded in our legal system.

Conversations and disagreements about the definition of competition and what role antitrust law should play in facilitating competition are encouraged. Vertical merger enforcement presents a unique opportunity for this debate: Whose harms and whose benefits should we recognize? How should we balance harms and benefits when they are borne by different groups? How should we balance long-term harms and short-term benefits? These are hard questions that require more vigorous discussions about the role of regulation in the markets and the normative values of economic liberty. Hopefully, this Note provides a forum for that.

GOTTA CATCH 'EM ALL: LEGISLATIVE OVERREACH IN FLORIDA AND TEXAS ANTI-MODERATION LAWS

Utkarsh Srivastava[†]

TABLE OF CONTENTS

I.	INTRODUCTION	1438
A.	THE ANTI-MODERATION DEBATE	1438
B.	THE ANTI-MODERATION LAWS AND THE PLATFORMS' APPELLATE CHALLENGE.....	1441
1.	<i>Florida Anti-Moderation Law</i>	1442
2.	<i>Texas Anti-Moderation Law</i>	1442
II.	ANTI-MODERATION LAWS ARE AIMED AT REGULATING THE DIGITAL PUBLIC SQUARE	1444
A.	ORIGIN OF THE ANTI-MODERATION LAWS.....	1444
B.	THE AIMS OF ANTI-MODERATION LAWS.....	1445
1.	<i>Legislative Findings and Lawmakers' Motivations Behind the Anti- Moderation Laws</i>	1446
2.	<i>The Idea of a Public Square, Sphere, or Space</i>	1448
3.	<i>Anti-Moderation Laws Aim to Regulate Platforms Which Affect Political Debate</i>	1451
III.	ANTI-MODERATION LAWS ARE UNCONSTITUTIONAL AS THEY INFRINGE THE FIRST AMENDMENT.....	1452
A.	ANTI-MODERATION LAWS IMPLICATE THE PLATFORMS' FIRST AMENDMENT RIGHTS.....	1452
B.	TESTING THE CONSTITUTIONALITY OF LAWS IMPLICATING THE FIRST AMENDMENT.....	1453
C.	GUIDANCE FROM <i>PACKINGHAM V. NORTH CAROLINA</i>	1454
D.	THE GOVERNMENTAL INTEREST INVOLVED	1455
E.	THE BROAD SWEEP OF FLORIDA AND TEXAS STATUTES.....	1456
1.	<i>Actions Regulated by the Anti-Moderation Laws</i>	1456

DOI: <https://doi.org/10.15779/Z38DF6K47R>

© 2023 Utkarsh Srivastava.

† LL.M. 2023, University of California, Berkeley, School of Law. Sincere thanks to Professors Talha Syed and Tejas Narechania, as well as Bogdan Belei, Meaghan Katz, Zhudi Huang, and the 2022 Law & Technology Writing Workshop at Berkeley Law.

a)	The Florida Law (SB 7072).....	1456
b)	The Texas Law (HB 20).....	1458
2.	<i>Entities Regulated by Anti-Moderation Laws</i>	1459
F.	ANTI-MODERATION LAWS FAIL INTERMEDIATE SCRUTINY.....	1462
G.	ANTI-MODERATION LAWS FAIL STRICT SCRUTINY.....	1463
IV.	ANTI-MODERATION LAWS CANNOT BE REDRAFTED IN A CONSTITUTIONAL MANNER.....	1464
A.	CHANGING THE DEFINITION OF SOCIAL MEDIA PLATFORMS TO INCORPORATE THE MODERN PUBLIC SQUARE WILL NOT SOLVE THE CONSTITUTIONALITY ISSUE.....	1464
B.	CHANGING THE DEFINITION OF SOCIAL MEDIA PLATFORMS TO TARGET SPECIFIC PLATFORMS WILL VIOLATE THE FIRST AMENDMENT.....	1465
V.	CONCLUSION.....	1465

I. INTRODUCTION

A. THE ANTI-MODERATION DEBATE

On January 7th and 8th, 2021, Facebook, X,¹ and YouTube “did what legions of politicians, prosecutors and power brokers had tried and failed to do for years: They pulled the plug on [former] President [Donald] Trump.”² After determining that two of Trump’s tweets might encourage another event like the January 6th storming of the Capitol, X banned Trump permanently.³ Facebook took note of Trump’s praise for Capitol rioters on January 6th and suspended him for 2 years.⁴ Within a few days, YouTube too shut down

1. Twitter was renamed to X in July 2023. Wes Davis, *Twitter is Being Rebranded as X*, VERGE (July 24, 2023), <https://www.theverge.com/2023/7/23/23804629/twitters-rebrand-to-x-may-actually-be-happening-soon>.

2. Kevin Roose, *In Pulling Trump’s Megaphone, Twitter Shows Where Power Now Lies*, N.Y. TIMES (Jan. 9, 2021), <https://www.nytimes.com/2021/01/09/technology/trump-twitter-ban.html>.

3. X, *Permanent Suspension of @realDonaldTrump*, X (Jan. 8, 2021), https://blog.twitter.com/en_us/topics/company/2020/suspension.

4. Nick Clegg, *In Response to Oversight Board, Trump Suspended for Two Years; Will Only Be Reinstated if Conditions Permit*, META (Jun. 4, 2021), <https://about.fb.com/news/2021/06/facebook-response-to-oversight-board-recommendations-trump/>.

Trump's channel.⁵ All three platforms have since reversed their bans on the former President.⁶

As the saying goes, no good deed goes unpunished. Republicans across the country swiftly took action against the platforms in an attempt to rein them in. Florida and Texas enacted laws (SB 7072⁷ and HB 20⁸ respectively) (collectively, the "Anti-Moderation Laws") which ostensibly seek to prevent censorship and interference with digital expression. The Anti-Moderation Laws primarily require tech platforms to carry certain speech without exercising content moderation. The laws also contain certain transparency measures which mandate that platforms publish content moderation policies and inform users of changes to rules, terms, and agreements.

With promises of protecting the First Amendment rights of their citizens, Governors Ron DeSantis and Greg Abbott signed off on laws which would require social media platforms to spend billions of dollars in infrastructure changes.⁹ However, given the sheer breadth of the laws, the platforms would still likely violate the laws despite their efforts to comply.¹⁰ Shortly after their passage, the platforms rushed to the courts asking for stays and scrapping of the laws. After going through the appellate system, the laws have been stayed for now by the courts. The parties asked the Supreme Court to hear the cases

5. Kari Paul, *YouTube Extends Ban on Trump Amid Concerns About Further Violence*, GUARDIAN (Jan. 26, 2021), <https://www.theguardian.com/us-news/2021/jan/26/youtube-trump-ban-suspension>.

6. Nick Clegg, *Ending Suspension of Trump's Accounts With New Guardrails to Deter Repeat Offenses*, META (Jan. 25, 2023), <https://about.fb.com/news/2023/01/trump-facebook-instagram-account-suspension/>; Clare Duffy & Paul LeBlanc, *Elon Musk Restores Donald Trump's Twitter Account*, CNN (Nov. 20, 2022), <https://www.cnn.com/2022/11/19/business/twitter-musk-trump-reinstate/index.html>; Adi Robertson, *Donald Trump Has Started Posting on YouTube Again*, VERGE (Mar. 17, 2023), <https://www.theverge.com/2023/3/17/23644748/donald-trump-youtube-suspension-lifted-presidential-campaign>.

7. FLA. STAT. § 501.2041 (2022).

8. TEX. CIV. PRAC. & REM. CODE ANN. § 143A.002 (West 2021).

9. Madlin Mekelburg, *Texas Social-Media Law Put on Hold Pending Supreme Court Review*, BLOOMBERG (Oct. 11, 2022), <https://www.bloomberglaw.com/ms/product/blaw/document/RJNQZPDWX2PS>.

10. See Charlie Warzel, *Is This the Beginning of the End of the Internet?*, ATLANTIC (Sept. 28, 2022), <https://www.theatlantic.com/ideas/archive/2022/09/netchoice-paxton-first-amendment-social-media-content-moderation/671574/> (discussing hypotheticals which show that the law might be unworkable); Daphne Keller (@daphnehk), X (Sept. 16, 2022, 8:49 PM), <https://twitter.com/daphnehk/status/1570983158665052163> (discussing how the platforms can try to comply with the Texas law); and Mike Masnick, *Just How Incredibly Fucked Up Is Texas' Social Media Content Moderation Law?*, TECHDIRT (May 12, 2022) <https://www.techdirt.com/2022/05/12/just-how-incredibly-fucked-up-is-texas-social-media-content-moderation-law/> (giving examples of how the laws will lead to tremendous amounts of wasteful litigation).

and the court in turn had asked the U.S. Solicitor General to weigh in on the issue.¹¹ The Solicitor General urged the Supreme Court to grant certiorari in the cases but also to only undertake a limited review and exclude the transparency mandates from their review.¹² The Supreme Court has agreed with the Solicitor General and will hear the cases while limiting their review to content moderation questions and not to the transparency mandates.¹³ The Supreme Court is scheduled to hear the oral arguments in the cases on February 26, 2024.¹⁴

Currently, there is a circuit split regarding the Anti-Moderation Laws. The Eleventh Circuit (ruling on the Florida law) and the Fifth Circuit (ruling on the Texas law) came to diametrically opposed answers to similar legal questions. The courts differ on the most basic aspects of First Amendment analysis regarding the laws, such as: whether editorial discretion being exercised by the social media platforms is speech; whether the laws are content-based or content-neutral; whether the platforms are common carriers; and how § 230 of the Communications Decency Act impacts the analysis.

However, there is one aspect of the Anti-Moderation Laws that has largely gone unnoticed. The Anti-Moderation Laws cover a vast ambit of entities within their sweep, given that their definitions of social media platforms are exceedingly broad.¹⁵ This Note will analyze the impact of these broad definitions and how that might impact the First Amendment analysis of the laws. First Amendment scrutiny typically involves the courts looking at the

11. Mike Masnick, *Supreme Court Punts on Florida and Texas Social Media Moderation Laws, Asks US Government to Weigh in*, TECHDIRT (Jan. 23, 2023), <https://www.techdirt.com/2023/01/23/supreme-court-punts-on-florida-and-texas-social-media-moderation-laws-asks-us-government-to-weigh-in/>.

12. Brief for the United States as Amicus Curiae, *Moody v. Netchoice, LLC* (No. 22-277); Brief for the United States as Amicus Curiae, *Netchoice, LLC v. Moody* (No. 22-393); Brief for the United States as Amicus Curiae, *Netchoice, LLC v. Paxton* (No. 22-555).

13. Mike Masnick, *Let's Go! Supreme Court Grants Cert to Hear Cases About Social Media Moderation Laws in Florida & Texas*, TECHDIRT (Sept. 29, 2023), <https://www.techdirt.com/2023/09/29/lets-go-supreme-court-grants-cert-to-hear-cases-about-social-media-moderation-laws-in-florida-texas/>.

14. Amy Howe, *Court schedules February argument session*, SCOTUSBLOG (Jan. 5, 2024), <https://www.scotusblog.com/2024/01/court-schedules-february-argument-session/>.

15. The issue of how unworkable the Anti-Moderation Laws would be for certain entities has been discussed in Mike Masnick, *Did the 5th Circuit Just Make It So That Wikipedia Can No Longer Be Edited in Texas?*, TECHDIRT (Sept. 23 2022), <https://www.techdirt.com/2022/09/23/did-the-5th-circuit-just-make-it-so-that-wikipedia-can-no-longer-be-edited-in-texas/>. This article was a major inspiration for this Note.

governmental interest in the law and whether the law has been drafted properly to further that purpose.¹⁶

Part II examines the aims of the Anti-Moderation Laws and how they inform the governmental interests at play. It looks at the idea of the public square¹⁷ that has been invoked in the drafting of the laws to explain how the aim of the laws is narrower than was originally perceived by everyone other than the legislators.

Part III looks at the First Amendment analysis itself. The First Amendment is implicated in the present case even though no speech (as it is traditionally understood) is being targeted by the laws. Based on whether the law is content-based or content-neutral, courts have used different tests for First Amendment cases. By analyzing *Packingham v. North Carolina*, which is a factually analogous case, this Part conducts First Amendment analysis of the Anti-Moderation Laws. This section specifically looks at the various entities which will be caught in the sweep of the Anti-Moderation Laws and argues that the vast sweep is not in line with the laws' purpose.

The remainder of this section lays out the governmental interest involved in the Anti-Moderation Laws. It proceeds to examine whether the provisions concerning the actions and entities governed by the laws are adequately tailored to pass a First Amendment analysis. The Note weighs the laws against both intermediate and strict scrutiny, and concludes that the laws fail both levels of scrutiny.

Finally, in Part IV, the Note argues that there are no changes to the Anti-Moderation Laws which would allow them to pass First Amendment scrutiny while allowing them to retain their essence.

B. THE ANTI-MODERATION LAWS AND THE PLATFORMS' APPELLATE CHALLENGE

While the two Anti-Moderation Laws are similar in nature, each has certain distinctive features. This sub-Part lays out the details of the laws as well as the judicial challenges they have faced.

16. See *Reed v. Town of Gilbert*, 576 U.S. 155, 171 (2015) (explaining the test for strict scrutiny); Ashutosh Bhagwat, *The Test That Ate Everything: Intermediate Scrutiny in First Amendment Jurisprudence*, 2007 U. ILL. L. REV. 783, 801 (2007) (explaining the test for intermediate scrutiny).

17. This Note refers to the concept of a public square, which has been referred elsewhere to as public sphere or public space. For the purposes of this Note, these terms are synonymous.

1. *Florida Anti-Moderation Law*

The Florida law prohibits social media platforms from restricting posts made by public officials or candidates for office, and further protects those persons from being deplatformed.¹⁸ It extends similar protections to most media organizations, as long as the posted content is not obscene.¹⁹ For all other users, the law prohibits platforms from deplatforming them or reducing their reach without notifying them.²⁰ Finally, it has certain transparency measures, such as requiring publication of standards, which inform users of changes to rules, terms and agreements.²¹ The Florida law applies to all information services, systems, search engines, and access software providers that provide multiple users with access to a server and that cross either of the law's specified revenue or user thresholds.²²

The District Court for the Northern District of Florida issued a preliminary injunction on the grounds that the Florida law was viewpoint-based, violated the First Amendment, and failed a strict scrutiny analysis.²³ The District Court's order was appealed before the Eleventh Circuit which affirmed the preliminary injunction as it applied to the anti-content moderation parts.²⁴ The Eleventh Circuit held that the Florida law triggered First Amendment scrutiny because it restricts the exercise of editorial judgment by the platforms and it would fail even an intermediate scrutiny analysis.²⁵ The Eleventh Circuit order has been appealed before the United States Supreme Court by the Florida Attorney General.²⁶

2. *Texas Anti-Moderation Law*

Unlike the Florida law, the Texas law does not protect users differently based on whether they are candidates for office or part of the media. It simply prohibits social media platforms from censoring users based on their viewpoint.²⁷ It has certain exceptions to the prohibition which involve sexual exploitation of children, incitement of criminal activity, and unlawful

18. FLA. STAT. § 501.2041(1)(c) (2022).

19. *Id.* §§ 501.2041(2)(j), (4).

20. *Id.* § 501.2041(2)(d).

21. *Id.* § 501.2041(2).

22. *Id.* § 501.2041(1)(g).

23. *NetChoice, L.L.C. v. Moody*, 546 F. Supp. 3d 1082 (N.D. Fla. 2021).

24. *NetChoice, L.L.C. v. Att'y Gen., Fla.*, 34 F.4th 1196 (11th Cir. 2022).

25. *Id.* at 1231, 1227.

26. Rebecca Kern, *Florida Appeals 11th Circuit Social Media Ruling to SCOTUS*, POLITICO (Sept. 21, 2022), <https://www.politico.com/news/2022/09/21/florida-appeals-11th-circuit-social-media-ruling-to-supreme-court-00058073>.

27. TEX. CIV. PRAC. & REM. CODE ANN. § 143A.002 (West 2021).

expression, among others.²⁸ Among other requirements, the platforms must disclose certain policies, how they are implemented, and maintain a complaint and appeals system.²⁹ The Texas law applies to websites or applications with over fifty million monthly active users that are open to the public, that allow a user to create an account, and that enable users to communicate with others primarily for posting information.³⁰

The District Court for the Western District of Texas issued a preliminary injunction against the Texas law on the grounds that it violates the platforms' First Amendment rights.³¹ The court found that the Texas law "imposes content-based, viewpoint-based, and speaker-based restrictions" and fails both strict and intermediate scrutiny.³² This preliminary injunction was stayed by the Fifth Circuit without providing any reasons, which would have allowed the law to go into effect.³³

However, the Fifth Circuit stay was then vacated by the Supreme Court, stopping the law from taking effect.³⁴ The majority did not provide any reasoning for their decision. However, Justice Alito wrote a dissent (joined by Justice Thomas and Justice Gorsuch) arguing: (1) the law concerns issues of great importance that should be reviewed by the Supreme Court; (2) it is unclear if the platforms will succeed in their lawsuit against the Texas law under existing constitutional law; (3) the law and the applicants' business models are novel; (4) the application of existing precedents (which predate the internet) to large social media companies is not obvious; and (5) Texas should not be required to seek preclearance from the federal courts before putting its laws into effect and the preliminary injunction was a "significant intrusion on state sovereignty."³⁵

The Fifth Circuit reviewed the Texas law again after the Supreme Court's ruling.³⁶ It found the law to be constitutional as it does not compel or obstruct the platforms' own speech in any way and the platforms had no First Amendment right to censor users.³⁷ In a First Amendment challenge, according to the Fifth Circuit, a plaintiff must show that the impugned law

28. *Id.* § 143A.006.

29. TEX. BUS. & COM. CODE ANN. §§ 120.051–120.053, 120.101 (West 2021).

30. *Id.* § 120.001(1).

31. *NetChoice, L.L.C. v. Paxton*, 573 F. Supp. 3d 1092 (W.D. Tex. 2021).

32. *Id.* at 1114.

33. *NetChoice, L.L.C. v. Paxton*, No. 21-51178, 2022 WL 1537249 (5th Cir. May 11, 2022).

34. *NetChoice, L.L.C. v. Paxton*, 142 S. Ct. 1715 (2022).

35. *Id.* at 1716.

36. *NetChoice, L.L.C. v. Paxton*, 49 F.4th 439 (5th Cir. 2022).

37. *Id.* at 494.

either compels the host to speak or restricts the host's speech, and the court found that the Texas law does neither.³⁸ The Fifth Circuit applied intermediate scrutiny applied because they found the law to be content-neutral.³⁹ According to the Fifth Circuit, the Texas law satisfies intermediate scrutiny because it advanced an important governmental interest in protecting the free exchange of ideas and information, and it did not burden substantially more speech than necessary to further the state's interests.⁴⁰ Accordingly, the Fifth Circuit vacated the preliminary injunction, and the case was remanded.⁴¹ The ruling was, however, put on hold while the parties ask the Supreme Court to hear the case.⁴² The law has now been challenged before the United States Supreme Court by the platforms.⁴³

II. ANTI-MODERATION LAWS ARE AIMED AT REGULATING THE DIGITAL PUBLIC SQUARE

A. ORIGIN OF THE ANTI-MODERATION LAWS

The Anti-Moderation Laws were conceived in the wake of major social media platforms banning former President Trump. However, there was an intervening step in the story involving the highest court in the land. In *Biden v. Knight First Amendment Institute at Columbia University*, which was eventually rendered moot due to Trump losing the 2020 presidential election,⁴⁴ Justice Clarence Thomas wrote a concurring opinion which set in motion ideas that were eventually heavily relied upon in the drafting of the Florida and Texas statutes.⁴⁵ The case involved President Trump blocking several users on his X account.⁴⁶ The Second Circuit found this blocking to violate the First Amendment as the President's X account was a public forum where he acted in a governmental capacity while blocking users.⁴⁷ The blocking was thus

38. *Id.* at 459.

39. *Id.* at 480.

40. *Id.* at 482–83.

41. *Id.* at 494.

42. Mike Masnick, *Texas' Ridiculous Content Moderation Bill Put on Hold Until The Supreme Court Can Consider It*, TECHDIRT (Oct. 13, 2022), <https://www.techdirt.com/2022/10/13/texas-ridiculous-content-moderation-bill-put-on-hold-until-the-supreme-court-can-consider-it/>.

43. Brian Fung, *Tech Groups Ask Supreme Court to Rule on Hot-Button Texas Social Media Law*, CNN (Dec. 15, 2022), <https://www.cnn.com/2022/12/15/tech/tech-groups-supreme-court-texas-social-media-law/index.html>.

44. 141 S. Ct. 1220.

