

**NEW YORK UNIVERSITY
ANNUAL SURVEY
OF AMERICAN LAW**

**VOLUME 68
ISSUE 2**

NEW YORK UNIVERSITY SCHOOL OF LAW
ARTHUR T. VANDERBILT HALL
Washington Square
New York City

New York University Annual Survey of American Law
is in its seventy-first year of publication.

L.C. Cat. Card No.: 46-30523
ISSN 0066-4413
All Rights Reserved

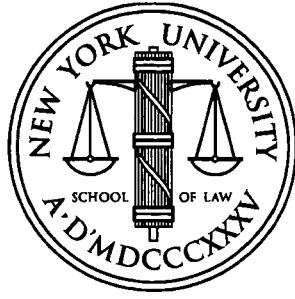
New York University Annual Survey of American Law is published quarterly at 110 West 3rd Street, New York, New York 10012. Subscription price: \$30.00 per year (plus \$4.00 for foreign mailing). Single issues are available at \$16.00 per issue (plus \$1.00 for foreign mailing). For regular subscriptions or single issues, contact the *Annual Survey* editorial office. Back issues may be ordered directly from William S. Hein & Co., Inc., by mail (1285 Main St., Buffalo, NY 14209-1987), phone (800-828-7571), fax (716-883-8100), or email (order@wshein.com). Back issues are also available in PDF format through HeinOnline (<http://heinonline.org>).

All articles copyright © 2012 by the *New York University Annual Survey of American Law*, except when otherwise expressly indicated. For permission to reprint an article or any portion thereof, please address your written request to the *New York University Annual Survey of American Law*.

Copyright: Except as otherwise provided, the author of each article in this issue has granted permission for copies of that article to be made for classroom use, provided that: (1) copies are distributed to students at or below cost; (2) the author and journal are identified on each copy; and (3) proper notice of copyright is affixed to each copy.

Manuscripts: The *Annual Survey* invites the submission of unsolicited manuscripts. Text and citations should conform to the 19th edition of *A Uniform System of Citation*. Please enclose an envelope with return postage if you would like your manuscript returned after consideration.

Editorial Office: 110 West 3rd Street, New York, N.Y. 10012
(212) 998-6540
(212) 995-4032 Fax
<http://www.law.nyu.edu/pubs/annualsurvey>



*For what avail the plough or sail
Or land or life, if freedom fail?*

EMERSON

NEW YORK UNIVERSITY ANNUAL SURVEY OF AMERICAN LAW

2012-2013 BOARD OF EDITORS

Editor-in-Chief
THEODORE KELLY

Managing Editors
JEFFREY R. BENDEL
ANDREW OSARCHUK

Development Editors
JOSEPH M. ENO
CARLY WEINREB

Note Editors
PAUL CHAFFIN
STEPHEN ELKIND

Executive Article Editors
MAX KAPLAN
DAVID KATZ
CLIFFORD MARK LANEY
ROSE S. LEE

KEVIN EUNHO PARK
KELLY STEPHANIE ROCHESTER PARKER
EMILY WINOGRAD
DENNY WON

Senior Articles Editor
E. DANIELLE BYAM

Symposium Editor
JASON SCHEFF

Article Editors

JOSE ACOSTA
TAYLOR ALEXANDER
NICHOLAS AXELROD
AYAL BAINVOILL
MADELEINE EDDY BARENHOLTZ
ASHLEY PARKS BELTON
JEREMY H. ERSHOW
KEVIN FRICK
ANDREW J. GERSHON
GANARAJ S. HEGDE
LAUREN M. ISAACSON

ELIZABETH B. JORDAN
ERIC KAMERMAN
LORETTA LAU
WILLIAM CLAYTON LAWRENCE
JUDD LINDENFELD
MEGAN C. MANFRED
AVA McALPIN
JAKE ORNSTEIN
MATTHEW ROTBART
JESSE B. SACKS
PAUL L. SANDLER

SOPHIE SAVRYN
P. DAVID SEIBERT
ELANA BROWNSTEIN SIEGEL
SARAH E. SIEGEL
YOAV SIMCHONI
JENNA E. SMALL
PAIGE CATHLEEN SPENCER
RIKKI STERN
CHARLES TANENBAUM
JORDAN WELLS
KARI WOHLSCHEGEL

Staff Editors

SIOBHAN C. ATKINS
PAUL S. BALIK
GABRIEL BRUNSWICK
MARY BRUST
MICHAEL A. CANENCIA
JOSEPH M. CASTELLI
WALTER A. CIACCI
ILYSSA L. COGHLAN
JOSHUA S. COHN
GRAHAM B. COLE
MATTHEW WARREN DEEVERS
DANIELLE ELIZABETH DEBOLD
JOSHUA M. DRAPEKIN
NICOLE A. ESCOBAR
MEAGAN L. FROEMMING
JULIAN J. GINOS
DAVID A. GIROUX

MICHAEL F. GOON
ALEX J. GORMAN
JORDANA L. HAVIV
PETER HUR
GAIL L. HYMAN
NICOLE N. IDOKO
CHRISTINE A. KUYEKE
HAROLD LAIDLAW
CHARLES MICHAEL LUPICA
LEAH S. MARTIN
CHIP MCCORKLE
ERIC A. MINTZ
BENJAMIN E. NOTTERMAN
GIL OFIR
JANE PARK

JULIA C. PILCER
JEFFREY A. RITHOLTZ
ARYEH L. ROSKIES
THEODORE D. SAMETS
ALEXANDRA SAMOWITZ
NICHOLAS C. SCHOLTEN
MELISSA B. SIEGEL
AMANDEEP SINGH
CHRISTOPHER BRODERICK SMITH
NAOMI R. SOSNER
GREGORY SPRINGSTED
MAX TIERMAN
PATRICK R. TOTARO
Yael TZIPORI
PETER VAN VALKENBURCH
DANIEL WOLFF
KYUNG EUN K. WON

SUMMARY OF CONTENTS

- RESTRUCTURING REGULATORY REVIEW OF ENDOCRINE-
DISRUPTING CHEMICALS UNDER CALIFORNIA'S PROPOSITION
65: LESSONS FROM THE REVIEW OF BPA
Rachael Rawlins 185
- WHEN BATHTUB CROCODILES ATTACK: THE TIMING AND
PROPRIETY OF CAMPAIGNING BY JUDICIAL RETENTION
ELECTION CANDIDATES
David W. Earley 239
- SCALING THE PATENT SYSTEM
Christina Mulligan & Timothy B. Lee 289
- DECODING FIRST AMENDMENT COVERAGE OF COMPUTER
SOURCE CODE IN THE AGE OF YOUTUBE, FACEBOOK, AND
THE ARAB SPRING
Jorge R. Roig 319
- THE BOUNDARY BETWEEN ZEALOUS ADVOCACY AND
OBSTRUCTION OF JUSTICE AFTER SARBANES-OXLEY
Greta Fails 397
- INTO THE PERFECT STORM: THE FAILURE OF TRUSTEE
ACTIONS AGAINST THIRD PARTIES
Nick Axelrod 441

RESTRUCTURING REGULATORY REVIEW OF ENDOCRINE-DISRUPTING CHEMICALS UNDER CALIFORNIA’S PROPOSITION 65: LESSONS FROM THE REVIEW OF BPA

RACHAEL RAWLINS*

ABSTRACT

This article proposes a redesign of the regulatory process, especially as it relates to the review of endocrine-disrupting chemicals, under California’s Proposition 65—a “public right to know law” of national significance. Drawing lessons from the Proposition 65 review of bisphenol A (BPA), this article proposes a redesign of the chemical listing process that would require the regulatory agency to adopt rules and require findings of fact to increase transparency and accountability. In the face of significant advocacy science fueling well-represented industry opposition, and without full disclosure of conflicts of interest, the current regulatory framework in California assigns a mountain of review work to an inadequately specialized, part-time committee. With no clear standards and little time, the committee is assigned mixed questions of law and science where significant policy decisions are quietly hidden behind purportedly scientific conclusions. Rules are needed to increase transparency by creating an honest demarcation between policy and science so that the public may take action as necessary to further public policy objectives. Rules are also needed to set standards by which to critically evaluate conflicts of interest and advocacy science, and to require that warning labels identify the specific chemical and potential exposure. This article proposes to open a public rulemaking process that would include highly trained and specialized scientists and ultimately create a more specialized review board.

Introduction	186	R
I. The Failing Regulatory System and Restoring the Promise of Proposition 65	193	R

* Faculty Fellow, Center for Sustainable Development Senior Lecturer, University of Texas. I would like to thank Dr. Frederick S. vom Saal and Professor Wendy Wagner for their foundational work, and their suggestions and comments on this article. I would also like to thank the editors at N.Y.U. Annual Survey of American Law, and especially Kelly Parker for her diligent and careful work.

A.	Proposition 65’s Potential to Unveil Dangerous Sources of Exposure	196	R
B.	The Evolving Role of the DART Identification Committee and the Review of Endocrine-Disrupting Chemicals	198	R
II.	Requirements for Transparency and Accountability ..	203	R
A.	Defining “Clearly Shown” Consistent with Scientific Standards and Societal Choices	204	R
B.	Defining “Reproductive Toxicity” Consistent with Advancing Science.....	209	R
III.	Research Design and Quality Standards to Assess and Evaluate Advocacy Science.....	216	R
IV.	A More Specialized Committee, Extended Review Period, and Clear Disclosure of Conflicts of Interest .	222	R
V.	Standards and Findings to Avoid Inappropriate Reliance on Unverified Claims of Adverse Consequences	228	R
Conclusion	233	R

INTRODUCTION

One in eight women will be diagnosed with breast cancer, twelve percent under the age of forty-four.¹ Breast cancer now strikes teens and tweens.² In 2009 a ten-year-old in California underwent a mastectomy,³ and in 2010 so did a four-year-old from Toronto.⁴ One third of adults⁵ and seventeen percent of all U.S.

1. See *SEER Stat Fact Sheets: Breast, Surveillance Epidemiology and End Results*, NAT’L. CANCER INST., <http://seer.cancer.gov/statfacts/html/breast.html> (last visited Dec. 27, 2012) (based on rates from 2005–2009).

2. Madison Park, *Tweens Challenged by Grown-Up Malady: Breast Cancer*, CNN HEALTH (Oct. 26, 2009, 9:38 AM), <http://www.cnn.com/2009/HEALTH/10/26/tweens.breast.cancer/index.html> (These cases are “extreme examples of a troubling trend emerging with breast cancer, medical experts say. Younger women are getting a disease that usually strikes around menopause—and no one knows why.”).

3. *Ten-Year-Old Bravely Battles Breast Cancer*, CBS NEWS (May 19, 2009, 10:32 AM), <http://www.cbsnews.com/stories/2009/05/19/earlyshow/health/main5024777.shtml>.

4. Greg McArthur, *Four Year Old Battles Breast Cancer*, THE GLOBE & MAIL, <http://www.theglobeandmail.com/life/health/four-year-old-battles-breast-cancer/article1704895/> (last updated Mar. 16, 2011, 11:43 AM).

5. *Overweight and Obesity: Adult Obesity Facts*, CTRS. FOR DISEASE CONTROL & PREVENTION, <http://www.cdc.gov/obesity/data/adult.html> (last updated Aug. 13, 2012).

children between the ages of two and nineteen are obese.⁶ One in six men will be diagnosed with prostate cancer in their lifetime.⁷ More than 30 million people in the United States have some type of thyroid dysfunction.⁸ Endocrine-disrupting chemicals may be contributing to these stunning statistics. Chemicals that interfere with endocrine function have been found to affect male and female reproduction, neuroendocrinology, thyroid function, metabolism and obesity, breast development, breast cancer, prostate cancer, and cardiovascular endocrinology.⁹

The Endocrine Society, the world's oldest, largest, and most active organization devoted to research on hormones and the clinical practice of endocrinology, recently issued a statement expressing concern that the public may be at risk because critical information about potential health effects of endocrine-disrupting chemicals is being overlooked in the development of federal health and safety guidelines and regulations.¹⁰ The current federal and state regulatory regimes are struggling in their attempts to deal appropriately with these chemicals. One such chemical, the herbicide Atrazine, effectively banned in Europe, is exported to the United States in massive quantities.¹¹ Tyrone Hayes, a professor at UC Berkeley, lecturing on the chemically castrating effects of the Atrazine on frogs, including female eggs growing in male testes, described EPA's stunning response—it was “unclear” as to whether this reproductive mutation qualified as an “adverse effect!”¹²

6. *Overweight and Obesity: Childhood Obesity Facts Data and Statistics*, CTRS. FOR DISEASE CONTROL & PREVENTION, <http://www.cdc.gov/obesity/data/childhood.html> (last updated Dec. 21, 2012).

7. *What Are the Key Statistics About Prostate Cancer?*, AM. CANCER SOC'Y, <http://www.cancer.org/Cancer/ProstateCancer/DetailedGuide/prostate-cancer-key-statistics> (last updated Dec. 14, 2012).

8. Marilyn Fuller Delong, *Thyroid Dysfunction (Course # 8431)*, CME RES., 3 (July 11, 2012), http://www.netce.com/582/Course_8431.pdf.

9. Evanthia Diamanti-Kandarakis et al., *Endocrine-Disrupting Chemicals: An Endocrine Society Scientific Statement*, 30 ENDOCRINE REVS. 293, 293 (2009), available at http://www.endo-society.org/journals/scientificstatements/upload/edc_scientific_statement.pdf.

10. See THE ENDOCRINE SOC'Y, POSITION STATEMENT: ENDOCRINE-DISRUPTING CHEMICALS (2009), available at <http://www.endo-society.org/advocacy/policy/upload/Endocrine-disrupting-chemicals-position-statement.pdf>.

11. See Tyrone Hayes, *What is Atrazine? And Why Do We Love It?*, OUR WORLD . . . OUR FUTURE, <http://atrazinelovers.com/ml.html> (last visited Dec. 27, 2012).

12. Tyrone Hayes, Professor of Integrative Biology at Univ. of Cal., Berkeley, From Silent Spring to Silent Night, Lecture at The Univ. of Tex. at Austin Envtl. Sci. Inst., PowerPoint Slide 17 (Notes) (Jan. 30, 2009), available at <http://www>.

Another endocrine-disrupting chemical similarly muddled and mired in politics is bisphenol A (BPA), a synthetic estrogen.¹³ The alarm bells have been sounding on BPA for quite some time. In 2006, an expert panel sponsored by the National Institutes of Health, the EPA, and Commonweal (a non-profit health and environmental research group) concluded that people are exposed to BPA at levels that cause problems in wildlife and laboratory animals, and that there is “great cause for concern” with regard to the potential for similar adverse effects in humans.¹⁴ The panel explained that recent trends in human diseases relate to adverse effects observed in experimental animals exposed to low doses of BPA.¹⁵ As specific examples, the panel noted the increase in hormonally mediated cancers, such as prostate and breast cancer; urogenital abnormalities in male babies; a decline in semen quality in men; early onset of puberty in girls; an increase in metabolic disorders, including insulin resistant (type 2) diabetes and obesity; and increases in neurobehavioral problems such as attention deficit hyperactivity disorder (ADHD).¹⁶

Despite the evidence, the Food and Drug Administration (FDA) has fallen into a state of regulatory malaise. In its review of BPA, the FDA’s Science Advisory Board, although operating under a standard where “safety” is defined as “reasonable certainty in the minds of competent scientists that the additive is not harmful to man or animal,”¹⁷ ultimately deferred instead to a political decision. The Science Board cautiously concluded:

Coupling together the available qualitative and quantitative information (including application of uncertainty factors) provides a sufficient scientific basis to conclude that the Margins of Safety defined by FDA as “adequate” are, in fact, inadequate.

esi.utexas.edu/k-12-a-the-community/hot-science-cool-talks/lecture-archives/from-silent-spring-to-silent-night.

13. BPA mimics the activity of estradiol and is similar in structure and efficacy to the estrogenic drug diethylstilbestrol (DES). Frederick vom Saal & Wade Welshons, *Large Effects from Small Exposures. II. The Importance of Positive Controls in Low-Dose Research on Bisphenol A*, 100 ENVTL. RES. 50, 50 (2006).

14. Frederick vom Saal et al., *Chapel Hill Bisphenol A Expert Panel Consensus Statement: Integration of Mechanisms, Effects in Animals and Potential to Impact Human Health at Current Levels of Exposure*, 24 REPROD. TOXICOLOGY 131, 131, 136 (2007) [hereinafter *Chapel Hill*].

15. *Id.* at 131.

16. *Id.*

17. Food Additives Amendment of 1958, Pub. L. No. 85-929, 72 Stat. 1784; S. REP. NO. 85-2422, at 2–3 (1958), *reprinted in* 1958 U.S.C.C.A.N. 5300, 5301–02 (describing the relevant standard as one of “reasonable certainty in the minds of competent scientists that the additive is not harmful to man or animal”).

This does not mean that the potential exposures are not “*acceptable*”. The latter is the subject of policy that appropriately rests with the Commissioner of the FDA.¹⁸

The FDA itself, although admitting “some concern” about the potential low-dose effects of BPA on the brain, behavior, and prostate gland in fetuses, infants and young children, has thus far deferred any significant regulatory response.¹⁹ Declining a petition by the Natural Resources Defense Council (NRDC) to initiate a rulemaking to prohibit the use of BPA in food and food packaging, the FDA determined, “as a matter of science and regulatory policy,” that the best course of action is to continue its review and study of emerging data on BPA.²⁰ In response to a request from the American Chemistry Council, the FDA is amending the food additive regulations to remove authorization for polycarbonate resins²¹ in baby bottles and spill-proof cups, but this action is based on abandonment (following movement in the retail market²²), not any finding concerning safety.²³

Although acknowledging that “there are still a lot of outstanding questions,” Linda S. Birnbaum, Director of the National Institute of Environmental Health Sciences at the National Institutes of Health recently reported: “Our grantees have published nearly 100 papers [on BPA] since January 2010. Nothing has been published

18. FDA SCI. BD. SUBCOMM. ON BISPHEENOL A, SCIENTIFIC PEER-REVIEW OF THE DRAFT ASSESSMENT OF BISPHEENOL A FOR USE IN FOOD CONTACT APPLICATIONS 4 (2008) (emphasis in original), *available at* <http://www.fda.gov/ohrms/dockets/ac/08/briefing/2008-4386b1-24.pdf> (as modified and submitted from the Science Board to the FDA).

19. *Bisphenol A (BPA): Use in Food Contact Application*, FDA, <http://www.fda.gov/NewsEvents/PublicHealthFocus/ucm064437.htm> (last updated Apr. 2, 2012) (“FDA is continuing to consider the low dose toxicity studies of BPA as well as other recent peer-reviewed studies related to BPA.”).

20. Letter from David Dorsey, Acting Assoc. Comm’r for Policy & Planning, FDA, to Sarah Janssen & Aaron Colangelo, Nat’l Res. Def. Council 15 (Mar. 30, 2010), *available at* <http://www.regulations.gov/#!documentDetail;D=FDA-2008-P-0577-0007> (finding the data presented in the petition insufficient to initiate rulemaking).

21. Polycarbonate resins are formed by the condensation of 4,4'-isopropyl enediphenol (i.e., Bisphenol A (BPA)), and carbonyl chloride or diphenyl carbonate. Indirect Food Additives: Polymers, 77 Fed. Reg. 41,899, 41,902 (July 17, 2012) (to be codified at 21 C.F.R. pt. 177), *available at* <https://www.federalregister.gov/articles/2012/07/17/2012-17366/indirect-food-additives-polymers#p-3>.

22. Ylan Q. Mui, *Wal-Mart to Pull Bottles Made with Chemical BPA*, WASH. POST, Apr. 18, 2008, <http://www.washingtonpost.com/wp-dyn/content/article/2008/04/17/AR2008041704205.html>.

23. Indirect Food Additives: Polymers, 77 Fed. Reg. at 41,900–01.

that says there isn't any problem here."²⁴ Meanwhile exposure to BPA, already detected in 92.6% of persons in the 2003/2004 U.S. National Health and Nutrition Examination Survey,²⁵ is on the rise. Chemists first created polycarbonate from BPA in 1952.²⁶ It is now a high-volume chemical present in a many products including the interior coating of food cans, wine storage vats, water carboys, milk containers, food storage vessels, baby formula bottles, water pipes, dental materials, automotive lenses, optical lenses, protective window glazing, compact discs, thermal paper, paper coatings, and dyes.²⁷ In the United States, production quantities increased from 521 million kilograms in 1990 to 736 million kilograms in 1995.²⁸ Estimated production in the United States in 2007 was one billion kilograms.²⁹

Several foreign, state, and local governments have taken action on BPA. In October 2008, Canada added BPA to its toxic substance list.³⁰ Since that time, the European Commission,³¹ the French Na-

24. Bettina Boxall & Eryn Brown, *FDA Decides Not to Ban BPA in Food Packaging*, L.A. TIMES, Mar. 30, 2012, <http://articles.latimes.com/2012/mar/30/nation/la-na-bpa-fda-20120331>.

25. See REPROD. & CANCER HAZARD ASSESSMENT BRANCH, OFFICE OF ENVTL. HEALTH HAZARD ASSESSMENT, CAL. ENVTL. PROT. AGENCY, EVIDENCE ON THE DEVTL. AND REPROD. TOXICITY OF BISPHENOL A 21 (2009) [hereinafter CAL. ENVTL. PROT. AGENCY], available at http://oehha.ca.gov/prop65/CRNR_notices/state_listing/data_callin/pdf/BPA050109.pdf (measured as urinary BPA).

26. Vom Saal & Welshons, *supra* note 13, at 51.

27. Ana M. Sota et al., *Does Breast Cancer Start in the Womb?*, 102 BASIC & CLINICAL PHARMACOLOGY & TOXICOLOGY 125, 127 (2008).

28. ENV'T CAN. & HEALTH CAN., CHEMICAL ABSTRACTS SERVICE REGISTRY NUMBER 80-05-7, SCREENING ASSESSMENT FOR THE CHALLENGE, PHENOL, 4,4'-(1-METHYLETHYLIDENE)BIS-(BISPHENOL A) at i (2008), available at http://www.ec.gc.ca/substances/ese/eng/challenge/batch2/batch2_80-05-7_en.pdf.

29. *Id.*

30. *Id.* at 76; see also *Chemical Substances: Bisphenol A*, GOV'T OF CAN., <http://www.chemicalsubstanceschimiques.gc.ca/challenge-defi/batch-lot-2/bisphenol-a/index-eng.php> (last updated Apr. 12, 2012). The *Canadian Environmental Protection Act (CEPA)* defines "toxic" substances as those that enter or may enter the environment at levels or conditions that have or may have a harmful effect on the environment; are or could be dangerous to the environment on which life depends; or are or could be dangerous to human life or health. Before the government can regulate these substances, they have to be added to the List of Toxic Substances. *Chemical Substances: The Canadian Environmental Protection Act, 1999*, GOV'T OF CAN., <http://www.chemicalsubstanceschimiques.gc.ca/about-apropos/cepa-lcpe-eng.php> (last updated Mar. 20, 2012).

31. Liz Szabo, *Europe Votes to Ban Chemical from Baby Bottles*, USA TODAY, http://www.usatoday.com/news/world/2010-11-29-BPA29_ST_N.htm (last updated Nov. 29, 2010).

tional Assembly,³² the United Arab Emirates,³³ and China's Ministry of Health³⁴ have all taken action to place restrictions on the use of BPA. In the United States, some state and local governments have passed laws banning BPA in beverage containers for young children.³⁵ Connecticut has gone even further and banned the use of BPA in reusable food and beverage containers.³⁶ Massachusetts issued an advisory against the use of such products for small children, pregnant women, and breast-feeding mothers, and additional legislation has been under consideration in several states.³⁷

In an action that may have significant consequences for the nation as a whole, BPA is currently under review pursuant to Proposition 65 by California's Office of Environmental Health Hazard Assessment (OEHHA).³⁸ Proposition 65, a "public right to know law" adopted by California voters in 1986,³⁹ requires warning labels on consumer products that contain certain chemicals identified as either carcinogens or reproductive toxicants.⁴⁰ Proposition 65 has

32. *Les députés votent l'interdiction du bisphénol A dans les emballages alimentaires*, LE MONDE, Oct. 12, 2011, http://www.lemonde.fr/planete/article/2011/10/12/les-deputes-votent-l-interdiction-du-bisphenol-a-dans-les-emballages-alimentaires_1586413_3244.html, translated in Abdelfattah60, *MEPs Vote to Ban BPA in Food Packaging*, THE PACKAGING SAFETY BLOG (Oct. 12, 2011), <http://safepackaging.eu/france-bans-bisphenol-food-contact-materials>.

33. *UAE to Ban BPA Baby Bottles*, EMIRATES 24—7 (Nov. 28, 2010), <http://www.emirates247.com/news/emirates/uae-to-ban-bpa-baby-bottles-2010-11-28-1.322183>.

34. Lu Feiran, *China to Ban Plastic Bottles to Feed Babies*, SHANGHAI DAILY.COM (Mar. 5, 2011), <http://www.shanghaidaily.com/nsp/National/2011/03/05/China%2Bto%2Bban%2Bplastic%2Bbottles%2Bto%2Bfeed%2Bbabies>.

35. Mark N. Duvall & Russell N. Fraker, *Bisphenol A: A Hot Topic at FDA, EPA, States, and the Courts*, CLIENT ALERT (Beveridge & Diamond, P.C., Wash. D.C.), Feb. 19, 2010, available at <http://www.bdlaw.com/assets/attachments/BD%20Client%20Alert%20-%20BPA%20A%20Hot%20Topic%20at%20EPA%20FDA%20States%20and%20Courts.pdf> (including Minnesota, Connecticut, the City of Chicago, and Suffolk, Albany and Schenectady Counties in NY).

36. *Id.*

37. *Id.*

38. Request for Relevant Information on a Chemical Being Considered for Listing by the Authoritative Bodies Mechanism: Bisphenol-A, 7-Z Cal. Regulatory Notice Reg. 252, 252–53 (Feb. 12, 2010), available at http://oehha.ca.gov/prop65/CRNR_notices/admin_listing/requests_info/callinBPA021210.html.

39. Clifford Rechtschaffen, *CPR Perspective: The Public Right to Know*, CENTER FOR PROGRESSIVE REFORM, <http://www.progressivereform.org/perspright.cfm> (last visited Dec. 28, 2012); *Frequently Asked Questions About Proposition 65*, CAL. OFFICE OF ENVTL. HEALTH HAZARD ASSESSMENT (OEHHA), <http://oehha.ca.gov/prop65/p65faq.html> (last visited Dec. 28, 2012).

40. See CAL. HEALTH & SAFETY CODE § 25249.6 (West 2006). Warnings are generally required unless the chemical is present in the product below a level that

been credited with stimulating significant consumer product reformulation, which, in some cases, has been close to industry-wide with a nationwide effect.⁴¹ For those products that remain on the market, current regulatory standards fall short of requiring full disclosure by allowing generic warning statements that fail to identify the chemical.⁴² With some regulatory adjustment, however, Proposition 65 has the potential to publicly expose products containing BPA that the federal government has thus far been unwilling or unable to identify.

Although known as the most ambitious attempt by any state to regulate hazardous chemical exposure through information dissemination,⁴³ as it stands today, Proposition 65 is failing to live up to its full regulatory potential. As previously proposed by Clifford Rechtschaffen, OEHHA should require labels to specifically expose the chemical and the source of exposure.⁴⁴ Reform is also needed in other areas. Difficulties have emerged in determining, defining, and communicating appropriate standards against which to access research, in dealing with conflicts of interest, in working across different regulatory regimes, in handling inappropriate and inadequate back-door cost benefits analysis, as well as responding to a proliferation of industry-sponsored studies creating an overwhelming environment of uncertainty. There are questions as to the role of scientists and advisory boards, appropriate standard setting, communication and implementation of those standards, and how to appropriately inform public opinion.

This article takes a close look at Proposition 65 in relation to the regulation of BPA and proposes a redesign of the regulatory process, especially as it relates to endocrine-disrupting chemicals. In the face of significant advocacy science fueling well-represented industry opposition, and without full disclosure of conflicts of inter-

poses “no significant risk,” that is, a level that causes no more than one excess lifetime case of cancer per 100,000 exposed individuals and, for reproductive toxicants, 1/1000th of the highest level at which the chemical has been shown to have no observable reproductive effect. CAL. HEALTH & SAFETY CODE § 25249.10(c) (West 2006); 27 CAL.CODE REGS. tit. 27, § 25703(b) (2012); *Proposition 65 in Plain Language!*, CAL. OFFICE OF ENVTL. HEALTH HAZARD ASSESSMENT (OEHHA), <http://oehha.ca.gov/prop65/background/p65plain.html> (last updated Mar. 2010). If a chemical is present in a product, the burden is on industry to show that it does not exceed the allowable level. See CAL. HEALTH & SAFETY CODE § 25249.10(c) (West 2006).

41. Clifford Rechtschaffen, *The Warning Game: Evaluating Warnings Under California's Proposition 65*, 23 *ECOLOGICAL L.Q.* 303, 341 (1996).

42. See *id.* at 363–64.

43. See *id.* at 305.

44. *Id.* at 363–64.

est, the current regulatory framework assigns a mountain of review work to an inadequately trained part-time committee. With no clear standards and little time, the committee is assigned mixed questions of law and science where significant policy decisions are quietly hidden behind purportedly scientific conclusions. Standards are particularly important in the context of endocrine-disrupting chemicals, where the science is rapidly evolving with increasing levels of complexity. Given the complexity of the science, a public rulemaking process should open discussion and allow for input from specialized scientists, as well as public consideration of the evolving policy issues. OEHHA should adopt standards to increase transparency and accountability, to expose conflicts of interest, and to require critical evaluation of research design.

The first section of this article discusses the nationwide importance of Proposition 65 in the context of our failing federal regulatory system. This section discusses the separate mechanisms for listing endocrine-disrupting chemicals under Proposition 65 and the increasing importance of review by the Developmental and Reproductive Toxicant (DART) Identification Committee. It also discusses the importance of regulatory reform to keep pace with our evolving understanding of endocrine-disrupting chemicals. The second section identifies the need for clear rules and mandatory fact findings to increase scientific transparency and accountability. This section discusses issues related to the “clear evidence” standard of review and the meaning of “reproductive toxicity” in the context of the evolving science. The third section identifies the need for research design and quality standards to weigh and effectively evaluate advocacy science. The fourth section identifies the need for a more specialized science review committee, an extended review period, and clear disclosure of conflicts of interest. Finally, the fifth section discusses the importance of standards to focus decisionmaking on appropriate scientific criteria, especially in the face of unsubstantiated claims of adverse consequences that may otherwise quietly threaten to disrupt the process.

I. THE FAILING REGULATORY SYSTEM AND RESTORING THE PROMISE OF PROPOSITION 65

It is clear that BPA is leaching from many products, including food and beverage packaging and containers, but the full range of

sources is unknown.⁴⁵ BPA is used in the production of epoxy resins and polycarbonate plastic, food and drink packaging, and resins used as lacquers to coat metal products such as food cans, bottle tops, and water supply pipes.⁴⁶ The European Union has identified wine as a significant source of exposure due to an epoxy resin used to line wine vats.⁴⁷ The coating on metal lids for glass jars and bottles has also been identified as a source of exposure.⁴⁸ BPA is used in commercial polyvinyl chloride (PVC) cling films and plastic sheeting bags,⁴⁹ certain “microwave-safe” containers, and frozen food packaging.⁵⁰ Other potentially important sources include sports and office cooler polycarbonate water bottles,⁵¹ credit card receipts (which reportedly have enormously high levels in an unbound form that may be transferred from fingers to food),⁵² and even building materials. One study reported significantly higher urinary levels of total BPA, along with significantly higher levels of

45. See *infra* text accompanying notes 49–52, 54–55.

46. CAL. ENVTL. PROT. AGENCY, *supra* note 25, at 10.

47. U.S. DEP’T OF HEALTH AND HUMAN SERVS. NAT’L TOXICOLOGY PROGRAM CTR. FOR THE EVALUATION OF RISKS TO HUMAN REPROD., NIH PUBL’N NO. 08-5994, NTP-CERHR MONOGRAPH ON THE POTENTIAL HUMAN REPRODUCTIVE AND DEVELOPMENTAL EFFECTS OF BISPHENOL A 4 (2008), *available at* <http://ntp.niehs.nih.gov/ntp/ohat/bisphenol/bisphenol.pdf>.

48. EUROPEAN FOOD SAFETY AUTH., OPINION OF THE SCIENTIFIC PANEL ON FOOD ADDITIVES, FLAVOURINGS, PROCESSING AIDS AND MATERIALS IN CONTACT WITH FOOD ON A REQUEST FROM THE COMMISSION RELATED TO 2,2-BIS(4-HYDROXY-PHENYL)PROPANE (BISPHENOL A), QUESTION NO. EFSA-Q-2005-100, 428 E.F.S.A. J. 1, 1 (Nov. 29, 2006), *available at* <http://www.efsa.europa.eu/en/scdocs/doc/s428.pdf>.

49. *Id.* at 10.

50. Siel Ju, *Plastic in the Microwave: Is BPA in Your Frozen Dinner?*, MOTHER NATURE NETWORK (Feb. 8, 2010, 3:14 PM), <http://www.mnn.com/health/fitness-well-being/blogs/plastic-in-the-microwave-is-bpa-in-your-frozen-dinner> (citing a test done by the Milwaukee Journal Sentinel finding that BPA leached from “microwave-safe” plastics, ranging from frozen food trays to plastic baby food packaging, *available at* <http://media.jsonline.com/documents/BPAstudy.pdf>).

51. Hoa Le et al., *Bisphenol A is Released from Polycarbonate Drinking Bottles and Mimics the Neurotoxic Actions of Estrogen in Developing Cerebellar Neurons*, 176 TOXICOLOGY LETTERS 149 (2008); Jennifer Grayson, *How to Avoid the Sneakiest Sources of BPA*, WebMD (Feb. 9, 2010), <http://blogs.webmd.com/health-home/2010/02/how-to-avoid-the-sneakiest-sources-of-bpa.html>.

52. Janet Raloff, *Concerned About BPA: Check Your Receipts*, SCIENCE NEWS (Oct. 7, 2009), http://www.sciencenews.org/view/generic/id/48084/description/Concerned_about_BPA_Check_your_receipts (interviewing John Warner, a former professor of Green Chemistry at the University of Massachusetts, who compared the nanogram quantities of BPA leaching out of polycarbonate water bottles to that of the average cash register receipt, which is not bound into a polymer like the BPA in polycarbonates, but is instead “individual molecules loose and ready for uptake”).

follicle stimulating hormones and significantly lower levels of testosterone, in workers applying paint consisting of ten to thirty percent epoxy resins.⁵³

Although BPA has been approved for multiple uses as a food contact substance,⁵⁴ the FDA has not attempted to identify all of the different types of food products that may be contaminated with BPA.⁵⁵ The FDA has made some attempt to evaluate BPA for use in food contact applications, but it looked only at a small sample of canned products as a source of adult exposure and only at canned formula and polycarbonate baby bottles as a source of infant exposure.⁵⁶ Under the circumstances, perhaps the job was just too difficult. Interpreting its regulatory authority, the FDA explains:

Current BPA food contact uses were approved under food additive regulations issued more than 40 years ago. . . . Once a food additive is approved, any manufacturer of food or food packaging may use the food additive in accordance with the regulation. There is no requirement to notify FDA of that use. For example, today there exist hundreds of different formulations for BPA-containing epoxy linings, which have varying characteristics. As currently regulated, manufacturers are not required to disclose to FDA the existence or nature of these formulations. Furthermore, if FDA were to decide to revoke one or more approved uses, FDA would need to undertake what could be a lengthy process of rulemaking to accomplish this goal.⁵⁷

53. The painters had “follicle stimulating hormone levels of 7.68 international units, which was significantly higher than the non-painter mean of 5.53 international units,” and the painters had a “testosterone level of 3.5 nanograms per milliliter, which was . . . ‘significantly lower’ than the non-painter level of 5.818 nanograms per milliliter.” *Meeting on Proposition 65 Before the Devtl. & Reprod. Toxicant Identification Comm. of the Office of Env’tl. Health Hazard Assessment of the State of Cal.*, 150–51 (July 15, 2009) [hereinafter *DART Meeting on Proposition 65* (July 15, 2009)], available at http://oehha.ca.gov/prop65/public_meetings/pdf/DARTIC-Transcript71509.pdf (statement of Julie Silas, Director of Healthcare Projects for the Healthy Building Network, quoting the authors of the study); CAL. ENVTL. PROT. AGENCY, *supra* note 25, at A1-17 to -18.

54. *E.g.*, 21 C.F.R. § 177.1555 (2012); 21 C.F.R. § 177.1595 (2012).

55. FDA, DRAFT ASSESSMENT OF BISPHENOL A FOR USE IN FOOD CONTACT APPLICATIONS 6 (2008), available at http://www.fda.gov/ohrms/dockets/AC/08/briefing/2008-0038b1_01_02_FDA%20BPA%20Draft%20Assessment.pdf (“FDA does not maintain a list of all the specific product manufactured from BPA nor does it maintain a list of the various processors for the BPA-containing products . . .”).

56. The FDA relied on sixteen samples of canned food for adults and fourteen samples of canned formula for infants. *Id.* at 7–10.

57. *Bisphenol A (BPA): Use in Food Contact Application*, *supra* note 19.

A. Proposition 65's Potential to Unveil Dangerous Sources of Exposure

Proposition 65 has the potential to unveil not only contaminated food products, but also other potentially dangerous sources of BPA nationwide. Given the importance of the California market and the cost of selling different forms of the same product, businesses often choose to include informational warnings mandated by California law on products sold throughout the United States.⁵⁸ However, under the current rules, even if the listing of BPA as a reproductive toxin is finalized under Proposition 65, warnings may do little to lift the curtain on BPA. Proposition 65 requires “clear and reasonable” warnings before exposing consumers to listed chemicals.⁵⁹ However, Proposition 65 regulations establish the following “safe harbor” warning messages that have been used on virtually all consumer product warnings: “Warning: This product contains a chemical known to the State of California to cause cancer” or “Warning: This product contains a chemical known to the State of California to cause birth defects or other reproductive harm.”⁶⁰ The warning statement informs individuals only that the product contains a chemical, not whether the product will expose them to the chemical, not the identity of the chemical, nor the source of the exposure.⁶¹ Allowing such a generic warning statement falls short of Proposition 65’s goal of allowing for informed consent.⁶² The statutory preamble declares the people’s right “to be informed about exposures to chemicals that cause cancer, birth defects or other reproductive harm.”⁶³ The intent of voters was to “receive warnings which will enable them to make informed

58. ZYGMUNT J.B. PLATER ET. AL., ENVIRONMENTAL LAW AND POLICY: NATURE, LAW AND SOCIETY 539–40 (3d ed. 2004).

59. CAL. HEALTH & SAFETY CODE § 25249.6 (West 2006) (“No person in the course of doing business shall knowingly and intentionally expose any individual to a chemical known to the state to cause cancer or reproductive toxicity without first giving clear and reasonable warning to such individual . . .”).

60. See CAL. CODE REGS. tit. 27, § 25603.2 (2012).

61. See Rechtschaffen, *The Warning Game: Evaluating Warnings Under California's Proposition 65*, *supra* note 41, at 326 (identifying these issues and concluding that the rules should be reformed to improve disclosure).

62. See *id.* at 307, 318, 319 n.78, 363.

63. CAL. HEALTH & SAFETY CODE § 25249.5 hist. n. § 1(a) (West 2006); see also CAL. HEALTH & WELFARE AGENCY, REVISED FINAL STATEMENT OF REASONS FOR CAL. CODE REGS. tit. 22, § 12601 at 22 (1988) [hereinafter CAL. HEALTH & WELFARE AGENCY], available at http://www.oehha.org/prop65/law/pdf_zip/12601FSORNov1988.pdf (official filing accompanying regulatory amendment) (analogizing purpose of warning requirement to purpose of informed consent doctrine).

2012] ENDOCRINE-DISRUPTING CHEMICALS AND PROP 65 197

choices.”⁶⁴ The ballot argument read: “Proposition 65 also tells businesses: Don’t expose us to any [listed] chemicals without first giving us a clear warning. We each have a right to know, and to make our own choices about being exposed to these chemicals.”⁶⁵ For individuals particularly concerned about BPA, a generic warning statement that fails to disclose the specific chemical or the potential for exposure does not allow for that choice.

This issue is especially important to pregnant mothers, parents of young children, and other population groups who may be particularly vulnerable to BPA. As acknowledged by the FDA, infants are particularly sensitive to exposure to BPA because their neurological and endocrine systems are developing, and because their hepatic system for detoxification and elimination of such substances as BPA is immature.⁶⁶ Although not officially announced by the FDA, its Science Board also discussed the possibility that sensitive populations may include patients with hormone sensitive cancers, including breast cancer.⁶⁷ According to the Endocrine Society, the significant increase in the incidence of breast cancer in the industrialized world in the last fifty years may be due to exposure to hormonally active chemicals like BPA that have been released into the environment from industrial and commercial sources.⁶⁸ Evidence is also emerging that BPA may also pose a serious risk to the now 2.9 million breast cancer survivors in the United States⁶⁹ by interfering with tamoxifen and chemotherapy treatment.⁷⁰

64. CAL. HEALTH & WELFARE AGENCY, *supra* note 63, at 3–4, 43 (“The apparent purpose of any warning under the Act is to permit the persons exposed to make choices about the exposure.”).

65. *Id.* at 43 (emphasis omitted).

66. See *Bisphenol A (BPA): Use in Food Contact Application*, *supra* note 19.

67. See *Meeting Before Science Board Advisory Committee to the FDA* 291 (2008) [hereinafter *Meeting Before Science Board Advisory Committee*], available at <http://www.fda.gov/ohrms/dockets/ac/08/transcripts/2008-4386t1-03.pdf> (statement of Dr. David Parkinson, Member, Science Board Advisory Comm., FDA); see, e.g., Shanaz H. Dairkee et al., *Bisphenol A Induces a Profile of Tumor Aggressiveness in High-Risk Cells from Breast Cancer Patients*, 68 *CANCER RES.* 2076 (2008); see also Martin Mittelstaedt, *Bisphenol A Can Alter Genes, Study Finds*, THE GLOBE & MAIL, <http://www.theglobeandmail.com/life/bisphenol-a-can-alter-genes-study-finds/article671016/> (last updated Mar. 13, 2009, 11:45 AM).

68. See Diamanti-Kandarakis et al., *supra* note 9, at 305.

69. *Breast Cancer Overview*, AM. CANCER SOC’Y, <http://www.cancer.org/Cancer/BreastCancer/OverviewGuide/breast-cancer-overview-key-statistics> (last updated Dec. 5, 2012) (reflecting 2012 estimates).

70. William H. Goodson III et al., *Activation of the mTOR Pathway by Low Levels of Xenoestrogens in Breast Epithelial Cells from High-Risk Women*, 32 *CARCINOGENESIS* 1724, 1724 (2011); Elizabeth W. LaPensee et al., *Bisphenol A at Low Nanomolar Doses Confers Chemoresistance in Estrogen Receptor- α -Positive and -Negative Breast Cancer Cells*,

The FDA has acknowledged that it “supports reasonable steps to reduce exposure of infants to BPA in the food supply,” has promised that it “will work with industry to support and evaluate manufacturing practices and alternative substances,” and has stated that it will support “the industry’s actions to stop producing BPA-containing bottles and infant feeding cups for the U.S. market.”⁷¹ Instead of just waiting and hoping for industry to change course, if Proposition 65 were reformed to truly allow for informed consent, consumers could choose to take more decisive precautionary action and encourage change through their own purchase decisions. Environmental and public health organizations are already starting to specifically identify products leaching BPA,⁷² but the task is overwhelming.

B. The Evolving Role of the DART Identification Committee and the Review of Endocrine-Disrupting Chemicals

There are four different mechanisms for listing chemicals under Proposition 65.⁷³ One option is through the state’s qualified experts: a chemical can be listed if either the Carcinogen Identifica-

117 ENVTL. HEALTH PERSP. 175, 176 (2009) (the chemotherapy drugs studied were doxorubicin, cisplatin, and vinblastine); Victoria Colliver, *Study: BPA, Methylparaben Block Breast Cancer Drugs*, S.F. CHRON., Sept. 13, 2011, http://articles.sfgate.com/2011-09-13/news/30147741_1_cancer-cells-breast-cells-bpa; see also *Meeting Before Science Board Advisory Committee*, *supra* note 67, at 291–92.

71. *Bisphenol A (BPA): Use in Food Contact Application*, *supra* note 19.

72. See, e.g., BREAST CANCER FUND, *BPA IN KIDS’ CANNED FOOD: A PRODUCT TESTING REPORT BY THE BREAST CANCER FUND* (2011), available at <http://www.breastcancerfund.org/assets/pdfs/publications/bpa-in-kids-canned-food.pdf>; *BPA-Free Canned Food Options*, THE SOFT LANDING, <http://guide.thesoftlanding.com/bpa-free-canned-food-options/> (last updated Oct. 25, 2012) (identifying BPA free options); *Chemicals and Points of Concern*, THE ALLIANCE FOR A HEALTHY TOMORROW (Oct. 31, 2007), <http://www.healthytomorrow.org/2007/10/chemicals.html> (posting information on studies identifying sources of BPA).

73. *Mechanisms for Listing and Delisting Chemicals Under Proposition 65*, CAL. OFFICE OF ENVTL. HEALTH HAZARD ASSESSMENT (OEHHA) (May 15, 2007), http://oehha.ca.gov/prop65/policy_procedure/listde051007.html (“The statute defines four mechanisms by which carcinogens and reproductive toxicants are listed. First, a chemical must be listed if one of the State’s Qualified Expert committees decides that a chemical has been clearly shown to cause cancer or reproductive toxicity through scientifically valid testing according to generally accepted principles. Second, a chemical must be listed if it is formally identified as a carcinogen or reproductive toxicant by a body considered authoritative under Proposition 65. Third, a chemical must be listed if a State or federal agency formally requires it to be identified or labeled as a carcinogen or reproductive toxicant. Fourth, a chemical must be listed if it is identified by reference in Labor Code section 6382(b)(1) or (d).”); see also CAL. HEALTH & SAFETY CODE § 25249.8 (West 2006).

tion Committee (CIC) or the Developmental and Reproductive Toxicant (DART) Identification Committee finds that the chemical has been “clearly shown to cause cancer or reproductive toxicity.”⁷⁴ On July 15, 2009, pursuant to procedures very much in need of reform as discussed below, the Proposition 65 DART Identification Committee voted not to list BPA as a reproductive toxicant.⁷⁵ However, BPA is still under review⁷⁶ through another mechanism for listing. A chemical may also be listed when an organization that has been designated as an “authoritative body” by the CIC or DART Identification Committee has already identified a chemical as causing cancer or birth defects or other reproductive harm.⁷⁷ Relevant here is the finding of the National Toxicology Program Center for Risks to Human Reproduction (NTP-CERHR)⁷⁸ that there is clear evidence of adverse developmental effects in laboratory animals at high levels of exposure to BPA.⁷⁹ Through the authoritative bodies mechanism, once the NTP-CERHR concludes that there is clear evidence of reproductive toxicity, the chemical must be listed unless scientifically valid data that were not considered by the authoritative body clearly establish that the sufficiency of evidence criteria were not met.⁸⁰ Based on the NTP-CERHR report and the references cited in that report, the OEHHA staff announced in February 2010 that the evidence appears sufficient for listing BPA and initi-

74. CAL. HEALTH & SAFETY CODE § 25249.8 (West 2006); CAL. CODE REGS. tit. 27, §§ 25102(c), 25302(a)–(c) (2009); CAL. CODE REGS. tit. 27, § 25305 (2012); *see also Mechanisms for Listing and Delisting Chemicals Under Proposition 65, supra* note 73.