45. Blake Ellis Reid, *Uncommon Carriage*, 76 STAN. L. REV. 89 (2024).

46. *Knight First Amend. Inst. at Columbia Univ. v. Trump*, 928 F.3d 226, 230 (2d Cir. 2019).

47. *Id.* at 238, 236.

unconstitutional viewpoint discrimination and could not stand in light of the First Amendment.⁴⁸

The case then reached the Supreme Court.⁴⁹ By this time however, X itself had banned Trump.⁵⁰ The Supreme Court remanded the case back to the Second Circuit with instructions to dismiss it as moot.⁵¹ However, Justice Thomas took the opportunity to express his views on X's⁵² actions.⁵³

Justice Thomas found the “private, concentrated control over online content and platforms available to the public” to be problematic.⁵⁴ He then laid out the beginnings of a “solution” which relied on doctrines that “limit the right of a private company to exclude” including “common carrier” and “places of public accommodation.”⁵⁵

While concluding his opinion, Justice Thomas stated: (1) if the aim is to ensure “speech that is not smothered, then the more glaring concern must perforce be the dominant digital platforms themselves”; and (2) private digital platforms hold “the right to cut off speech . . . most powerfully”⁵⁶ He ended his opinion by stating that the relevance of this power for First Amendment purposes and “the extent to which [it] can lawfully be modified” were interesting and important questions.⁵⁷

B. THE AIMS OF ANTI-MODERATION LAWS

The questions raised by Justice Thomas ceased to be hypothetical as Florida and Texas enacted Anti-Moderation Laws that same year.⁵⁸ These laws restricted the ability of social media companies to moderate content on their platforms.⁵⁹ This Section of the Note will lay out the motivations behind the Anti-Moderation Laws and elaborate on the concept of a public square. It then argues that the laws were meant to regulate only those platforms which affect political debate.

48. *Id.* at 239.

49. *Biden v. Knight First Amendment Inst. at Columbia Univ.*, 141 S. Ct. 1220.

50. X, *supra* note 3.

51. 141 S. Ct. at 1220.

52. X was not a party to the action before the court.

53. 141 S. Ct. at 1221–27.

54. *Id.* at 1222.

55. *Id.* at 1222–23.

56. *Id.* at 1227.

57. *Id.*

58. FLA. STAT. § 501.2041 (2022); TEX. CIV. PRAC. & REM. CODE ANN. §143A.002. (West 2021).

59. *Id.*

1. *Legislative Findings and Lawmakers' Motivations Behind the Anti-Moderation Laws*

The two Anti-Moderation Laws are primarily aimed at protecting the First Amendment rights of citizens and are very cognizant of the importance of social media platforms in this regard. The lawmakers wanted to protect citizens from the platforms' private censorship. However, both the legislative findings and the lawmakers' statements show that the concern is only about certain specific platforms and not all social media platforms.

Florida's SB 7072 starts with the following declarations: (1) "Floridians increasingly rely on social media platforms to express their opinions"; (2) "Social media platforms have transformed into the new public town square"; (3) "Social media platforms have become as important for conveying public opinion as public utilities are for supporting modern society"; and (4) "Social media platforms hold a unique place in preserving First Amendment protections for all Floridians and should be treated similarly to common carriers."⁶⁰ These declarations show that law is concerned only with those social media platforms which are equivalent to public utilities and are essential to safeguarding Floridians' First Amendment rights.

The official press release accompanying the Florida law states that the law aims to ensure protection against "Silicon Valley elites" by taking back the "virtual public square."⁶¹ The Florida law became a priority for Governor Ron DeSantis after X and Facebook blocked Trump from their platforms.⁶² The sponsor of the House version of the bill admitted that he started pursuing the bill after X and Facebook's response to the *New York Post* story on Hunter Biden.⁶³ This further illustrates the focus of the laws on only the biggest platforms like X and Facebook.

Texas's HB 20 starts with declarations similar to the Florida law: (1) "[S]ocial media platforms function as common carriers, are affected with a public interest, are central public forums for public debate"; and (2) "[S]ocial media platforms with the largest number of users are common

60. S.B. 7072, 2021 Leg., Reg. Sess. (Fla. 2021).

61. *Governor Ron DeSantis Signs Bill to Stop the Censorship of Floridians by Big Tech* (May 24, 2021), <https://www.flgov.com/2021/05/24/governor-ron-desantis-signs-bill-to-stop-the-censorship-of-floridians-by-big-tech/>.

62. Jim Saunders & Tom Urban, *Social Media Crackdown Clears Florida Senate, Giving Gov. Ron DeSantis One of His Top Priorities*, S. FLA. SUN SENTINEL (Apr. 27, 2021), <https://www.sun-sentinel.com/news/politics/fl-ne-nsf-florida-senate-approves-social-media-crackdown-20210426-hcbykznsca4ngpyaa5a667nwm-story.html>.

63. Mary Ellen Klas, *Florida Lawmakers Advance Bill to Penalize Social Media Companies*, TAMPA BAY TIMES (Apr. 27, 2021), <https://www.tampabay.com/news/florida-politics/2021/04/27/florida-lawmakers-advance-bill-to-penalize-social-media-companies/>.

carriers by virtue of their market dominance.”⁶⁴ Here too, the focus is on the biggest social media platforms, i.e., the dominant ones and the ones who are central public forms for debate.

Governor Abbott, while signing the Texas bill into law, called social media websites the “modern-day public square.”⁶⁵ The statement of intent for the Texas law notes that the need for protection from private censorship stems from “the nearly universal adoption of a few sites.”⁶⁶ Further, during the Texas Senate debates over the law, the bill’s author (Representative Briscoe Cain) cited Justice Kennedy’s opinion in *Packingham v. North Carolina*⁶⁷ to say that the public square the law sought to regulate is a few dominant websites which “for many are the principal sources for knowing current events, checking ads for employment, speaking, and listening.”⁶⁸ Cain also noted Elizabeth Warren’s criticism of Facebook⁶⁹ wherein the former presidential candidate said, “we must ensure that today’s tech giants do not . . . wield so much power that they can undermine our democracy.”⁷⁰ Focusing in on Big Tech’s outsized influence, Cain emphasized that “a small handful of social media sites drive the national narrative and have massive influence over the progress and developments of medicine and science, social justice movements, election outcomes, and public thought.”⁷¹ The politicians’ intent behind the law is therefore to target the few dominant websites which can undermine democracy if their ability to drive the narrative goes unchecked.

While Cain did say that the bill doesn’t target specific companies, only those big enough to be a public square,⁷² both the law and his statements indicate that not all big websites are meant to be regulated. Indeed, the only platforms mentioned in the entire Senate debate are Facebook, X, and Instagram.⁷³

64. H.B. 20, 87th Leg., 2d Called Sess. (Tex. 2021).

65. *Governor Abbott Signs Law Protecting Texans from Wrongful Social Media Censorship*, OFF. TEX. GOVERNOR (Sept. 9, 2021), <https://gov.texas.gov/news/post/governor-abbott-signs-law-protecting-texans-from-wrongful-social-media-censorship>.

66. BRISCOE CAIN, SELECT COMM. ON CONST. RTS. & REMEDIES, BILL ANALYSIS, H.B. 20, 2021 Leg., 87th Sess. (Tex. 2021).

67. 137 S. Ct. 1730 (2017).

68. House Journal, 87th Leg., 2d Called Sess., Fourth Day Supplement S157, S175 (Tex. 2021), <https://journals.house.texas.gov/hjrnl/872/pdf/87C2DAY04SUPPLEMENT.pdf>.

69. Elizabeth Warren, *Here’s How We Can Break Up Big Tech*, MEDIUM (Mar. 8, 2019), <https://medium.com/@teamwarren/heres-how-we-can-break-up-big-tech-9ad9e0da324c>.

70. House Journal, 87th Leg., 2d Called Sess., Fourth Day Supplement S157, S175 (Tex. 2021), <https://journals.house.texas.gov/hjrnl/872/pdf/87C2DAY04SUPPLEMENT.pdf>.

71. *Id.*

72. *Id.* at S177.

73. *Id.*

The legislative findings and the statements by the politicians involved in the enactment of the Anti-Moderation Laws clearly show that the aim of the laws is not to target all or even most social media companies, but only those that directly affect the public's political opinions and electoral outcomes.

2. *The Idea of a Public Square, Sphere, or Space*

The claim that the internet is the “modern public square” has been repeated so often that it has now become conventional wisdom.⁷⁴ However, neither Justice Kennedy (when he called internet the modern public square in *Packingham*⁷⁵) nor any of the lawmakers involved in the enactment of the Anti-Moderation Laws have explained what they mean by it. Therefore, to understand why and how the internet should be regulated in its capacity as a modern public square, one must understand what a “public square” means in this context.

Public squares are a concept that has long been discussed in a sociological sense. Hannah Arendt in her seminal work *The Human Condition* discussed the theme of the common place for public discussions.⁷⁶ Arendt discusses the concept as a space of appearance where it provides the “widest possible publicity” to individuals and the option of being seen and heard by everyone.⁷⁷ This is a realm that “is common to all of us and distinguished from our privately owned place in it.”⁷⁸ Given that Arendt was writing in a time before the internet, her articulation of the public realm is more spatial, as it was an improvised place that arose from the actions and words of people who came together to undertake common activities.⁷⁹

While Arendt may have started the discussion on public spaces, the concept of public square is most commonly associated with another German, Jürgen Habermas. Habermas introduced the concept of a “public sphere” which is a discursive arena where people discuss matters of common concern.⁸⁰ It is not part of the state and “is ideally the site of free, unrestricted, rational communication.”⁸¹ It is a “site for the production and circulation of

74. Mary Anne Franks, *Beyond the Public Square: Imagining Digital Democracy*, 131 YALE L.J. F. 427, 427 (2021–2022).

75. *Packingham v. North Carolina*, 137 S. Ct. 1730, 1732 (2017).

76. HANNAH ARENDT, *THE HUMAN CONDITION* 50 (1958).

77. *Id.*

78. *Id.*

79. Alexey Salikov, *Hannah Arendt, Jürgen Habermas, and Rethinking the Public Sphere in the Age of Social Media*, 17 RUSS. SOC. REV., no. 4, 2018, at 88, 93.

80. Nancy Fraser, *The Theory of the Public Sphere*, in *THE HABERMAS HANDBOOK* 245 (Hauke Brunkhorst, Regina Kreide & Ristina Lafont eds., 2018).

81. *Id.*

discourses that can in principle be critical of the state.”⁸² It is not part of a market, as it is “a theater for debating and deliberating rather than for buying and selling.”⁸³ It allows the people to scrutinize and hold state officials accountable, as well as rein in the operation of private power.⁸⁴ In the internet context, this excludes government websites and ecommerce websites from the discussion.

Public spheres are thus places where people “establish common goals in pursuit of our common good,” “promote the general welfare,” and weave together a “democratic political culture.”⁸⁵ Further, the public sphere is linked to political communication.⁸⁶ Citizens meet in this space of political communication and form political opinions.⁸⁷

The initial conception of the Habermasian public sphere drew criticism by scholars who pointed out that he had not adequately incorporated the effects of structural inequalities which prevented many from participating in the public sphere at a level which was on par with others.⁸⁸ His critics also said that he had not considered the full force of structural issues that choke “the flow of public opinion from society to the state” and thereby deprive it of “political muscle.”⁸⁹ Essentially, the initial idea of a public sphere was deemed far too idealistic and failed to account for power dynamics in society, both between the people themselves and between the people and the state. In the current context, this would mean that websites with little viewership should not be seen as a public square as they would have little “political muscle” to effect any real change.

Habermas responded to some of these criticisms by providing a revised idea of the public sphere as a decentralized network of multiple, overlapping communicative spaces.⁹⁰ He reiterated this in a later work where he described the public sphere as “an intermediary system of communication between

82. Franks, *supra* note 74, at 446 (citing JÜRGEN HABERMAS, *THE STRUCTURAL TRANSFORMATION OF THE PUBLIC SPHERE: AN INQUIRY INTO A CATEGORY OF BOURGEOIS SOCIETY* 36 (Thomas Burger trans., 1991)).

83. *Id.*

84. Fraser, *supra* note 80.

85. Leo Casey, *Why We Defend the Public Square*, ALBERT SHANKER INST. (May 7, 2015), <https://www.shankerinstitute.org/blog/why-we-defend-public-square>.

86. Christian Fuchs, *The Digital Commons and the Digital Public Sphere: How to Advance Digital Democracy Today*, 16 WESTMINSTER PAPERS COMM. & CULTURE, no. 1, 2021, at 9, 13, <https://doi.org/10.16997/wpcc.917>.

87. *Id.*

88. Fraser, *supra* note 81, at 249.

89. *Id.*

90. *Id.* at 249–50.

formally organized and informal face-to-face deliberations in arenas at both the top and the bottom of the political system.”⁹¹

While the internet has been argued to not be a public sphere,⁹² such arguments are outdated in light of the massive electoral changes brought about by the power of the internet in recent years. Starting with the Arab Spring in 2011 and right up to the 2022 U.S. midterm elections, the internet has been a crucial part of deciding who wins elections. However, it is not the entirety of the internet that comprises this public square, given that smaller websites are unable to effect real change on the political stage. The three biggest websites that affect democratic debate are Facebook, X, and YouTube—the “dominant platforms in global content sharing.”⁹³ The importance of these three platforms can be gauged by the fact that their actions lead to calls for regulating social media platforms.⁹⁴ Therefore, the public square in this context is limited to these three websites.

The Arendtian and Habermasian ideas discussed above are at the root of the general understanding of a “public square” in the context of social media platform regulation. Politicians, judges, and critics associate greater importance to the moderation of content on certain platforms because they believe that this moderation is leading to changes in electoral outcomes, something that is associated with discussion in a “public square.” This also falls within the definition Habermas proposed⁹⁵ because it is an “intermediary system of communication” in arenas across the political system. As a result, regulation

91. Jürgen Habermas, *Political Communication in Media Society*, COMMUNICATION THEORY 16, 415 (2006). Reading his works together, Habermas’s argument for a desirable public sphere is as follows: (1) the public sphere must remain independent as it has its own code of rational-critical debate; and (2) this independence is required both from state and private actors. Lewis A. Friedland, Thomas Hove & Hernando Rojas, *The Networked Public Sphere*, 13 JAVNOST - PUB., no. 4, 2006, at 5, 12. In case of Anti-Moderation Laws, this creates an issue as they are ostensibly state action which results in independence from private actors in the public square. However, this does not affect the constitutional analysis of the laws.

92. Stuart Jeffries, *What The Philosopher Saw*, FIN. TIMES (May 1, 2010) <https://www.ft.com/content/eda3bcd8-5327-11df-813e-00144feab49a> (arguing that the web cannot produce public spheres because users are dispersed and form opinions simultaneously); Zizi Papacharissi, *The Virtual Sphere*, 4 NEW MEDIA & SOC. 1, 9 (2002) (arguing that the internet is merely a “new public space for politically oriented conversation” and has not ascended to the level of a public sphere).

93. Kate Klonick, *The New Governors: The People, Rules, and Processes Governing Online Speech*, 131 HARV. L. REV. 1598, 1603 (2018).

94. See Dawn Carla Nunziato, *Protecting Free Speech and Due Process Values on Dominant Social Media Platforms*, 73 HASTINGS L.J. 1255, 1262–69 (2022) (tracing the call for regulation of social media platforms notes to actions only by X, Facebook and YouTube); *infra* Part II.B.4 (establishing that Anti-Moderation Laws aim to regulate platforms which affect political debate).

95. Habermas, *supra* note 91.

of the modern public square is informed by our collective understanding of which platforms are consequential.

3. *Anti-Moderation Laws Aim to Regulate Platforms Which Affect Political Debate*

While it is hard to attribute a clear purpose to the lawmakers in Florida and Texas, it certainly appears that they intended to free up the public square (comprised of Facebook, X, and YouTube) from the control of private companies rather than regulate the internet as an expansive virtual public space. It is extremely important to clarify this purpose because First Amendment scrutiny of the laws requires a clear purpose in place to weigh the laws against.

The Florida law is specifically aimed at protecting political candidates from moderation by the social media platforms,⁹⁶ while the Texas law is aimed at prohibiting all moderation based on “viewpoint.”⁹⁷ These laws stem from actions by only the major platforms, for example, X, Facebook, and YouTube.⁹⁸ Even scholars who are proponents of laws prohibiting social media platforms from discriminating on the basis of viewpoint favor legislation which targets discrimination on the basis of political views.⁹⁹ Indeed, they note that such prohibiting provisions “would be narrowly tailored because [they require] only that platforms refrain from censoring speech on the basis of its political content.”¹⁰⁰

Both statutes are aimed at regulating the dominant social media platforms which bear some resemblance to public utilities and are central for public debate. The analogy to public squares further strengthens this idea as those were meant to affect general welfare and political issues.¹⁰¹ Given the laws themselves and the circumstances surrounding their enactment, they are meant to regulate only the platforms which influence political debates and the narrative around government. In the current scenario, these are limited to Facebook, X, and YouTube.

96. FLA. STAT. § 501.2041(2)(h) (2022).

97. TEX. CIV. PRAC. & REM. CODE ANN. § 143A.002 (West 2021).

98. See Nunziato, *supra* note 94, at 1262–69 (2022) (tracing the call for regulation of social media platforms notes to actions only by X, Facebook and YouTube).

99. Prasad Krishnamurthy & Erwin Chemerinsky, *How Congress Can Prevent Big Tech from Becoming the Speech Police*, HILL (Feb. 18, 2021), <https://thehill.com/opinion/judiciary/539341-how-congress-can-prevent-big-tech-from-becoming-the-speech-police>.

100. *Id.*

101. *Supra* Section II.B.3 (discussing the idea of a public square/sphere/space).

III. ANTI-MODERATION LAWS ARE UNCONSTITUTIONAL AS THEY INFRINGE THE FIRST AMENDMENT

A. ANTI-MODERATION LAWS IMPLICATE THE PLATFORMS' FIRST AMENDMENT RIGHTS

The First Amendment prohibits the enactment of any laws which abridge the freedom of speech or of the press.¹⁰² This prohibition on state action in turn provides a right to people and organizations to exercise their freedom of speech. This right has been read to include the right of private organizations to exercise editorial discretion.¹⁰³

In *Miami Herald Publishing Co. v. Tornillo*, the Supreme Court struck down a Florida statute which required newspapers to give political candidates free space to reply to columns which attacked the candidate.¹⁰⁴ The case upheld the right of a newspaper to decide what it publishes and held that the exercise of editorial control and judgment comprises “the choice of material to go into a newspaper . . . and treatment of public issues and public officials—whether fair or unfair.”¹⁰⁵

More recently, in *Manhattan Community Access Corp. v. Halleck*, the Supreme Court has held that a corporation operating public access channels had the right to exclude certain speakers.¹⁰⁶ Since the corporation was a private actor, it was not limited by the First Amendment with regard to how it exercised “editorial discretion over the speech and speakers on its public access channels.”¹⁰⁷ Further, that case supported the idea that corporations have First Amendment rights.¹⁰⁸ The entity operating public access channels in that case was a corporation and was held to have First Amendment rights to exercise editorial discretion.¹⁰⁹ This position had been clearly recognized earlier in *Citizens United v. FEC*, where the Supreme Court held: “First Amendment protection extends to corporations.”¹¹⁰

Since corporations have First Amendment rights, and editorial discretion is a form of speech protected under the First Amendment, the restrictions put

102. U.S. CONST. amend. I.

103. Evelyn Douek & Genevieve Lakier, *Rereading “Editorial Discretion,”* KNIGHT FIRST AMEND. INST. COLUM. U. (Oct. 24, 2022), <https://knightcolumbia.org/blog/rereading-editorial-discretion>.

104. 418 U.S. 241 (1974).

105. *Id.* at 258.

106. 139 S. Ct. 1921 (2019).

107. *Id.* at 1933.

108. *Id.*

109. *Id.*

110. 558 U.S. 310, 342 (2010).

in place by the Anti-Moderation Laws implicate the First Amendment rights of social media platforms. Therefore, these Anti-Moderation Laws must be analyzed in light of the tests laid down by the courts for constitutionality of laws implicating the First Amendment.

B. TESTING THE CONSTITUTIONALITY OF LAWS IMPLICATING THE FIRST AMENDMENT

While evaluating cases concerning the free speech clause of the First Amendment, the Supreme Court has held that content-based restrictions must satisfy strict scrutiny while content-neutral laws only need to satisfy intermediate scrutiny.¹¹¹ Content-based laws are those which “suppress, disadvantage, or impose differential burdens upon speech because of its content.”¹¹² Content-neutral laws are those which “are unrelated to the content of speech.”¹¹³

This Note does not discuss whether the Anti-Moderation Laws are content-based or content-neutral because it posits that the laws fail the lower standard of intermediate scrutiny and will therefore automatically fail the higher standard of strict scrutiny. Therefore, regardless of whether the laws are content-based or content-neutral, the Anti-Moderation Laws fail the First Amendment analysis.