75. *DART Meeting on Proposition 65* (July 15, 2009), *supra* note 53, at 253–55.

76. The process was initiated following a petition from the Natural Resources Defense Council. *See Request for Relevant Information on a Chemical Being Considered for Listing by the Authoritative Bodies Mechanism: Bisphenol-A, 7-Z* Cal. Regulatory Notice Reg. 252, 252 (Feb. 12, 2010), *available at* http://oehha.ca.gov/prop65/CRNR_notices/admin_listing/requests_info/callinBPA021210.html.

77. *See* CAL. HEALTH & SAFETY CODE § 25249.8(b) (West 2006); CAL. CODE REGS. tit. 27, § 25102(c) (2009); *Mechanisms for Listing and Delisting Chemicals Under Proposition 65, supra* note 73.

78. Designated as an authoritative body by the DART Board in 2002. *Meeting on Proposition 65 Before the Devtl. & Reprod. Toxicant Identification Comm. of the Office of Env'tl. Health Hazard Assessment of the State of Cal.* 188 (July 12, 2010), *available at* http://oehha.ca.gov/prop65/public_meetings/pdf/DARTIC071211trans.pdf (statement of Caroline Cox, Center for Environmental Health).

79. *Request for Relevant Information on a Chemical Being Considered for Listing by the Authoritative Bodies Mechanism: Bisphenol-A, 7-Z* Cal. Regulatory Notice Reg. at 253.

80. CAL. CODE REGS. tit. 27, § 25306(a), (g), (h), (l)(3) (2012).

ated the listing process, which will include a review of public comments.⁸¹

BPA may ultimately be successfully listed through the authoritative bodies mechanism; however, NTP-CERHR decisions may no longer be available for listing other endocrine-disrupting chemicals in the future. Although the DART Identification Committee declined a 2011 request from the American Chemistry Council (ACC)⁸² to rescind the designation of the NTP-CERHR,⁸³ at least insofar as it concerns the listing of future chemicals, the NTP-CERHR has nevertheless ceased to be an authoritative body. Recognizing the need for a more holistic approach, the National Toxicology Program Center for the Evaluation of Risks to Human Reproduction recently changed its regulatory structure and its name; it is now called the National Toxicology Program's Office of Health Assessment and Translation.⁸⁴ NTP representatives explain:

A strict focus on reproductive and developmental end points evaluated in the context of current human exposures may not result in the most health protective levels of concern, and could be confusing to the public. From a public health perspective, understanding the implications of current human ex-

81. *Request for Relevant Information on a Chemical Being Considered for Listing by the Authoritative Bodies Mechanism: Bisphenol-A*, CAL. OFFICE OF ENVTL. HEALTH HAZARD ASSESSMENT (Feb. 2, 2010), http://oehha.ca.gov/prop65/CRNR_notices/admin_listing/requests_info/callinBPA021210.html.

82. See Letter from Stanley W. Landfair, Counsel for the Polycarbonate/Global Grp. of the Am. Chemistry Council, et al., to Dorothy Burk, Chairperson, Cal. Devtl. & Reprod. Toxicant Identification Comm. (Oct. 14, 2010), available at http://oehha.ca.gov/prop65/public_meetings/pdf/101810accletter.pdf; Letter from Caroline Silveira, Dir., State Affairs, Grocery Mfrs. Ass'n, to Dorothy Burk, Chairperson, Cal. Devtl. & Reprod. Toxicant Identification Comm., & Joan Denton, Dir., Cal. Office of Env'tl. Health Hazard Assessment (Oct. 20, 2010), available at http://oehha.ca.gov/prop65/public_meetings/pdf/GMA_CCHR110110.pdf.

83. *Meeting on Proposition 65 Before the Devtl. & Reprod. Toxicant Identification Comm. of the Office of Env'tl. Health Hazard Assessment of the State of Cal.* 200 (July 12, 2011) [hereinafter *DART Meeting on Proposition 65* (July 12, 2011)], available at http://oehha.ca.gov/prop65/public_meetings/pdf/DARTIC071211trans.pdf; *Meeting Synopsis and Slide Presentations: Developmental and Reproductive Toxicant Identification Committee Meeting Held on July 12 and 13, 2011*, CAL. OFFICE OF ENVTL. HEALTH HAZARD ASSESSMENT (OEHHA) (Aug. 8, 2011), http://oehha.ca.gov/prop65/public_meetings/2011DARTsynop.html.

84. See John R. Bucher et al., *The Office of Health Assessment and Translation: A Problem-Solving Resource for the National Toxicology Program*, 119 ENVTL. HEALTH PERSP. A 196, A 196 (2011).

posures should include consideration of all relevant health effects.⁸⁵

Concerned that their regulatory interests may no longer be sufficiently aligned, on July 12, 2011, the DART Identification Committee deferred consideration as to whether to identify the new Office of Health Assessment and Translation as an authoritative body.⁸⁶

The regulatory jurisdiction of DART Identification Committee is limited to “reproductive toxicity,”⁸⁷ which OEHHA guidelines define to include “developmental toxicity” (including “adverse effects on the products of conception”),⁸⁸ “female reproductive toxicity,” and “male reproductive toxicity.”⁸⁹ The guidelines broadly define female and male “reproductive toxicity” to include “impaired or altered endocrine function;”⁹⁰ however, as discussed in the following section, it is not clear that the guidelines are being understood or interpreted to include consideration of the full breadth of possible detrimental effects on the endocrine system. Historically, relatively few chemicals have been listed under Proposition 65 in the absence of developmental toxicity.⁹¹

Given new scientific information as to the breadth of action of endocrine-disrupting chemicals, the DART Identification Committee should consider following the approach of the NTP-CERHR and adopt the broadest possible interpretation of “reproductive toxicity.” Although the data is still limited, according to the Endo-

85. *Id.*

86. See *DART Meeting on Proposition 65* (July 12, 2011), *supra* note 83, at 203–04.

87. CAL. CODE REGS. tit. 27, § 25305(b) (2012).

88. CAL. OFFICE OF ENVTL. HEALTH HAZARD ASSESSMENT, CRITERIA FOR RECOMMENDING CHEMICALS FOR LISTING AS “KNOWN TO THE STATE TO CAUSE REPRODUCTIVE TOXICITY” 1–2 (1993) [hereinafter OEHHA CRITERIA FOR RECOMMENDING CHEMICALS FOR LISTING] (emphasis added), available at http://oehha.ca.gov/prop65/policy_procedure/pdf_zip/dartCriteriaNov1993.pdf.

89. *Id.* at 1.

90. *Id.* at 2–3.

91. Of the 302 chemicals that have been listed as “reproductive toxins” through all the listing processes under Proposition 65, only 32 are associated with reproductive effects alone; the overwhelming majority, 208, have been listed for developmental toxicity alone. Cal. Office of Env'tl. Health Hazard Assessment, Proposition 65 Listing Mechanisms (Informational Agenda Item), Staff Presentation at the Meeting on Proposition 65 Before the Dev'tl. & Reprod. Toxicant Identification Comm. of the Office of Env'tl. Health Hazard Assessment of the State of Cal., PowerPoint Slides 3–4 (July 12–13, 2011) [hereinafter OEHHA Staff Presentation], available at http://oehha.ca.gov/prop65/public_meetings/pdf/071211DARTIClisting.pdf.

crine Society, the increased incidence of testicular cancer and malformations of the male genital tract and the decrease in quantity and quality of human sperm may be linked to the introduction of endocrine-disrupting chemicals into the environment.⁹² The increase in breast cancer also correlates with increased exposure to endocrine-disrupting chemicals,⁹³ and these chemicals have been linked through laboratory studies to many female reproductive disorders, including polycystic ovarian syndrome, aneuploidy, premature ovarian failure, reproductive tract anomalies, uterine fibroids, endometriosis, and ectopic gestation.⁹⁴

Considering just one of these disorders, endometriosis (an estrogen-dependent gynecological disorder associated with pelvic pain and infertility), the estimated health care costs for diagnosis and treatment totaled approximately \$22 billion in 2002, and there has been only limited success in achieving successful treatment of endometriosis-related pain.⁹⁵ In addition to reproductive tract disorders, new research suggests that exposure to endocrine-disrupting chemicals may play a role in both the diabetes and the obesity epidemics in the United States.⁹⁶ In 2008, the medical care costs of obesity in the United States totaled about \$147 billion.⁹⁷

Given these staggering costs, even while recognizing the possibility of other causal and contributing factors, California should consider a broad focus on all health effects related to endocrine disruption. Such reform may also bring harmony between the practice and interests of the DART Identification Committee and the National Toxicology Program's new Office of Health Assessment and Translation and encourage continued designation of the Na-

92. Diamanti-Kandarakis et al., *supra* note 9, at 305.

93. *See id.*

94. *See id.* at 300–01 (discussing laboratory studies with rodents, ungulates, and nonhuman primates and explaining that many of the mechanisms by which the disorders are caused by endocrine-disrupting chemicals are understood and, moreover, are conserved between animals and humans).

95. *See id.* at 304–05 (citing S. Simoens et al., *Endometriosis: Cost Estimates and Methodological Perspective*, 13 HUMAN REPRODUCTION UPDATE 395, 401 (2007)).

96. Jennifer Lee, *Child Obesity Is Linked to Chemicals in Plastics*, N.Y. TIMES, (Apr. 17, 2009, 1:31 PM), <http://cityroom.blogs.nytimes.com/2009/04/17/child-obesity-is-linked-to-chemicals-in-plastics/>; Thaddeus Schug, *NTP Workshop Investigates Links Between Chemicals and Obesity*, ENVTL. FACTOR (Feb. 2011), <http://www.niehs.nih.gov/news/newsletter/2011/february/science-ntp-workshop>.

97. *Adult Overweight and Obesity: Causes and Consequences*, CTRS. FOR DISEASE CONTROL & PREVENTION, <http://www.cdc.gov/obesity/causes/economics.html> (last updated Apr. 27, 2012) (citing Eric A. Finkelstein et al., *Annual Medical Spending Attributable to Obesity: Payer- and Service-Specific Estimates*, 28 HEALTH AFF. W822, w822 (2009)).

tional Toxicity Program as an authoritative body, reopening this important mechanism as a vehicle for listing. Historically many more chemicals have been listed as reproductive toxins through the authoritative bodies mechanism than through the state's qualified experts.⁹⁸ Failing designation of the new office of the NTP as an authoritative body, more responsibility will fall to the DART Identification Committee to list endocrine-disrupting chemicals.⁹⁹ Reform of the DART Identification Committee review process may thus assume increasing importance. There is a need for new rules to increase the availability and reliability of this vehicle for listing. Regulatory reform is needed to increase transparency and accountability as well as allow for consideration of a broader spectrum of health effects.

II. REQUIREMENTS FOR TRANSPARENCY AND ACCOUNTABILITY

As discussed in detail in the following sections, the 2009 Proposition 65 DART Identification Meeting on BPA¹⁰⁰ encountered serious difficulties due to vague definitions and a dearth of interpretive guidance. Proposition 65 regulates chemicals "clearly shown" to cause cancer or "reproductive toxicity."¹⁰¹ However, there are only very limited guidelines (the "Criteria for Recommending Chemicals for Listing as 'Known to the State to Cause Reproductive Toxicity'" (Guidelines))¹⁰² and no regulations defining these statutory standards. There was considerable debate at the public hearing as to the scope of adverse effects that fall within the realm of "reproductive toxicity."¹⁰³ However, the OEHHA staff, although present and participating at the hearing,¹⁰⁴ did not discuss the Guidelines or take a position as to whether there was sufficient evidence of relevant adverse effects. At the conclusion of the hearing, it was unclear whether the DART Identification Committee's decision was influenced by industry's arguments in favor of a nar-

98. See OEHHA Staff Presentation, *supra* note 91, at 5, 7.

99. NTP-CERHR is one of four Authoritative Bodies for Reproductive Toxicity. *Id.* at 5, 17.

100. *DART Meeting on Proposition 65* (July 15, 2009), *supra* note 53.

101. CAL. HEALTH & SAFETY CODE § 25249.8(b) (West 2006).

102. OEHHA CRITERIA FOR RECOMMENDING CHEMICALS FOR LISTING, *supra* note 88.

103. See, e.g., *DART Meeting on Proposition 65* (July 15, 2009), *supra* note 53, at 222-24.

104. *Id.* at i.

row focus as to what qualifies as “reproductive toxicity” or if it was based instead entirely on the question of the adequacy of the evidence.

As reflected in the discussion below, the Guidelines were not clearly referenced at the hearing in a way that would suggest that they were consistently guiding the process. Moreover the Guidelines were “not intended to limit the scope of the committee’s consideration”¹⁰⁵ and the Guidelines themselves may need revision to conform to the evolving science. In the absence of clear regulatory standards, the Committee members were left to come up with their own varying interpretations of critical terms. The hearing transcript reflects a need for a discussion that involves both the public and the scientific community to consider appropriate and transparent definitions of both “clearly shown” and “reproductive toxicity.” OEHHA should open a rulemaking process.

A. *Defining “Clearly Shown” Consistent with Scientific Standards and Societal Choices*

One DART Identification Committee member interpreted the statutory requirement that reproductive toxicity be “clearly shown” to require conclusive evidence, an especially difficult standard here, where there have been serious reports of advocacy science.¹⁰⁶ Committee Member Roberts reasoned, “At least, in my perspective, there are not clear effects on the low-dose levels, because we have seen situations where some studies are positive and some studies are negative.”¹⁰⁷ Committee Member White explained, “I didn’t quite feel like there was conclusive and clear evidence”¹⁰⁸ Chairperson Burk noted, “[W]e all have . . . probably our own definition of clear.”¹⁰⁹

An interpretation of “clearly shown” that would require conclusive evidence is discordant with the precautionary public policy disclosure objectives of Proposition 65. Regardless of any safety determination, Proposition 65 aimed to allow consumers to make

105. OEHHA CRITERIA FOR RECOMMENDING CHEMICALS FOR LISTING, *supra* note 88, at 1.

106. Jane Houlihan et al., *Timeline: BPA from Invention to Phase-Out*, ENVTL. WORKING GRP., <http://www.ewg.org/reports/bpatimeline> (last updated Mar. 2011).

107. *DART Meeting on Proposition 65* (July 15, 2009), *supra* note 53, at 238.

108. *Id.* at 229.

109. *Id.* at 230.

their own choices about chemical exposure.¹¹⁰ Moreover asking a group of scientists to find that a chemical has been conclusively shown to cause reproductive toxicity is inconsistent with modern scientific theory.¹¹¹ Science is based on “generating hypotheses and testing them to see if they can be falsified.”¹¹² Certainty is elusive if not impossible to establish. Reducing one type of error, inevitably increases another—“Type I’ errors are created by accepting hypotheses that are ultimately shown to be wrong, whereas ‘Type II’ errors are created by rejecting hypotheses that are ultimately shown to be true.”¹¹³ Regulatory agencies protecting the public interest should logically be most concerned with false negatives (Type II) errors, whereas industry is generally most concerned with false positives (Type I errors).¹¹⁴ As William R. Freudenburg et al. explain in their article, *Scientific Certainty Argumentation Methods (SCAMs): Science and the Politics of Doubt*:

In environmental and technological controversies, a Type II error is not merely an abstract possibility, but a risk that innocent people will get sick or die. In light of this reality, it is difficult to believe that anyone who believes in truly balanced or “sound science”—or for that matter, any well-informed person of good will—could seriously contend that the “proper” balance involves a decision to focus exclusively on Type I errors while deciding to ignore Type II errors completely. That, however, is nevertheless the net effect of successful efforts to argue for full “scientific certainty” before a regulation can be said to be “justified”—and that, in short is a SCAM.¹¹⁵

In the context of science, the most we could possibly ask for would be clear *evidence* of reproductive toxicity. Even then, we would still have to consider what is meant by “clear” and what is meant by “reproductive toxicity;” that is, how much certainty and

110. CAL. HEALTH & WELFARE AGENCY, *supra* note 63, at 43 (“The apparent purpose of any warning under the Act is to permit the persons exposed to make choices about the exposure.”).

111. Vom Saal & Welshons, *supra* note 13, at 69 (“In experimental research scientists test whether the hypothesis that the observed results come from the same distribution (the null hypothesis) can be rejected with a specific level of confidence. . . . The hypothesis that results all come from the same distribution (or the same population) can be disproved or falsified only at some specified level of confidence, it can never be proven to be correct.”).

112. *Daubert v. Merrell Dow Pharm., Inc.*, 509 U.S. 579, 593 (1993).

113. William R. Freudenburg et al., *Scientific Certainty Argumentation Methods (SCAMs): Science and the Politics of Doubt*, 78 SOC. INQUIRY 2, 7 (2008).

114. Vom Saal & Welshons, *supra* note 13, at 69–70.

115. Freudenburg et al., *supra* note 113, at 31.

what type of evidence is sufficient. These questions are not purely questions of science. OEHHA's staff, however, failed to acknowledge the complexity. Responding to a complaint about OEHHA's failure to provide any clarification to correct confusion about the charge,¹¹⁶ OEHHA staff explained:

We agree that the "clearly shown" [sic] standard in the statute and regulations has become the subject of much debate in public comments in recent years. This standard is not a legal determination; it is instead a scientific judgment in which the state's qualified experts are expected to apply their own knowledge and expertise to determine if a chemical has been "clearly shown by scientifically valid testing according to generally accepted principals to cause reproductive toxicity."¹¹⁷

Yet there must be more than only "scientific judgment" at play. Carefully dissecting this issue in *The Myth of Science as a "Neutral Arbiter" for Triggering Precautions*, Vern Walker explains:

Numerous non-scientific decisions are necessarily involved in both making and warranting findings that a triggering risk exists. Making a finding of risk involves decisions about the meaning of "risk of harm," about the meaning of any qualitative or quantitative modifiers, and about the truth modality of (or degree of confidence in) the finding as a whole. Moreover, every determination that the available scientific evidence warrants a finding of risk involves decisions about the acceptable degree of various types of uncertainty: conceptual uncertainty, measurement uncertainty, sampling uncertainty, modeling uncertainty, and causal uncertainty.¹¹⁸

The lack of rules defining a more articulate standard allows the DART Identification Committee members to create a decisionmaking process that lacks transparency, allows policy decisions to hide behind the cloak of "science," and encourages deferred decisionmaking. In her article, *The Science Charade in Toxic Risk Regulation*,

116. Letter from Dr. Sarah Janssen, Staff Scientist, Natural Res. Def. Council, et al., to Joan Denton, Dir., Cal. Office of Env'tl. Health Hazard Assessment 2-3 (July 22, 2009) [hereinafter Letter from Dr. Sarah Janssen et al.], available at http://oehha.ca.gov/prop65/public_meetings/pdf/DARTIC072209.pdf (letter joined by Breast Cancer Fund, Clean Water Action, Environment California, Healthy Building Network, and Science & Environmental Health Network).

117. Letter from Joan Denton, Dir., Cal. Office of Env'tl. Health Hazard Assessment, to Dr. Sarah Janssen, Staff Scientist, Natural Res. Def. Council, et al. 2-3 (Sept. 1, 2010) [hereinafter Letter from Joan Denton], available at http://oehha.ca.gov/prop65/public_meetings/pdf/090110OEHHA.pdf.

118. Vern R. Walker, *The Myth of Science as a "Neutral Arbiter" for Triggering Precautions*, 26 B.C. INT'L & COMP. L. REV. 197, 228 (2003).

Wendy Wagner captures this, apparently common, situation when there is a “statutory mandate that appears to require protective standards to be based at least in part on science, coupled with a deficient understanding of the science-policy nature of risk assessment.”¹¹⁹ She explains:

Once given responsibility for setting a single, quantitative standard, agency scientists generally take one of two approaches: 1) they continue indefinitely to look to science to resolve the trans-scientific questions; or 2) they substitute their own values for the policy choices needed at the trans-scientific junctures and characterize the final science-policy decisions as the result of scientific experimentation and scientific judgment. In either case, the results are disturbing.¹²⁰

Without publicly accessible standards to guide the process, there is also a lack of transparency that interferes with the proper functioning of our political system. The public must be able to discover and understand the policy decisions hidden within the science in order for the political process to work. In *Using Science in a Political World: The Importance of Transparency in Natural Resource Regulation*, Holley Doremus explains:

[T]he technical complexities of science must not be allowed to obscure the political judgments that are ultimately at the heart of regulatory decisions. . . . Ultimately, where the burden of proof should lie and how strong that burden should be are societal choices that will depend upon societal judgments about the costs of different types of error. In a democracy, the public must be the final arbiters of the relative importance of goals that may be in tension with one another.¹²¹

To the extent that the basis for the DART Identification Committee decisions, both scientific and policy decisions, can be separated and made accessible to the public, the public would then have the opportunity to respond as necessary to encourage corrective action through the political and legal system.

No findings of fact were issued at the end of the DART Identification Committee hearings on BPA, and the Committee’s decision-making process lacked transparency. In a response to a letter

119. Wendy E. Wagner, *The Science Charade in Toxic Risk Regulation*, 95 COLUM. L. REV. 1613, 1632 (1995) (footnotes omitted).

120. *Id.* (footnotes omitted).

121. Holly Doremus, *Using Science in a Political World: The Importance of Transparency in Natural Resource Regulation*, in *RESCUING SCIENCE FROM POLITICS: REGULATION AND THE DISTORTION OF SCIENTIFIC RESEARCH* 143, 153 (Wendy Wagner & Rena Steinzor eds., 2006).

submitted by several advocacy organizations,¹²² the OEHHA staff indicated that they would, in the future, provide a copy of the Guidelines to each Committee member prior to meetings, organize written and oral presentations with the goal of focusing Committee members on each endpoint of concern, and identify the studies that OEHHA staff feels are most important to the Committee's evaluation of the chemical.¹²³ Although undoubtedly helpful, these actions alone are not enough to resolve problems with the review process. A stronger approach would be to use the rulemaking process to open a public discussion with the scientific community concerning the appropriate scope of the Guidelines given contemporary understanding of the breadth of action of endocrine-disrupting chemicals, and the necessary level of detail to guide the decisionmaking process as to the sufficiency of the evidence.

Resolving issues through a rulemaking proceeding would also create an opportunity to add requirements for findings of fact that could help to focus the Committee's attention on relevant parameters of defensible and unbiased science and encourage the staff to propose findings and conclusions for Committee review. This technique forces the regulatory agency to reflect carefully on what should be the appropriate basis for its decision. As described by the California Supreme Court: "Among other functions, a findings requirement serves to conduce the administrative body to draw legally relevant subconclusions supportive of its ultimate decision; the intended effect is to facilitate orderly analysis and minimize the likelihood that the agency will randomly leap from evidence to conclusions."¹²⁴

A more transparent process with findings of fact may also help inform the public in the face of the inevitably misleading spin presented by opponents to a listing. When the DART Identification Committee declined to list BPA as a reproductive toxicant, it was concluding only that it felt that the evidence fell short of the standard, "clearly shown to cause reproductive toxicity;" it was not concluding that BPA is proven "safe." However, that did not stop blogger Kerri Toloczko from declaring that the DART Committee had indeed made such a finding of safety.¹²⁵ Nor did it prevent

122. Letter from Dr. Sarah Janssen et al., *supra* note 116.

123. Letter from Joan Denton, *supra* note 117, at 2-3.

124. *Topanga Ass'n for a Scenic Cmty. v. Cnty. of L.A.*, 522 P.2d 12, 18 (Cal. 1974).

125. Kerri Toloczko, *Junk Science Has Consequences: Environmental Lobby Shows No Concern for California's Financial Woes*, BREITBART (Apr. 29, 2010), <http://www.breitbart.com>.

Nick Kump of Elmets Communications from reporting that the “state’s top panel of independent experts found no particular risk in BPA” and referring to a “mountain of evidence showing BPA’s benign safety profile.”¹²⁶ More of the same was reported in a blog of the National Association of Manufacturers (NAM):

“We can now add California to the growing list of agencies that have concluded that BPA does not pose a risk to the general public,” said Dr. John M. Rost, NAMPA [North American Metal Packaging Alliance] Chairman. “It’s important to note that when politics and media interference are taken out of the process, and safety decisions are made by qualified, independent scientific experts, we see the same conclusion time and time again—that BPA is safe.”¹²⁷

B. Defining “Reproductive Toxicity” Consistent with Advancing Science

There was significant discussion and confusion in the hearing as to the scope of adverse effects within the realm of “reproductive toxicity.” Industry took the position that many of the adverse effects identified by the opposition at the hearing are not within the realm of “reproductive toxicity” and thus fall outside regulatory jurisdiction.¹²⁸ According to an article authored by thirty-six sponsoring scientists, the effects of BPA excluded from consideration in industry-sponsored studies include: altered metabolism related to metabolic syndrome; altered adiponectin secretion (a condition predicting heart disease and type 2 diabetes); altered epigenetic programming leading to precancerous lesions of the prostate; differential growth patterns in the developing prostate; abnormal growth, gene expression, and precancerous lesions of the mammary glands; adverse effects on the female reproductive system, in-

breitbart.com/Big-Government/2010/04/29/Junk-Science-Has-Consequences—Environmental-Lobby-Shows-No-Concern-for-California—s-Financial-Woes. Kerri Toloczko is a Senior Fellow at Let Freedom Ring specializing in policy analysis, issue advocacy, and coalition building for the public policy community. *Kerri Toloczko*, THE HUFFINGTON POST, <http://www.huffingtonpost.com/kerri-toloczko> (last visited Dec. 30, 2012).

126. Nick Kump, *Unsuspecting Californians Footing Bill for NRDC Scare Campaign*, ELMETS COMM’NS (July 12, 2011, 12:18 PM), <http://www.elmets.com/blog/unsuspecting-californians-footing-bill-for-nrdc-scare-campaign>.

127. Carter Wood, *Scientific Integrity: Calif. Board Dismisses BPA as Health Threat*, SHOPFLOOR (July 17, 2009), <http://shopfloor.org/tag/proposition-65>. Shopfloor is a blog of the National Association of Manufacturers. *About*, SHOPFLOOR, <http://www.shopfloor.org/about-nam> (last visited Dec. 30, 2012).

128. See *DART Meeting on Proposition 65* (July 15, 2009), *supra* note 53, at ii, 135–38 (statement of Dr. Jay Murray, Murray & Associates).

cluding uterine fibroids, paraovarian cysts, and chromosomal abnormalities in oocytes; and neurochemical and behavioral abnormalities.¹²⁹

One difficulty during the hearing involved the division of responsibility between the Developmental and Reproductive Toxicant (DART) Identification Committee and the Carcinogen Identification Committee (CIC). Although Proposition 65 simply requires the “opinion of the state’s qualified experts,”¹³⁰ regulations divide the task: review of toxicity as a carcinogen is completed by the CIC, whereas review of toxicity as a reproductive toxin is completed by the DART Identification Committee.¹³¹ It is evident from the hearing that existing procedures and Guidelines are inadequate to prevent critical issues from falling between the DART Identification Committee and the CIC, and others from potentially being excluded entirely from the review process.

In the Guidelines, the DART Identification Committee reserved for itself the question of “transplacental carcinogenesis.”¹³² However, in the hearing on BPA, the Committee clearly had difficulty parsing out that issue. The Committee specifically asked its chief counsel whether neoplastic lesions that are attributed to exposure neonatally would be under the DART Identification Committee or under the CIC, and whether it was necessary that the lesion impact the reproductive potential of the animals.¹³³ She responded by simply turning the question back on the Committee: “[I]t’s probably not as clear as it could be, but it is somewhat in your area of expertise whether you think that is an effect or not.”¹³⁴

Although there was significant discussion concerning carcinogenicity, in the end, this issue was largely ignored by the DART Identification Committee members. Dr. vom Saal, a leading BPA

129. John Peterson Myers et al., *Why Public Health Agencies Cannot Depend on Good Laboratory Practices as a Criterion for Selecting Data: The Case of Bisphenol A*, 117 ENVTL. HEALTH PERSP. 309, 309, 310 (2009).

130. CAL. HEALTH & SAFETY CODE § 25249.8(b) (West 2006).

131. CAL. CODE REGS. tit. 27, § 25102(c) (2009); CAL. CODE REGS. tit. 27, § 25305(a), (b) (2012).

132. OEHHA CRITERIA FOR RECOMMENDING CHEMICALS FOR LISTING, *supra* note 88, at 2.

133. *DART Meeting on Proposition 65* (July 15, 2009), *supra* note 53, at 222–24; *see also* Letter from Dr. Sarah Janssen et al., *supra* note 116, at 3 (expressing concern that the Committee, in its confusion, inappropriately avoided the question as to whether BPA was a “transplacental” carcinogen).

134. *DART Meeting on Proposition 65* (July 15, 2009), *supra* note 53, at 224.

university research scientist¹³⁵ testifying in favor of the listing, mentioned that Dr. Huff from the National Cancer Institute wrote an article drawing the conclusion that if BPA were evaluated using current standards,¹³⁶ it would be deemed a carcinogen.¹³⁷ Also addressing the issue of carcinogens, Dr. Wu of the OEHHA staff reported on effects of BPA in mice that are “typically associated with carcinogenesis,” including significant increases in “the number of terminal end buds,” “maturation of cells comprising the fat pad,” “altered localization of collagen,” and “cell cycle alteration” in the mammary gland.¹³⁸

During the final DART Identification Committee discussion, OEHHA staff member Dr. Zeise noted that although they “didn’t look in detail at the cancer endpoint, there are a number of early-in-life studies and in utero studies that show precursor lesions.”¹³⁹ Committee Member White also noted “the possibility of mammary gland alterations and lesions,” which he considered “very significant” in relation to “breast cancer lesions.”¹⁴⁰ However, Dr. Zeise reiterated that the Committee “didn’t evaluate the carcinogenicity studies.”¹⁴¹

“Developmental effects” that result entirely or predominantly from postnatal exposure were also excluded from consideration.¹⁴² According to OEHHA Chief Counsel Monahan-Cummings, this narrow interpretation was based on the Preamble to Proposition 65, which identifies the chemicals of concern as those that “cause cancer, *birth defects*, or other reproductive harm.”¹⁴³ She explained, “So our interpretation of that has been that we are looking at prenatal exposures that may cause, you know, developmental effects after birth, but we’re not looking at exposures after birth that may cause [developmental] effects later.”¹⁴⁴ “Transplacental carcinogenesis,”

135. *Id.* at ii, 56 (Dr. vom Saal is from the University of Missouri and has been conducting BPA research funded by the National Institute of Environmental Health Sciences for 13 years).

136. *Id.* at 234 (based on the studies discussed by an industry representative, Dr. Hentges from the American Chemistry Council).

137. *Id.*

138. *Id.* at i, 26.

139. *Id.* at 251.

140. *DART Meeting on Proposition 65* (July 15, 2009), *supra* note 53, at 233.

141. *Id.* at 234.

142. *DART Meeting on Proposition 65* (July 12, 2011), *supra* note 83, at i, 117 (statement of Dr. Jim Donald, Chief, Reproductive Toxicology & Epidemiology Section, California Office of Environmental Health Hazard Assessment).

143. *Id.* at 126 (emphasis added) (citing to page 53 of the preamble to Proposition 65).

144. *Id.* at 126–27.

for example, is listed only under “developmental toxicity” (one component of “reproductive toxicity”),¹⁴⁵ and not under the other components, “female reproductive toxicity,” or “male reproductive toxicity.”¹⁴⁶ “Developmental toxicity” concerns “adverse effects *on the products of conception*,”¹⁴⁷ in other words, effects due to prenatal exposure.

This fragmentation of the review process that excludes consideration of postnatal exposure in evaluating “developmental toxicity,” complicates the review of rodent studies. In order to evaluate the period of prenatal exposure in humans, the corresponding time period for analysis of exposure in rodents includes a period of postnatal exposures. As Dr. Woodruff explained in her testimony:

[W]hile most of the studies on BPA are from rodents or mice . . . the period of development of mice is somewhat different than the period of development for humans in terms of the actual timing of birth. So human gestation goes up to about 40 weeks. This is equivalent to both prenatal gestation for the mice and also postnatal growth up to about Day 50. So any experiments done in mice from prenatal or *up to postnatal day 50* is equivalent to prenatal experiments in humans.¹⁴⁸

Confusion clearly reigned on this issue. Committee Member Roberts noted, “[W]hat I tried to limit myself to are where the exposure in the animal studies occurred in what would be considered equivalent to prenatal exposure in the human, which is pretty much the gestational period in a rodent, plus maybe *a few days afterwards*.”¹⁴⁹ According to Dr. vom Saal, this narrowing of the scope of relevant studies is significant. Many studies have been done on the neonatal rodent, where the researcher can directly control ex-

145. OEHHA CRITERIA FOR RECOMMENDING CHEMICALS FOR LISTING, *supra* note 88, at 1 (“For purposes of these criteria, ‘reproductive toxicity’ includes ‘developmental toxicity’, ‘female reproductive toxicity’, and ‘male reproductive toxicity’.”).

146. *Id.* at 2–3.

147. *Id.* at 1–2 (emphasis added) (including but not limited to: “(1) Embryo/fetal mortality (including resorption, miscarriage/spontaneous abortion, or still-birth), malformations, structural abnormalities and variations, altered fetal growth, and change in gestational age at delivery. (2) Postnatal parameters including growth and development, physiological deficits and delay, neurological, neurobehavioral and psychological deficits, altered sex ratio, abnormal sexual development or function, and morbidity or mortality. (3) Transplacental carcinogenesis. (4) Somatic or genetic (germ cell) mutations in the conceptus.”).

148. *DART Meeting on Proposition 65* (July 15, 2009), *supra* note 53, at 81–82 (emphasis added) (statement of Dr. Tracey Woodruff, University of California, San Francisco) (partially reiterating earlier statements by Dr. vom Saal).

149. *Id.* at 235 (emphasis added).

posure instead of trying to control exposure to the fetus through exposure to the mother.¹⁵⁰ By evaluating exposures to the neonatal rodent, researchers can use biomonitoring of chemical blood levels to compare exposure to that of the human fetus.¹⁵¹

To simplify matters and align Committee deliberations with the intent of the voters, the issue of “transplacental” carcinogenesis should be included not only under the definition of “developmental toxicity,” but also under a broad regulatory interpretation of “reproductive toxicity.” “Reproductive toxicity” could be defined to include both prenatal and postnatal exposures that seriously and adversely affect the endocrine system and reproductive organs. The voters broadly stated that they were interested not only in cancer and birth defects, but also “other reproductive harm,” which may be understood to include cancer and precancerous conditions in the reproductive organs, regardless of whether those effects were due to prenatal or postnatal exposure. Moreover, treatment for cancer in reproductive organs, both breast and prostate cancer, may include chemotherapy and hormone suppression, the effects of which include “chemical castration” and infertility.¹⁵² Broadly defining “reproductive toxicity” to include consideration of carcinogenicity to the extent that it is related to endocrine disruption would allow for more holistic consideration of endocrine disrupting chemicals and avoid the possibility of significant issues falling between the DART Identification Committee and the CIC.

Even aside from the question of carcinogenicity, there was also considerable confusion in the discourse as to the realm of “repro-

150. Telephone Interview with Dr. Frederick S. vom Saal, Curators’ Professor, Div. of Biological Sci., Coll. of Arts & Sci., Univ. of Mo.-Columbia (Sept. 22, 2011); see e-mail from Dr. Frederick S. vom Saal, Curators’ Professor, Div. of Biological Sci., Coll. of Arts & Sci., Univ. of Mo.-Columbia, to author (Sept. 22, 2011, 10:22 AM) [hereinafter e-mail from Dr. Frederick S. vom Saal (Sept. 22, 2011, 10:22 AM)] (on file with author).

151. Telephone Interview with Dr. Frederick S. vom Saal, *supra* note 150; e-mail from Dr. Frederick S. vom Saal (Sept. 22, 2011, 10:22 AM), *supra* note 150.

152. See, e.g., Tiina Saarto et al., *Chemical Castration Induced by Adjuvant Cyclophosphamide, Methotrexate, and Fluorouracil Chemotherapy Causes Rapid Bone Loss that Is Reduced by Clodronate: A Randomized Study in Premenopausal Breast Cancer Patients*, 15 J. CLINICAL ONCOLOGY 1341, 1341 (1997); *Living with Prostate Cancer: Loss of Fertility*, PROSTATE CANCER FOUND., http://www.pcf.org/site/c.leJRIRORepH/b.5837043/k.B194/Loss_of_Fertility.htm (last visited Jan. 13, 2013); cf. *Hormone Blockade Makes Radiation Therapy More Effective for Medium and High-Risk Localized Prostate Cancer*, PSA RISING MAGAZINE (Sept. 13, 2000), <http://psa-rising.com/medicalpike/ebr/plus-hormone-df-0900.shtml> (reporting that radiation in combination with androgen suppression therapy helps block the production or action of male hormones that have been shown to fuel prostate cancer).

ductive toxicity” generally. Industry representatives repeatedly attempted to dismiss studies that revealed negative effects of BPA by arguing that those effects were outside regulatory purview. Dr. Murray (a former DART Identification Committee member introduced by the lawyer representing the American Chemistry Council)¹⁵³ explained in his testimony: “most of these studies do not take it out to a reproductive endpoint. A lot of them focus on unique endpoints. Some of them look at molecular approaches and there’s nothing wrong with that, but they’re not tied to an adverse effect.”¹⁵⁴ Dr. Tyl, also speaking on the side of industry, focused the Committee’s attention on the most obvious of endpoints:

So we’re looking at endpoints that indicate adverse outcome, okay. We’re not looking at the early molecular biochemical kinds of markers. Not that they’re not interesting and fascinating and not that they shouldn’t be pursued, but we’re looking at endpoints. What is the—is there an adverse consequence to these early changes?¹⁵⁵

Remarking on one study, she noted, “[T]he animals bred. They got pregnant. They had babies. They developed the babies. They went through puberty. They grew up. They had babies.”¹⁵⁶

The OEHHA staff presentation, although evading any conclusions as to whether the regulatory standard had been satisfied, discussed more sophisticated studies and appeared to support reliance on more subtle endpoints as the foundation for a conclusion as to reproductive toxicity. Dr. Moran, a staff toxicologist reporting on the endocrine activity of BPA, explained that BPA interferes with reproductive hormones as well as glucose and insulin in a way that can both increase or decrease production, and that BPA interferes with metabolism.¹⁵⁷ As to female reproductive toxicity, another staff toxicologist, Dr. Wu, concluded that there were “limited data on reproductive effects of Bisphenol A in women,” that “recurrence of miscarriage in women is possibly consistent with the perturbation of the meiotic cell cycle and the chromosome misalignment in oocytes noted in laboratory animals,” that “[n]umerous female animal studies showed effects on the female reproductive system from Bisphe-

153. *DART Meeting on Proposition 65* (July 15, 2009), *supra* note 53, at ii, 94 (statement of Stanley Landfair, American Chemistry Council).

154. *Id.* at ii, 135–36 (statement of Dr. Jay Murray, Murray & Associates).

155. *Id.* at ii, 126 (statement of Dr. Rochelle W. Tyl, RTI International).

156. *Id.* at ii, 121–22 (statement of Dr. Rochelle W. Tyl, RTI International).

157. *See id.* at i, 44 (statement of Dr. Francisco Moran-Messen, Staff Toxicologist, Reproductive Toxicology and Epidemiology Section, California Office of Environmental Health Hazard Assessment).

nol A,” and that “[a]lterations to the uterus, ovary, follicles and oocytes, estrous cycle, vagina and mammary gland were notable.”¹⁵⁸ Assuming the relevance of these endpoints, limited evidence of the negative effects of BPA in humans, supported by sufficient experimental animal data, would fall within the criteria for a finding of “reproductive toxicity” under the Guidelines.¹⁵⁹

The Guidelines broadly define female and male “reproductive toxicity” to include “genetic damage to the ovum [or spermatozoon] or its precursors” and “impaired or altered endocrine function.”¹⁶⁰ However, it is not clear that they were understood or interpreted so as to allow for consistency with modern science. As Dr. vom Saal explains, if understood in the context of modern science, “endocrine function” includes not only effects on hormones transmitted in the blood, but all methods of signaling between and within cells, including neurotransmitters that act as endocrine signals.¹⁶¹ According to the Endocrine Society, although endocrine-disrupting chemicals were originally thought to exert actions primarily through nuclear hormone receptors,¹⁶² it is now understood that the mechanisms are much broader than originally recognized: endocrine disruptors act via nuclear receptors, nonnuclear steroid hormone receptors,¹⁶³ nonsteroid receptors,¹⁶⁴ enzymatic pathways involved in steroid biosynthesis and/or metabolism, and numerous other mechanisms that involve both the endocrine and the reproductive systems.¹⁶⁵

Given advances in science, OEHHA should follow the National Toxicology Program’s new, more holistic approach and pursue the broadest possible regulatory interpretation consistent with the language of Proposition 65. “Reproductive toxicity” should be clearly defined and interpreted so as to include all serious adverse effects

158. *Id.* at i, 27–28 (statement of Dr. Lily Wu, Staff Toxicologist, Reproductive Toxicology and Epidemiology Section, California Office of Environmental Health Hazard Assessment).

159. OEHHA CRITERIA FOR RECOMMENDING CHEMICALS FOR LISTING, *supra* note 88, at 4.

160. *Id.* at 2–3.

161. See e-mail from Dr. Frederick S. vom Saal (Sept. 22, 2011, 10:22 AM), *supra* note 150.

162. Diamanti-Kandarakis et al., *supra* note 9, at 294 (including estrogen receptors, androgen receptors, progesterone receptors, thyroid receptors, and retinoid receptors).

163. *Id.* (e.g., membrane ERs).

164. *Id.* (e.g., neurotransmitter receptors, such as the serotonin receptor, dopamine receptor, and norepinephrine receptor, and orphan receptors, such as the aryl hydrocarbon receptor (AhR)).

165. *Id.*

related to the functioning of the endocrine system, including all effects that cause people pain, cost society money, and lead to other serious health conditions, like precancerous conditions and obesity. As explained above, expanding the realm of relevant endpoints would also allow the DART Identification Committee to remain aligned with the new National Toxicology Program Office of Health Assessment and Translation and may encourage the designation of this reincarnation of the NTP-CERHR as an authoritative body. California's OEHHA should open a rulemaking proceeding and encourage the participation of scientists highly specialized in the effects of endocrine chemicals. According to Dr. vom Saal, the concept of "reproductive toxicity" should include all adverse developmental effects where there is permanent adverse change caused by a chemical at the genetic, epigenetic, molecular, cellular, tissue, organ, organism, or population level.¹⁶⁶ Such a definition would necessarily also include endocrine disruption that leads to precancerous conditions and allow for more holistic review of endocrine-disrupting chemicals.¹⁶⁷

III. RESEARCH DESIGN AND QUALITY STANDARDS TO ASSESS AND EVALUATE ADVOCACY SCIENCE

Another difficulty that arose in the DART hearing on BPA was the question of how to weigh and evaluate conflicting studies. The only mention of any standard came from Stanley Landfair, the lawyer representing industry, who urged the Committee to use the "weight-of-the-evidence approach."¹⁶⁸ He clarified that the Committee knew better than he what that meant, but that it did not mean "simply to count up the studies."¹⁶⁹ The Guidelines state the following as "weight of evidence" considerations:

- (1) Data on a single species from a well conducted developmental or reproduction study may be sufficient to classify an agent as a reproductive toxicant provided there are not *equally well conducted studies* which do not show an effect and which

166. See e-mail from Dr. Frederick S. vom Saal, Curators' Professor, Div. of Biological Sci., Coll. of Arts & Sci., Univ. of Mo.-Columbia, to author (Sept. 22, 2011, 9:58 AM) (on file with author) (noting that considerations at the "population level" should include, for example, changes such as those seen in fish, where populations are found to have fifty percent intersex members).

167. *Id.*

168. *DART Meeting on Proposition 65* (July 15, 2009), *supra* note 53, at ii, 95 (statement of Stanley Landfair, American Chemistry Council).

169. *Id.* (statement of Stanley Landfair, American Chemistry Council).

2012] ENDOCRINE-DISRUPTING CHEMICALS AND PROP 65 217

have sufficient power to call into questions [sic] the repeatability of the observation in the positive study.

(2) Data on more than one species or from more than a single study increase the confidence for classification of an agent as a reproductive toxicant.¹⁷⁰

Missing are any detailed criteria by which to evaluate conflicting studies or to consider issues of appropriate research design or conflicts of interest. The lack of such criteria was particularly important in the review of BPA, where, as may be expected when significant economic interests are involved,¹⁷¹ there have been serious allegations of advocacy science.¹⁷² A 2005 analysis of BPA literature revealed that the funding source correlated perfectly with the findings.¹⁷³ Of the 115 studies on health effects of BPA, 94 were government-funded studies conducted in domestic and international academic laboratories.¹⁷⁴ All of these government-funded studies found adverse effects at low-dose exposure, yet not a single industry-funded study reported adverse effects.¹⁷⁵

Industry involvement with the regulatory review of BPA has been persistent and extensive. In 2006, the National Toxicology Program's Center for the Evaluation of Risks to Human Reproduction (NTP-CERHR) published a draft advisory panel report that was allegedly written largely by an outside consultant who was fired the next year after public disclosure of its conflicts of interest with the regulated industry.¹⁷⁶ Still building on this report as its foundation document,¹⁷⁷ the FDA relied on only two studies, both sponsored by the American Plastics Council, as the basis of its initial decision

170. OEHHA CRITERIA FOR RECOMMENDING CHEMICALS FOR LISTING, *supra* note 88, at 5 (emphasis added).

171. See generally THOMAS O. MCGARITY & WENDY E. WAGNER, BENDING SCIENCE: HOW SPECIAL INTERESTS CORRUPT PUBLIC HEALTH RESEARCH (2008).

172. See Frederick S. vom Saal & Claude Hughes, *An Extensive New Literature Concerning Low-Dose Effects of Bisphenol A Shows the Need for a New Risk Assessment*, 113 ENVTL. HEALTH PERSP. 926, 926 (2005).

173. See *id.*

174. *Id.*

175. See *id.* at 928.

176. Houlihan et al., *supra* note 106 (citing CTR. FOR THE EVALUATION OF RISKS TO HUMAN REPROD., NAT'L TOXICOLOGY PROGRAM, U.S. DEP'T OF HEALTH & HUMAN SERVS., DRAFT NTP-CERHR REPORT ON THE REPRODUCTIVE AND DEVELOPMENTAL TOXICITY OF BISPHENOL A (2006), available at http://ntp.niehs.nih.gov/ntp/ohat/bisphenol/Bisphenol_A_Draft_Report.pdf).