Under strict scrutiny, the Government must prove that the restriction furthers a compelling interest and is narrowly tailored to achieve that interest.¹¹⁴ In contrast, the test for intermediate scrutiny is more nuanced (or in some ways more incoherent). Ashutosh Bhagwat has traced eight different kinds of free speech cases in which intermediate scrutiny was applied to First Amendment cases.¹¹⁵ Based on this analysis, Bhagwat concludes that the Supreme Court appears to have landed on the following test for intermediate scrutiny: “laws will be upheld so long as they serve some sort of a significant/substantial/important governmental interest and are reasonably well tailored to that purpose (i.e., not unreasonably overbroad).”¹¹⁶

The First Amendment analysis of the Anti-Moderation Laws thus takes the following form:

111. *Turner Broad. Sys. v. FEC*, 512 U.S. 622, 642 (1994).

112. *Id.*

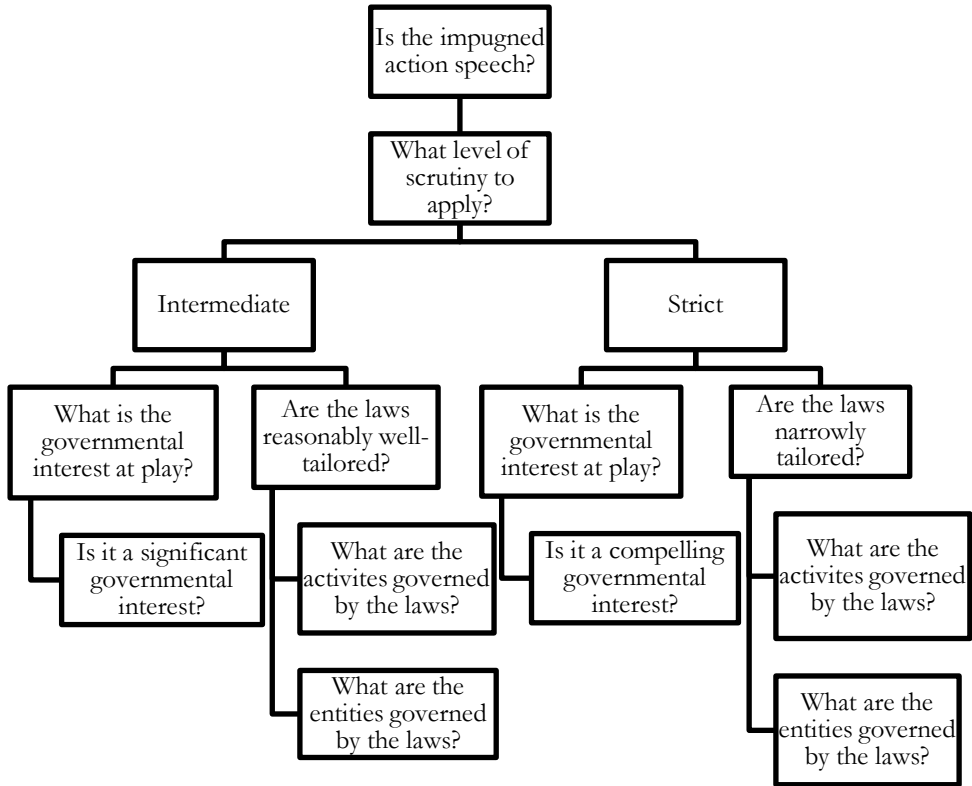
113. *Id.*

114. *Reed v. Town of Gilbert*, 576 U.S. 155, 171 (2015).

115. Bhagwat, *supra* note 16, at 788–800.

116. *Id.* at 801.

Figure 1: Flowchart of Steps in Supreme Court's First Amendment Analysis



The constitutional analysis of the Anti-Moderation Laws should be instructed by *Packingham*, a case where a state statute was struck down because its scope implicated a vast number of websites.¹¹⁷

C. GUIDANCE FROM *PACKINGHAM V. NORTH CAROLINA*

The defendant in *Packingham v. North Carolina* was a registered sex offender who expressed happiness on Facebook when his traffic ticket was dismissed.¹¹⁸ A lower court held that the defendant's action violated a North Carolina statute which criminalized the access of most social media websites by registered sex offenders.¹¹⁹ Specifically, the law prohibited registered sex offenders from accessing a "commercial social networking Web site where the sex offender [knew] that the site permits minor children to become members or to create or maintain personal Web pages."¹²⁰

117. *Packingham v. North Carolina*, 137 S. Ct. 1730, 1730 (2017)

118. *Id.* at 1734.

119. *Id.* at 1731.

120. *Id.* at 1733.

Aimed at protecting children from sexual abuse,¹²¹ the law covered within its sweep websites which fulfilled all the following criteria:¹²² (1) the operator of the website derived revenue from the website; (2) the website facilitated “social introduction” between people; (3) the website allowed users to create profiles; and (4) the website provided users/visitors “mechanisms to communicate with other users.”

The Supreme Court struck down the statute as unconstitutional on First Amendment grounds and noted that the statute would “bar access not only to commonplace social media websites but also to websites as varied as Amazon.com, Washingtonpost.com, and Webmd.com.”¹²³ Noting that North Carolina had not been able to show that the “sweeping law is necessary or legitimate to serve” the government’s purpose,¹²⁴ the court ruled that the statute would fail even intermediate scrutiny.¹²⁵

Justice Alito in his concurring opinion explained the application of intermediate scrutiny in greater detail.¹²⁶ Applying the statute’s criteria (enumerated above), he showed how it would cover almost any website.¹²⁷ Then using examples, he showed that the statute bars “access to [many] websites which are most unlikely to facilitate the commission of a sex crime against a child.”¹²⁸ Since the statute had a broad reach, and barring registered sex offenders from such a large number of websites did “not appreciably advance the State’s goal of protecting children from . . . sex offenders,” the law was unconstitutional.¹²⁹

D. THE GOVERNMENTAL INTEREST INVOLVED

The first aspect of a First Amendment analysis (under either strict scrutiny or intermediate scrutiny) is to determine the governmental interest at play.

Before the Eleventh Circuit, Florida failed to offer a governmental interest in its anti-moderation law, which left the court to theorize as to the potential governmental interest.¹³⁰ The court came up with two such interests to carry out a First Amendment analysis: (1) “counteracting unfair private censorship that privileges some viewpoints over others on social-media platforms”; and

121. *Id.* at 1740.

122. *Id.* at 1733–34.

123. *Id.* at 1736.

124. *Id.* at 1737.

125. *Id.* at 1736.

126. *Id.* at 1740–43.

127. *Id.* at 1740–41.

128. *Id.* at 1741.

129. *Id.* at 1743.

130. *NetChoice, L.L.C. v. Att’y Gen., Fla.*, 34 F.4th 1196, 1228 (11th Cir. 2022).

(2) “promoting the widespread dissemination of information from a multiplicity of sources.”¹³¹ In its own First Amendment analysis, the Fifth Circuit considered “protecting the free exchange of ideas and information” in Texas as the relevant governmental interest.¹³²

However, in light of the analysis undertaken above in Part II, this Note posits that the aim of the Anti-Moderation Laws is not to regulate all websites. Instead, the legislative aim is to regulate the modern public square (Facebook, X, and YouTube).¹³³ This renders the governmental interests discussed by the Eleventh and Fifth Circuits of little use, and one must instead consider the alternative governmental interest in regulating the modern public square. Regulation of the modern public square can be considered a valid government interest as governments should be allowed to act in the interest of preserving democracy and fair electoral practices. In *Citizens United v. FEC*, Justice Stevens in a partly concurring opinion found there to be a compelling government interest in “preserving the integrity of the electoral process . . . sustaining the active, alert responsibility of the individual citizen in a democracy for the wise conduct of the government and maintaining the individual citizen’s confidence in government.”¹³⁴

Therefore, the Anti-Moderation Laws (as far as they regulate the modern public square) can be considered as furthering a compelling government interest which is required for a strict scrutiny analysis. Since the governmental interest is a compelling one, it is also a significant/substantial/important governmental interest (as required for an intermediate scrutiny analysis) because a compelling government interest is surely a significant/substantial/important governmental interest as well.

E. THE BROAD SWEEP OF FLORIDA AND TEXAS STATUTES

1. *Actions Regulated by the Anti-Moderation Laws*

a) The Florida Law (SB 7072)

The Florida law largely protects candidates for office and journalistic enterprises from editorial discretion. It applies to a broader class of social media platforms (compared to the Texas law) but requires them to comply with it only with regards to certain classes of users.

131. *Id.*

132. *NetChoice, L.L.C. v. Paxton*, 49 F.4th 439, 482 (5th Cir. 2022).

133. *Supra* Section II.B.4 (arguing that Anti-Moderation Laws aim to regulate platforms affecting political debate).

134. *Citizens United v. Fed. Election Comm’n*, 558 U.S. 310, 440 (2010).

It prohibits “social media platforms” from deplatforming¹³⁵ or shadowbanning¹³⁶ a candidate for office who the platform knows to be a candidate.¹³⁷ The platform also cannot prioritize or de-prioritize posts¹³⁸ by such candidates.¹³⁹ These are all actions that have been taken by social media platforms to respond to speech on their platforms that they find undesirable.¹⁴⁰ While this list does not encompass the vast universe of actions that platforms can take,¹⁴¹ it does cover the most often-discussed remedies for violation of platform rules.

The protections are available to the candidates only between the date of qualification and the date they cease to be a candidate.¹⁴² The platform must provide each user a way to identify themselves as a qualified candidate in a manner that allows the platform to confirm their candidature on the relevant election website.¹⁴³ Further, the platform cannot censor,¹⁴⁴ deplatform, or

135. FLA. STAT. § 501.2041(1)(c) (defining deplatforming as the “action or practice by a social media platform to permanently delete or ban a user or to temporarily delete or ban a user from the social media platform for more than 14 days”).

136. *Id.* § 501.2041(1)(f) (defining shadow-banning as the “action by a social media platform, through any means, whether the action is determined by a natural person or an algorithm, to limit or eliminate the exposure of a user or content or material posted by a user to other users of the social media platform,” including “acts of shadow banning by a social media platform which are not readily apparent to a user”).

137. *Id.* § 501.2041(2)(h).

138. *Id.* § 501.2041(1)(e) (defining post-prioritization as the “action by a social media platform to place, feature, or prioritize certain content or material ahead of, below, or in a more or less prominent position than others in a newsfeed, a feed, a view, or in search results. the term does not include post-prioritization of content and material of a third party, including other users, based on payments by that third party, to the social media platform”).

139. *Id.* § 501.2041(2)(h).

140. *See* Eric Goldman, *Content Moderation Remedies*, 28 MICH. TECH. L. REV. 1, 23–40 (2021) (discussing the various remedies exercised by platforms for violations of their rules).

141. *See id.* (noting that the remedies the platforms can employ include editing/redacting content, adding warnings to the content, disabling comments, removing credibility badges, forfeiting earnings etc. Many of these will not be covered by the Florida law).

142. FLA. STAT. § 501.2041(2)(h) (2022).

143. *Id.*

144. *Id.* § 501.2041(1)(b) (defining “censor” as “any action taken by a social media platform to delete, regulate, restrict, edit, alter, inhibit the publication or republication of, suspend a right to post, remove, or post an addendum to any content or material posted by a user. the term also includes actions to inhibit the ability of a user to be viewable by or to interact with another user of the social media platform”).

shadow ban a journalistic enterprise¹⁴⁵ based on the content of its publication or broadcast unless the content is obscene.¹⁴⁶

The law allows for moderating the content of most users but adds an obligation on the platforms as it prohibits them from censoring content, shadow banning, or deplatforming a user without notifying the user.¹⁴⁷ The exception to this notification requirement is if the content being censored is obscene.¹⁴⁸

The Florida law also has certain transparency measures which are not relevant for the purpose of this Note.

b) The Texas Law (HB 20)

The Texas law prohibits social media platforms from censoring¹⁴⁹ users based on one or more of the following criteria: (1) the user's viewpoint; (2) the viewpoint represented in the user's experience; or (3) the user being in Texas.¹⁵⁰ The prohibition applies regardless of whether the viewpoint is expressed on the platform or off of it.¹⁵¹ However, the platform may censor content that fulfills one or more of the following criteria: (1) federal law has specifically authorized censoring of that content; (2) the content is the subject of a request from an organization for preventing sexual exploitation of children and protecting sexual abuse survivors from ongoing harassment; (3) it directly incites criminal activity or threatens a "person or group because of their race, color, disability, religion, national origin or ancestry, age, sex, or status as a peace officer or judge"; or (4) it is "unlawful expression."¹⁵²

The Texas law also has certain transparency measures which are not relevant for the purpose of this Note.

145. *Id.* § 501.2041(1)(d) (defining "journalistic enterprise" as "an entity doing business in Florida that: 1. publishes in excess of 100,000 words available online with at least 50,000 paid subscribers or 100,000 monthly active users; 2. publishes 100 hours of audio or video available online with at least 100 million viewers annually; 3. operates a cable channel that provides more than 40 hours of content per week to more than 100,000 cable television subscribers; or 4. operates under a broadcast license issued by the Federal Communications Commission").

146. *Id.* § 501.2041(2)(j).

147. *Id.* § 501.2041(2)(d).

148. *Id.* § 501.2041(4).

149. TEX. CIV. PRAC. & REM. CODE ANN. § 143A.001(1) (West 2021) (defining "censor" as to "block, ban, remove, deplatform, demonetize, de-boost, restrict, deny equal access or visibility to, or otherwise discriminate against expression").

150. *Id.* § 143A.002(a). The last part regarding the user being in Texas appears to have been added to prevent platforms from stopping the provision of their services in Texas.

151. *Id.* § 143A.002(b).

152. *Id.* § 143A.006(a).

2. *Entities Regulated by Anti-Moderation Laws*

The Florida law applies to “social media platforms” which are “information service[s], system[s], internet search engine[s], or access software provider[s]” doing business in Florida which: (1) “provides or enables computer access by multiple users to a computer server, including an internet platform or a social media site”; and (2) have annual gross revenues of over \$100 million or at least 100 million monthly global users.¹⁵³

The Texas law applies to “social media platforms” which are websites or internet applications that are “open to the public, [allow] a user to create an account, and [enable] users to communicate with others . . . [primarily] for posting information, comments, messages or images.”¹⁵⁴ The law exempts from its application internet service providers, emails, as well as online services, applications, and websites which consist primarily of information or content that is not user-generated but is pre-selected by service providers.¹⁵⁵ The Texas law’s effect is also limited to social media platforms that have over fifty million monthly active users in the United States.¹⁵⁶

The definition of social media platforms is extremely broad under both statutes. Under Florida law, *any* website that crosses the revenue or user thresholds will be required to comply with the provisions of the law. This means that among others, the following websites will fall within the ambit of the law: Amazon (an online marketplace),¹⁵⁷ Netflix (an online movie and TV streaming platform),¹⁵⁸ Wikipedia (an online encyclopedia),¹⁵⁹ Pornhub (a pornographic website),¹⁶⁰ Eventbrite (an online ticketing service),¹⁶¹ Bit.ly (a

153. FLA. STAT. § 501.2041(1)(g) (2022).

154. TEX. BUS. & COM. CODE ANN. § 120.001(1) (West 2021).

155. *Id.*

156. *Id.* § 120.002(b).

157. Amazon.com, Inc., Annual Report (Form 10-K), at 65 (Feb. 4, 2022), <https://d18rn0p25nwr6d.cloudfront.net/CIK-0001018724/f965e5c3-fded-45d3-bbdb-f750f156dcc9.pdf> (showing that sales from only the online stores in 2021 was \$222 billion).

158. Netflix, Inc., Annual Report (Form 10-K), at 20 (Jan. 27, 2022), https://s22.q4cdn.com/959853165/files/doc_financials/2021/q4/da27d24b-9358-4b5c-a424-6da061d91836.pdf (showing that revenue in 2021 was \$29 billion).

159. WIKIMEDIA STAT., [https://stats.wikimedia.org/#/en.wikipedia.org/reading/unique-devices/normal|line|1-month|\(access-site\)~mobile-site*desktop-site|monthly](https://stats.wikimedia.org/#/en.wikipedia.org/reading/unique-devices/normal|line|1-month|(access-site)~mobile-site*desktop-site|monthly) (last visited Aug. 7, 2023) (showing that almost 800 million unique devices visited the English Wikipedia site in August 2022).

160. SIMILARWEB, <https://www.similarweb.com/website/pornhub.com/#overview> (last visited Aug. 7, 2023) (showing that Pornhub had 2.5 billion visits in August 2022).

161. EVENTBRITE, Q4 2021 SHAREHOLDER LETTER 15, (Feb. 10, 2022), https://s22.q4cdn.com/238770421/files/doc_financials/2021/q4/Q4-2021-Earnings-Shareholder-Letter-vFINAL.pdf (showing that Eventbrite had net revenue of \$187.1 million in 2021).

URL shortening website),¹⁶² Steam (an online games marketplace),¹⁶³ and Etsy (an online marketplace).¹⁶⁴

It can be argued that when read practically, the Florida law will not affect entities with little to no “user-generated content,” such as Netflix and Steam, as they were not taking content moderation decisions anyway. However, a couple of hypotheticals show how the law will be of concern to such entities as well. For instance, consider a moviemaker whose film is on Netflix decides to run for a political position. As per Florida law, Netflix now may not remove the film from the platform till the time the filmmaker remains a candidate. Further, it may not even be able to downrank the movie as it is not allowed to de-prioritize posts by candidates. Similarly, if a media house produces a movie that is picked up by Netflix, every action by Netflix regarding that movie will be subject to Florida law. Further, all such websites will have to comply with the transparency provisions by having standards for censorship, deplatforming, and shadow banning, even if they do not do any of those things. The websites will also be unable to change their user-facing rules any earlier than once in 30 days. This will require resources, and the legal teams at such entities will have to weigh many of their respective company’s actions against the obligations in the Florida law.

While the Texas law does add more criteria to its definition, it still catches within its ambit a lot of websites which have user-generated content. The following entities (among others) would fall within the ambit of the Texas law: Wikipedia (an online encyclopedia);¹⁶⁵ Shopify (an online service to create

162. *Most Popular Websites Worldwide as of November 2021, By Total Visits*, STATISTA, <https://www.statista.com/statistics/1201880/most-visited-websites-worldwide/> (last visited Aug. 7, 2022) (showing that 2.11 billion users visited Bit.ly in November 2021; assuming 10% of those users were from the United States, that is 211 million users from the United States).

163. *Steam – 2021 Year in Review*, STEAM (Mar. 8, 2022), <https://store.steampowered.com/news/group/4145017/view/3133946090937137590> (stating that Steam had 132 million monthly active players globally).

164. Etsy, Inc., Annual Report (Form 10-K), at 73 (Feb. 24, 2022), <https://d18rn0p25nwr6d.cloudfront.net/CIK-0001370637/619701ee-f7dc-4baa-9463-4374cfcef85e.pdf>. (showing that Etsy had a total revenue of \$2.3 billion in 2021).

165. WIKIMEDIA STAT., [https://stats.wikimedia.org/#/en.wikipedia.org/reading/unique-devices/normal|line|1-month|\(access-site\)~mobile-site*desktop-site|monthly](https://stats.wikimedia.org/#/en.wikipedia.org/reading/unique-devices/normal|line|1-month|(access-site)~mobile-site*desktop-site|monthly) (last visited Aug. 7, 2023) (showing that almost 800 million unique devices visited the English Wikipedia site in August 2022; assuming 10% of that is United States users, then that is 80 million monthly active users from the United States).

shopping platform);¹⁶⁶ Pornhub (a pornographic website);¹⁶⁷ Pinterest (an image-sharing website);¹⁶⁸ and Indeed (an online jobs board).¹⁶⁹ Similar to the discussion under the Florida law, a few hypotheticals will show how the Texas law will require entities to devote resources to ensure compliance with the law.

Consider the most ridiculous example first. The law will require Pornhub to not discriminate between users based on their viewpoint. Since many pornographic movies have some semblance of a story which could be depicting a viewpoint, a user could ask Pornhub to rank their content on the first page, otherwise Pornhub might be discriminating against them on the basis of viewpoint. Further, Wikipedia will be required to produce a transparency report on how it moderated content. Given that it is a non-profit largely run by volunteers, it would be extremely difficult for it to muster the resources for a team which can compile all the content moderation decisions taken on the website and present them to the Texas government in the specified manner.