177. Susanne Rust et al., *Plastics Industry Behind FDA Research on Bisphenol A, Study Finds*, JOURNAL SENTINEL (Oct. 23, 2008), <http://www.jsonline.com/watchdog/watchdogreports/34469194.html>.

concerning the safety of BPA in February 2008.¹⁷⁸ In March of 2008, as part of its investigation of BPA, the U.S. House of Representatives Committee on Energy and Commerce and its Subcommittee on Oversight and Investigations sent an inquiry to The Weinberg Group asking questions about case studies reported on its website that “tout its successes in certain scientific and regulatory matters.”¹⁷⁹ Many of the case studies reflected involvement in advocacy science, noting such objectives as “delay[ing] cancellation of a new drug,” combining “epidemiological [sic] expertise with across-the-board strategic thinking,” development of a “defensible message,” and identifying “a national team of expert scientists” who “also prepared reviews for publication.”¹⁸⁰ Later, in October of 2008, after the FDA appointed a subcommittee to review its Draft Assessment of BPA, a research institute founded and co-directed by the subcommittee’s chairman was reported to have received five million dollars from an outspoken opponent of BPA regulations,¹⁸¹ who had reportedly expressed his views that BPA was “perfectly safe” to the chairman on several occasions.¹⁸²

Given the advocacy effort, it is not surprising that industry-funded studies managed to find their way to the forefront of regulatory review. According to an article authored by John Peterson Myers and Frederick S. vom Saal and signed onto by thirty-four scientists, most of whom were employed by national or international universities:

Despite strong evidence of aberrations caused by low doses of BPA in animals exposed during fetal and neonatal life in studies conducted by the world’s leading academic and government experts in the fields of endocrine disruption,

178. Letter from Steven R. Mason, Acting Assistant Comm’r for Legislation, FDA, Dep’t of Health & Human Servs., to John D. Dingell, Chairman, Comm. on Energy & Commerce, U.S. House of Representatives 2–3 (Feb. 25, 2008), *available at* http://www.bdlaw.com/assets/attachments/FDA_Letter_to_House_Energy_Committee_re_BPA.pdf.

179. Letter from John D. Dingell, Chairman, Comm. on Energy & Commerce, U.S. House of Representatives, & Bart Stupak, Chairman, Subcomm. on Oversight & Investigations, Comm. on Energy & Commerce, U.S. House of Representatives, to Matthew R. Weinberg, Chief Exec. Officer, The Weinberg Grp. at 1 (Mar. 6, 2008), *available at* http://democrats.energycommerce.house.gov/images/stories/Documents/investigations/public_health/Bisphenol.030608.Weinberg.pdf.

180. *Id.* at 1–2.

181. Susanne Rust & Meg Kissinger, *FDA Looks into Bisphenol A Advocate’s Donation to Science Center*, JOURNAL SENTINEL (Oct. 15, 2008), <http://www.jsonline.com/watchdog/watchdogreports/34469724.html>; Rust et al., *supra* note 177.

182. Rust et al., *supra* note 177.

endocrinology, neurobiology, reproductive biology, genetics, and metabolism, a relatively small number of studies reporting no adverse effects at low doses of BPA have continued to be promoted by the chemical industry and used by regulatory agencies.¹⁸³

Myers and vom Saal reported that the chemical industry managed to secure a position of superiority for their studies by pointing out that other studies did not conform to “good laboratory practices” (GLP).¹⁸⁴ Not only were industry representatives touting compliance with these GLP standards at the DART hearing,¹⁸⁵ but industry studies also rode this coattail through the FDA review process,¹⁸⁶ which created persuasive precedent for the industry-friendly decision by the DART Identification Committee. The protesting scientists explain, however, that “good laboratory practices,” are merely the name given to regulatory standards that involve certain record keeping and related requirements that are not generally the standard at small university laboratories.¹⁸⁷ The GLP rules were issued to address potential conflicts of interest and outright fraud by vested interests, and are arguably inappropriate in a university setting, where studies are publicly funded with no apparent conflicts of interest.¹⁸⁸ Moreover these scientists maintain that reliance on GLP confuses and merges the question of reliability (whether the

183. Myers et al., *supra* note 129, at 309–10.

184. *See id.* at 309–10, 314.

185. *See, e.g., DART Meeting on Proposition 65* (July 15, 2009), *supra* note 53, at 115, 123, 125–26 (statements of Dr. Rochelle W. Tyl, representing RTI International) (“[T]his was a guideline study. We did it under the U.S. EPA OPPTS testing guidelines, Office of Prevention—OPP, Pesticides and Toxic Substances.” “This study exceeded the OECD regulatory guidelines.” “So in conclusion based on our study guidelines—and our studies are guideline studies under good laboratory practices.”).

186. Myers et al., *supra* note 129, at 309. The FDA and the European Food Safety Authority (EFSA) each took actions that deemed two industry-funded GLP studies to be superior to hundreds of other publicly funded studies. *Id.*

187. *Id.* at 309. The requirements concern the “care and feeding of laboratory animals, standards for facility maintenance, calibration and care of equipment, personnel requirements, inspections, study protocols, and collection and storage of raw data.” *Id.*

188. *See id.* at 309–10. The rules were first issued by the U.S. FDA in 1978 in response to a situation of sloppy laboratory practices that were ultimately discovered to involve outright fraud. The discovery of these practices led the EPA to require reexamination of more than 4000 tests conducted by one of the largest private laboratories, and brought into question fifteen percent of the pesticides brought into use in the United States. *Id.*

results are replicable) with the separate question as to whether the methods used result in finding the truth.¹⁸⁹

Myers and vom Saal identify many flaws in the industry sponsored studies, including reporting an impossibly high prostate weight for control animals¹⁹⁰ (thus creating an inappropriate point of comparison for the BPA treated animals), selecting inappropriate animal models (use of a species of rat insensitive to estrogen), ignoring the failure of the positive control to show an effect (which “indicates the experiment failed”), ignoring an inordinately high dose required for the positive control (estradiol) to cause an effect (indicating that the system is insensitive to exogenous estrogens and thus inappropriate for studying BPA), ignoring test systems likely contaminated with estrogen (where responses of the negative control animals did not differ from the responses of animals given significant doses of the known estrogenic chemical, DES), and using “outdated and insensitive assays” incapable of detecting low dose endocrine-disrupting effects of BPA.¹⁹¹

The sheer number of reputable scientists reporting serious research design problems with industry sponsored studies suggests the need for new rules to create standards for research quality and design. Borrowing from another context, it may be instructive to reflect on the criteria that the Supreme Court identified in *Daubert v. Merrell Dow Pharmaceuticals, Inc.*, to address the question of whether potentially dubious evidence may be put before a jury.¹⁹² These standards are not generally applicable in the context of administrative law, where the agency is thought to have sufficient expertise to sort out the quality of the science.¹⁹³ However, here, in the context of BPA regulation, where there are questions about the

189. See Myers et al., *supra* note 129, at 309–10.

190. *Id.* at 311 (exceeding by seventy percent the prostate weights reported by other studies).

191. *Id.* at 310–12; see also Vom Saal & Welshons, *supra* note 13, at 63–66. (“When the positive control does not show a positive effect, one has to decide whether the system being studied is completely unresponsive to estrogenic stimulation . . . or whether there was contamination by estrogen that interfered with detection of an estrogenic response The purpose of including negative and positive controls for estrogenic activity and making comparisons to historic data on negative and positive control values from prior experiments is to be able to make this determination”) (citation omitted).

192. *Daubert v. Merrell Dow Pharm., Inc.*, 509 U.S. 579 (1993).

193. See Bayliss v. Barnhart, 427 F.3d 1211, 1218 n.4 (9th Cir. 2005) (explaining that *Daubert* does not govern the admissibility of evidence in an administrative proceeding where Federal Rule of Evidence 702 is inapplicable). *But see* Niam v. Ashcroft, 354 F.3d 652, 660 (7th Cir. 2004) (applying the “spirit of *Daubert*” to administrative proceedings).

expertise of the Committee (discussed in the next section) and especially given the complexity of endocrine-disrupting chemicals, some guidelines may be particularly useful.

The *Daubert* inquiry is flexible, but there are several questions the judge must consider. The first is whether the “theory or technique has been subjected to peer review and publication,” since submission to the scrutiny of the scientific community is considered “a component of ‘good science,’ in part because it increases the likelihood that substantive flaws in methodology will be detected.”¹⁹⁴ However, publication is not always a marker of good science. As Thomas McGarity and Wendy Wagner show in their book, *Bending Science*, the peer review process has proven incapable of consistently identifying and filtering out bent science.¹⁹⁵ In the ongoing battle over BPA, twenty-four scientists signed on as authors of a letter to the editor of *Toxicological Sciences* requesting that the journal adopt guidelines to screen out flawed research.¹⁹⁶ Identifying flaws in a study on BPA accepted for publication, the authors explained that the study did not establish the sensitivity of the test animal (the “LE rat”) to the positive control ethinylestradiol before determining what dose of BPA to test in their study.¹⁹⁷ They explained that the lowest effect dose of ethinylestradiol for the LE rat was “2500-fold higher than the maternal dose required to stimulate effects on offspring in mice,” and that the study reported no effect of ethinylestradiol at doses “sufficient to cause temporary sterility in 99.7% of women who properly use oral contraceptives.”¹⁹⁸ They further explained that a potential contributor to the low sensitivity to estrogen was the use of polycarbonate cages made from BPA.¹⁹⁹

Given the prevalence of advocacy science, rules establishing standards for appropriate research design may be more effective than relying on the peer review process alone. Following *Daubert*, additional considerations might also include the known or potential rate of error, the existence and maintenance of standards controlling the technique’s operation (which in this case could include the appropriate use of positive and negative controls), and “general acceptance,” meaning “explicit identification of a relevant scien-

194. *Daubert*, 509 U.S. at 593.

195. MCGARITY & WAGNER, *supra* note 171, at 10.

196. Frederick S. vom Saal et al., *Flawed Experimental Design Reveals the Need for Guidelines Requiring Appropriate Positive Controls in Endocrine Disruption Research*, 115 *TOXICOLOGICAL SCI.* 612, 612–13 (2010).

197. *Id.* at 612.

198. *Id.* at 612–13.

199. *Id.* at 613.

tific community and an express determination of a particular degree of acceptance within that community.’”²⁰⁰ Widespread acceptance is an important factor: “‘a known technique which has been able to attract only minimal support within the community’ . . . may properly be viewed with skepticism.”²⁰¹ Given the fierce debate between industry scientists on the one hand and government and university scientists on the other, a rule requiring consideration of general acceptance in the scientific community might have led to an entirely different outcome and a significant regulation of BPA.

IV. A MORE SPECIALIZED COMMITTEE, AN EXTENDED REVIEW PERIOD, AND CLEAR DISCLOSURE OF CONFLICTS OF INTEREST

Rules are also needed to establish more specialized qualifications for the review committee, to allow for a longer review period, and to require disclosure of conflicts of interest to assist the Committee in evaluating the “weight of the evidence” in this complicated and specialized area of science.

Endocrine disruptors are particularly complicated not only due to their breadth of action, but also the very different consequences they may have depending upon the age of exposure.²⁰² There is a lag between the time of exposure and the manifestation of a disorder, and different classes of endocrine disruptors may be additive or even synergistic.²⁰³ There may be transgenerational effects, affecting not just the exposed individual, but also children and subsequent generations.²⁰⁴ Endocrine disruptors can cause adverse effects at infinitesimally low levels of exposure, and may exert more potent effects at low doses than at higher doses; they have been known to have nontraditional dose-response curves, such as inverted-U or U-shaped curves.²⁰⁵

200. *Daubert v. Merrell Dow Pharm., Inc.*, 509 U.S. 579, 594 (1993) (quoting *United States v. Downing*, 753 F.2d 1224, 1238 (3d Cir. 1985)).

201. *Id.* (citation omitted) (quoting *Downing*, 753 F.2d at 1238).

202. See POSITION STATEMENT: ENDOCRINE-DISRUPTING CHEMICALS, *supra* note 10, at 1–2.

203. See *id.* at 1, 3.

204. *Id.* at 2.

205. See POSITION STATEMENT: ENDOCRINE-DISRUPTING CHEMICALS, *supra* note 10, at 2; Diamanti-Kandarakis et al., *supra* note 9, at 296.

The current standards do not require the DART review board to have any particular expertise in endocrinology. The regulatory standard for the composition of the review board requires only that they be “experts from among the following areas of specialization: epidemiology, developmental toxicology, reproductive toxicology, teratology, medicine, public health, biostatistics, biology, toxicology, and related fields.”²⁰⁶ The DART Identification Committee meets as infrequently as once a year, and it reviews a broader class of chemicals than just endocrine disruptors.²⁰⁷ The DART Identification Committee members at the time of the BPA hearing included a toxicologist from Chevron, a family practitioner who has since served as a tobacco industry spokesperson, and representatives from the following university departments: anatomy, epidemiology and preventative medicine, obstetrics and gynecology, pediatrics, nutrition, and family and preventative medicine.²⁰⁸ Responding to concerns about the adequacy of expertise of Committee members, the OEHHA Director has acknowledged that “the appointment of additional members to the committee with backgrounds in areas such as male reproductive hazards would benefit the committee’s overall review of certain chemicals for possible listing.”²⁰⁹

At the DART Identification Committee’s hearing, the discussion on BPA reflected confusion about basic principles of endocrinology. Dr. Keen, rephrasing statements by Dr. vom Saal, identified vom Saal’s hypothesis as identifying the possibility of a trimodal response where there are “very bad effects potentially at parts per trillion,” at “parts per billion it gets a little bit better,” and then at “parts per million maybe it gets worse again.”²¹⁰ At least one of the

206. CAL. CODE REGS. tit. 27, § 25302(b)(2) (2009).

207. CAL. CODE REGS. tit. 27, § 25302(c) (2009); CAL. CODE REGS. tit. 27, § 25305(b) (2012).

208. See *DART Meeting on Proposition 65* (July 15, 2009), *supra* note 53, at i, 2; e-mail from Monet Vela, Public Records Act Coordinator, Cal. Office of Env’tl. Health Hazard Assessment (OEHHA), to author (May 27, 2011, 11:42 AM) [hereinafter e-mail from Monet Vela] (on file with author) (containing resumes of DART Committee members); see also Letter from Dr. Sarah Janssen et al., *supra* note 116, at 1 (expressing concern regarding the lack of expertise of the DART Committee); Dan Morain, *Big Tobacco’s Unlikeliest Ally*, SACRAMENTO BEE, <http://www.sacbee.com/2012/05/06/4467266/big-tobaccos-unlikeliest-allyphysician.html> (last updated Sept. 11, 2012).

209. Letter from Joan Denton, *supra* note 117, at 2.

210. *DART Meeting on Proposition 65* (July 15, 2009), *supra* note 53, at 73–74 (statement of Dr. Carl Keen, Member, Developmental and Reproductive Toxicant Comm., California Office of Environmental Health Hazard Assessment) (rephrasing statements from Dr. Frederick vom Saal’s presentation).

Committee members seemed incredulous, noting, “[I]f it’s at high-dose levels, anything—particularly with meds, as we know, the higher you go with respect to dose, then you’re going to start to see some effects.”²¹¹ Another Committee member mentioned the possibility of a trimodal response, reflected on the need for additional studies and concluded: “I do have a fear. My fear is, is that we are—because we’re looking at the data the way we’re supposed to, and it’s as a whole that . . . we could be missing a clear and present danger.”²¹²

The science in this area is highly sophisticated and the language is difficult for someone outside the field to comprehend. Take for example, the following explanation from Dr. vom Saal at the hearing:

And the fact that Bisphenol A can alter epigenetic programming was demonstrated by Dolinoy a couple years ago, where they took a mouse with a retrotransposon, a gene spliced into the animal, that if it’s demethylated, and therefore active, then this gene causes obesity and a coat color change. What they demonstrated was that Bisphenol A led to a gene where there were no methyl groups available as opposed to the gene being normally silenced by being methylated. So this is a clear example of epigenetic programming and permanent silencing or activation of genes that totally alter the life history of the animal.²¹³

Simplifying the matter somewhat, he did show a picture of a rather fat mouse.²¹⁴ As confirmed later via e-mail, the above quoted language means that the study provided evidence that BPA alters genetic programming in such a way as to create a predisposition to obesity.²¹⁵

Even the most basic question as to the applicability of animal studies was at issue. Dr. vom Saal explained:

[T]hese cell culture studies indicate that at the cellular level, there’s no difference in response to Bisphenol A between rat,

211. *Id.* at 230 (statement of Dr. La Donna White, Member, Developmental and Reproductive Toxicant Comm., California Office of Environmental Health Hazard Assessment).

212. *Id.* at 252 (statement of Dr. Carl Keen, Member, Developmental and Reproductive Toxicant Comm., California Office of Environmental Health Hazard Assessment).

213. *Id.* at 71.

214. See e-mail from Dr. Frederick S. vom Saal, Curators’ Professor, Div. of Biological Sci., Coll. of Arts & Sci., Univ. of Mo.-Columbia, to author (May 29, 2011, 6:05 PM) (on file with author).

215. *Id.*

2012] ENDOCRINE-DISRUPTING CHEMICALS AND PROP 65 225

mouse, and human cells. There are some pharmacokinetic differences, but they're not anywhere near great enough to account for the effects that you're seeing down in the four part per trillion range.²¹⁶

However, Dr. Hentges, testifying for the American Chemistry Council, later provided another perspective, stating that there is "a significantly longer half-life for BPA in rodents and significantly greater systemic bioavailability" and that "[b]ecause of these key differences, extrapolation of any effects in rodent studies to humans would be tenuous. And in particular, effects that are observed in rodent studies are likely to over predict what could happen in humans."²¹⁷ In the final discussion, Committee Member White stated, "I can't see the extrapolation of the animals into human data."²¹⁸ Yet an international group of over thirty experts at a 2010 joint meeting of the World Health Organization (WHO) and the Food and Agriculture Organization of the United Nations (FAO) recently came to the opposite conclusion, finding that, given similarities in BPA metabolism, animal studies can be appropriately used for extrapolation to humans.²¹⁹

In addition to exposing the need for a more specialized review board, the discussion during the BPA hearing also suggests that a longer review period is warranted. With no proposed findings of fact or conclusions prepared by the full time OEHHA staff scientists, OEHHA asked far too much, especially of a Committee whose members hold full time professional positions and where their work on the Committee is outside their primary areas of expertise. At the end of the hearing, the DART Identification Committee was clearly overwhelmed. Committee Member Roberts noted the "huge num-

216. *DART Meeting on Proposition 65* (July 15, 2009), *supra* note 53, at 74–75.

217. *Id.* at 98, 103.

218. *Id.* at 230.

219. "Despite some differences between BPA metabolism and disposition in rodents and primates, internal exposures to aglycone BPA are remarkably similar for adult rodents, non-human primates and humans. This apparent lack of requirement for allometric scaling is atypical in the therapeutic drug and general chemical literature and suggests that a specific adjustment for interspecies differences in toxicokinetics is not required." FOOD & AGRIC. ORG. OF THE UNITED NATIONS & WORLD HEALTH ORG., JOINT FAO/WHO EXPERT MEETING TO REVIEW TOXICOLOGICAL AND HEALTH ASPECTS OF BISPHENOL A SUMMARY REPORT 12, 40–42 (2010), *available at* ftp://ftp.fao.org/ag/agn/agns/BPA_Summary_Report.pdf. According to Dr. vom Saal, this statement leaves no room for discussion as to the applicability of animal studies. e-mail from Dr. Frederick S. vom Saal, Curators' Professor, Div. of Biological Sci., Coll. of Arts & Sci., Univ. of Mo.-Columbia, to author (May 28, 2011, 1:14 AM) (on file with author).

ber of studies”²²⁰ and that he had to review “this entire binder of information.”²²¹ Dr. Hentges, testifying on behalf of the American Chemistry Council, made reference to NTP and European Union documents which hit the Committee members’ mailboxes “with a very heavy thud about a month ago.”²²² Committee Member Keen said, “I think the materials that we got were—I’ll use the word ‘overwhelming.’”²²³ As a point of reference, in preparation for the 2006 meeting on the state of knowledge on BPA, thirty-eight of the world’s leading scientific experts on BPA, organized as five panels of experts from different disciplines, prepared extensive working documents over a six month period reviewing different aspects of the BPA literature, covering in all over 700 published studies.²²⁴ There are now approximately 1000 articles relating to BPA.²²⁵

An inadequate period of study preceding the DART hearing also sets up a situation where the hearing itself may have assumed predominant importance that is particularly troubling in this setting where there is industry-funded opposition. After a neutral presentation by staff, the format set up the illusion of two sides battling in a fair playing field, where the interests of one side should be balanced against the other. There was no discussion of conflicts of interest, and industry representatives failed to clearly identify themselves. The lawyer representing the American Chemistry Council introduced Jay Murray of Murray and Associates without clarifying whether Mr. Murray was representing a client:

I don’t think you need any introduction to Jay Murray. But for the audience, Jay was many years ago a member of this scientific advisory panel. He’s an authority in this field. And he will speak to you about, what we call, the non-conventional studies, why they shouldn’t be relied upon to support a regulatory conclusion.²²⁶

220. *DART Meeting on Proposition 65* (July 15, 2009), *supra* note 53, at 235.

221. *Id.* at 238.

222. *Id.* at 98–99.

223. *Id.* at 238.

224. The meeting was sponsored by the National Institutes of Health, the EPA, and Commonweal (a non-profit health and environmental research group). *Chapel Hill*, *supra* note 14, at 131–32, 138 (these experts concluded that people are exposed to BPA at levels that cause problems in wildlife and laboratory animals and that there is “great cause for concern” with regard to the potential for similar adverse effects in humans); Pete Myers, ENVTL. HEALTH NEWS (Aug. 11, 2007), <http://www.environmentalhealthnews.org/newscience/2007/2007-0803chapelhillconsensus.html> (synopsis of *Chapel Hill*, *supra* note 14).

225. Myers et al., *supra* note 129, at 310.

226. *DART Meeting on Proposition 65* (July 15, 2009), *supra* note 53, at ii, 94.

2012] ENDOCRINE-DISRUPTING CHEMICALS AND PROP 65 227

Dr. Tyl, also speaking on the side of industry, stressed that she was not speaking as an “advocate,” identified herself as working for a “nonprofit contract research organization” that was eighty percent federally funded, and declined to mention whether the particular studies she was presenting were funded by industry.²²⁷

The issue of undisclosed conflicts of interest was discussed in the aftermath of the BPA hearing. According to a report by representatives from environmental and public health organizations who spoke with Committee members after the hearing, two panel members had expressed their belief that industry had not even been present at the hearing, and that the American Chemistry Council (ACC) was just a non-profit organization,²²⁸ presumably not understanding that it is a trade association representing chemical manufacturers whose stated mission is to improve the public image of the chemical industry and “deliver business value through exceptional advocacy.”²²⁹ Publicly discussing the issue of undisclosed conflicts of interest at a later hearing, the Committee was advised by its Chief Counsel that “the Open Meeting Act specifically says you cannot require someone to state their name, affiliation, or any other information if they want to speak in front of the group.”²³⁰ However, she also added that “[i]t doesn’t say you can’t ask.”²³¹ The Committee took no additional action to resolve the issue.

Assuming the accuracy of this interpretation of California’s Open Meeting Act,²³² another option would be to require OEHHA

227. *Id.* at ii, 112 (“What I thought I would do is go over the five studies that my staff and I have done at RTI with Bisphenol A. I’d like to just indicate that RTI is a nonprofit contract research organization. . . . We’re about 80 percent funded federally and about 20 percent funded commercially.”).

228. Letter from Dr. Sarah Janssen et al., *supra* note 116, at 2.

229. *About ACC*, AM. CHEMISTRY COUNCIL, <http://www.americanchemistry.com/About> (last visited Jan. 2, 2013).

230. *Meeting on Proposition 65 Before the Devtl. & Reprod. Toxicant Identification Comm. of the Office of Env’tl. Health Hazard Assessment of the State of Cal.* 103–04 (Oct. 21, 2010), available at http://oehha.ca.gov/prop65/public_meetings/pdf/DARTIC102110Transcript.pdf (statement of Carol Monahan-Cummings, Chief Counsel, California Office of Environmental Health Hazard Assessment).

231. *Id.* at 104 (statement of Carol Monahan-Cummings, Chief Counsel, California Office of Environmental Health Hazard Assessment).

232. The Act itself only specifically refers to conditions placed on “attendance at a meeting,” not speaking. CAL. GOV’T CODE § 11120 (West 2005); CAL. GOV’T CODE § 11124 (West 2005). However, in its interpretation of the Act, the California Attorney General’s Office provides an example that refers to both attendance and speaking: “For example, while the Act does not prohibit use of a sign-in sheet, notice must be clearly given that signing-in is voluntary and not a pre-requisite to either attending the meeting or speaking at the meeting.” OFFICE OF THE ATTORNEY GEN., CAL. DEP’T OF JUSTICE, A HANDY GUIDE TO THE BAGLEY-KEENE OPEN

staff to research and clearly identify issues concerning conflicts of interest in all studies that are considered and relied upon by the agency prior to any final decision. As has been recommended by Thomas McGarity and Wendy Wagner in the context of all policy-relevant research, there should be rules that require full disclosure of conflicts of interest, including the level of sponsor of any studies; disclosure of all affiliations, funding sources, and financial or management relationships; certification that all authors have agreed to be listed and have approved the manuscript; disclosure of the role that any sponsors played in study design, in the collection, analysis and interpretation of data, or in the writing of the report, or in the decision to submit the report for publication.²³³

V.

STANDARDS AND FINDINGS TO AVOID INAPPROPRIATE RELIANCE ON UNVERIFIED CLAIMS OF ADVERSE CONSEQUENCES

Another troubling issue at the DART hearing on BPA was the potential influence of testimony that brought evidence before the Committee that was both unverified and outside its statutory and regulatory purview. Industry representatives brought forward testimony consistent with their reported public relations strategy to highlight the costs of any restrictive regulatory measures by touting the benefits of BPA.²³⁴ As reported by the *Washington Post*, based on internal notes of a private meeting:

Industry representatives weighed a range of ideas, including “using fear tactics [e.g. “Do you want to have access to baby food anymore?” as well as giving control back to consumers (e.g. you have a choice between the more expensive product that is frozen or fresh or foods packaged in cans) as ways to dissuade people from choosing BPA-free packaging,” the notes said. The attendees estimated it would cost \$500,000 to craft a message for a public relations campaign, according to the notes. “Their ‘holy grail’ spokesperson would be a ‘pregnant

MEETING ACT 9 (2004), available at http://ag.ca.gov/publications/bagleykeene2004_ada.pdf. An analysis of any First Amendment concerns is outside the scope of this paper.

233. MCGARITY & WAGNER, *supra* note 171, at 237–38.

234. See Lyndsey Layton, *Strategy Being Devised to Protect Use of BPA*, WASH. POST (May 31, 2009), http://articles.washingtonpost.com/2009-05-31/politics/36808376_1_bpa-canned-goods-notes.

young mother who would be willing to speak around the country about the benefits of BPA,” the notes said.²³⁵

This strategy was in full force at the DART hearing on BPA. Although the testimony related to matters outside its charge, the DART Identification Committee nevertheless heard testimony about the benefits of BPA from both Dr. Hoyle from the North American Metal Packaging Alliance and Caroline Silveira from the Grocery Manufacturers Association.²³⁶ Ms. Silveira testified as to BPA’s “critical function in protecting the integrity of certain metal-packaging components” and its importance in helping foods to retain nutrition, quality, and consumer acceptability.²³⁷ She stated that listing would compromise the availability of safe, affordable and nutritious foods.²³⁸ Dr. Hoyle testified as to the “*potential health hazards*” that may arise from listing BPA as a Proposition 65 reproductive toxicant and from requiring warning labels.²³⁹ He said that BPA is “critical” to maintaining the sterility of canned food and eliminates the problem of swelled cans.²⁴⁰ Dr. Hoyle further stated that there would be public health consequences to deselecting epoxy coatings on metal cans, that they protect against botulism, and that they affect the most needy in our society.²⁴¹ He explained that the WIC Program (Special Supplemental Nutrition Program for Women, Infants, and Children)²⁴² and food pantries both rely on these epoxies, and that if products are required to be labeled, production will cease due to liability concerns.²⁴³ He talked about how metal-packaged products are important when there are disasters and also for the military.²⁴⁴

Dr. Hoyle further stated that the alternatives are untested, and that “there is no readily available, suitable, fully tested material that you can drop in as an alternative.”²⁴⁵ Alluding to Dr. vom Saal’s previous testimony referencing published research and multiple data sets showing a fifty percent drop in BPA exposure in Japan after changing the can lining away from BPA to polyethylene ter-

235. *Id.*

236. *DART Meeting on Proposition 65* (July 15, 2009), *supra* note 53, at 87, 139.

237. *Id.* at 88, 89.

238. *Id.* at 90.

239. *Id.* at 139–40 (emphasis added).

240. *Id.* at 141.

241. *Id.* at 142.

242. *About WIC*, USDA FOOD & NUTRITION SERV., <http://www.fns.usda.gov/wic/> (last visited Dec. 29, 2012).

243. *DART Meeting on Proposition 65* (July 15, 2009), *supra* note 53, at 142–44.

244. *Id.* at 144.

245. *Id.* at 146.

epththalate,²⁴⁶ Dr. Hoyle stated that even the “PET” that was used in Japan relies on epoxies and that PET is not a barrier to the epoxy migrating, so “the same amount of BPA go[es] through whether the PET is there or not.”²⁴⁷

Particularly troubling about this part of the hearing on BPA was the lack of any supporting evidence or evaluations of these statements. There was no indication that any of these statements as to the necessity of can linings, the extent of any public health hazard, or the feasibility and effectiveness of switching to PET were based on any study, scientific review, or analysis. Presumably because this testimony was irrelevant to the question of whether BPA has been clearly shown to be a reproductive toxicant and thus outside the purview of the Committee, no studies appear to have been submitted on this issue. There was no staff review or analysis of these issues in the staff report or presentation. The information is suspiciously incomplete and seemingly inaccurate given that some companies have now made public statements that they are selling BPA-free cans.²⁴⁸ Kathleen Roberts, a lobbyist for the North American Metal Packaging Alliance, which represents the makers of metal cans and their customers, acknowledged that “alternatives are available but not for all uses currently in the marketplace.”²⁴⁹

One practical option would be to explicitly require OEHHA staff to evaluate and discuss all information presented to the Board, including any unverified claims of adverse consequences. However, staff review and discussion of the availability of alternatives may effectively elevate the importance of this discussion and suggest that it is legally relevant to the decision to list a chemical under Proposition 65. As discussed by Daniel Farber in his article *Rethinking the Role of Cost-Benefit Analysis*, “Importing legally irrelevant factors into a decision violates the basic precepts of modern administrative law,”²⁵⁰ and it’s unlikely that OEHHA could successfully interpret the Proposition 65 standard, “clearly shown to cause reproductive

246. *Id.* at 75.

247. *Id.* at 146.

248. See, e.g., *Eden Foods Bisphenol-A (BPA) Free Pioneer*, EDEN ORGANIC, http://www.edenfoods.com/articles/view.php?articles_id=178 (last visited Jan. 3, 2013); Alice Wessendorf, *Desperately Seeking BPA-Free Canned Foods*, HEALTHIER TALK (Dec. 7, 2010, 8:00 AM), <http://www.healthiertalk.com/desperately-seeking-bpa-free-canned-foods-2948> (listing seven companies which offer BPA-free canned foods).

249. Layton, *supra* note 234.

250. Daniel A. Farber, *Rethinking the Role of Cost-Benefit Analysis*, 76 U. CHI. L. REV. 1355, 1378 (2009) (book review); see also *Citizens to Preserve Overton Park, Inc. v. Volpe*, 401 U.S. 402, 416 (1971) (requiring the reviewing court to consider “whether the decision was based on a consideration of the relevant factors”).

toxicity,” to legally require considerations extraneous to the question of toxicity. In the context of the Clean Air Act, for example, in *Whitman v. American Trucking Associations*, the Supreme Court interpreted a standard similarly focused on toxicity to exclude cost considerations. The statutory standard in *Whitman* required that the National Ambient Air Quality Standards (NAAQS) be standards “‘which in the judgment of the Administrator, based on [the] criteria . . . and allowing an adequate margin of safety, are requisite to protect the public health.’”²⁵¹ The Supreme Court found implausible industry’s argument that the “terms ‘adequate margin’ and ‘requisite’ leave room to pad health effects with cost concerns.”²⁵² The Court reasoned that the cost of implementation is “both so indirectly related to public health and so full of potential for canceling the conclusions drawn from direct health effects” that it would have to have been mentioned in the statute if Congress had wanted it to be taken into consideration.²⁵³

Responding to a delegation challenge, the Supreme Court in *Whitman* explained that it has upheld agency implementation of other equally indeterminate standards²⁵⁴ and that it has never demanded a “determinate criterion” for saying “how much [of the regulated harm] is too much.”²⁵⁵ The practical reality, however, is that if there are no determinate criteria, the decision will still be made with reference to some sort of context, and the cost to industry is unlikely to be ignored. According to credible observers, in actually setting the standards at issue in *Whitman*, “the EPA had in fact considered costs, although tacitly and without public supervision.”²⁵⁶ In his concurring opinion, Justice Breyer acknowledged and condoned at least some consideration of context. It was his opinion that the words “‘requisite to protect the public health’ with ‘an adequate margin of safety’ . . . do not describe a world that is

251. *Whitman v. Am. Trucking Ass’ns*, 531 U.S. 457, 466 (2001) (alteration in original) (quoting 42 U.S.C. § 7409(b)(1) (2006)).

252. *Id.* at 468.

253. *Id.* at 469 (emphasis omitted).

254. *Id.* at 473–74 (including “necessary to avoid an imminent hazard to public safety” and “set the standard which most adequately assures, to the extent feasible, on the basis of the best available evidence, that no employee will suffer any impairment of health” (internal quotation marks omitted)).

255. *Id.* at 475 (alteration in original) (internal quotation marks omitted). “The constitutional question is whether the statute has delegated legislative power to the agency.” *Id.* at 472. The Court of Appeals had reasoned that the EPA “‘lack[ed] any determinate criteria for drawing lines. It had failed to state intelligibly how much is too much.’” *Whitman*, 531 U.S. at 472 (alteration in original) (quoting *Am. Trucking Ass’ns v. EPA*, 175 F.3d 1027, 1034 (D.C. Cir. 1999)).

256. Cass R. Sunstein, *Regulating Risks After ATA*, 2001 SUP. CT. REV. 1, 11.

free of all risk—an impossible and undesirable objective.”²⁵⁷ He explained that the EPA can consider background circumstances when deciding “‘what risks are acceptable in the world in which we live.’”²⁵⁸ He felt that the statute permitted consideration of “comparative health risks,” such as health risks that may stem from reducing “tropospheric ozone (which, it is claimed, helps prevent cataracts and skin cancer).”²⁵⁹ Perhaps quietly acknowledging the reality that cost considerations must have some place in the process, the majority opinion recognized that cost could be taken into consideration at a later point in the administrative process, that is, when regulators determine how to implement ambient air quality standards.²⁶⁰

In the case of Proposition 65, however, there is no point later in the process for consideration of consequences. Granted the statute requires only a warning, not a restriction or a ban, still Proposition 65 has been known to have a significant effect on the market and related costs are unlikely to be entirely ignored by decision makers. In this context, it is particularly important that rules create reasonably clear standards. Although it may be impossible to keep extraneous considerations completely outside consideration (especially when faced with an ongoing public relations campaign), clear guidelines and mandatory written findings of fact would create a more transparent process and help focus the Committee’s attention on factors relevant to the scientific questions concerning the sufficiency of the evidence on toxicity.

Another option to keep inflated claims of adverse consequences outside consideration would be to adopt an amendment to Proposition 65 that would explicitly allow for staff evaluation of all information before the Board, including information extraneous to the question of toxicity. However, any such law would have to be carefully written so as not to override Proposition 65’s ultimate ob-

257. *Whitman v. Am. Trucking Ass’ns*, 531 U.S. 457, 494 (2001).

258. *Id.* at 495 (Breyer, J., concurring) (quoting *Natural Res. Def. Council, Inc. v. EPA*, 824 F.2d 1146, 1165 (D.C. Cir. 1987)).

259. *Id.* at 495 (Breyer, J., concurring).

260. *See id.* at 466–67 (discussing CAA statutory provisions specifically allowing for waiver of compliance deadlines and consideration of economic costs in setting standards of performance for new sources, setting compliance deadlines for emissions standards for automobiles, fuel additives, and aircraft emission standards, and in performing various other duties); *see also id.* at 493 (Breyer, J., concurring) (explaining that “[s]tates may consider economic costs when they select the particular control devices used to meet the standards” and that “industries experiencing difficulty in reducing their emissions can seek an exemption or variance from the state implementation plan”).

jective to enable the purchaser to make an informed decision as to whether there are better alternatives and allow the market to respond accordingly. The focus should remain on the question of toxicity.

CONCLUSION

In the DART hearing on BPA, the outcome might well be expected given the lack of any special expertise on the part of the board, very complex material, a very short review period, no clear understanding of “reproductive toxicity,” inadequately disclosed conflicts of interest, and a dearth of regulatory guidance that left the door open to an interpretation that would require conclusive evidence. Given the prevalence of advocacy science, there was, of course, a body of conflicting evidence.

For Proposition 65 to fully meet its regulatory objectives and truly allow for public disclosure of endocrine-disrupting chemicals, OEHHA needs to adopt clear regulatory standards with requirements for findings of fact that increase transparency and accountability. Standards must be adopted that will separate as clearly as possible the science from the policy decisions. Without better standards, an appointed Committee is left to quietly substitute its policy judgments for those of the public. What is needed is an honest and public admission of what is science, and what is policy, so that the public may take corrective action as may be necessary to further public policy objectives.

In contrast to the decision from the FDA’s Science Board on BPA, there was no attempt by the DART Identification Committee to separate the scientific decisions from the policy decisions. Despite a standard that calls for “reasonable certainty in the minds of competent scientists,”²⁶¹ the FDA’s Science Board on BPA recognized that the decision was really one of mixed science and policy and that they alone were not the arbitrators of the regulatory decision.²⁶² Standards as to acceptable margins of safety were clearly guiding the process.²⁶³ However, given the regulatory malaise that

261. Food Additives Amendment of 1958, Pub. L. No. 85-929, 72 Stat. 1784; S. REP. NO. 85-2422, at 2–3 (1958), *reprinted in* 1958 U.S.C.C.A.N. 5300, 5301–02 (describing the relevant standard as one of “reasonable certainty in the minds of competent scientists that the additive is not harmful to man or animal”).

262. *See* FDA SCIENCE BOARD SUBCOMMITTEE ON BISPHENOL A, *supra* note 18, at 4.

263. *Id.* (“[T]he Margins of Safety defined by [the] FDA as ‘adequate’ are, in fact, inadequate. This does not mean that the potential exposures are not ‘acceptable.’”)

followed, standards are still needed to guide the policy decisions themselves.

OEHHA has publicly taken the position that the ultimate question under Proposition 65 is purely a question of “science,”²⁶⁴ yet in its refusal to assist the Committee by proposing any conclusions as to the sufficiency of the evidence, the OEHHA staff appear to quietly recognize themselves that there is considerable room for judgment that may fall outside the realm of science: “OEHHA staff avoid making specific arguments for or against the listing of any given chemical since this decision is entirely within the purview of the expert committees.”²⁶⁵ Yet a full time staff with over a year of study on the issue²⁶⁶ would seem to be in a better position than a part-time committee with no particularly specialized expertise and a very short review period to make at least draft recommendations on the science.

The problems with clarity and transparency identified in this article are not unique to Proposition 65 but reflect a larger problem in regulatory decisionmaking involving scientific assessments. In its ongoing review of the EPA’s *Draft IRIS Assessment of Formaldehyde*, a committee of the National Research Council identified problems “similar to those which have been reported over the last decade by other NRC committees,” including “problems with clarity and transparency of the methods,” the role of guidelines in the “preparation of the assessment,” lack of consistency with no clear “underlying conceptual framework,” and inadequate “documentation of methods and criteria for identifying evidence . . . [and] critically evaluating . . . studies.”²⁶⁷

With minor regulatory reform, ideally through a publicly accessible and transparent rulemaking process involving the scientific

264. See Letter from Joan Denton, *supra* note 117, at 2–3 (“This standard is not a legal determination; it is instead a scientific judgment in which the state’s qualified experts are expected to apply their own knowledge and expertise to determine if a chemical has been ‘clearly shown by scientifically valid testing according to generally accepted principals to cause reproductive toxicity.’”).

265. *Id.* at 2.

266. CAL. ENVTL. PROT. AGENCY, *supra* note 25, at 8 (“OEHHA had selected BPA through its prioritization process as a candidate for consideration by the DART IC, and substantial staff work on preparation of hazard identification materials had already occurred, before the NTP-CERHR Monograph was published [in September 2008].”).

267. REVIEW OF THE ENVIRONMENTAL PROTECTION AGENCY’S DRAFT IRIS ASSESSMENT OF FORMALDEHYDE, Comm. to Review EPA’s Draft IRIS Assessment of Formaldehyde, Bd. on Envtl. Studies and Toxicology, Div. on Earth & Life Studies, Nat’l Research Council 4 (2011), available at http://www.nap.edu/catalog.php?record_id=13142.

2012] ENDOCRINE-DISRUPTING CHEMICALS AND PROP 65 235

community, Proposition 65 could effectively provide for public disclosure and allow for a precautionary approach to endocrine-disrupting chemicals. OEHHA should adopt rules to define “reproductive toxicity” broadly and consistently with the evolving science recognizing the breadth of action of endocrine-disrupting chemicals. The rules should be broad enough to include all serious adverse health effects relating to the functioning of the endocrine system, including all effects that cause people pain, cost society money, and lead to serious health problems, like precancerous conditions and obesity. The proposed reform is particularly important today following the dissolution of the NTP-CERHR and the increased importance of the DART Identification Committee review as a listing mechanism. OEHHA recognition of a broad definition of “reproductive toxicity” may also align the practice and goals of the DART Identification Committee with the National Toxicology Program’s new Office of Health Assessment and Translation, and encourage future designation of the latter as an authoritative body.

OEHHA should also adopt rules to define “clearly shown” that include standards to assess and weigh advocacy science. Such standards should consider the basics of appropriate research design and general acceptance of the work within the scientific community. There should also be clear and detailed standards to require disclosure of all conflicts of interest, related both to the studies relied upon, and to the testimony at hearings to consider the listing of chemicals under Proposition 65.

Clear standards coupled with required findings of fact would help focus the Committee on appropriate issues in the face of any exaggerated claims of adverse consequences due to the alleged unavailability of alternatives. Rules recognizing and sorting the policy questions from the science questions and requiring the DART Identification Committee to issue findings of fact may encourage the staff to provide more comprehensive assistance, ideally including proposed findings of fact, without fear of overstepping into the role of policy-making.

Ideally there should be a transparent discussion of existing alternatives including a serious evaluation of any exaggerated claims of adverse consequences. However, it would be difficult to institutionalize this discussion, especially under the current legal and regulatory regime. The question of alternatives may ultimately be best addressed pursuant to California’s Green Chemistry Initiative. In 2008, California Senate Bill 509 established a Toxics Information

Clearinghouse,²⁶⁸ and California Assembly Bill 1879 required California's Department of Toxic Substances to adopt regulations that would establish a process to identify and prioritize chemicals of concern, identify alternatives, and consider requirements for labeling restrictions or prohibitions.²⁶⁹ However, this effort too has fallen subject to criticism for failing to identify sufficient legal standards. Commenting on the informal draft rule, Joseph Guth of the University of California, Berkeley Center for Green Chemistry requested that the agency articulate a transparent standard as to how the conflict between the "interests in environmental health and economic factors are 'best' balanced": "Without an articulated standard, there is no hope of . . . decisions being transparent, consistent or accountable to the public."²⁷⁰ The final regulatory structure of California's Green Chemistry initiative remains to be seen.²⁷¹

Meanwhile Americans may receive some protection from the transnational reach of the European Union's toxic substance law: Registration, Evaluation, and Authorization of Chemicals (REACH).²⁷² For authorization to use chemicals of "very high concern," REACH requires applications to include an "analysis of alternatives, considering their risks and the technical and economic feasibility of substitution."²⁷³

268. S.B. No. 509, 2007–2008 Sess. (Cal. 2008) (codified at CAL HEALTH & SAFETY CODE §§ 25251, 25256–25257.1 (West Supp. 2012)).

269. See Assemb. B. No. 1879, 2007–2008 Sess. (Cal. 2008) (codified at CAL. HEALTH & SAFETY CODE §§ 25252–25255, 25257 (West Supp. 2012)).

270. Letter from Joseph H. Guth, Ph.D., J.D., Member, Green Ribbon Sci. Panel, Sci. & Envtl. Health Network, Univ. of Cal., Berkeley Ctr. for Green Chemistry, to Debbie Raphael, Dir., Cal. Dep't of Toxic Substances Control 2 (Dec. 31, 2011), available at <http://www.dtsc.ca.gov/LawsRegsPolicies/Regs/upload/SCPInformalComments201201Web1.pdf> (no. (IC)19 of Cal. Dep't of Toxic Substances Control compilation of comments on the informal draft rule).

271. Assembly Bill 1879 required that regulations be adopted by January 1, 2011; however, the Safer Consumer Products regulations are yet to be finalized. Assemb. B. No. 1879; see *Safer Consumer Products Regulations*, CAL. DEP'T OF TOXIC SUBSTANCES CONTROL, <http://www.dtsc.ca.gov/SCPRegulations.cfm> (last visited Jan. 3, 2013) (indicating several different notice and comment periods for the regulations, including one ending January 22, 2013); Safer Consumer Products Proposed Regulations, R-2011-02 (proposed July 2012) (to be codified at CAL. CODE REGS. tit. 22, §§ 69501–69599), available at <http://www.dtsc.ca.gov/upload/SCPProposedRegulationsNoUnderlineJuly2012.pdf>.

272. See Noah M. Sachs, *Jumping the Pond: Transnational Law and the Future of Chemical Regulation*, 62 VAND. L. REV. 1815, 1819, 1822 (2009).

273. Regulation (EC) 1907/2006 of the European Parliament and of the Council of 18 December 2006 Concerning the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH), Establishing a European Chemicals Agency, Amending Directive 1999/45/EC and Repealing Council Regulation

Finally, at least in so far as its work involves the review of endocrine-disrupting chemicals, OEHHA should adopt rules that require the DART Identification Committee, or a subcommittee, to have a high level of expertise in the area of endocrinology and that allow for a longer review period as necessary to consider the complexity of the science, the number of studies under review, and any time demands that may be involved in identifying and dealing appropriately with advocacy science.