The definitions and the hypotheticals clearly show that even if they are read practically, the laws cover within their ambit a large number of entities which have not been part of the discussion around these laws.

166. *Shopify Usage Statistics*, BUILTWITH, <https://trends.builtwith.com/shop/Shopify> (last visited Aug. 7, 2023) (showing that 2.5 million websites in the United States use Shopify; assuming those websites have an average of 20 monthly active users, then that is 50 million monthly active users from the United States).

167. *Porn Sites Collect More User Data Than Netflix or Hulu. This Is What They Do With It*, QUARTZ, <https://qz.com/1407235/porn-sites-collect-more-user-data-than-netflix-or-hulu-this-is-what-they-do-with-it> (last visited Aug. 18, 2023) (stating that Pornhub has over 100 million daily visits, which works out to 3 billion monthly visits). The United States is by far the country with the highest daily traffic to Pornhub. *2021 Year in Review*, PORNHUB: INSIGHTS (Dec. 14, 2021), <https://www.pornhub.com/insights/yir-2021#Countries-by-Traffic>.

168. *Social Media Usage in the United States*, STATISTA, <https://www.statista.com/study/40227/social-social-media-usage-in-the-united-states-statista-dossier/> (last visited Aug. 7, 2023) (showing that Pinterest had 98.77 million users in 2021; assuming that about half of them are active monthly users, Pinterest is likely to cross the 50 million monthly active users threshold).

169. *Worldwide Visits to Indeed.com From November 2022 to April 2023*, STATISTA, <https://www.statista.com/statistics/1259806/number-of-unique-visitors-to-indeed/> (last visited Aug. 7, 2023) (showing that Indeed had over 650 million unique global users in May 2022; even if 10% of those visitors are from the United States, that will cross the fifty million monthly active users threshold).

F. ANTI-MODERATION LAWS FAIL INTERMEDIATE SCRUTINY

To survive intermediate scrutiny, the Anti-Moderation Laws must serve “a significant/substantial/important governmental interest and [be] reasonably well tailored to that purpose (i.e., not unreasonably overbroad).”¹⁷⁰

The governmental interest at play here is to regulate the modern public square to ensure that democracy and the electoral process are preserved. This is a significant/substantial/important governmental interest in light of *Citizens United*.¹⁷¹

While judging the laws against the second part of intermediate scrutiny (i.e., that they are reasonably well tailored to the governmental interest) there are two aspects to consider: (1) whether the actions regulated by the laws are reasonably well tailored to the governmental interest; and (2) whether the broad sweep of entities covered by the laws still leave the laws reasonably well tailored to the governmental interest.

On the first aspect, the laws provide specific checks and obligations on social media platforms which will reduce the discretion they have in moderating content. Since the aim of the laws is to open up the modern public square and ensure that conversation flows freely with little intervention by private actors, the laws might further that aim. While it is certainly debatable what the practical effects of the law will be,¹⁷² for the purposes of an analysis that is taking place before they go into effect, the provisions of the laws do appear to further free discussion on social media platforms.

However, the laws fail the second part of the test. As discussed above in Section III.E.2, the definitions of social media platforms are extremely broad and will implicate a large number of entities that have nothing to do with the modern public square and political discussions.

While it can be argued that the biggest social media platforms such as Facebook, X, and YouTube are the modern public square, given the immense impact they have on politics and democracy, the same cannot be said for the other websites that will come under the sweep of the Anti-Moderation Laws. The Anti-Moderation Laws cover within their ambit websites that are in no way linked to “political communication.”¹⁷³ These websites do not “establish common goals in pursuit of our common good” or “promote the general

170. Bhagwat, *supra* note 16, at 801.

171. *Citizens United v. Fed. Election Comm’n*, 558 U.S. 310, 440 (2010).

172. *See* Warzel, *supra* note 10 (discussing hypotheticals which show that the law might be unworkable); Keller, *supra* note 10 (discussing how the platforms can try to comply with the Texas law); Masnick, *supra* note 10 (giving examples of how the laws will lead to tremendous amounts of wasteful litigation).

173. Fuchs, *supra* note 86, at 13.

welfare” to weave together a “democratic political culture” as required in a public square.¹⁷⁴ Netflix, Pornhub, Eventbrite, and Steam might be useful services in their own right, but they do not contribute to the public discourse in a manner such that governments need to regulate them through Anti-Moderation Laws.

Regulating Netflix and Pornhub would not further even the expansive governmental interests that were considered by the Eleventh¹⁷⁵ and Fifth¹⁷⁶ Circuits. When measured against the actual government interest (regulation of the modern public square), the aims and effects of these laws are in no way linked, and these laws cannot be said to be reasonably well tailored to the governmental respect.

In *Packingham*, Justice Alito used the fact that the statute barred access to many websites unrelated to the state’s purpose to deduce that the impugned statute had a broad reach and did not appreciably advance the state’s goal.¹⁷⁷ Similarly, the implication under Anti-Moderation Laws of multiple websites which have little relation to the states’ purposes is that the statute is overbroad. Since the Anti-Moderation Laws are not reasonably well-tailored to the governmental interest, they fail intermediate scrutiny and are therefore unconstitutional.

G. ANTI-MODERATION LAWS FAIL STRICT SCRUTINY

Given the tiered system of analysis, a law that fails intermediate scrutiny will inevitably fail strict scrutiny.¹⁷⁸ This is true in the present case as well.

Under the strict scrutiny test, the restriction implicated must further a compelling government interest and be narrowly tailored to achieve that interest.¹⁷⁹ Here, the governmental interest is a compelling one as discussed in Section III.D. The second part of the test (the laws being narrowly tailored to achieve the compelling government interest) can again be split into two: (1)

174. Casey, *supra* note 85.

175. The Eleventh Circuit considered ‘counteracting unfair private censorship that privileges some viewpoints over others on social-media platforms and promoting the widespread dissemination of information from a multiplicity of sources’ as a potential governmental interest. *NetChoice, L.L.C. v. Att’y Gen., Fla.*, 34 F.4th 1228 (11th Cir. 2022).

176. The Fifth Circuit considered ‘protecting the free exchange of ideas and information in Texas’ as a governmental interest. *NetChoice, L.L.C. v. Paxton*, 49 F.4th 439, 482 (5th Cir. 2022).

177. *Packingham v. North Carolina*, 137 S. Ct. 1730, 1743 (2017)

178. See Dan V. Kozlowski & Derigan Silver, *Measuring Reed’s Reach: Content Discrimination in the U.S. Circuit Courts of Appeals After Reed v. Town of Gilbert*, 24 COMM. L. & POLY 191, 196 (2019) (noting that “it is still the case that it is much easier for a law to pass intermediate scrutiny than strict scrutiny”).

179. Bhagwat, *supra* note 16, at 171.

the actions regulated by the laws; and (2) the entities regulated by the laws. While the actions may be argued to be narrowly tailored to ensure free discussion on the platforms, the broad universe of the entities regulated prevents the laws from being narrowly tailored. The laws regulate entities which have nothing to do with the governmental interest at play. Therefore, the laws fail strict scrutiny and are therefore unconstitutional.

IV. ANTI-MODERATION LAWS CANNOT BE REDRAFTED IN A CONSTITUTIONAL MANNER

A. CHANGING THE DEFINITION OF SOCIAL MEDIA PLATFORMS TO INCORPORATE THE MODERN PUBLIC SQUARE WILL NOT SOLVE THE CONSTITUTIONALITY ISSUE

Given the thesis of this paper that the Anti-Moderation Laws are unconstitutional simply because of bad definitions of social media platforms, one's first instinct might be to simply redraft the definitions themselves. However, that is far easier said than done.

If the definition is simply amended to include websites and apps which comprise the modern public square, the definition would be far too broad and be suspect to a vagueness challenge. While all laws have some vagueness, the Supreme Court has clarified that greater precision is required when laws regulate speech.¹⁸⁰ Laws can be challenged as being facially unconstitutional for being unduly vague, and a successful facial challenge will result in the law being entirely invalidated.¹⁸¹ In *Baggett v. Bullitt*, the Supreme Court held that a state law requiring state employees to swear that they were not a “subversive person” was invalid as its language was unduly vague, uncertain, and broad.¹⁸² The reasoning for the court's decision was that the ambiguities inherent in the term “subversive” gave people little guidance as to what exactly was proscribed.¹⁸³

Practically, entities would be confused as to whether they fall within the modern public square or not. They would thus not know whether to comply with the laws. Therefore, changing the definition to incorporate the modern public square would not solve the constitutionality issue of the laws.

180. ERWIN CHEMERINSKY, *THE FIRST AMENDMENT* 59 (2021).

181. *Id.* at 57.

182. 377 U.S. 360, 366 (1964).

183. *Id.* at 371.

B. CHANGING THE DEFINITION OF SOCIAL MEDIA PLATFORMS TO TARGET SPECIFIC PLATFORMS WILL VIOLATE THE FIRST AMENDMENT

Another option could be to change the definition of social media platforms to name specific platforms which comprise the modern public square according to the state. As discussed above in Part II.B.3, the modern public square comprises Facebook, X, and YouTube for now. However, this would be akin to targeting specified speakers for their speech, which would be unconstitutional. The Supreme Court has held that in the context of political speech, the government may not “impose restrictions on certain disfavored speakers.”¹⁸⁴ The court there observed that restrictions “based on the identity of the speaker are all too often simply a means to control content.”¹⁸⁵

Speaker-based discrimination infringes the First Amendment because by regulating those who may speak, the government can control the content of what is said because personal identity usually correlates with political opinions.¹⁸⁶ Further, the speaker’s identity shapes how the content is received and interpreted.¹⁸⁷ Therefore, in light of *Citizens United*, a law that imposes restrictions on certain speakers in the political context will not stand.¹⁸⁸ The Anti-Moderation Laws, even if amended to enumerate certain platforms in the definition of social media platforms, would be unconstitutional.

V. CONCLUSION

The Anti-Moderation Laws are part of a backlash against tech companies. Lawmakers have used various areas of laws to regulate tech platforms, including privacy, antitrust, and free speech. Given the harms these companies have (intentionally or unintentionally) brought into the world, the clamor for regulating them has steadily increased. However, it is important to balance the need for free expression and the right to access information with the need to protect individuals and society from harmful or malicious content. Finding the right balance is rarely easy, and there are always competing interests at play.

Overall, these types of laws may be seen as problematic because they could potentially interfere with the ability of social media companies to enforce their own terms of service and moderate content on their platforms in a way that

184. *Citizens United v. Fed. Election Comm’n*, 558 U.S. 310, 341 (2010).

185. *Id.* at 340.

186. Michael Kagan, *Speaker Discrimination: The Next Frontier of Free Speech*, 42 FLA. ST. U. L. REV. 765, 816 (2015).

187. *Id.*

188. *Citizens United*, 558 U.S. at 341.

they see fit. Ironically—given the laws’ aim to free up online speech—these laws could lead to less online speech and a less diverse and open online environment as bots and extremists take over the online space while platforms are helpless to control them. Social media companies will be more hesitant to moderate content out of fear of facing legal consequences, which could result in more harmful or malicious content remaining on the platform. It is worth noting that the platforms may still be able to take action to remove certain types of content that are not protected by the Anti-Moderation Laws. However, the specific provisions of the laws may limit the discretion of social media companies to make these types of decisions.

Ultimately, the question of whether the Texas and Florida social media laws are bad will depend on one’s perspective and values. Some may view these laws as necessary protections for free speech, while others may see them as harmful interference with the ability of social media companies to regulate content on their platforms.

However, none of the above discussions change the fact that the Anti-Moderation Laws are extremely broad statutes which cover within their scope websites which do not contribute to the states’ goal of having an unfettered modern public square. The origin of these laws is clearly in the political arena and the lawmakers’ motivations are to largely protect political speech. However, the drafting of these laws has led to the scenario where many entities who have no effect on political speech are implicated under these laws. Given their overbroad nature, these laws should be struck down as unconstitutional. There is also no way of redrafting the laws in a way that allows them to apply to social media platforms that are actually the modern public square that the laws aim to regulate.

These laws are likely just the beginning of state action against platforms given the deadlock at the federal level regarding tech legislation. If the Elon Musk-X saga has shown us anything, it is that there is a need to regulate the power that has landed in the hands of a few technocrats. However, given the exponential effects of such laws on the entire internet ecosystem, the drafting of such laws with anticipation of the future effects of these laws becomes extremely important. Laws moderating platforms can very easily stifle innovation because the largest players are the most well-placed to comply with the onerous obligations that such laws bring along. Further, even the largest platforms need protection from the partisan actions of lawmakers in fiercely red or blue states.

Tech regulation is hard work, and the Anti-Moderation Laws show how difficult it is to draft laws which only have the intended effect and no more.

However, that is no excuse for putting overbroad laws in the books which harm free speech far more than they promote it.

THE FORGOTTEN PUBLIC INTEREST STANDARD

Bogdan Belei[†]

TABLE OF CONTENTS

I.	INTRODUCTION	1469
II.	THE ORIGINS OF THE PUBLIC INTEREST STANDARD.....	1472
	A. THE HISTORY OF MEDIA REGULATION IN THE UNITED STATES..	1472
	B. THE GREAT COMPROMISE: COMMERCIAL BROADCASTERS AND PUBLIC INTEREST GROUPS	1474
	C. THE PUBLIC INTEREST FACTORS AND CROSS-OWNERSHIP	1476
	1. <i>Diversity</i>	1478
	2. <i>Localism</i>	1480
	3. <i>Competition</i>	1483
III.	THE FIGHT OVER AMERICA’S PUBLIC INTEREST	1486
	A. REGULATORY PURPOSE: DEMOCRACY VS. EFFICIENCY	1486
	1. <i>The Democracy Model</i>	1486
	2. <i>The Efficiency Model</i>	1490
	B. ERA OF DEREGULATION	1491
IV.	REVIEWING <i>FCC V. PROMETHEUS RADIO PROJECT</i>.....	1494
V.	REVIVING THE PUBLIC INTEREST STANDARD.....	1499
VI.	CONCLUSION.....	1503

I. INTRODUCTION

In 2017, Federal Communications Commission (FCC) Chairman Ajit Pai issued an order to revoke the Commission’s long-standing rules against media cross-ownership. The move allowed broadcasters to increase the number of television and radio stations they could own. Less than a month later, Sinclair Broadcast Group—the second-largest television station broadcaster in the

DOI: <https://doi.org/10.15779/Z38M32NB6N>

© 2023 Bogdan Belei.

† Associate, Skadden, Arps, Slate, Meagher & Flom LLP and J.D. 2023, University of California, Berkeley, School of Law. Sincere thanks to Professor Talha Syed, as well as Will Kasper, Yuhan Wu, Shih-wei Chao, and the *Berkeley Technology Law Journal* Editors. All views expressed herein, and all errors, are my own.

United States¹—took advantage of the FCC’s newfound leniency. In a landmark \$3.9 billion deal, Sinclair proposed to buy a rival competitor, Tribune Media.² Sinclair sought to own or control stations televising to more than 73% of all households with a television set in the United States through the merger.³ Less than a month after the deal was proposed, Sinclair was accused of forcing dozens of its local news anchors to recite an identical script in newsrooms all across America.⁴ The company was not only gaining corporate control of a supermajority of America’s television stations, but it was ensuring that American viewers were hearing a uniform message from a singular source. While the merger ultimately failed to materialize due to competition concerns, the potential ramifications would have affected even more foundational aspects of America’s democracy.⁵ The 2017 rule change and Sinclair’s attempt to consolidate the industry was only the latest struggle over the future of media regulation.

Media broadcasting has been governed by the public interest standard for nearly one hundred years. First introduced in the Radio Act of 1927, the public interest standard requires broadcast licensees to operate in the “public interest, convenience and necessity.”⁶ The policy emerged from a compromise between commercial broadcasters and public interest groups.⁷ The federal government established a licensing regime for broadcasters but required them to uphold the public interest. The term was never statutorily defined but it adopted long-held principles reflective of independent media and the freedom of press—namely, diversity, localism, and competition.⁸

1. Neil Macker, *New Coverage of TV Station Owners*, MORNINGSTAR (Jan. 1, 2020), <https://www.morningstar.com/articles/961093/new-coverage-of-tv-station-owners>.

2. Sydney Ember & Michael J. de la Merced, *Sinclair Unveils Tribune Deal, Raising Worries It Will Be Too Powerful*, N.Y. TIMES (May 8, 2017), <https://www.nytimes.com/2017/05/08/business/media/sinclair-tribune-media-sale.html>.

3. Klint Finley, *FCC Wants to Ease Rules to Benefit Broadcast Giant Sinclair*, WIRED (Oct. 27, 2017), <https://www.wired.com/story/fcc-wants-to-ease-rules-to-benefit-broadcast-giant-sinclair/>.

4. Jacey Fortin & Jonah Engel Bromwich, *Sinclair Made Dozens of Local News Anchors Recite the Same Script*, N.Y. TIMES (Apr. 2, 2018), <https://www.nytimes.com/2018/04/02/business/media/sinclair-news-anchors-script.html>.

5. Reuters, *Tribune Media Sues Sinclair for \$1 Billion in Damages After Terminating \$3.9 Billion Acquisition Deal*, CNBC (Aug. 9, 2018), <https://www.cnbc.com/2018/08/09/tribune-media-terminates-deal-to-be-bought-by-sinclair.html>.

6. Radio Act of 1912, ch. 287, § 1, 37 Stat. 302.

7. See *The Public Interest Standard in Television Broadcasting*, BENTON INST. FOR BROADBAND & SOC., https://www.benton.org/initiatives/obligations/charting_the_digital_broadcasting_future/sec2 (last visited Nov. 10, 2023) [hereinafter BENTON INST.].

8. See *infra* Section II.C.

The purpose of this Note is to remember the forgotten public interest standard and reverse course on the last thirty years of harmful deregulation in the broadcasting industry. In Part II, this Note traces the origins of media regulation in the United States and how the public interest standard emerged as an important mechanism for democratic governance. Born out of the fear of oligopolies in media ownership, the public interest standard was designed to protect against a concentrated media environment. After its founding, it was enforced to this end for the next fifty years. The last thirty years have been a departure from the original purpose of the law. In Part III, this Note traces how the public interest standard has been interpreted and enforced by two separate political camps: proponents of the democracy model and proponents of the efficiency model. This Part aligns the purposes of the public interest standard with the democracy model, while describing the efficiency model as an aberration promoted by corporate interests at the expense of a vibrant, diverse, and representative democracy.

The following Parts focus on recent developments and the future of the public interest standard. In Part IV, the article analyzes *FCC v. Prometheus Radio Project*—the most recent Supreme Court case that reviewed the FCC’s administrative authority and allowed the Commission to revoke media cross-ownership rules. The Court ignored the normative issues concerning the public interest standard. However, Justice Clarence Thomas wrote a concurring opinion where he objected to the Third Circuit imposing a procedural requirement for the FCC to consider minority and female ownership during their rule review process. Justice Thomas described diversity ownership merely as a proxy for viewpoint diversity, and thus unwarranted. By setting this distinction, Thomas attempted to define the FCC’s regulatory target as consumers, rather than producers. However, this distinction is irrelevant. First, the FCC has continuously pursued diversity ownership through rules and regulations over the course of decades. Second, it is unlikely that the FCC could ever achieve viewpoint diversity with respect to minorities and women without promoting diversity ownership.

Finally, Part V charts a path for reversing the current trajectory of media deregulation. The FCC must revitalize enforcement of the public interest standard and interpret it as designed—by prioritizing democratic safeguards ahead of efficiency and economic competition. In practice, this means that the FCC should reinvoke ownership rules to prevent market concentration and only relax them in small- to mid-sized markets where there is substantial evidence of market failure. If a local market cannot sustain competition among multiple broadcasters, then the FCC should allow mergers that will ensure that consumers are receiving quality information. To avoid cyclical rulemaking,

Congress should pass a revised Communications Act that provides greater protections to the public interest and takes account of technological changes since 1996.

II. THE ORIGINS OF THE PUBLIC INTEREST STANDARD

A. THE HISTORY OF MEDIA REGULATION IN THE UNITED STATES

The media and its influence on the public have always been vital to American democracy. While the Constitution was written in secrecy, it was reprinted by almost all newspapers and vigorously debated.⁹ In 1804, Thomas Jefferson wrote, “Our first object should therefore be, to leave open to him all the avenues to truth. The most effectual hitherto found, is the freedom of the press.”¹⁰ Likewise, James Madison opined that, “A popular government, without popular information, or the means of acquiring it, is but a Prologue to a Farce or a tragedy; or perhaps both.”¹¹ The freedom of the press and access to independent sources were at the root of the Founders’ concerns.¹² These principles have driven the purpose of media regulation ever since.

Like the early United States, most democracies viewed concentrated media ownership as a threat to press freedom and democracy.¹³ As a result, media diversity became a guiding principle for regulators. At the federal level, since its founding in the 1700s, the Postal Service heavily subsidized postage rates to support a growing newspaper industry.¹⁴ Likewise, state and local governments took legislative action to ensure that their communities did not fall victim to market capture and were serviced by varied interests. In 1821, the New York State constitution required that “every citizen may freely speak,

9. See ANTHONY FELLOW, *AMERICAN MEDIA HISTORY* 12 (2012).

10. Letter from Thomas Jefferson to John Tyler (June 28, 1804), *in* 11 *THE WRITINGS OF THOMAS JEFFERSON*, at 33 (Albert Ellery Bergh ed., 1907).

11. Letter from James Madison to W. T. Barry (Aug. 4, 1822), *in* 9 *THE WRITINGS OF JAMES MADISON*, at 103 (Gaillard Hunt ed., 1910).

12. Eugene Volokh, *Freedom for the Press as an Industry, or for the Press as a Technology? From the Framing to Today*, 160 U. PA. L. REV. 459, 469–70 (2012).

13. See C. EDWIN BAKER, *MEDIA CONCENTRATION AND DEMOCRACY: WHY OWNERSHIP MATTERS* 2 (2007).

14. See RICHARD R. JOHN, *SPREADING THE NEWS: THE AMERICAN POSTAL SYSTEM FROM FRANKLIN TO MORSE* (1995); RICHARD B. KIELBOWICZ, *NEWS IN THE MAIL: THE PRESS, POST OFFICE, AND PUBLIC INFORMATION, 1700–1860s* (1989); RICHARD D. BROWN, *THE STRENGTH OF A PEOPLE: THE IDEA OF AN INFORMED CITIZENRY IN AMERICA, 1650–1870* (1996); PAUL H. STARR, *THE CREATION OF THE MEDIA: POLITICAL ORIGINS OF MODERN COMMUNICATIONS* (2004).

write and publish his sentiments on all subjects.”¹⁵ Toward the end of the century, New York explicitly sought to promote competition and diversity among the newspaper industry by requiring local governments to advertise in at least two local papers of different parties.¹⁶ Legislators became even more concerned with concentrated ownership as industrialization consolidated the national economy.

Beginning with the American Industrial Revolution, the growth of the media industry rapidly expanded beyond local operations managed under local ownership. The march westward to the Pacific was matched by a rapid modernization in technology and a natural lean toward growth-oriented businesses and economies of scale. New technology—such as the steam-powered “double-press”—had a profound impact on the industry’s capabilities, allowing newspapers to increase production tenfold overnight.¹⁷ Later, the introduction of the telegraph and radio outgrew the local business models of newspapers and expanded their reach and content to suit more regional and national audiences. As the communications industry evolved, industry founders adopted the idea of enlightened monopolies characterized by concentrated ownership.¹⁸

For the first time in history, mere individuals had control over an instantaneous and massive information industry. In 1926, Texas Democrat Representative Luther Alexander Johnson warned that “American thought and American politics will be largely at the mercy of those who operate [broadcast] stations.”¹⁹ This sentiment was not only pervasive among political observers worried about democratic decline, but also among cultural critics which recognized the power of media in shaping social patterns. In an essay titled “The Outlook for American Culture,” writer Aldous Huxley criticized the media’s newfound efficiency: “Mass production is an admirable thing when applied to material objects; but when applied to things of the spirit it is not so

15. Heming Nelson, *A History of Newspaper: Gutenberg’s Press Started a Revolution*, WASH. POST (Feb. 11, 1998), <https://www.washingtonpost.com/archive/1998/02/11/a-history-of-newspaper-gutenbergs-press-started-a-revolution/2e95875c-313e-4b5c-9807-8bcb031257ad>.

16. See BAKER, *supra* note 13, at 2.

17. Nelson, *supra* note 15.

18. See TIM WU, *THE MASTER SWITCH: THE RISE AND FALL OF INFORMATION EMPIRES* 7–8 (2010).

19. Steve Rendall, *The Fairness Doctrine: How We Lost it and Why We Need It Back*, SISYPHUS (July 2018), <https://sisyphuslitmag.org/2018/07/the-fairness-doctrine-how-we-lost-it-and-why-we-need-it-back/>.

good.”²⁰ In government and in social circles, the independence and diversity of media was widely considered sanctimonious.

B. THE GREAT COMPROMISE: COMMERCIAL BROADCASTERS AND PUBLIC INTEREST GROUPS

As the national communications industry grew larger and broadcasting technology became sufficiently pervasive, there was a pressing need for federal government oversight. Initially, Congress passed the Radio Act of 1912 and authorized the Department of Commerce to regulate the distribution of radio licenses.²¹ It was illegal to transmit on radio without a license;²² however, due to the broad availability of spectrum frequency, the Commerce Secretary had no authority to deny licenses.²³ By the mid-1920s, this decentralized approach ran into interference issues as there was no mechanism to coordinate frequencies and power levels.²⁴ Congress sought to prevent market failure and protect the value of wireless services by establishing a system of regulatory control. However, in doing so, Congress needed to balance two separate goals: fostering commercial development of the industry and ensuring that broadcasting served the informational needs of American citizens.²⁵

Commercial broadcasters and public interest groups needed to reach a compromise.²⁶ The commercial broadcasters, represented by the National Association of Broadcasters (NAB), worried that signal interference thwarted the development of broadcasting and preferred a certain level of administrative coordination. At the same time, the industry was adamant about retaining editorial control over programming and the ability to organize individual

20. Aldous Huxley, *The Outlook for American Culture: Some Reflections in a Machine Age*, HARPER'S MAG. (Aug. 1927), <https://harpers.org/archive/1927/08/the-outlook-for-american-culture/>.

21. See Radio Act of 1912, ch. 287, 37 Stat. 302.

22. See *id.*

23. See BENTON INST., *supra* note 7.

24. By 1916, there were approximately 500 radio stations operating in the United States with only 89 available wave-length channels. There were approximately 400 stations applying for broadcasting licenses, yet no more than 331 stations could operate on the spectrum without significant interference. See James Patrick Taugher, *The Law of Radio Communication with Particular Reference to a Property Right in a Radio Wave Length*, 12 MARQ. L. REV. 179, 181 (1928); see also Jennifer Davis, *Anniversary of the Radio Act of 1927, The Beginning of Broadcast Regulation*, LIBR. CONGRESS BLOG (Feb. 23, 2016), <https://blogs.loc.gov/law/2016/02/anniversary-of-the-radio-act-of-1927-the-beginning-of-broadcast-regulation/>.

25. See BENTON INST., *supra* note 7.

26. See *id.*

broadcasting stations into national networks.²⁷ Meanwhile, public interest groups feared that a national licensing system would give preference to commercial interests and suppress free speech interests.²⁸ A number of free speech advocates—including politicians, educators, labor activists, and religious groups—argued for a common carriage regime that would prohibit broadcasters from denying public interest groups access to their channels and allow anyone to buy airtime.²⁹ By resolving these competing interests, the federal government could encourage innovation in the broadcasting industry while retaining the public benefits of these technologies.

With the passage of the Radio Act of 1927, and later the Communications Act of 1934, Congress resolved the broadcasting dispute. First, Congress banned common carrier regulation and mandated a government-sanctioned licensing regime.³⁰ The FRC, and later the FCC, was authorized to assign licensees designated channels in the electromagnetic spectrum. Without common carriage, Congress limited free speech rights to broadcasters with a valid license. However, this exclusionary licensing regime was justified when Congress simultaneously introduced a requirement that broadcast licensees must operate in the “public interest, convenience and necessity.”³¹ Broadcasters were entrusted with spectrum allocation in return for guarantees that they would serve the public interest by adhering to certain factors. The Supreme Court has referred to broadcasters’ role as public “fiduciaries” under this arrangement,³² and the FCC has stated that a “station itself must be operated as if owned by the public . . . as if people of a community should own a station and turn it over to the best man in sight with this injunction: ‘Manage this station in our interest’”³³ The purpose of the public interest was generally resolved, however the standard itself remained relatively vague.

Despite its deep reverence for the media as a democratic governing institution, the FCC never defined the “public interest” after its inception in

27. Stuart N. Brotman, *Revisiting the Broadcast Public Interest Standard in Communications Law and Regulation*, BROOKINGS INST. (Mar. 23, 2017), <https://www.brookings.edu/articles/revisiting-the-broadcast-public-interest-standard-in-communications-law-and-regulation/>.

28. See BENTON INST., *supra* note 7.

29. See *id.*

30. See *id.*

31. See *id.*

32. *Red Lion Broad. Co. v. FCC*, 395 U.S. 367, 389 (1969).

33. John W. Willis, *The Federal Radio Commission and the Public Service Responsibility of Broadcast Licensees*, 11 FED. COMM. B. J. 5, 14 (1950) (citing to a 1930 Federal Radio Commission decision).

the Radio Act of 1927.³⁴ Both the Radio Act and the Communications Act of 1934 refer to the “public interest” in various forms without providing an explicit statutory definition.³⁵ As such, it has been difficult to institutionalize the public interest standard; its interpretation and enforcement has changed over time to reflect the contemporary doctrinal mainstream or the political leanings of the revolving Executive Branch.³⁶ Former FCC Commissioner Ervin Duggan once opined that “[s]uccessive regimes at the FCC have oscillated wildly between enthusiasm for the public interest standard and distaste for it.”³⁷ Despite certain administrations showing distaste, both sides have invoked their interpretation of “public interest.” The FCC has never done away with the public interest standard—instead, courts and the Commission’s leadership have shaped policy through administrative orders and precedent. While the standard applies to all FCC rulemaking, it has been ardently disputed in the context of media ownership.

C. THE PUBLIC INTEREST FACTORS AND CROSS-OWNERSHIP

Since its founding, the FCC has been concerned with ownership concentration and its influence on viewpoint diversity.³⁸ In 1938, the FCC adopted a presumption against granting radio licenses that would create duopolies—common ownership or control of stations with overlapping signal contours—specifically to uphold the “diversification of service.”³⁹ A few years later, the Commission instated a television duopoly rule which barred a single entity from owning two or more broadcast television stations that “would substantially serve the same area.”⁴⁰ Both in its approach to radio and

34. Becky Chao, *The Value of the FCC’s Public Interest Mandate in Empowering Community Voices*, NEW AM. (Dec. 14, 2017), <https://www.newamerica.org/millennials/dm/value-fccs-public-interest-mandate-empowering-community-voices/>.

35. See, e.g., 47 U.S.C. §§ 201(b), 215(a), 319(c), 315(a) (“public interest”); §§ 214(a), 214(c) (“public convenience and necessity”); § 214(d) (“interest of public convenience and necessity”); §§ 307(a), 309(a), 319(d) (“public interest, convenience and necessity”); § 307(a) (“public convenience, interest or necessity”); §§ 311(b), 311(c)(3) (“public interest, convenience or necessity”).

36. See J. Roger Wollenberg, *The FCC as Arbiter of “The Public Interest, Convenience, and Necessity,”* in A LEGISLATIVE HISTORY OF THE COMMUNICATIONS ACT OF 1934, at 61, 77–78 (Max Paglin ed., 1989).

37. *Public Interest and Localism: Hearing Before the Comm. on Commerce, Sci., & Transp.*, 108th Cong. 18 (2003) (prepared statement of Robert Corn-Revere, Partner, Davis Wright Tremaine LLP).

38. See Christa Corrine McLintock, *The Destruction of Media Diversity, or: How the FCC Learned to Stop Regulating and Love Corporate Dominated Media*, 22 J. MARSHALL J. COMPUTER & INFO. L. 569, 585 (2004).

39. Genesee Radio Corp., 5 F.C.C. 183 (1938).

40. Part 4—Broadcast Services Other Than Standard Broadcast, 6 Fed. Reg. 2282, 2284–85 (May 6, 1941).

television ownership, the FCC favored policies that promoted diverse ownership and preserved viewpoint diversity across media markets.

Beginning in the 1960s, the FCC adopted three ownership rules concerning newspaper, broadcast, radio, and television.⁴¹ In 1964, the agency adopted the Local Television Ownership Rule that restricts the number of local television stations that an entity may own in a single market. The rationale behind the FCC's decision was to have the rule "act indirectly to curb regional concentrations of ownership as well as overlap itself."⁴² The Radio/Television Cross-Ownership Rule was implemented in 1970 and limited the number of combined radio stations and television stations that an entity may own in a single market. And finally, in 1975, the FCC adopted the Newspaper/Broadcast Cross-Ownership Rule that prohibits a single entity from owning a radio or television broadcast station and a daily print newspaper in the same media market. At the time, the agency implemented these rules to protect against media concentration.⁴³

Under the Communications Act, each ownership rule needed to be justified in serving the public interest. The FCC sought to meet this standard by addressing three public interest factors: diversity, localism, and competition. First, in pursuit of diversity, the FCC targeted a variety of goals including a diversity of viewpoints, programming, and outlets, as well as increased diversity in ownership.⁴⁴ Critics have disputed which 'type' of diversity is most impactful to achieve the public interest and which type the FCC is required to consider when rulemaking.⁴⁵ Second, by restricting the quantity of media outlets that a company could own or control within a geographic market, the new rules allowed the agency to promote localism.⁴⁶ Healthy measures around competition were expected to stimulate localism as broadcasters compete for local viewers. However, critics have pointed to localism as an ill-defined and unjustified principle that limits political debate.⁴⁷ Finally, the new ownership

41. See *FCC v. Prometheus Radio Project*, 141 S. Ct. 1150, 1155 (2021).

42. Part 73—Radio Broadcast Services, 29 Fed. Reg. 7535, 7537 (June 12, 1964).

43. DOUGLAS GOMERY, *THE FCC'S NEWSPAPER-BROADCAST CROSS-OWNERSHIP RULE: AN ANALYSIS 1* (2002).

44. See DANA A. SCHERER, CONG. RSCH. SERV., R45338, *FEDERAL COMMUNICATIONS COMMISSION (FCC) MEDIA OWNERSHIP RULES 1* (2021), <https://crsreports.congress.gov/product/pdf/R/R45338/3>.

45. See *Prometheus*, 141 S. Ct. at 1161–62 (Thomas, J., concurring).

46. See SCHERER, *supra* note 44, at 26.

47. John Samples, *Broadcast Localism and the Lessons of the Fairness Doctrine*, CATO INST. (May 27, 2009), <https://www.cato.org/sites/cato.org/files/pubs/pdf/pa639.pdf>.

rules promoted fair competition and aimed to protect against abusive exercises of market power.⁴⁸

However, the relationship between fairness and competition is dynamic and complicated—it involves both normative and procedural challenges for the future of competition law. The following Sections briefly characterize the three public interest factors and explain how the Court and the FCC has interpreted them throughout the last century.

1. *Diversity*

The benefit of diversity to the public interest stems from its benefit to democracy. In a 1919 dissent, Justice Oliver Wendell Holmes wrote that “the ultimate good desired is best reached by free trade in ideas.”⁴⁹ The free trade of ideas promises unimpeded exchange of information, dissent, accountability, and freedom of expression. These benefits recede when a dearth of diverse voices, sources, or content leads to a limited range of ideas. The FCC pursued this theory by passing the Financial Interest and Syndication (“FinSyn”) Rules in 1970.⁵⁰ The FinSyn rules intended to “limit network control over television programming and thereby encourage the development of a diversity of programs through diverse and antagonist sources of program services.”⁵¹ By the early 1990s, the FinSyn rules were repealed as critics argued that they “undermined the role of independent producers rather than enhanced them” due to the financial barriers of entering and financing national broadcasting networks.⁵² Nevertheless, their passage and surrounding debate evidences how diversity has always been a staple value of media regulation and consumption.

However, diversity has been seldom defined for the public interest.⁵³ In 1999, Duke Professor Phillip Michael Napoli produced a typology including the varieties of diversity.⁵⁴ Among the three main groups, Napoli included: source diversity, content diversity, and exposure diversity.⁵⁵ Source diversity is intended to produce a diversity of content in theory and provide viewers with

48. *See id.* at 1.

49. *Abrams v. United States*, 250 U.S. 616, 630 (1919).

50. Jennifer Gonzalez, *Syndication Regulation and TV’s Big Three: Broadcasting Regulations and 1970s Television*, LIBR. CONGRESS BLOG (Jan. 31, 2023), <https://blogs.loc.gov/law/2023/01/syndication-regulation-and-tvs-big-three-broadcasting-regulations-and-1970s-television/>.

51. *See* Phillip Napoli, *Deconstructing the Diversity Principle*, 49 J. COMM. 7, 10 (1999).

52. Matthew P. McAllister, *Financial Interest and Syndication Rules*, in *ENCYCLOPEDIA OF TELEVISION* 875, 875 (Horace Newcomb ed., 2d ed. 2004).

53. *See supra* Section II.A.

54. *See* Napoli, *supra* note 51, at 1.

55. *See id.* at 10. Source diversity can be broken down into three separate categories according to Napoli: (a) ownership diversity of content or programming; (b) ownership diversity of media outlets; and (c) workforce diversity at media outlets.

options.⁵⁶ Content diversity is intended to expose consumers to new types of information that reflects the demographic diversity of the population and, ultimately, the different ideas and viewpoints that they represent. As such, it can be segmented into: (1) program-type format (e.g., comedy, drama, news program); (2) demographic diversity (i.e., portraying racially, ethnically, and gender diverse people in programming); and (3) idea-viewpoint diversity.⁵⁷ Finally, exposure diversity refers to the content that consumers ultimately are exposed to and which enables their participation in the marketplace of ideas.⁵⁸ The Supreme Court has suggested that regulators' pursuit of policies that encourage exposure to diverse sources and diverse content are in line with free speech principles and promote the public interest.⁵⁹ When the FCC has promulgated new regulations or the Court has interpreted the public interest, they have considered one or several of these factors with varying levels of specificity.

For instance, in *FCC v. Prometheus Radio Project*, the issue of minority ownership was a crucial dispute.⁶⁰ Industry respondents rejected minority ownership from the FCC's consideration under § 202(h).⁶¹ While minority ownership was not a consideration by the FCC prior to 1973,⁶² this changed when the D.C. Circuit Court of Appeals held that race was a "relevant and substantial" factor in the FCC's evaluation of radio license applicants.⁶³ Shortly thereafter, the FCC extended their diversity ownership consideration to women as well.⁶⁴ The D.C. Circuit affirmed the importance of minority

56. Source diversity has been the focus of merger proceedings. However, in 2002 the FCC could not conclude that source diversity should be a policy goal of the agency's broadcast ownership rules. *See* 2002 Biennial Regulatory Review—Review of the Commission's Broadcast Ownership Rules and Other Rules Adopted Pursuant to Section 202 of the Telecommunications Act of 1996, 18 FCC Rcd. 13620, 13633 (2003), <https://www.fcc.gov/document/2002-biennial-regulatory-review-review-commissions-broadcast-3> [hereinafter 2002 Review I].

57. *See* Napoli, *supra* note 51, at 11.

58. *See id.* at 24–25.

59. *See* *New York Times v. Sullivan*, 376 U.S. 254 (1964); *Red Lion Broadcasting Co. v. Federal Communications Comm.*, 395 U.S. 367 (1969).

60. 141 S. Ct. at 1155.

61. *See* Reply Brief for Industry Petitioners at 4, *FCC v. Prometheus Radio Project*, 141 S. Ct. 1150 (2021) (Nos. 19-1231 & 19-1241) [hereinafter Reply Brief for Industry Petitioners] (arguing "Section 202(h) does not expressly direct the FCC to consider minority and female ownership, and 'the public interest' cannot be understood as implicitly requiring the Commission to do so.").

62. Robert B. Horwitz, *On Media Concentration and the Diversity Question*, 21 INFO. SOC'Y 181, 190 (2005).

63. *TV 9, Inc. v. FCC*, 495 F.2d 929, 942 (1973).