The issue of adequate expertise of boards and scientific peer review panels to address endocrine-disrupting chemicals has been recognized at the national level. In 2011, a bill was introduced in the House of Representatives that would establish an Endocrine Disruption Expert Panel to guide federal regulatory decisions regarding endocrine-disrupting chemicals in the context of several existing federal health and environmental statutory programs.²⁷⁴ Members of the panel would have “established expertise in the field of endocrine disruption research by publishing research in peer-reviewed literature and have received Federal endocrine-research-related funding within the 2 years preceding appointment.”²⁷⁵ California should consider a similar standard for the DART Identification Committee. The federal bill also recognizes the “need to educate the public on the results of research on endocrine-disrupting chemicals so that manufacturers, processors, retailers, and individual consumers can make informed decisions about potential exposures to harmful chemicals.”²⁷⁶ The bill declares that “[p]eople should be protected from chemicals that are found to have endocrine-disrupting effects.”²⁷⁷

With a little reform to the regulatory framework, Proposition 65 could live up to the expectations of the California voters and create an opportunity for people to protect themselves from endocrine-disrupting chemicals. To fully inform the public, require-

(EEC) 793/93 and Commission Regulation (EC) 1488/94 as well as Council Directive 76/769/EEC and Commission Directives 91/155/EEC, 93/67/EEC, 93/105/EC and 2000/21/EC, 2006 O.J. (L396) 1, 25 (EC), *available at* <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2006:396:0001:0849:EN:PDF>.

274. *See* Endocrine-Disrupting Chemicals Exposure Elimination Act of 2011, H.R. 2521, 112th Cong. §§ 4(2), 101(a)(2) (2011).

275. *Id.* § 101(b)(1)(A).

276. *Id.* § 3(6). The bill would also establish a research program and a hazard classification system for endocrine-disrupting chemicals. *See id.* § 101(a)(1), (2). It would require federal agencies to develop strategies for reducing exposure, and require exposure pathways to be mitigated where the expert panel has identified a chemical as being of a high level of concern. *See id.* §§ 101(a)(4), 201(a), (c)(1).

277. *Id.* § 3(7).

238 NYU ANNUAL SURVEY OF AMERICAN LAW [Vol. 68:185

ments for warning labels must also be reformed to specifically identify the chemical and the source of potential exposure.

WHEN BATHTUB CROCODILES ATTACK: THE TIMING AND PROPRIETY OF CAMPAIGNING BY JUDICIAL RETENTION ELECTION CANDIDATES

DAVID W. EARLEY*

“I did not seek this fight, but I will not shrink from it.”
—California Chief Justice Rose Bird, 1986¹

“We’re not forming campaign committees. We’re not going to
become politicians.”
—Iowa Chief Justice Marsha Ternus, 2010²

Introduction	240	R
I. Retention Elections: Then and Now	243	R
A. Rubber Stamp Electorate: Historical Trends in Judicial Retention Elections	243	R
1. History of Merit Selection	245	R
2. Historical Trends in Retention Election Outcomes	247	R
3. California 1986	249	R
B. Rubber Stamp No More: Judicial Retention Elections in 2010	252	R
1. Iowa 2010	253	R
2. Illinois 2010	256	R
3. Importance, or Unimportance, of 2010	258	R
II. Reconsideration of Retention Election Electioneering Rules	259	R
A. Competing Goals Surrounding Retention Elections	259	R
1. Judicial Impartiality	260	R
2. Judicial Accountability	262	R

* Counsel, Brennan Center for Justice at New York University School of Law. The author is immensely grateful to J. Adam Skaggs for his invaluable advice in the drafting of this article in addition to Jonathan Backer for his research assistance. The views expressed in this article are solely those of the author.

1. *CBS Evening News* (CBS television broadcast May 21, 1986), available at <http://www.youtube.com/watch?v=Kd162US36to>.

2. Rod Boshart, *Ternus: Iowa Judges Won't Become Politicians*, E. IOWA GOV'T (Aug. 28, 2010, 10:12 PM), <http://easterniowagovernment.com/2010/08/28/ternus-iowa-judges-wont-become-politicians/>.

240	NYU ANNUAL SURVEY OF AMERICAN LAW	[Vol. 68:239]	
	3. Informed Electorate	264	R
	B. Current Rules Governing Retention Elections....	266	R
	1. Judicial Codes of Conduct	266	R
	2. Three Types of Retention Election Campaign Rules	266	R
	3. History of the ABA Model Rules	271	R
	4. Evaluation of Electioneering Rules.....	274	R
	III. Should Judges Facing Active Opposition Campaign? .	280	R
	Conclusion	287	R

INTRODUCTION

Judicial retention elections have historically been low-key affairs. Voters, often largely ignorant of the candidates for whom they're voting,³ usually give a rubber stamp “yes” vote in sufficient numbers to outnumber those who voted “no” instead. Many of those who physically make it to the polls fail to vote in the retention election at all.⁴ In the vast majority of cases, the judge is retained and moves on to another multi-year term.

Times are changing, however. While most judicial retention elections are still relatively unnoticed by the public, some recent retention contests have become money-soaked, political maelstroms—and signs suggest that this trend will continue.⁵ If the 2010 retention elections in Iowa and Illinois are any example, future retention elections may feature a public bombarded by advertisements decrying “activist”⁶ and “out of control” judges who threaten

3. See G. Alan Tarr, *Do Retention Elections Work?*, 74 MO. L. REV. 605, 625 (2009) (citing JUSTICE AT STAKE CAMPAIGN, JUSTICE AT STAKE FREQUENCY QUESTIONNAIRE 4 (2001), available at http://www.justiceatstake.org/media/cms/JASNationalSurveyResults_6F537F99272D4.pdf).

4. See *infra* Part I.A.2.

5. See *infra* note 230 and accompanying text.

6. For a discussion on “activist” judges, see *A Conversation About Judicial Independence and Impartiality*, 89 JUDICATURE 339, 343 (2006) [hereinafter *Conversation*] (statement of Shirley Abrahamson, C.J.) (“None of us can pull or change or temper a decision because we are concerned that somebody might say—and this is a code word—‘you are an activist judge.’ ‘Activist judge’ means the person doesn’t like the decision. If you understand that, you understand everything that’s happening in the United States, I think, to judges.”).

2012] CAMPAIGNING BY JUDICIAL RETENTION CANDIDATES 241

the rights of the people.⁷ Some judges will fight back,⁸ while others may choose not to do so.⁹

The pressure that elected judges feel when deciding a controversial case was framed by the late Justice Otto Kaus of the Supreme Court of California as the “crocodile in the bathtub” dilemma: “[I]t [is] like finding a crocodile in your bathtub when you go in to shave in the morning. You know it’s there, and you try not to think about it, but it’s hard to think about much else while you’re shaving.”¹⁰ To extend Kaus’s metaphor, a judge today has two options should the crocodile attack. He can either dull his razor fighting off the beast or instead hope to survive the attack without giving up the ultra-fine edge of his blade so that, afterward, he can still get that perfect shave.¹¹ In other words, the dilemma becomes: should judges campaign, risking the possibility that they’ll be cynically viewed as “politicians in robes,”¹² or should they allow public opinion to take its course, for good or ill?

After retention elections involving extensive electioneering, regardless of the outcome, the public is often left with a lack of confidence in the judiciary out of fear either that judges are “legislating from the bench” or that they’re bought and paid for by powerful special interests. In both cases, the judiciary may be seen as a political entity—a result seemingly inevitable once campaigning heats up. Be that as it may, states that currently hold judicial elections are unlikely to shift away from them in the near future,¹³ raising the

7. See, e.g., IowaForFreedom, *TV Ad: Send Them a Message*, YOUTUBE (Sept. 13, 2010), <http://www.youtube.com/watch?v=Y0Or8tGuleY> [hereinafter *Send Them a Message*].

8. See Larty Aspin, *The 2010 Judicial Retention Elections in Perspective: Continuity and Change from 1964 to 2010*, 94 JUDICATURE 218, 227–30 (2011) (providing an example of a judge who “ran an active and expensive campaign for retention”).

9. See Mark Curriden, *Judging the Judges: Landmark Iowa Elections Send a Tremor Through the Judicial Retention System*, A.B.A. J., Jan. 2011, at 56–57 (explaining that the Iowa justices facing retention in 2010 did not campaign or raise money).

10. Gerald F. Uelman, *Crocodiles in the Bathtub: Maintaining the Independence of State Supreme Courts in an Era of Judicial Politicization*, 72 NOTRE DAME L. REV. 1133, 1133 (1997).

11. Male pronouns are used only in the interest of extending the simile; no disrespect is intended toward female members of the judiciary and the legal profession.

12. Sandra Day O’Connor, *The Essentials and Expendables of the Missouri Plan*, Speech Before the University of Missouri School of Law (Feb. 27, 2009), in 74 MO. L. REV. 479, 489 (2009) (modified for publication purposes).

13. Tiffany L. Carwile, Note, *Stop Restricting Speech and Educate the Public: A Review of the ABA’s Proposed Campaign Activity Canon of the Model Code of Judicial Conduct*, 15 WM. & MARY BILL RTS. J. 1053, 1085 (2007); Ofer Raban, *Judicial Impartiality and the Regulation of Judicial Election Campaigns*, 15 U. FLA. J.L. & PUB. POL’Y. 205,

need for an exploration of appropriate campaign conduct in retention elections. Given the events of the 2010 Iowa and Illinois elections, the promise of engaged opposition campaigns in 2012, and the vast unawareness regarding retention elections generally, the time is ripe for reevaluating the rules and norms governing retention contest electioneering¹⁴ to determine whether they are serving societal goals. While state constitutions generally rigidly fix the states' methods of judge selection, the rules governing judicial electioneering are fairly malleable due to their relative ease of amendment and the lack of public attention that such rules receive.

This Article focuses on two related issues: what set of rules should be adopted to govern judicial electioneering and whether judges should campaign in response to opposition. Part I places current trends in retention elections in context by providing a historical overview. Part II examines the different fundraising and campaigning rules that exist among different retention election systems to determine which are most desirable. Every state has adopted some form of one of three rule sets to govern when retention election candidates may campaign, and each rule set permits electioneering under different circumstances: the Active Opposition Rule (after active opposition appears), the Candidacy Rule (after the judge declares his or her candidacy), and the Fixed Time Rule (during a set time window before the election). After considering each rule's impact upon judicial impartiality, judicial accountability, and an informed electorate, the Article concludes that the Fixed Time Rule is the most desirable. Part III addresses the new crocodile in the bathtub dilemma: whether candidates should campaign in response to an opposition effort. While there are reasonable arguments supporting the idea that judges facing active opposition should not campaign,¹⁵ such activity is unlikely to further politicize an election already politicized by the opposition. Such campaigning can actually help protect the impartiality of the judiciary.

Note that this Article does not attempt to evaluate and compare different methods of judicial selection, nor does it endorse any

208 (2004) (“[N]o state currently holding judicial elections is likely to do away with them in the near future.”). *But see* ADAM SKAGGS ET AL., *THE NEW POLITICS OF JUDICIAL ELECTIONS 2009–10*, at 23–26 (Charles Hall ed., 2011), <http://www.brennancenter.org/page/-/Democracy/NewPolitics2010.pdf> (describing recent legislative efforts to modify or abolish merit selection).

14. The term “electioneering” is used throughout this article to refer collectively to the acts of fundraising and campaigning, which are sometimes regulated by different timing rules within the same state.

15. *See infra* Part III.

particular selection system.¹⁶ Instead, the focus is solely on issues within the preexisting retention election framework, based on the theory that those states that use and will continue to use such contests may benefit from a thorough analysis of the rules that govern the timing of electioneering during retention elections. This Article also provides guidance to judges deciding whether to campaign in response to opposition.

I.

RETENTION ELECTIONS: THEN AND NOW

A. *Rubber Stamp Electorate: Historical Trends in Judicial Retention Elections*

Retention elections are generally the final step in the method of judicial selection commonly referred to as merit selection¹⁷ or the Missouri plan.¹⁸ In the typical merit selection system, an independent judicial selection commission, comprised of lawyers and non-lawyers appointed by a variety of public and private officials, compiles a list of three to five candidates and presents it to the gov-

16. Examples of different forms of selection methods include partisan elections, nonpartisan elections, merit selection, and pure appointment. There is already a vast literature elsewhere exploring the concerns of choosing among the methods. See, e.g., Michael S. Kang & Joanna M. Shepherd, *The Partisan Price of Justice: An Empirical Analysis of Campaign Contributions and Judicial Decisions*, 86 N.Y.U. L. REV. 69 (2011); David E. Pozen, *The Irony of Judicial Elections*, 108 COLUM. L. REV. 265 (2008); Joanna M. Shepherd, *Money, Politics, and Impartial Justice*, 58 DUKE L.J. 623 (2009); Roy A. Schotland, *New Challenges to States' Judicial Selection*, 95 GEO. L.J. 1077 (2007); Ryan L. Souders, Note, *A Gorilla at the Dinner Table: Partisan Judicial Elections in the United States*, 25 REV. LITIG. 529 (2006).

17. Some, perhaps not unreasonably, take exception to this term, preferring "commission selection." See, e.g., Peter D. Webster, *Selection and Retention of Judges: Is There One "Best" Method?*, 23 FLA. ST. U. L. REV. 1, 29 n.186 (1995) ("It is unclear from where this label first came. However, what is clear is that this value-laden label is, today, a deliberate choice of most proponents of the system."). Most of the current literature, however, appears to prefer the traditional "merit selection" and "Missouri plan" labels, so that terminology is used here for simplicity and clarity.

18. "Although not the first state to adopt a commission plan, Missouri's system has been the one most frequently looked to as a model." Webster, *supra* note 17, at 29 n.185. The predecessor California plan differed from the Missouri plan, however.

They approved a proposal calling for nominations of candidates to fill vacancies on the supreme court and the courts of appeal by the governor, subject to confirmation by a commission composed of the attorney general, the chief justice of the supreme court, and a presiding justice of the courts of appeal.

Once confirmed, the judge was required to stand for retention at the next gubernatorial election and, thereafter, at regular intervals.

Id. at 29–30 (footnotes omitted).

ernor.¹⁹ The governor then selects one of those candidates for the bench; some states also require confirmation by one or both houses of the legislature.²⁰ Depending upon the state, after a full term²¹ or after a shorter probationary period,²² the citizens vote on whether to retain the judge.

The only issue in a retention election is whether the judge will continue to serve; the judge runs for retention alone and faces no opposing candidates. The retention election question is usually framed along the lines of “Shall Judge Zimmerman be retained in office?” with accompanying “yes” and “no” choices. Most states with retention elections require only a simple majority in order for a judge to be retained.²³ Only New Mexico (57%)²⁴ and Illinois (60%)²⁵ require higher affirmative votes.²⁶ If the judge is retained, he or she is given a full term, at the end of which there is another retention election.²⁷ If the judge is not retained, the process starts over with the selection commission presenting a new list of candidates to the governor.

19. LARRY C. BERKSON & RACHEL CAUFIELD, *AM. JUDICATURE SOC'Y, JUDICIAL SELECTION IN THE UNITED STATES: A SPECIAL REPORT 2* (2004), available at <http://www.ajs.org/selection/docs/Berkson.pdf>.

20. See Thomas R. Phillips, *The Merits of Merit Selection*, 32 HARV. J.L. & PUB. POL'Y 67, 76 (2009). Some states have measures for handling a situation where a governor fails to make a selection. See, e.g., COLO. CONST. art. VI, § 20(1) (providing for the Chief Justice to make the selection if the governor fails to do so within 15 days of having been presented the nomination list); IND. CONST. art. VII, § 10 (same, but within 60 days); UTAH CONST. art. VIII, § 8(1) (same, but within 30 days); MO. CONST. art. V, § 25(a) (providing for selection by the nonpartisan judicial nominating commission after 60 days); TENN. CODE ANN. § 17-4-112 (2012) (allowing governor to request an additional list of nominees from the nominating commission).

21. See, e.g., CAL. CONST. art. VI, § 16(a), (d)(1) (providing for appointed candidates to serve for remainder of current 12-year term).

22. BERKSON & CAUFIELD, *supra* note 19, at 2; see also, e.g., ALASKA CONST. art. IV, § 6 (providing for an approximately three year probationary period followed by ten year term after retention); UTAH CONST. art. VIII, § 9 (same).

23. Aspin, *supra* note 8, at 224.

24. N.M. CONST. art. VI, § 33(A).

25. ILL. CONST. art. VI, § 12(d).

26. Aspin, *supra* note 8, at 224 n.16.

27. The length of a full term varies by state but lasts between six and twelve years for state supreme courts. See *Judicial Selection in the States*, AM. JUDICATURE SOC'Y, <http://www.judicialselection.us/> (last visited Feb. 28, 2012). Kansas Court of Appeals judges have the shortest terms of any state judges who must face retention, serving only four-year terms. See *id.*; KAN. STAT. ANN. § 20-3006(b)(2) (2011).

2012] CAMPAIGNING BY JUDICIAL RETENTION CANDIDATES 245

While sixteen of the nineteen states with appellate court judicial retention elections use the merit selection system,²⁸ three states utilize a hybrid system for selecting judges.²⁹ In these states, whenever a vacancy is created on the court, that vacancy is filled through a partisan election.³⁰ The winner of that election then sits for a full term and faces a retention election at the end of that term and all subsequent terms.³¹ If a judge leaves the bench before the expiration of the term, the seat is temporarily filled by appointment until the next election,³² at which point the process starts over with a partisan election.

1. History of Merit Selection

Merit selection is the result of a compromise in the early 20th century. While the federal constitution adopted in 1787 provides for the lifetime appointment of judges,³³ by the mid-1800s, populist concerns about control of judicial appointments by political machines led most states to select judges through contested elections.³⁴ This shift did little to quell complaints about the

28. Those sixteen states are Alaska, Arizona, California, Colorado, Florida, Indiana, Iowa, Kansas, Maryland, Missouri, Nebraska, Oklahoma, South Dakota, Tennessee, Utah, and Wyoming. *Methods of Judicial Selection*, AM. JUDICATURE SOC'Y, http://www.judicialselection.us/judicial_selection/methods/selection_of_judges.cfm (last visited Mar. 27, 2012). In Tennessee, if a judge does not receive a recommendation for retention by the judicial evaluation commission, that judge must face a contested election rather than an uncontested retention election. TENN. CODE ANN. § 17-4-114(c) (2011). However, all 68 judges evaluated by the commission since it was created in 1994 have been recommended for retention. See Brian T. Fitzpatrick, *Election As Appointment: The Tennessee Plan Reconsidered*, 75 TENN. L. REV. 473, 484 (2008); see also, e.g., TENN. JUDICIAL EVALUATION COMM'N, TENNESSEE APPELLATE JUDGES EVALUATION REPORT (2010), available at http://www.tncourts.gov/sites/default/files/docs/jpec_evaluations_2010.pdf. Montana has a retention election when an incumbent is unopposed, but because of the unusual nature of this arrangement, the state is excluded from the analysis. *Methods of Judicial Selection*, supra. Idaho has merit selection for magistrate judges and is also excluded. See IDAHO CODE ANN. §§ 1-2220, 1-2205(b) (2011).

29. Illinois, Pennsylvania, and New Mexico. *Methods of Judicial Selection*, supra note 28.

30. ILL. CONST. art. VI, § 12(a); N.M. CONST. art. VI, § 33(A); PA. CONST. art. V, § 13(a).

31. ILL. CONST. art. VI, § 12(d); N.M. CONST. art. VI, § 33(A)-(D); PA. CONST. art. V, § 15(b).

32. ILL. CONST. art. VI, § 12(c); N.M. CONST. art. VI, § 35; PA. CONST. art. V, § 13(b).

33. U.S. CONST. art. III, § 1 ("The Judges, both of the supreme and inferior Courts, shall hold their Offices during good Behaviour . . .").

34. Hallie Sears, *A New Approach to Judicial Retention: Where Expertise Meets Democracy*, 24 GEO. J. LEGAL ETHICS 871, 873 (2011); see also Jed Handelsman

politicization of the judiciary, however, and individuals on both sides of the ideological spectrum expressed discontent.³⁵ Left-leaning advocates believed that judges had a class bias against the lower classes, while those on the right felt that judicial elections “enmeshed the judiciary in politics, undermined respect for the courts, and discouraged the selection of highly qualified jurists.”³⁶

The discontent over purely electoral selection methods led to the consideration of alternatives.³⁷ Professor Albert Kales at Northwestern University School of Law first proposed the merit selection system in 1914.³⁸ Under merit selection, apolitical judicial nominating commissions, instead of party leaders, selected candidates. This replacement “ensured that qualifications, not party service, were the criteria for elevation to the bench.”³⁹ Hypothetically, the judicial selection process could thus be insulated from politicians and the public.⁴⁰

Rather than being an essential part of the original merit selection plan, retention elections were “originally offered only to quiet the fears of devotees of the elective method.”⁴¹ These elections were expected to result in the removal of a judge from the bench only rarely, if ever. In fact, many backers of merit selection expected that the elections would eventually be removed from the process altogether.⁴² This never came to pass.

Over the course of the past century, many states implemented the merit selection system. Merit selection received some initial formal support when the American Judicature Society recommended its use in 1920.⁴³ In 1934, California became the first state to adopt any kind of merit selection model,⁴⁴ under which gubernatorial nominations were subject to confirmation by a commission.⁴⁵ Three years later, the American Bar Association endorsed Califor-

Shugerman, *Economic Crisis and the Rise of Judicial Elections and Judicial Review*, 123 HARV. L. REV. 1061, 1080, 1092–93 (2010) (explaining how New York triggered a wave of states adopting judicial elections).

35. See Tarr, *supra* note 3, at 606–07.

36. *Id.*

37. See Sears, *supra* note 34, at 873.

38. See Tarr, *supra* note 3, at 608 & nn.17–18.

39. *Id.* at 611.

40. *Id.* at 608.

41. *Id.* at 609 (quoting Glenn Winters, president of the American Judicature Society).

42. *Id.*

43. *Id.*

44. Sears, *supra* note 34, at 873–74.

45. See *supra* note 18.

R

R

R

2012] CAMPAIGNING BY JUDICIAL RETENTION CANDIDATES 247

nia's move,⁴⁶ adding powerful institutional support for merit selection.⁴⁷ In 1940, Missouri enacted its own version of merit selection that became a popular model for other states.⁴⁸ By 1960, three states employed merit selection to choose state supreme court justices,⁴⁹ and by 1980 eighteen did so.⁵⁰ As of 2011, sixteen states use a variant of the model for at least some of their appellate courts.⁵¹

2. Historical Trends in Retention Election Outcomes

Historically, judges facing a retention election have almost always been retained. Out of 7,689 judicial retention election candidates between 1964 and 2010, only sixty-seven judges—less than one-hundredth of one percent—were rejected by voters.⁵² The mean affirmative vote nationwide in retention elections was 69.5% in 2010, a decline from the 75% that had remained fairly steady since 1998.⁵³ In the typical retention election, individual voters tend to treat candidates as a single bloc, voting the same way for all judges facing retention in the election.⁵⁴ Those instances in which judges are removed are usually the result of targeted efforts against particular judges; other judges facing retention in the same election are mostly unaffected.⁵⁵ Clearly voters are “not indiscriminately throwing all the rascals out.”⁵⁶

Retention election voters tend to be largely uninformed about the candidates. A study by the Justice at Stake Committee⁵⁷ found that 73% of voters reported having only some or a little information

46. Sears, *supra* note 34, at 874.

47. Tarr, *supra* note 3, at 609.

48. Sears, *supra* note 34, at 874.

49. Alaska, Kansas, and Missouri. See Tarr, *supra* note 3, at 605.

50. *Id.*

51. See *supra* note 28.

52. Aspin, *supra* note 8, at 225 fig. 6. Note that these figures only include states that have retention elections for the major trial courts and appellate courts. See *id.* at 218 n.1. Over half of those rejections came from Illinois, which requires a 60% affirmative vote for a judge to be retained; only one of those Illinois judges rejected had a vote total of less than 50%. Larry Aspin, *Judicial Retention Election Trends 1964–2006*, 90 JUDICATURE 208, 210 (2007) [hereinafter Aspin 2007].