64. *Gainesville Media, Inc.*, 70 F.C.C.2d 143, 149 (Rev. Bd. 1978).

ownership in 1983 because “our society benefits from exposure to a broad diversity of ideas and perspectives.”⁶⁵

However, before *Prometheus*, diversity ownership also faced several challenges from a set of Justice Sandra Day O’Connor dissents in the early 1990s.⁶⁶ In *Metro Broadcasting*, the majority reasoned that equal employment opportunities would increase minority employment and “contribute significantly toward reducing and ending discrimination in other industries.”⁶⁷ O’Connor wrote that the FCC’s claim “that members of certain races will provide superior programming” should not be legitimized and upheld to a strict scrutiny standard.⁶⁸ Similarly in *Turner Broadcasting*, the Court held that cable broadcasters must carry local broadcast signals.⁶⁹ Once again, O’Connor stressed the importance of maintaining “constitutional requirements” for any interest in diversity of viewpoint or localism that preferences certain speech and restricts other.⁷⁰ In 1995, O’Connor was finally able to write a majority opinion in *Adarand Constructors, Inc. v. Peña* to overrule intermediate scrutiny for race-based ownership regulations.⁷¹ Nevertheless, the Court has never prohibited the use of race-neutral ownership regulation as a means to achieve racial diversity. The lengthy history of the FCC’s diversity regulation, and particularly its diversity ownership consideration, underscores its role in promoting the public interest.

2. *Localism*

Localism has been a core mission and policy goal of the FCC.⁷² Under Title III of the 1934 Communications Act, broadcasters must serve the public interest and must air programming that is “responsive to the interests and needs of their communities of license.”⁷³ Section 307(b) requires the Commission to “make such distribution of licenses, frequencies, hours of operation, and of power among the several States and communities as to provide a fair, efficient, and equitable distribution of [radio] service to each of the same.”⁷⁴ The FCC

65. *W. Mich. Broad. Co. v. FCC*, 735 F.2d 601 (D.C. Cir. 1984), *cert. denied*, 470 U.S. 1027 (1985).

66. *Metro Broad. v. FCC*, 497 U.S. 547 (1990); *Turner Broad. Sys., Inc. v. FCC*, 512 U.S. 622 (1994).

67. *Metro*, 497 U.S. at 555.

68. *Id.* at 620.

69. *Id.* at 637.

70. *Id.* at 680–81, 685.

71. *Adarand Constructors, Inc. v. Peña*, 515 U.S. 200 (1995).

72. *See, e.g.*, *Deregulation of Radio*, 84 F.C.C.2d 968, 994 ¶ 58 (1981) (“The concept of localism was part and parcel of broadcast regulation virtually from its inception.”).

73. *Broadcast Localism*, 19 FCC Rcd. 12425 (2004).

74. 47 U.S.C. § 307(b).

has respected and enforced the concept of localism because “every community of appreciable size has a presumptive need for its own transmission service..”⁷⁵ However, there is no specific statutory basis for a localism requirement nor an explicit mandate; the Commission has interpreted the concept of localism as a derivative of Title III’s broad authority and a factor within the Communications Act’s public interest standard.⁷⁶

The history of localism and § 307(b) explains how an informal principle became a regulatory obligation.⁷⁷ Beginning with the Federal Radio Act of 1927 (“1927 Act”), there has been no explicit reference to serve “specific” or “local” communities.⁷⁸ The 1927 Act provided that, the FRC, when granting or renewing licenses, must consider “a distribution of licenses, bands of frequency of wave lengths, periods of time for operation, and of power among the different States and communities as to give fair, efficient and equitable radio service to each of the same.”⁷⁹ According to the legislative history, allotment on an equitable basis “among States” was core to the provision.⁸⁰ A year later, Congress passed the Davis Amendment to amend § 9 of the 1927 Act to distribute broadcast services among five geographical zones, where licenses were allocated to specific states or zones.⁸¹ When the 1934 Communications Act was passed and the FCC replaced the FRC, § 307(b) was nearly identical to § 9 of the 1927 Act.⁸² Further, the Davis Amendment was repealed due to difficulties in administering the zone system.⁸³

In the succeeding decades of the FCC’s existence, there was no forceful localism obligation, but the Commission referenced the importance of broadcast localism. As part of the *Report on Chain Broadcasting* in 1941, the Commission stated that “[l]ocal program service is a vital part of community life. A station should be ready, able, and willing to serve the needs of the local community by broadcasting such outstanding local events as community concerts, civic meetings, local sports events, and other programs of local

75. Pac. Broad. of Mo. L.L.C., 18 FCC Rcd. 2291 (2003) (quoting Pub. Serv. Broad. of W. Jordan, Inc., 97 F.C.C.2d 960, 962 (Rev. Bd. 1984)).

76. Harry Cole & Patrick Murck, *The Myth of the Localism Mandate: A Historical Survey of How the FCC’s Actions Belie the Existence of a Governmental Obligation to Provide Local Programming*, 15 COMM.LAW CONSP. 339, 341–42 (2007).

77. *Id.* at 343–60.

78. *Id.* at 343.

79. Federal Radio Act of 1927, Pub. L. No. 69-632, § 9, 44 Stat. 1162, 1166.

80. 3 F.R.C. Ann. Rep. 1, 82 (1928).

81. See Cole & Murck, *supra* note 76, at 344–45.

82. See *id.* at 346.

83. See TYLER BERRY, COMMUNICATIONS BY WIRE AND RADIO 134 (1937) (citations omitted); Cole & Murck, *supra* note 76, at 347.

consumer and social interest.”⁸⁴ Further, the FCC held that “programs of local self-expression” were vital to a broadcaster’s “full function.”⁸⁵ Later in 1955, the Commission’s *En Banc Programming Inquiry* focused on network television practices once again reiterated that a “significant element of the public interest is the broadcaster’s service to the community.”⁸⁶ The Inquiry held that “[t]he principal ingredient of such [localism] obligation consists of a diligent, positive, and continuing effort by the licensee to discover and fulfill the tastes, needs, and desires of his service area.”⁸⁷ Increasingly, the Commission was stressing the importance of broadcast localism but nevertheless remained apprehensive about establishing concrete requirements—either due to its limited authority or due to fears over administration issues.

Beginning in the 1960s, the Commission pursued a regulatory system that incentivized broadcasters to advance localism even without a statutory obligation and without triggering First Amendment programming issues.⁸⁸ The FCC established several considerations for broadcasters seeking licenses or renewals, including: (1) maintaining a main studio in the community of license, and originating a majority of its content from that station;⁸⁹ (2) maintaining a local public inspection file with information about the station’s operations;⁹⁰ (3) maintaining detailed logs that describe a station’s local programming;⁹¹ (4) establishing lines of communication between community representatives and the station;⁹² (5) collecting public comments on a station’s renewal application based on their performance to serve the local

84. FED. COMM’NS COMM., REPORT ON CHAIN BROADCASTING, FCC Order No. 37, Docket No. 5060, at 63, 65 (1941).

85. *Id.* at 4.

86. Report and Statement of Policy Res: Commission En Banc Programming Inquiry, 44 F.C.C. 2303 (1960) [hereinafter *En Banc Programming Inquiry*].

87. *Id.* at 2312.

88. See Cole & Murck, *supra* note 76, at 358.

89. See, e.g., Amendment of Parts 1 and 73 of the Commission’s Rules and Regulations Pertaining to the Main Studio Location of FM and Television Broadcast Stations, Report and Order, 27 F.C.C.2d 851 (1971); Reiteration of Policy Regarding Enforcement of Main Studio Rule, 55 Rad. Reg. 2d (P & F) 1178 (1984); Amendment of Sections 73.1125 and 73.1130 of the Commission’s Rules, the Main Studio and Program Origination Rules for Radio and Television Stations, Report and Order (Proceeding Terminated), 2 FCC Rcd. 3215, 3216 (1987).

90. See, e.g., 47 C.F.R. § 73.3526 (2006); *Office of Comm’n of United Church of Christ v. FCC*, 707 F.2d 1413, 1427–28 (D.C. Cir. 1983).

91. See, e.g., Reregulation of Radio and TV Broadcasting, Order, 69 F.C.C.2d 979, 1002–08 (1978); *Office of Comm’n of United Church of Christ*, 707 F.2d at 1422.

92. See, e.g., *En Banc Programming Inquiry*, *supra* note 86; Primer on Ascertainment of Community Problems by Broadcast Applicants, Part I, Sections IV-A and IV-B of FCC Forms, Report and Order, 27 F.C.C.2d 650 (1971); Ascertainment of Community Problems by Broadcast Applicants, Memorandum Opinion and Order, 61 F.C.C.2d 1 (1976).

community.⁹³ Despite establishing these regulatory mechanisms, the Commission rarely denied licensing to a broadcaster that failed to adhere to public interest—and specifically, localism—programming.⁹⁴ The FCC approved thousands of licenses despite no proven record of broadcast localism and serious concerns about stations' programming performance.⁹⁵ By the 1970s, the Commission eliminated requirements to maintain program logging and program reporting.⁹⁶

Broadcast localism, as some critics argue, has become a mere virtue and hardly an obligation. Throughout its history, the Commission has debated whether it should use its licensing renewal process or rely on market forces and programming rules to incentivize broadcasters to further localism.⁹⁷ Without proper policies to assess a broadcaster's performance in providing quantity and quality content to a local community, localism has largely been an unenforced factor of the public interest standard.

3. *Competition*

To evaluate competition, the FCC considers whether stations have adequate incentives to produce diverse news and public interest programming within their communities.⁹⁸ However, the history of competition in the United

93. See Amendment of Section 1.580(m)(1)(iii) of the Rules, Governing Text of Licensee Notice to Public of Broadcast Renewal Application Filings, Memorandum Opinion and Order, 36 F.C.C.2d 685, 3 (1972).

94. See, e.g., Applications of Moline Television Corp. (WQAD-TV), Moline, Ill. For Renewal of License of WQAD-TV; Community Telecasting Corp., Moline, 11. For Construction Permit, Decision, 31 F.C.C.2d 289 (1971); Application of National Broadcasting Company, Inc. For Renewal of License of Station WRC-TV, Washington, D.C., Memorandum Opinion and Order, 52 F.C.C.2d 273 (1975); Application of Talton Broadcasting Company For Renewal of License of Station WHBB, Selma, Alabama, Memorandum Opinion and Order, 58 F.C.C.2d 169 (1976); Application of Vogel-Hendrix Corporation For Renewal of License of Station WAMA, Selma, Alabama, Memorandum Opinion and Order, 58 F.C.C.2d 495 (1976); Applications of Leflore Broadcasting Company, Inc. (WSWG-AM) Greenwood, Mississippi Dixie Broadcasting Company, Inc. (WSWG-FM) Greenwood, Mississippi For Renewal of Licenses, Decision, 65 F.C.C.2d 556 (1977).

95. See Cole & Murck, *supra* note 76, at 360.

96. See, e.g., Deregulation of Radio, 84 F.C.C.2d 968, 975 (1981); see also Deregulation of Radio, Memorandum Opinion and Order, 87 F.C.C.2d 796 (1981); *Office of Comm'n of United Church of Christ*, 707 F.2d at 1413; Deregulation of Radio, Second Report and Order (Proceeding Terminated), 96 F.C.C.2d 930 (1984); Office of Comm'n of the United Church of Christ v. FCC, 779 F.2d 702, 704 (D.C. Cir. 1985); Deregulation of Radio, Memorandum Opinion and Order (Proceeding Terminated), 104 F.C.C.2d 505 (1986).

97. Report on Broadcast Localism and Notice of Proposed Rule Making, 73 Fed. Reg. 8255 (Jan. 24, 2008).

98. 2014 Quadrennial Regulatory Review, 31 FCC Rcd. 9864, 9873 (2016) [hereinafter 2016 Second Report and Order].

States is a complex story characterized by cyclical and abrupt ideological shifts. While there are two agencies responsible for overseeing antitrust enforcement—namely, the Federal Trade Commission (FTC) and the Department of Justice (DOJ)—many other federal agencies regulate competition through their own rules. Over the one-hundred-and-thirty-year history of U.S. antitrust policy, the purpose of competition doctrine has oscillated between preserving democratic and social institutions and efficiently allocating economic resources.⁹⁹ It is no surprise that these competing doctrines closely reflect the dichotomy seen in the public interest models.¹⁰⁰

At the end of the 19th century, Congress passed the first antitrust law in the United States—the Sherman Act of 1890. Born out of popular resentment for concentrated and unfettered monopoly power, the bill passed nearly unanimously in both chambers; only one senator voted against it.¹⁰¹ The Sherman Act, as noted by the Supreme Court in 1958 and supported by one school of antitrust thought, was premised on the idea that “the unrestrained interaction of competitive forces will yield the best allocation of our economic resources, the lowest prices, the highest quality and the greatest material progress, while at the same time providing an environment conducive to the preservation of our democratic political and social institutions.”¹⁰² The delicate balancing of the Sherman Act’s complementary goals—economic prosperity and democracy—indicates that early competitive regulation intended to quell private concentrations of economic power from having a detrimental impact on political and social institutions.¹⁰³ In Justice Louis Brandeis’s words, antimonopoly laws intended to prevent “a power in this country of a few men so great as to be supreme over the law.”¹⁰⁴

99. Sergei Boris Zaslavsky & Melissa H. Maxman, *Too Political or Not Political Enough? A Debate on the Relationship Between Antitrust Enforcement and Democracy*, A.B.A. (May 22, 2023), https://www.americanbar.org/groups/antitrust_law/resources/podcasts/our-curious-amalgam/too-political-or-not-political-enough/; Greg Ip, *Antitrust’s New Mission: Preserving Democracy, Not Efficiency*, WALL ST. J. (July 7, 2021), <https://www.wsj.com/articles/antitrusts-new-mission-preserving-democracy-not-efficiency-11625670424>.

100. See *infra* Section III.A.

101. William L. Letwin, *Congress and the Sherman Antitrust Law: 1887–1890*, 23 U. CHI. L. REV. 221, 222 (1956).

102. N. Pac. Ry. Co. v. United States, 356 U.S. 1 (1958).

103. Maurice E. Stucke & Ariel Ezrachi, *The Rise, Fall, and Rebirth of the U.S. Antitrust Movement*, HARV. BUS. REV. (Dec. 15, 2017), <https://hbr.org/2017/12/the-rise-fall-and-rebirth-of-the-u-s-antitrust-movement>.

104. Louis D. Brandeis, Bos. Bar, Address to the Economic Club of New York: The Regulation of Competition Versus the Regulation of Monopoly (Nov. 1, 1912), in 3 YEARBOOK OF THE ECONOMIC CLUB OF NEW YORK 7 (1913).

While competition laws went unenforced for decades due to administrative negligence and judicial aversion, competition was largely understood as an issue of political economy until the 1970s. Throughout the mid-20th century, Congress proceeded to pass the Federal Trade Commission Act to ban “unfair methods of competition” and “unfair or deceptive practices.”¹⁰⁵ Fairness was prominently a characteristic of competition policy. Congress then passed the Clayton Act to address anticompetitive mergers and interlocking directorates.¹⁰⁶ The pursuit of the *competitive ideal*—an equitable dispersion of economic and political power to promote competition in line with democratic principles, as some scholars have defined it—characterized the “golden era” of competition enforcement.¹⁰⁷

Beginning in the late 1970s, American competition doctrine experienced a profound change. The Chicago School, advanced by the work of Robert Bork, shifted the traditional understanding of antitrust toward a theory dominated by conservative economics.¹⁰⁸ In 1979, the Supreme Court held that “Congress designed the Sherman Act as a ‘consumer welfare prescription’” and the consumer welfare standard became the doctrinal consensus for the next three decades.¹⁰⁹ While scholars have disagreed on aspects of consumer welfare, such as whether the analysis should end at price effects or total welfare, the Chicago School has prioritized efficiency and relied on the market to settle.¹¹⁰ Critiques have challenged the consumer welfare standard as non-interventionist, and blamed that lax standard for increasing levels of inequality and market concentration.¹¹¹

If the trajectory of competition doctrine sounds familiar, it is because the FCC’s media ownership rules have largely followed along in parallel. The FCC

105. 15 U.S.C. § 41.

106. 15 U.S.C. § 12.

107. Stucke & Ezrachi, *supra* note 103.

108. Tim Wu, *After Consumer Welfare, Now What? The “Protection of Competition” Standard in Practice* (Columbia Pub. Law Research Paper No. 14-608, 2018), <https://ssrn.com/abstract=3249173>.

109. *Reiter v. Sonotone Corp.*, 442 U.S. 330, 343 (1979) (citing Robert H. Bork, *THE ANTITRUST PARADOX: A POLICY AT WAR WITH ITSELF* 66 (1978)).

110. Daniel A. Crane, *Four Questions for the Neo-Brandeisians*, *ANTITRUST CHRONICLE* (April 2018), at 1, <https://www.competitionpolicyinternational.com/wp-content/uploads/2018/04/CPI-Crane.pdf>.

111. See, e.g., *Crack Down on Corporate Monopolies & the Abuse of Economic and Political Power*, *BETTER DEAL*, <http://abetterdeal.democraticleader.gov/wp-content/uploads/2017/10/A-BetterDeal-on-Competition-and-Costs.pdf> [<https://perma.cc/J8CL-XJQL>] (last visited Nov. 11, 2023) (“The extensive concentration of power in the hands of a few corporations hurts wages, undermines job growth, and threatens to squeeze out small businesses, suppliers, and new, innovative competitors.”).

began with a presumption against concentrated ownership in the 1930s¹¹² and later instituted more stringent media ownership rules in the 1960s.¹¹³ The rationales behind these rules were underpinned by a stringent commitment to the public interest and the media's role as a sociopolitical institution.¹¹⁴ While these tradeoffs were not seen as counterintuitive to competition, the emergence of law and economics, as well as the consumer welfare standard, revolutionized competition doctrine. Because fairness was no longer perceived as necessary for markets to function well, any factors that would impede the efficiency model, such as diversity or localism, were considered anti-competitive.

III. THE FIGHT OVER AMERICA'S PUBLIC INTEREST

A. REGULATORY PURPOSE: DEMOCRACY VS. EFFICIENCY

Without an explicit definition, different FCC administrations have enforced the public interest standard to achieve their own political objectives. In his 2006 article "Antitrust Law as Mass Media Regulation: Can Merger Standards Protect the Public Interest?," Georgetown Law Professor Howard Shelanski described two distinct public interest regimes that FCC administrations have pursued: the democracy model and the market-efficiency model.¹¹⁵

While each model claims to advance the public interest and prioritize the needs of American citizens, they envision the regulatory purpose of the law differently. The democracy model combines sociopolitical factors that prevent against concentrated ownership and promote local service and community.¹¹⁶ The efficiency model relies on market mechanisms to produce quality broadcasting which in turn aims to provide viewers with better quality information.¹¹⁷

1. *The Democracy Model*

Under the democracy model, media regulation is intended to preserve the ideals of localism, multiple voices, and access.¹¹⁸ While proponents advance

112. *Genesee*, 5 F.C.C. at 183.

113. See *supra* note 42 and accompanying text.

114. See Wu, *supra* note 108, at 11.

115. Howard A. Shelanski, *Antitrust Law as Mass Media Regulation: Can Merger Standards Protect the Public Interest?*, 94 CALIF. L. REV. 371, 371 (2006).

116. *Id.* at 384.

117. *Id.* at 383–84.

118. See Benjamin M. Compaine & Douglas Gomery, WHO OWNS THE MEDIA? COMPETITION AND CONCENTRATION IN THE MASS MEDIA INDUSTRY 554 (2000).

these ideals for slightly different reasons, they generally seek to promote a well-informed citizenry through the means of independent media. For instance, Yale Professor Robert Post has argued that “democracy requires a public forum in which all policy goals are open for discussion and none . . . is taken as given.”¹¹⁹ Similarly, Edwin Baker makes the point that self-determination is critical to democracy; in order to self-govern, citizens must be able to form public opinion within an egalitarian media structure.¹²⁰ Others have argued that the democracy model achieves other benefits such as viewpoint, source, and racial diversity in ownership.¹²¹ As such, proponents argue that diversity is critical among any media regulation objective.¹²²

Arguably, the democracy model aligns with how the founders envisioned the development of the free press and how media regulation developed up until recent decades.¹²³ The FCC, Congress, and courts overwhelmingly aligned with the democracy model for most of the 20th century. In 1931, the Supreme Court first ruled on “public interest” in *KFKB Broadcasting Ass’n v. Federal Radio Commission*. The Court granted the FRC discretion to limit licensing based on the “character and quality of the service rendered.”¹²⁴ A year later, in *Trinity Methodist Church v. Federal Radio Commission*, the Court allowed the FRC to deny a radio station broadcasting rights because it “obstructed the administration of justice, offended the religious susceptibilities of thousands, inspired political distrust and civic discord . . . and offended youth and innocence by the free use of words suggestive of sexual immortality.”¹²⁵ In the early days of the public interest standard, the Court ensured that broadcast media was operating with a sense of decency and with civic purpose.

A few years later, Congress adopted the Communications Act of 1934 (“1934 Act”). The “equal-time rule,” also known as § 315, required radio and television stations and cable systems to “afford equal opportunities” for

119. See Shelanski, *supra* note 115, at 387.

120. See BAKER, *supra* note 13, at 6–7.

121. See MCGANNON CTR., FORDHAM UNIV., THE CASE AGAINST MEDIA CONSOLIDATION: EVIDENCE ON CONCENTRATION, LOCALISM AND DIVERSITY 77, 201, 331 (Mark N. Cooper ed., 2017).

122. Geoffrey Starks, Commissioner, Fed. Comm’n’s Comm’n, Remarks of Commissioner Geoffrey Starks at the FCC Communications Equity and Diversity Council’s Media Ownership Diversity Symposium (Feb. 7, 2023), <https://docs.fcc.gov/public/attachments/DOC-391014A1.pdf>.

123. Cf. Shelanski, *supra* note 115, at 387 (noting that in 1940 the Supreme Court limited that mandate, declaring in *FCC v. Sanders Brothers Radio Station* that “the field of broadcasting is one of free competition . . . The Commission is given no supervisory control of the programs, of business management or of policy.”)

124. *KFKB Broad. Ass’n v. Fed. Radio Comm’n*, 47 F.2d 670 (D.C. Cir. 1931).

125. *Trinity Methodist Church, S. v. Fed. Radio Comm’n*, 62 F.2d 850 (D.C. Cir. 1932).

airtime to all legally qualified candidates for any public office.¹²⁶ This provision was explicitly enacted to protect against broadcasters abusing their political power and to ensure an informed public.¹²⁷ A decade after Congress passed the 1934 Act, the FCC issued the Public Service Responsibility of Broadcast Licensees, its first major guidelines on broadcast programming.¹²⁸ The document identified fourteen major elements of programming necessary to serve the public interest, including:

- (1) opportunity for local self-expression, (2) the development and use of local talent, (3) programs for children, (4) religious programs, (5) educational programs, (6) public affairs programs, (7) editorialization by licensees, (8) political broadcasts, (9) agricultural programs, (10) news programs, (11) weather and market reports, (12) sports programs, (13) service to minority groups, (14) entertainment programs.¹²⁹

The great variety was intentionally set to ensure that the public received a diversity of content, otherwise it might not be covered due to market failures.

For instance, the development of local talent or service to minority groups was deemed important to the public even if there was no overwhelming consumer demand.¹³⁰ In 1943, the Supreme Court then once again affirmed the FCC's important role in regulating the public interest and upheld their authority to enforce the Chain Broadcasting Regulations.¹³¹ The Court upheld the FCC's exercise of its statutory authority as constitutional because the "public interest" was not a "a mere general reference to public welfare without any standard to guide determinations" and "[t]he purpose of the [1936] Act, the requirements it imposes, and the context of the provision in question show the contrary."¹³²

After the Second World War, the Commission on Freedom of the Press (also known as the Hutchins Commission and led by the famed Robert

126. 47 U.S.C. § 315.

127. Richard G. Singer, *The FCC and Equal Time: Never-Neverland*, 27 MD. L. REV. 221, 236 (1967).

128. STEVEN WALDMAN, FED. COMM'NS COMM'N, *THE INFORMATION NEEDS OF COMMUNITIES: THE CHANGING MEDIA LANDSCAPE IN A BROADBAND AGE* 281 (2011).

129. *Id.* at 281.

130. *Id.*

131. *Nat'l Broad. Co. v. United States*, 319 U.S. 190 (1943).

132. *Id.* at 226 (quoting *N.Y. Cent. Sec. Corp. v. United States*, 287 U.S. 12, 24–25 (1932)); *see* *Radio Comm'n v. Nelson Bros. Co.*, 289 U.S. 266, 285 (1933); *FCC v. Pottsville Broad. Co.*, 309 U.S. 134, 137–38 (1940). *Compare* *Panama Refining Co. v. Ryan*, 293 U.S. 388, 428 (1935), *with* *Intermountain Rate Cases*, 234 U.S. 476, 486–89 (1914), *and* *United States v. Lowden*, 308 U.S. 225 (1939).

Maynard Hutchins) reaffirmed the media's public interest role.¹³³ The Hutchins Commission concluded that the press was a "conveyer of information, government watchdog, and educator."¹³⁴

Soon after, the FCC encountered First Amendment challenges to its public interest objectives and the Supreme Court once again upheld its authority. In 1949, the FCC introduced the now-defunct Fairness Doctrine which required broadcasters to present balanced coverage for controversial issues of public importance. The Fairness Doctrine intended to expose viewers to diverse information and prevent broadcasters from monopolizing the airwaves with biased coverage.¹³⁵ It wasn't until *Red Lion Broadcasting Co. v. Federal Communications Commission* that the Supreme Court addressed broadcasters' First Amendment rights and upheld the constitutionality of the Fairness Doctrine.¹³⁶ Referring to the legislative record for the Radio Act of 1927, the Court pointed to Congressman Byron R. White's reasoning for granting licenses "only to those stations whose operation would render a benefit to the public, are necessary in the public interest, or would contribute to the development of the art."¹³⁷ The First Amendment challenge, as a matter of the public interest, became a recurring factor where the Court has remained sensitive but largely deferential to administrative authority.¹³⁸

On several occasions, the Supreme Court explicitly recognized the importance of diversity within the public interest mandate in ways that align with the democracy model of media regulation. In 1972, the Court noted in *United States v. Midwest Video Corp.* that "it has long been a basic tenet of national communications policy that 'the widest possible dissemination of information from diverse and antagonistic sources is essential to the welfare of the public.'"¹³⁹ More recently in *Turner Broadcasting System, Inc. v. FCC*, Justice O'Connor quoted from the Cable Television Consumer Protection and Competition Act of 1992 to emphasize the role of diversity in media regulation: "[t]here is a substantial governmental and First Amendment

133. See COMM'N ON FREEDOM OF THE PRESS, A FREE AND RESPONSIBLE PRESS (Robert D. Leigh ed., 1947).

134. See Horwitz, *supra* note 62, at 182.

135. KATHLEEN ANN RUANE, CONG. RSCH. SERV., R40009, FAIRNESS DOCTRINE: HISTORY AND CONSTITUTIONAL ISSUES 10 (2011).

136. *Red Lion Broad. Co. v. FCC*, 395 U.S. 367 (1969).

137. *Id.* at 40.

138. See *Columbia Broad. Sys., Inc. v. Democratic Nat'l Comm.*, 412 U.S. 94 (1973).

139. 406 U.S. 649, 668 n.27 (1972) (plurality opinion).

interest in promoting a diversity of views provided through multiple technology media.”¹⁴⁰

From its emergence in the 1930s and up until the 1980s, the public standard doctrine was interpreted broadly and largely supported the democracy model of mass media regulation. The FCC used regulations, such as the cross-ownership rules, and its licensing authority to promote “localism, diversity of ownership, and diversity of programming.”¹⁴¹

2. *The Efficiency Model*

Approximately fifty years ago, media regulation advocates softened their adherence to the democracy model and embraced the efficiency-oriented model.¹⁴² The efficiency model seeks to serve consumer demand with greater efficiency by focusing on quality and responsiveness.¹⁴³ Free market conditions, proponents argue, can sufficiently supply the public with necessary information to make informed decisions without governmental intervention.¹⁴⁴ Further, efficiency advocates argue that the advent of the internet and the proliferation of other technologies have expanded access to different media sources.¹⁴⁵ Despite the alleged superiority of efficient enterprise and the abundance of outlets to choose from, these arguments did not prevail in media regulation policy until law and economics theories gained broader influence among policy circles.

Between the 1930s and the 1980s, there were only a few instances where the courts or the FCC used efficiency model rationales to support their decision-making.¹⁴⁶ In 1933, the Supreme Court held in *Federal Radio Commission v. Nelson Brothers Bond & Mortgage Co.* that the Commission’s requirement to act as “public convenience, interest or necessity requires” did not equate to a “setting up a standard so indefinite as to confer an unlimited power.”¹⁴⁷ In particular, the Court explicitly listed the factors which the Commission was required to consider, including: “its context, by the nature of

140. *Turner*, 512 U.S. at 676 (O’Connor, J., concurring in part and dissenting in part) (quoting Pub. L. No. 102-385, § 2(a)(6), 106 Stat. 1460 (1992)).

141. Shelanski, *supra* note 115, at 387.

142. *See id.* at 383.

143. *See id.*

144. Michael O’Rielly, *Defending Capitalism in Communications*, FED. COMM. COMMISSION: FCC BLOG (Feb. 12, 2016), <https://www.fcc.gov/news-events/blog/2016/02/12/defending-capitalism-communications>.

145. *Prometheus*, 141 S. Ct. at 1155 (“By the 1990s, however, the market for news and entertainment had changed dramatically. Technological advances led to a massive increase in alternative media options, such as cable television and the Internet.”).

146. *See* Shelanski, *supra* note 115, at 387.

147. *Nelson Bros.*, 289 U.S. at 285.

radio transmission and reception, by the scope, character and *quality of services*, and, where an equitable adjustment between States is in view, by the *relative advantages in service* which will be enjoyed by the public through the distribution of facilities.”¹⁴⁸ A few years later, in *FCC v. Sanders Brothers Radio Station*, the Court more forcefully held that “the field of broadcasting is one of free competition The Commission is given no supervisory control of programs, of business management or of policy.”¹⁴⁹ This departure seemed at odds with other similar cases years prior, such as *Associated Press, KFKB*, and *Trinity Methodist Church*, where the Court relied on the agency to secure the public interest with broad discretion and oversight.¹⁵⁰ Despite this aberration, the Supreme Court persistently protected the democratic ideals of localism, diversity, and access throughout the mid-century.

It was only during the 1980s that deregulation and free market solutions came to dominate political thought in government.¹⁵¹ In 1981, FCC Chairman Charles Ferris led a broadscale repeal of radio regulations because the public interest would be best served by eliminating “unnecessarily burdensome regulations of uniform applicability that fail to take into account local conditions, tastes or desires.”¹⁵² The Commission eliminated license-renewal guidelines requiring stations to offer non-entertainment programming, eliminated ascertainment requirements to evaluate community needs, removed restrictions on the number of commercials that could be aired, and abandoned requirements to keep public programming logs.¹⁵³ By 1984, President Ronald Reagan appointed FCC Chairman Mark Fowler who essentially transposed each of the radio rules on the television broadcasting stations.¹⁵⁴

B. ERA OF DEREGULATION

The final blow to FCC’s public interest deregulation came with the passage of the Telecommunications Act of 1996. A Republican-controlled Congress passed the law with overwhelming support—414 to 16 in the House and 91 to

148. *Id.* (emphasis added).

149. *FCC v. Sanders Bros. Radio Station*, 309 U.S. 470, 474–75 (1940).

150. *Associated Press v. United States*, 326 U.S. 1 (1945); *KFKB*, 47 F.2d at 670; *Trinity Methodist*, 62 F.2d at 850.

151. *See, e.g.*, Anthony E. Varona, *Out of Thin Air: Using First Amendment Public Forum Analysis to Redeem American Broadcasting Regulation*, 39 U. MICH. J.L. REFORM 149, 158–59 (2006); Harrison Donnelly, *Broadcasting Deregulation*, in EDITORIAL RESEARCH REPORTS 1987, at 629–44 (Hoyt Gimlin ed., 1988); Kevin M. Kruse & Julian Zelizer, *How Policy Decisions Spawned Today’s Hyperpolarized Media*, WASH. POST (Jan. 17, 2019), <https://www.washingtonpost.com/outlook/2019/01/17/how-policy-decisions-spawned-todays-hyperpolarized-media/>.

152. WALDMAN, *supra* note 128, at 283.

153. *See id.* at 283–84.

154. *See id.* at 284.

5 in the Senate—and President Bill Clinton signed it into law.¹⁵⁵ The purpose of the act could be found explicitly in its long title: “An Act to promote competition and reduce regulation in order to secure lower prices and higher quality services for American telecommunications consumers and encourage the rapid development of new telecommunications technologies.”¹⁵⁶ With significant pressure from the broadcasting lobby, the 1996 Act dismantled the FCC’s authority to regulate the public interest through license renewals.

First, Congress extended each license term from three to eight years for television and radio stations. Given the FCC’s weak licensing enforcement, this ensured that broadcasters would maintain licenses for nearly a decade without much scrutiny before renewal. Second, Congress prohibited the FCC from considering competing applications before an incumbent’s licenses could be revoked. The resulting outcome would disadvantage new competitors and thus, likely limit historically underrepresented media ownership. Finally, § 202(h) required the Commission to review its media ownership rules every four years.¹⁵⁷ As part of this process, the Commission must review any proposed rule change and, importantly, assess whether it is “necessary in the public interest as the result of competition.”¹⁵⁸ Bringing up rules for quadrennial reviews created a more politicized and litigious FCC rulemaking process. Together, these new rules set the stage for abandoning the public interest standard and deregulating media ownership.

Media companies opposed the cross-ownership rules since their inception.¹⁵⁹ However, it was not until the FCC’s 2002 Biennial Regulatory Review that the agency began to review and relax its rules governing market concentration and cross-ownership.¹⁶⁰ By June 2003, the Commission adopted a Report and Order which stated that, “neither an absolute prohibition on common ownership of daily newspapers and broadcast outlets in the same market (the ‘newspaper/broadcast cross-ownership rule’) nor a cross-service restriction on common ownership of radio and television outlets in the same market (the ‘radio-television cross-ownership rule’) [remain] necessary [for] the

155. *Congress Puts Finishing Touches on Major Industry Overhaul*, CONG. Q. ALMANAC (1995), <https://library.cqpress.com/cqalmanac/document.php?id=cqa195-1100302>.

156. Telecommunications Act of 1996, Pub. L. No. 104-104, § 202(b), 110 Stat. 56, 110 (1996).

157. *See id.* at 111–12.

158. *Id.* at 112.

159. GOMERY, *supra* note 43, at 1.

160. *See* 2002 Biennial Regulatory Review—Review of the Commission’s Broadcast Ownership Rules and Other Rules Adopted Pursuant to Section 202 of the Telecommunications Act of 1996, 17 FCC Rcd. 18503 (2002).

public interest.”¹⁶¹ The rule changes allowed the Commission to abandon the enforcement of market concentration to general antitrust laws rather than subject it to its own more stringent regulation.¹⁶²

Following the FCC’s move to deregulate the ownership rules, Prometheus Radio Project—a non-profit advocacy group with a mission to resist corporate media consolidation—and other public interest groups embarked on a nearly twenty-year journey to uphold the prior ownership rules.¹⁶³ Between 2003 and 2019, the Third Circuit reviewed four separate challenges to the FCC’s rule changes.¹⁶⁴ The Commission’s general position was that new technologies changed the industry and that the prior rules “inadequately [accounted] for the competitive presence of cable, [ignored] the diversity-enhancing value of the internet, and [lacked] any sound basis for a national audience reach cap.”¹⁶⁵ Meanwhile, the Third Circuit consistently held that the Commission failed to provide reasoned analysis for its numerical limits on common ownership, consider the effects of its new rules on minority ownership, or justify market share metrics and assumptions.¹⁶⁶ In each case, the Supreme Court denied certiorari for all relevant appeals.¹⁶⁷ In the most recent successful *Prometheus* challenge in 2016, the Third Circuit concluded that the Commission’s rule changes were arbitrary and capricious because they did not adequately assess the deregulatory effect on media ownership diversity—particularly minority and female ownership.¹⁶⁸ Under the Administrative Procedure Act, an agency violates the arbitrary and capricious standard when it “entirely fail[s] to consider an important aspect of the problem.”¹⁶⁹

Simultaneously in 2016, with the *Prometheus* litigation saga ongoing, FCC Chairman Tom Wheeler proposed to retain the original cross-ownership rules with slight modifications.¹⁷⁰ While the core rules remained intact, the FCC

161. 2002 Review I, *supra* note 56, ¶ 2.

162. *See* Shelanski, *supra* note 115, at 375.

163. *See* Prometheus Radio Project v. FCC (*Prometheus IV*), 939 F.3d 567 (3d Cir. 2019); Prometheus Radio Project v. FCC (*Prometheus III*), 824 F.3d 33 (3d Cir. 2016); Prometheus Radio Project v. FCC (*Prometheus II*), 652 F.3d 431 (3d Cir. 2011); Prometheus Radio Project v. FCC (*Prometheus I*), 373 F.3d 372 (3d Cir. 2004).

164. *See* cases cited *supra* note 163.

165. 2002 Biennial Regulatory Review, *supra* note 56.

166. *See* cases cited *supra* note 163.

167. *See, e.g., Prometheus II*, 652 F.3d at 431 (3d Cir. 2011), *cert. denied*, 567 U.S. 951 (2012); *Prometheus I*, 373 F.3d at 372, *cert. denied*, 545 U.S. 1123 (2005).

168. *Prometheus III*, 824 F.3d at 54 n.13.

169. *Motor Vehicle Mfrs. Ass’n v. State Farm Mut. Auto. Ins. Co.*, 463 U.S. 29, 43 (1983).

170. 2016 Second Report and Order, *supra* note 98; *see also* FCC Chair Proposes Retaining Most U.S. Media Ownership Rules, REUTERS (June 27, 2016), <https://www.reuters.com/article/us-usa-media-rules/fcc-chair-proposes-retaining-most-u-s-media-ownership-rules-idUSKCN0ZD2QC>.

created an exception which allowed “failed or failing newspapers” to receive investment from a broadcast television or radio station in the same market.¹⁷¹ The Newspaper Association of America reacted negatively to Chairman Wheeler’s proposal saying it was “stunned that any policymaker in the internet era would propose to keep a 1970s-era law that prevents broadcast stations and newspapers from being owned by the same company.”¹⁷² While the Democrat-controlled FCC attempted to preserve the ownership rules and with them, the public interest, the effort was short lived. The rule changes were once again challenged—this time by deregulation advocates and revenue-losing media companies—in the fourth iteration of the *Prometheus* saga.¹⁷³

By the end of 2016, American voters elected Donald Trump as President and subsequently the FCC’s political leadership changed with the appointment of Chairman Ajit Pai. The new chairman reinvigorated the campaign to deregulate the FCC ownership rules with significant overhauls in 2017 and 2018. In 2017, the Commission revoked the 2016 rule changes and eliminated the original cross-ownership rules.¹⁷⁴ In 2018, the Commission established an incubator program to promote the entry of new and diverse voices into the broadcast industry.¹⁷⁵ Both orders were challenged and in 2019, the Third Circuit ruled that the FCC had not “adequately considered the effects” of the new rules on “diversity in broadcast media ownership.”¹⁷⁶ This time, after seventeen years, the Supreme Court granted certiorari in 2020.¹⁷⁷

IV. REVIEWING *FCC V. PROMETHEUS RADIO PROJECT*

On April 1, 2021, the Supreme Court unanimously ruled in favor of the FCC’s deregulatory change to repeal or modify three media ownership rules—the Newspaper/Broadcast Cross-Ownership Rule, the Radio/Television

171. Roger Yu, *FCC Retains Media Cross-Ownership Rules*, USA TODAY (Aug. 11, 2016), <https://www.usatoday.com/story/money/2016/08/11/fcc-retains-media-cross-ownership-rules/88584310/>.

172. *U.S. FCC Votes to Keep Most Media Ownership Rules*, REUTERS (Aug. 11, 2016), <https://www.cnbc.com/2016/08/11/us-fcc-votes-to-keep-most-media-ownership-rules.html>.

173. *Prometheus IV*, 939 F.3d at 567.

174. *See* 2014 Quadrennial Regulatory Review—Review of the Commission’s Broadcast Ownership Rules and Other Rules Adopted Pursuant to Section 202 of the Telecommunications Act of 1996, Order on Reconsideration and Notice of Proposed Rulemaking, 32 FCC Rcd. 9802 (2017) [hereinafter 2017 Order].

175. *See* Rules and Policies to Promote New Entry and Ownership Diversity in the Broadcasting Services, Report and Order, 33 FCC Rcd. 7911 (2018) [hereinafter 2018 Rules and Policies].

176. *Prometheus IV*, 939 F.3d at 584–88.

177. *Prometheus*, 141 S. Ct. at 1157.

Cross-Ownership Rule, and the Local Television Ownership Rule.¹⁷⁸ In arguments, the FCC relied on its conclusion in the 2017 annual review, where the Commission found that the cross-ownership rules were “no longer necessary to serve the agency’s public interest goals of competition, localism, and viewpoint diversity.”¹⁷⁹ The agency argued that it had the administrative authority to make such rule changes after basing its decision on record evidence, public comments, and with consideration for media industry developments since the 1960s.¹⁸⁰

The parties sharply disagreed about the weight and scope of each public interest factor.¹⁸¹ The FCC and industry petitioners argued that § 202(h) authorized them to forego minority ownership analysis because the legislative intent prioritized competition. Industry petitioners claimed that § 202(h) required the FCC only to consider competition, rather than minority and female ownership—and that “the public interest” cannot be understood as implicitly requiring the Commission to [consider diversity ownership].¹⁸² According to them, Congress intended “competition to play a starring role, not second fiddle, in regulatory reform reviews” when drafting the Telecommunications Act of 1996.¹⁸³

Prometheus Radio Project and other media advocacy organizations opposed this characterization. The group argued that the FCC’s decision to change the rules was not made in the public interest because it was likely to harm minority and female ownership¹⁸⁴—factors that both Congress and the Supreme Court have recognized as “essential” to the public interest.¹⁸⁵ To support its factual conclusions, Prometheus relied on several studies conducted by Free Press, a media reform group.¹⁸⁶ The studies showed that past deregulation of ownership rules led to increases in media market

178. *Id.* at 1152–53.

179. *Id.* at 1158.

180. *Id.*; see, e.g., 2016 Second Report and Order, *supra* note 98, at 9803, 9807, 9825, 9834.

181. See, e.g., 2018 Quadrennial Regulatory Review—Review of the Commission’s Broadcast Ownership Rules and Other Rules Adopted Pursuant to Section 202 of the Telecommunications Act of 1996, Notice of Proposed Rulemaking, 33 FCC Rcd. 12111, 12116, 12128, 12140, ¶¶ 9, 40, 77 (2018); 2016 Second Report and Order, *supra* note 98, at 9865, ¶ 3.

182. Reply Brief for Industry Petitioners, *supra* note 61, at 25.

183. *Id.* at 4.

184. *Prometheus*, 141 S. Ct. at 1159.

185. Briefs of Members of Congress as *Amici Curiae* in Support of Respondents at 3, FCC v. Prometheus Radio Project, 141 S. Ct. 1150 (2021) (Nos. 19-1231 & 19-1241) (citing the Cable Television Consumer Protection and Competition Act of 1992; *Turner*, 512 U.S. at 676; and *United States v. Midwest Video Corp.*, 406 U.S. 649, 668 n.27 (1972), among others).

186. *Prometheus*, 141 S. Ct. at 1159.

concentration and ultimately decreased minority and female ownership levels.¹⁸⁷ According to the media advocacy groups, these negative results were significant because the FCC has “long acted on the theory that diversification of mass media ownership serves the public interest by promoting diversity of program and service viewpoints, as well as by preventing undue concentration of economic power.”¹⁸⁸

Writing for a unanimous court, Justice Brett Kavanaugh shied away from defining or balancing the public interest. The Court ruled that the FCC was “reasonable and reasonably explained for purposes of the APA’s deferential arbitrary-and-capricious standard” its interpretation of countervailing evidence.¹⁸⁹ First, upon evaluating the evidence, the FCC concluded that “no record evidence suggesting that past changes to the ownership rules had *caused* minority ownership levels to increase.”¹⁹⁰ Second, the FCC explained that the ownership rules no longer fit the reality of today’s media industry and that “permitting efficient combinations among radio stations, television stations, and newspapers would benefit consumers.”¹⁹¹ Succinctly, the Court held that “[t]he APA requires no more.”¹⁹²

After seventeen years, the Court’s opinion was relatively short. By skirting around the public interest standard, the Court avoided taking a normative stance on media regulation in America. Instead, the case focused on administrative authority and the burdens of agency rulemaking. If the evidentiary gap indeed favored the FCC’s discretion,¹⁹³ then it makes sense that none of the judges wrote a dissenting opinion. However, according to the Court, the factual gap mattered only in so much that the FCC gathered public comments and considered them; beyond that, the Commission is wholly justified in its interpretation of countervailing evidence, seemingly with little regard for the merits.¹⁹⁴ The result of this decision will be a wholly politicized agency rule-making process. If the last eighteen years are any example, the Third Circuit might be the public interest’s sole line of defense.

In a short concurrence, Justice Clarence Thomas weighed in to criticize the Third Circuit for improperly imposing a procedural requirement on the

187. *Id.*

188. *FCC v. Nat’l Citizens Comm. for Broad.*, 436 U.S. 775, 780 (1978).

189. *Prometheus*, 141 S. Ct. at 1160.

190. *Id.* at 1159.

191. *Id.* at 1157.

192. *Id.* at 1160.

193. *See id.* The FCC argued that there was a lack of predictive data to show that the rule changes would lead to fewer minority and female owners.

194. *See id.* at 1159–60.

FCC to consider ownership diversity.¹⁹⁵ According to Thomas, the FCC was only required to consider the “public interest as the result of competition” and it had “no obligation to consider minority and female ownership.”¹⁹⁶ The concurrence further stated that the FCC’s ownership rules were “never designed to foster ownership diversity” and thus, it does not matter that the agency considered it as a factor in its prior policy.¹⁹⁷ However, Thomas conceded that diversity ownership was, in fact, prior policy but only as a proxy for viewpoint diversity.¹⁹⁸ By setting this distinction, Thomas attempted to define the FCC’s regulatory target as consumers, rather than producers.¹⁹⁹ Citing the Supreme Court’s 1940 decision in *FCC v. Pottsville Broadcasting Co.*, Thomas highlighted that the Commission clarified that “emphasis must be first and foremost on the interest, the convenience, and the necessity of the listening public, and not on the interest, convenience, or necessity of the individual broadcaster.”²⁰⁰

This formalist approach ignores the purposes of the public interest standard. To make his arguments, Justice Thomas relied heavily on the public interest standard’s disputed history.²⁰¹ Since its adoption in the 1930s, the “public interest” has not been defined in any formal statutory manner. Because of this, Thomas insisted that the Third Circuit cannot inject a requirement to consider ownership diversity where one does not exist. According to Thomas, there is no “freestanding goal of promoting ownership diversity” and that promoting minority and female ownership only serves the core goal of maximizing the diversity of viewpoints.²⁰² Because the ownership rules were “never designed to foster ownership diversity,” Thomas argued that the FCC is only required to consider the effects of any rule change on viewpoint diversity.²⁰³

The purpose of the public interest standard, as evidenced by its origins and longstanding history, was to serve the informational needs of a well-informed citizenry. The FCC has consistently held that this goal should be achieved through a diversity of voices.²⁰⁴ While there are many ways to achieve a

195. *Id.* at 1160 (Thomas, J., concurring).

196. *Id.* at 1161.

197. *Id.*

198. *Id.* at 1162.

199. *See id.* at 1161 (“From its infancy, the FCC has generally focused on consumers, not producers.”).

200. *Pottsville*, 309 U.S. at 138 n.2 (quoting a 1928 agency document).

201. *Prometheus*, 141 S. Ct. at 1161–62 (Thomas, J., concurring).

202. *Id.* at 1162.

203. *Id.* at 1161.

204. *See* 2016 Second Report and Order, *supra* note 98 (stating that the FCC “has a long history of promulgating rules and regulations intended to promote diversity of ownership

diversity of voices, diverse ownership is one obvious and valid approach. The FCC recognized this when it adopted a presumption against media duopolies in 1938²⁰⁵ and even in 2018, when Commissioner Pai created an incubator program to promote the new and diverse voices entering the broadcast industry.²⁰⁶ To claim that the FCC has focused on consumers and not producers, as Justice Thomas did, is ahistorical.²⁰⁷

One likely reason why minority and female ownership was not formally recognized in neither congressional legislation nor the FCC's rulemaking is because these groups have been historically excluded. In 1971, only 10 of the 7,500 radio stations (0.13%), and none of the 1,000 television stations in the United States, were minority-owned.²⁰⁸ In 2019, when Black Americans made up roughly 14% of the U.S. population,²⁰⁹ still only 1.3% of U.S. full-power commercial TV stations were Black-owned.²¹⁰ Similarly, only 2% of commercial FM stations were Black-owned. Despite people of color (POC) making up 43% of the U.S. population, only 6% of the nation's full-power TV stations, 7% of commercial FM radio stations, and 13% of commercial AM radio stations were POC-owned.²¹¹ While it is impossible to tell whether the ownership rules were responsible for the modest increase in minority ownership as opposed to other factors, it is clear that deregulation and media consolidation produces the opposite result. According to a study from Free Press, the FCC's era of deregulation in the 1990s led to the loss of over 40% of minority-owned stations by 1998.²¹²

among broadcast licensees, and thereby foster a diversity of voices"); *Metro*, 497 U.S. at 566–68 (upholding “minority ownership policies” because they were “substantially related to the achievement of . . . broadcast diversity”).

205. *Genesee*, 5 F.C.C. at 183 (calling for a “diversification of service”).

206. *See* 2018 Rules and Policies, *supra* note 175.

207. *See Prometheus*, 141 S. Ct. at 1161 (Thomas, J., concurring).

208. Beth Brodsky & Daniel A. Hanley, *The FCC Seeks to Hinder Female and Minority Broadcast Ownership for Policies Favoring Concentrated Corporate Ownership*, COMMON DREAMS (Jan. 28, 2021), <https://www.commondreams.org/views/2021/01/28/fcc-seeks-hinder-female-and-minority-broadcast-ownership-policies-favoring>.

209. Christine Tamir, *The Growing Diversity of Black America*, PEW RES. CTR. (Mar. 25, 2021), <https://www.pewresearch.org/social-trends/2021/03/25/the-growing-diversity-of-black-america/>.

210. *FCC Media-Ownership Report Underscores the Agency's Historical Exclusion of Black People*, FREE PRESS (Sept. 7, 2021), <https://www.freepress.net/news/press-releases/fcc-media-ownership-report-underscores-agencys-historical-exclusion-black>.

211. *Id.*

212. Brodsky & Hanley, *supra* note 208.

There is a serious concern that the latest *Prometheus* ruling could lead to a similar outcome in the coming years.²¹³ While the Court in *Prometheus* acknowledges the FCC's examination of evidence concerning minority ownership, it does not afford the issue proper importance.²¹⁴ The strong connection between media diversity and democracy is indispensable and a lack of "empirical or statistical data" is an insufficient reason to forego ownership restrictions given the likelihood of long-term repercussions from a concentrated and homogenous media environment.

The FCC has long understood that diverse ownership has a profound effect on diverse viewpoints. During the Johnson Administration, the Kerner Commission was a group mandated to uncover the causes of civil unrest in 1967 and social conditions which foment riots.²¹⁵ The group found that television coverage gave the impression that the riots were confrontations between African Americans and whites, rather than the responses of African Americans to underlying "slum problems."²¹⁶ A separate report published in 1977, "Window Dressing on the Set: Women and Minorities in Television," found that "[f]orty percent of the white children attributed their knowledge about how blacks look, talk, and dress to television . . ."²¹⁷ These anecdotes and more underpin the notion that the diversity of broadcasters directly impacts the content that is produced and consumed by viewers. If the FCC directed its public interest regulation solely at consumers, it would be unable to achieve its objectives.

V. REVIVING THE PUBLIC INTEREST STANDARD

When Commissioner Ajit Pai repealed the 1975 Newspaper/Broadcast Cross-Ownership Rule in 2017, he claimed to promote the broadcasting industry's interests. In his order, Pai stated "By ending this entirely arbitrary test, we allow efficient combinations that can help television stations thrive."²¹⁸ Similarly, industry petitions in *Prometheus* claimed that Congress intended "competition to play a starring role, not second fiddle, in regulatory reform

213. *Supreme Court Awards the FCC for Long Neglecting Its Mandate to Promote Media Diversity in the United States*, FREE PRESS (Apr. 1, 2021), <https://www.freepress.net/news/press-releases/supreme-court-awards-fcc-long-neglecting-its-mandate-promote-media-diversity>.

214. *Prometheus*, 141 S. Ct. at 1160.

215. NAT'L ADVISORY COMM'N, REPORT OF THE NATIONAL ADVISORY COMMISSION ON CIVIL DISORDERS (1967).

216. *Id.* at 204.

217. U.S. COMM'N ON CIVIL RIGHTS, WINDOW DRESSING ON THE SET: WOMEN AND MINORITIES IN TELEVISION 46 (1977).

218. 2017 Order, *supra* note 174.

reviews.”²¹⁹ However, when the original rules were created, the Commission specifically stated that between its twin goals of viewpoint diversity and economic competition, viewpoint diversity was the “higher” policy.²²⁰ In its creation of the Radio/Television Cross-Ownership Rule in 1970, the Commission likewise said that the “principal purpose” was “promot[ing] diversity of viewpoints” and a secondary purpose is “promot[ing] competition.”²²¹ Today, the media industry and regulators seem to have forgotten the Commission’s mandate to serve the public interest.

If an agency is mandated to promote competition in parallel with sociopolitical factors such as diversity and localism, it cannot coherently do so without some acknowledgement of fairness. In the words of Professor Sandra Marco Colino: “It makes little sense to defend a competition policy that develops with its back purposefully turned to the attainment of moral and social justice.”²²² Unlike the consumer welfare standard, a competition policy involving fairness goes beyond a competitive playing field that exists only for efficient competitors. For instance, an interpretation that acknowledges fairness—rather than unfettered competition—would appreciate the historic disadvantages faced by minority broadcasters and their value to the public interest.²²³

Today, the biggest proponents of reincorporating fairness into competition policy and putting away the consumer welfare standard are Neo Brandeisians. This group, including members such as National Economic Advisor Tim Wu and FTC Chairwoman Lina Khan, advocates for a return to the “protection of competition” by focusing on structures and processes, rather than outcomes.²²⁴ Unlike the Chicago School, the Neo Brandeisians reject the promise of market forces and advocate for government law and policy to protect markets from being captured by private concentrations of

219. Reply Brief for Industry Petitioners, *supra* note 61, at 4.

220. Amendment of Sections 73.34, 73.249, and 73.636 of the Commission’s Rules Relating to Multiple Ownership of Standard, FM, and Television Broadcast Stations, 50 F.C.C.2d 1046, 1074 (1975).

221. *Id.*; see also Multiple Ownership of Standard, FM and TV Broadcast Stations, 22 F.C.C.2d 306, 313, ¶ 25 (1970) (stating that the “principal purpose” of the Radio/Television Cross-Ownership Rule is “promot[ing] diversity of viewpoints” and a secondary purpose is “promot[ing] competition”).

222. Sandra Marco Colino, *The Antitrust F Word: Fairness Considerations in Competition Law* 18 (Chinese Univ. of H.K. Faculty of Law Research Paper No. 2018-09, 2019), <https://ssrn.com/abstract=3245865>.

223. *Id.*

224. Lina Khan, *The New Brandeis Movement: America’s Antimonopoly Debate*, 9 J. EUR. COMPETITION L. & PRAC. 131, 132 (2018).

power.²²⁵ As Khan wrote in her seminal *Amazon's Antitrust Paradox* article, “[w]e cannot cognize the potential harms to competition posed . . . if we measure competition primarily through price and output.”²²⁶

By focusing on structure and process, the Neo Brandeisians seek to promote a system that eliminates abuses against competition. The approach follows Justice Brandeis’ concern with distinguishing behaviors (a merger or conduct) that promote the process of competition and behaviors that suppress or even destroy competition and encourage concentrated ownership.²²⁷ By maintaining sociopolitical considerations, this approach would protect competition and advance fairness using existing analyses.²²⁸ Some consumer welfare proponents claim that Neo Brandeisian analysis foregoes economics and could potentially overcorrect with “form-based” political interference to maximize democracy.²²⁹ Other critics argue that Neo Brandeisians propose a non-administrable system with no objective principles and many competing interests.²³⁰ For instance, Michigan State Law Professor Adam Candeub has argued that the FCC’s regulations have failed because they have applied antitrust law to the “marketplace of ideas.”²³¹ As a result, this system is criticized for “confus[ing] social and economic goals, creating an incoherent regulatory standard ripe for judicial reversal.”²³²

If fairness is incorporated into future analysis for competition, the Commission will need to deprioritize efficiency and economic competition. Unlike the DOJ and FTC, which focus on economic competition broadly, the FCC has a narrow, specific mandate to regulate communications. The original ownership rules should be maintained due to both the dearth of new entrants into the broadcasting industry and, particularly, the lack of diverse ownership. However, the ownership rules can be relaxed in small to mid-sized markets where there is substantial evidence of market failure. If a local market lacks the conditions for multiple broadcasters to compete for revenue or viewership, the FCC should ensure that consumers have access to quality information at the expense of diversity ownership. In these situations, diversity ownership is unlikely to be achieved regardless.

225. *Id.*

226. Lina M. Khan, Note, *Amazon's Antitrust Paradox*, 126 YALE L.J. 710 (2016).

227. *Chi. Bd. of Trade v. United States*, 246 U.S. 231, 238 (1918).

228. *See* Wu, *supra* note 108, at 11.

229. *See* Crane, *supra* note 110, at 4.

230. *See id.* at 3–5.

231. Adam Candeub, *Media Ownership Regulation, the First Amendment, and Democracy's Future*, 41 U.C. DAVIS L. REV. 1547 (2008).

232. *Id.*

Ownership restrictions are only one out of many ways to promote the public interest. While blocking mergers may be the best approach against media concentration, advocacy groups should explore news ways to promote the ideals of diversity, localism, and competition in today's contemporary media environment. The vague public interest standard has devolved beyond its original meaning and intent, and Congress should reconsider the current direction of media regulation. By updating the Communications Act, the legislature can reinvigorate America's commitment to its citizens to provide valuable, civic-minded information. Further, internet platforms have gained outsized influence in the media production industry since the 1990s. Internet companies do not face any of the requirements that broadcasters are beholden to. While the broadcasting industry may view this as a good reason to deregulate all media, media advocacy groups should push for more stringent compliance from internet content providers toward the ends of promoting the public interest.

Finally, the FCC should assuage the fears that proponents of deregulation have raised over the years and investigate them further. For instance, some critics of the public interest standard have argued that despite the FCC's intentions, ownership diversity will have little impact on the public interest because evidence suggests that media content is driven by demand (i.e., consumers) rather than supply (i.e., owners).²³³ These findings, however, go against countervailing evidence such as the Kerner Commission report.²³⁴ Others, such as the National Association of Broadcasters (NAB), claim that the FCC fails to account for the fact that broadcasters now compete with giant technology companies for advertising revenue while bearing high capital and operating costs.²³⁵ Outdated rules, the NAB says, "no longer enable broadcasters to viably operate in a competitive market or effectively serve the public interest."²³⁶ The FCC should invest more resources toward surveys that would gather adequate data on how Americans consume their information and what type of information they consume.

233. Matthew Gentzkow & Jesse M. Shapiro, *What Drives Media Slant? Evidence from U.S. Daily Newspapers*, 78 *ECONOMETRICA* 35, 38 (2010) (finding "little evidence that the identity of a newspaper's owner affects its slant").

234. See NAT'L ADVISORY COMM'N, *supra* note 215.

235. *Media Ownership Rules Are Detrimental to Competition, Loyalism, and Diversity*, NAB Says, NAB (Sept. 3, 2021), <https://www.nab.org/documents/newsRoom/pressRelease.asp?id=6190>.

236. *Id.*

VI. CONCLUSION

Since this country's founding, the media has been a core institution of an American democracy. Media regulation, accordingly, has been a critical function of democratic governance. Thus, the American experiment has relied on access to an egalitarian media structure where citizens have the ability for self-determination and self-governance. As the media industry became more complex, America's political leaders never abandoned these ideals. The Communications Act of 1933 established the public interest standard to protect against concentrated ownership and promote diversity, localism, and competition. For the last ninety years, the FCC has followed this mandate to balance commercial development and democratic values.

Until the 1980s, the consensus in Washington upheld the public interest as initially intended: the diversity of viewpoints took precedent over economic competition. But as political forces changed and market mechanisms won over, competition rose to the center stage. The new competition doctrine gaining prominence at the time was different from how competition was first envisioned during the turn of the 20th century. Efficiency and econometrics left little room for fairness or sociopolitical factors, such as diversity or localism. Both Democrats and Republicans adopted efficiency policies and worked to deregulate the media industry. The result has led to less protections against corporate concentration, and likely, a less-representative media environment for America's citizenry.

To correct this trend, Congress and the FCC should remember the public interest standard's democratic roots. Technocrats will be disappointed with any policy that seeks to maximize an intangible social value. True, the public interest is an intangible and incalculable social value, amenable to multiple competing, or even conflicting, interpretations. However, in the context of ownership, the media industry should broadly reflect the country. Some changes—such as revitalizing the notion of fairness in competition doctrine—may require insurmountable shifts in legal doctrine. Others, including more abundant and precise data collection on the FCC's part and public interest standards for internet platforms, could catalyze a movement for gradual reform. Regardless of the means, a media industry with diverse owners would mean that the content and direction of broadcasting serves the interests and needs of all Americans, not only those who are profitable and privileged.