53. Aspin, *supra* note 8, at 219.

54. Since 1988, the average absolute difference from the district mean affirmative vote has been 2.2% or less. Aspin, *supra* note 8, at 222 fig. 3.

55. Aspin 2007, *supra* note 52, at 210.

56. *Id.* This was particularly true in Iowa in 2010. See *infra* p. 255–56 (discussing retention of Judge Hanson).

57. “Justice at Stake is the only national organization that focuses exclusively on keeping courts fair and impartial. Justice at Stake leads a nonpartisan national partnership of more than 50 organizations, protecting our justice system through public education, litigation and reform.” *Justice at Stake’s Mission*, JUSTICE AT STAKE

about judicial candidates, while 14% reported that they had none at all.⁵⁸ A 1992 study found that more than two-thirds of Florida voters admitted “some confusion about retention elections,” 40% believed that judges appeared on the ballot because they “had done something wrong,” and only 30% understood that retention elections were a normal event.⁵⁹

Perhaps as a result of this lack of knowledge, many voters fail to vote in judicial retention elections at all—a phenomenon known as “rolloff.”⁶⁰ While rolloff has historically hovered around 34% for retention elections nationwide,⁶¹ in recent years it has declined to an average of about 24%,⁶² with a low of 18.6% in Alaska and a high of 42.9% in Arizona.⁶³ The recent rise in opposition campaigns was likely a significant contributor to this decline.⁶⁴

Judges are much less likely to be retained when they receive negative attention from multiple sources.⁶⁵ For example, in April 2010, the Alaska Commission on Judicial Conduct found “probable cause” that Judge Richard Postma suffered from “mental health difficulties” that were or could become permanent and that made him

CAMPAIGN, <http://www.justiceatstake.org/about/index.cfm> (last visited Mar. 26, 2012).

58. JUSTICE AT STAKE CAMPAIGN, JUSTICE AT STAKE FREQUENCY QUESTIONNAIRE 4 (2001), *available at* http://www.justiceatstake.org/media/cms/JASNationalSurveyResults_6F537F99272D4.pdf. The survey even included voters in states with voter guides. Tarr, *supra* note 3, at 625.

59. Webster, *supra* note 17, at 35.

60. The “rolloff” of a race is the percentage of those who went to the polls to vote on election day, but failed to vote in that particular race. MARK LAWRENCE KORNBLOH, WHY AMERICA STOPPED VOTING: THE DECLINE OF PARTICIPATORY DEMOCRACY AND THE EMERGENCE OF MODERN AMERICAN POLITICS 16 (2000). For example, if 1000 people cast a ballot, but 200 did not vote in a certain retention election, the rolloff for that race would be 20%. Precise explanations for rolloff in retention elections vary and are far from definitive. *See* Aspin, *supra* note 8, at 220–21 (stating that there is contradictory evidence regarding whether increased retention election awareness and the propagation of voter guides affect rolloff).

61. *See* Aspin 2007, *supra* note 52, at 209 fig. 1, 212.

62. *See* Aspin, *supra* note 8, at 219 fig. 1.

63. *Id.* at 220. However, it would seem that some of this can be accounted for by the fifty-three retention races voters were asked to weigh in on in Maricopa County, Arizona. *See id.* at 222. This would likely tax even the most diligent voter.

64. *Id.* at 221. *But see id.* (“Campaigns against judges, however, were neither necessary nor sufficient to reduce rolloff. Rolloff increased 2.5 percent in Alaska despite the late campaign against Supreme Court Justice Dana Fabe and the Alaska Judicial Council recommendation against the retention of Superior Court Judge Richard Postma. On the other hand, rolloff decreased 11.8 points in Tennessee and 9.5 points in Indiana in the absence of any campaigns against the judges standing for retention.”).

65. *Id.* at 225.

“unable to fulfill the duties of his office.”⁶⁶ When this was coupled with a later recommendation against retention by the Alaska Judicial Council, Postma was unable to win retention, receiving only 46% of the vote.⁶⁷

Though most judges seeking retention have been retained, there have been incidents where major campaigns were organized in opposition to particular judges’ retentions. Massive efforts were undertaken in Iowa and Illinois in 2010,⁶⁸ but they were not unprecedented—their most notable forerunner was the 1986 California retention election, still the most expensive retention election of all time.⁶⁹

3. California 1986

In 1986, the death penalty was a significant issue nationwide,⁷⁰ but perhaps nowhere more so than California. Proponents of the death penalty had grown upset with Chief Justice Rose Bird of the California Supreme Court, as well as with two of her colleagues, Justices Joseph Grodin and Cruz Reynoso. Bird had voted to reverse every one of the sixty-one death penalty cases that had come before the court in her nine years on the bench, and Grodin and Reynoso had typically followed her lead.⁷¹

Employing slogans such as “bye bye Birdie,”⁷² opponents of the justices spent about \$6.6 million in their “strenuous, emotionally

66. *Id.*

67. *Id.*

68. *See infra* Parts I.B.1–2.

69. SKAGGS ET AL., *supra* note 13, at 20.

70. Precisely why the death penalty has been a significant issue in the United States is a very complex issue. *See, e.g.*, Carol S. Steiker, *Capital Punishment and American Exceptionalism*, in AMERICAN EXCEPTIONALISM AND HUMAN RIGHTS 57, 61 (Michael Ignatieff ed., Princeton University Press 2005) (offering ten reasons why the death penalty has been an exceptional issue relative to other western democracies). As a 1995 article noted, “Judicial campaigns in which the death penalty is an issue can degenerate to almost Orwellian levels of absurdity, raising serious questions about the ability of judges to remain fair and impartial.” Stephen B. Bright & Patrick J. Keenan, *Judges and the Politics of Death: Deciding Between the Bill of Rights and the Next Election in Capital Cases*, 75 B.U. L. REV. 759, 785 (1995); *see also* Uelmen, *supra* note 10, at 1135–37.

71. B. Michael Dann & Randall M. Hansen, *Judicial Retention Elections*, 34 LOY. L.A. L. REV. 1429, 1432 (2001); *see also* Nicholas L. Georgakopoulos, *Judicial Reaction to Change: The California Supreme Court Around the 1986 Elections*, 13 CORNELL J.L. & PUB. POL’Y 405, 410 (2004).

72. *CBS Evening News*, *supra* note 1.

charged” effort to oust them.⁷³ Having already formally announced his opposition to Bird, California Governor George Deukmejian publicly warned Justices Grodin and Reynoso that he would oppose their retention bids unless they voted to uphold more death sentences.⁷⁴ Deukmejian eventually carried out his threat and opposed the retention of the two associate justices as well.⁷⁵

Bird actively campaigned for retention, explaining that she “did not seek this fight, but . . . [would] not shrink from it.”⁷⁶ True to her word, Bird’s largely positive campaign “stress[ed] the traditional independence of the judiciary—the one message that pollsters had determined would not work.”⁷⁷ According to a CBS Evening News broadcast, the 50-year-old Bird employed “plunging neck lines, glittering jewels, and all the help from Hollywood she [could] get,” including personal appearances at campaign events by Academy Award winning director Warren Beatty.⁷⁸ Proponents of the justices collectively spent about \$4 million.⁷⁹

None of the three justices were retained, with Bird garnering only 34% in affirmative votes, Reynoso, 40%, and Grodin, 43%.⁸⁰ Some observers predicted that the California election would have broader consequences outside the state. Before election day, Dan Rather noted that “this big money, high profile fight could . . . make judges nationwide think twice about politics, pressures, and principles.”⁸¹ Gerald Uelmen, then-Dean of the Santa Clara University School of Law, warned, “The California events of 1986 should

73. Robert S. Thompson, *Judicial Retention Elections and Judicial Method: A Retrospective on the California Retention Election of 1986*, 61 S. CAL. L. REV. 2007, 2038, 2040 (1988).

74. Bright & Keenan, *supra* note 70, at 760.

75. *Id.* When asked why he chose to run for governor rather than a second term as attorney general, Deukmejian explained that “Attorneys General don’t appoint judges—Governors do.” *California Governor George Deukmejian*, NAT’L GOVERNORS’ ASSOC., http://www.nga.org/cms/home/governors/past-governors-bios/page_california/col2-content/main-content-list/title_deukmejian_george.html (last visited Dec. 7, 2011).

76. *CBS Evening News*, *supra* note 1.

77. Todd E. Pettys, *Letter from Iowa: Same-Sex Marriage and the Ouster of Three Justices*, 59 U. KAN. L. REV. 715, 740 (2011) (internal quotation marks omitted). For samples of ads supporting Bird, see tt4jd, *1986 Rose Elizabeth Bird Campaign 30 Sec Spots.wmv*, YOUTUBE (Oct. 25, 2011), <http://www.youtube.com/watch?v=YB8xFq3lGXI>.

78. *CBS Evening News*, *supra* note 1. Bird specifically decided not to run a joint campaign with the other two justices up for retention. Thompson, *supra* note 73, at 2037.

79. Thompson, *supra* note 73, at 2038.

80. Dann & Hansen, *supra* note 71, at 1432.

81. *CBS Evening News*, *supra* note 1.

R

R

R

2012] CAMPAIGNING BY JUDICIAL RETENTION CANDIDATES 251

not be dismissed as an aberration. Political forces have been unleashed that will return to haunt us.”⁸²

Contrary to the warnings, there was little nationwide fallout from the California races. Generally, retention election candidates continued to be retained as a matter of course, with only an occasional contested retention election. One notable contest involved Tennessee Supreme Court Justice Penny White, who was not retained in 1996 after she voted to overturn a single death penalty conviction.⁸³ Another notable incident arose from the Pennsylvania Supreme Court’s retention elections in 2005, when Justice Russell Nigro was defeated and his colleague Justice Sandra Schultz Newman was narrowly retained after the public reacted negatively to the so-called “Pay Grab” controversy, involving a sudden pay increase for government officials.⁸⁴ Because the justices benefited from the new legislation and were the only statewide candidates on the ballot at the election, they suffered the electorate’s wrath, even though neither justice had played a role in passing the pay increase.⁸⁵

The infrequency of contested retention elections nationally may not accurately reflect the full impact of the California 1986 races, however. It is possible that judges facing retention, aware of what happened in California, started aligning their judicial decisions more closely with public opinion to avoid a similar fate. Indeed, after the 1986 election, “California’s Supreme Court had one of the highest rates of upholding death sentences in the nation.”⁸⁶ In 1988, ousted Justice Grodin explained that neither he nor former Justice Otto Kaus, who resigned from the bench of his own accord in 1985,⁸⁷ could be certain that their votes in important cases were not subconsciously affected by a fear of not being retained.⁸⁸ Still, the events of California in 1986 seem like an aberra-

82. Gerald F. Uelmen, *Commentary: Are We Reprising a Finale or an Overture?*, 61 S. CAL. L. REV. 2069, 2073 (1988) (responding to Thompson, *supra* note 73).

83. See Stephen B. Bright, *Political Attacks on the Judiciary: Can Justice Be Done amid Efforts to Intimidate and Remove Judges from Office for Unpopular Decisions?*, 72 N.Y.U. L. REV. 308, 310 (1997).

84. Thomas A. Donovan, *Why Do We Elect Judges?*, FED. LAW. Feb. 2009, at 6, 8.

85. *Id.* (“Although the state’s chief justice had publicly defended the need to raise judicial salaries, only the legislative and executive branches had been responsible for enacting the pay raise.”).

86. Kenneth J. Aulet, *It’s Not Who Hires You but Who Can Fire You: The Case Against Retention Elections*, 44 COLUM. J.L. & SOC. PROBS. 589, 609 (2011).

87. Maura Dolan, *Otto Kaus Dies; Former Justice on State High Court*, L.A. TIMES, Jan. 13, 1996, http://articles.latimes.com/1996-01-13/local/me-24215_1_state-high-court/.

88. Joseph R. Grodin, *Developing a Consensus of Constraint: A Judge’s Perspective on Judicial Retention Elections*, 61 S. CAL. L. REV. 1969, 1980 (1988); see also Republi-

tion. In California's next judicial election, no organized opposition formed to unseat the five candidates up for retention, and none of the candidates had to spend much beyond the \$2300 filing fee to run.⁸⁹ All five justices were retained,⁹⁰ and the fear that retention elections would routinely be opposed, in California or elsewhere, proved to be unfounded.⁹¹

B. Rubber Stamp No More: Judicial Retention Elections in 2010

While retention elections generally remain "rubber stamp" affairs, national and state politics have become extremely polarized in recent years, and partisans on both sides of the aisle have grown comfortable using nastier tactics. This is as true in judicial elections as other races, and is reflected by the flood of money pouring into judicial races. From 2000 through 2009, state supreme court candidates raised a combined \$206.9 million, more than double the previous decade's total.⁹² This increase has translated into mudslinging attack ads sponsored by powerful special interests who are playing a significantly greater role.⁹³

Against this backdrop of increasingly costly and increasingly partisan judicial elections, from 2000 to 2009, retention elections

can Party of *Minn. v. White*, 536 U.S. 765, 791 (2002) (O'Connor, J., concurring) (citing Grodin and stating that merit selection "obviously reduces threats to judicial impartiality, even if it does not eliminate all popular pressure on judges").

89. Dann & Hansen, *supra* note 71, at 1432.

90. See MARCH FONG EU, SECRETARY OF STATE, STATEMENT OF VOTE, GENERAL ELECTION NOVEMBER 6, 1990, at 48, available at <http://www.sos.ca.gov/elections/sov/1990-general/1990-general-sov.pdf>. As a contemporary news report explained, however, the reason for the judges' retention and the lack of opposition were likely precisely because those judges acted in line with the 1986 opposition and "[v]oter approval of the five . . . would ensure that a conservative majority—all appointees of Gov. George Deukmejian—will continue to lead the seven-member court as it has since 1987." Philip Hager, *No Opposition, Little Notice for 5 State Justices Up for Election*, L.A. TIMES (Nov. 4, 1990), http://articles.latimes.com/1990-11-04/news/mn-5654_1_supreme-court-justices. Since the departures of the defeated justices in 1986, the court had "upheld 76 death sentences and reversed only 26." *Id.*

91. Outside of California that year, "10 judges were defeated and 19 others were within 5 percentage points of their state's retention threshold." Aspin, *supra* note 8, at 224. From 1988 through 2006, 28 judges nationwide were not retained. See Aspin 2007, *supra* note 52, at 211 tbl. 2.

92. JAMES SAMPLE, ADAM SKAGGS, JONATHAN BLITZER & LINDA CASEY, *THE NEW POLITICS OF JUDICIAL ELECTIONS 2000-2009: DECADE OF CHANGE 1* (2010). Note that this figure includes both retention elections and standard judicial elections involving opposing candidates. See *id.* at 2.

93. SKAGGS ET AL., *supra* note 13, at 3 (indicating that independent spending made up 29.8% of all spending in state high court elections, a percentage significantly higher than the last four years).

2012] CAMPAIGNING BY JUDICIAL RETENTION CANDIDATES 253

were the one type of judicial election that defied the pattern—retention candidates raised only \$2.2 million during that period, less than one percent of the total raised by supreme court candidates in any type of election.⁹⁴ In 2010, however, a number of states saw concerted efforts to remove particular judges from office. Spending skyrocketed in retention elections to almost \$4.9 million in 2010 alone, with candidates raising \$2.8 million and independent groups spending \$2.1 million.⁹⁵ Iowa and Illinois accounted for the lion's share of this spending.⁹⁶

1. Iowa 2010

The most salient retention election race of 2010 resulted in the rejection of all three Iowa Supreme Court justices up for retention—Chief Justice Marsha Ternus, Justice Michael Streit, and Justice David Baker—over their votes in a controversial same-sex marriage case. Just a year earlier, in *Varnum v. Brien*, the Iowa Supreme Court, in a unanimous 7-0 vote, held that the state's "statute limiting civil marriage to a union between a man and a woman violate[d] . . . the equal protection clause of the Iowa Constitution."⁹⁷

Opposition to the decision and the justices emerged immediately.⁹⁸ "Bob Vander Plaats, a former Republican candidate for governor, organized an anti-retention effort called 'Iowa for Freedom.'⁹⁹ Commercials advocating that the judges be rejected emphasized the "activist"¹⁰⁰ nature of the decision and questioned what the justices would "do to other long-established Iowa traditions and rights" should the public retain them.¹⁰¹ The ads included pictures of a church, a boy scout, hunters, and children pledging allegiance to the American flag.¹⁰² The idea of "sending a

94. SAMPLE ET AL., *supra* note 92, at 1–2. Concern about particular judicial races being expensive is not a new phenomenon. The ABA Special Committee on Standards of Judicial Conduct charged with drafting the original 1972 Model Rules of Judicial Conduct apparently was "informed of campaigns for judicial office in which costs ran into the tens of thousands of dollars." E. WAYNE THODE, REPORTER'S NOTES TO CODE OF JUDICIAL CONDUCT 99 (1973).

95. SKAGGS ET AL., *supra* note 13, at 7.

96. *Id.* at 7–8.

97. *Varnum v. Brien*, 763 N.W.2d 862, 872 (Iowa 2009).

98. Curriden, *supra* note 9, at 57.

99. *Id.*

100. *See supra* note 6.

101. *See Send Them a Message, supra* note 7; *see also* IowaForFreedom, *Iowa for Freedom*, YOUTUBE (Oct. 20, 2010), <http://www.youtube.com/watch?v=MBEynTYEyPY> [hereinafter *Iowa for Freedom* Oct.].

102. *See Send Them a Message, supra* note 7; *Iowa for Freedom* Oct., *supra* note 101. One seemingly Internet-only video even featured a picture of a home with a

message” was a recurring theme, both during¹⁰³ and after¹⁰⁴ the campaign. The anti-retention campaign cost about \$1 million,¹⁰⁵ with almost all of the money coming from out of state¹⁰⁶—a mere \$10,000 originated from within Iowa.¹⁰⁷

In spite of this flurry of opposition, “[t]he justices raised no money[,] . . . declined to campaign,”¹⁰⁸ and gave no interviews¹⁰⁹ beyond a single commentary by Ternus regarding the election. Ternus explained, “We’re not forming campaign committees. We’re not going to become politicians.”¹¹⁰ She added, “I hope it’s not a one-sided debate. I hope that people who understand the system and the role of the court will speak out more and more and I believe that’s beginning to happen.”¹¹¹ In-state supporters of the justices created their own organization, Fair Courts for US, and spent nearly \$400,000 in an effort to retain the justices.¹¹² However,

white picket fence. See IowaForFreedom, *Iowa for Freedom*, YouTube (Sept. 24, 2010), <http://www.youtube.com/watch?v=WZG-eKFP8IA> [hereinafter *Iowa for Freedom* Sept.].

103. See *Send Them a Message*, *supra* note 7 (telling viewers to “send them a message”). CBS Evening News similarly mentioned before the 1986 California retention election that “the Bird race is sending a message to elected judges across the nation, putting them on notice that they may no longer be immune to political pressure.” *CBS Evening News*, *supra* note 1.

104. Grant Schulte, *Iowans Dismiss Three Justices*, DES MOINES REGISTER (Nov. 3, 2010), <http://www.desmoinesregister.com/article/20101103/NEWS09/11030390/Iowans-dismiss-three-justices> (quoting Vander Platts as saying “The people of Iowa stood up in record numbers and sent a message . . . that it is ‘We the people,’ not ‘We the courts.’”).

105. Linda Casey, *Independent Expenditure Campaigns in Iowa Topple Three High Court Justices*, NAT’L INST. ON MONEY IN STATE POLITICS (Jan. 10, 2011), <http://www.followthemoney.org/press/ReportView.phtml> (indicating that \$990,651 was spent opposing the justices’ retention).

106. *Id.* “Iowa for Freedom” spent \$171,025 and was based in Iowa, but was funded by the Mississippi-based American Family Association’s AFA Action Inc. See *id.*

107. *Id.* (indicating that the lone Iowa-based organization spent \$10,178 against retention).

108. Curriden, *supra* note 9, at 57.

109. Ryan C. Cicoski, *Judicial Independence and the Rule of Law: A Warning from Iowa*, DEL. LAW., Summer 2011, at 19.

110. Boshart, *supra* note 2.

111. *Id.* After the election, Streit reflected on his decision not to campaign, explaining, “When you get involved in politics, you get labeled. You label yourself, or you let other people label you. . . . You have expectations that you’ve raised in other people’s minds on how you’re going to behave, and you will try to reach those expectations.” Patrick Caldwell, *Disorder in the Court*, in JUSTICE FOR SALE: A SPECIAL REPORT FROM THE AMERICAN PROSPECT MAGAZINE 14, 20 (2011) (internal quotation marks omitted).

112. Cicoski, *supra* note 109, at 19.

R

R

2012] CAMPAIGNING BY JUDICIAL RETENTION CANDIDATES 255

Fair Courts bought no television or newspaper advertisements, unlike the justices' opposition, which also organized a highly visible bus tour¹¹³ that visited twenty cities.¹¹⁴

All three justices were rejected by voters, with each receiving only about 45% in affirmative votes.¹¹⁵ In a statement posted on the court's website after the election, the justices said:

We hope Iowans will continue to support Iowa's merit selection system for appointing judges. This system helps ensure that judges base their decisions on the law and the Constitution and nothing else. Ultimately, however, the preservation of our state's fair and impartial courts will require more than the integrity and fortitude of individual judges, it will require the steadfast support of the people.¹¹⁶

In an interview several months after the election, sitting Iowa Supreme Court Justice David Wiggins commented that the ousted judges "took the position that judges should not get involved in politics. They maintained their integrity. . . . And sometimes you lose your job by doing the right thing."¹¹⁷ Wiggins, who joined the *Varnum* opinion, is up for retention in 2012.¹¹⁸

Surprisingly, Judge Robert Hanson, the Iowa trial court judge who initially ruled in the same case that limiting marriage to a man and a woman violated the state constitution,¹¹⁹ was retained in the same election.¹²⁰ No major opposition campaign was waged against Hanson's retention;¹²¹ he won easily with a 66% affirmative vote.¹²² Reacting to the vote, Hanson said, "I'm extremely grateful for the support from Polk County voters. Unfortunately, I'm also totally disheartened with what's happened to the three supreme court jus-

113. *Id.*

114. See Jason Hancock, *Iowans Vote to Oust All Three Supreme Court Justices*, IOWA INDEP. (Nov. 2, 2010), <http://iowaindependent.com/46917/iowans-vote-to-oust-all-three-supreme-court-justices>.

115. Aspin, *supra* note 8, at 228–29 tbl. 3.

116. Schulte, *supra* note 104.

117. Caldwell, *supra* note 111, at 20 (internal quotation marks omitted).

118. See O. Kay Henderson, *Branstad: Justice Wiggins to be Targeted in '12*, RADIO IOWA (Apr. 16, 2011), <http://www.radioiowa.com/2011/04/16/branstad-justice-wiggins-to-be-targeted-in-12/>.

119. See *Varnum v. Brien*, No. CV5965, 2007 WL 2468667 (Iowa Dist. Ct. Aug. 30, 2007), *aff'd*, 763 N.W.2d 862 (Iowa 2009).

120. Curriden, *supra* note 9, at 58.

121. See Aspin, *supra* note 8, at 228–29 tbl. 3 (lacking any reference to an opposition campaign against Hanson).

122. Curriden, *supra* note 9, at 58.

tics. It's a shocking, SHOCKING lack of appreciation of our judicial system."¹²³

According to some, the justices' refusal to campaign was an important factor in their defeat.¹²⁴ Albert Klumpp, a leading researcher of retention elections, explained that the justices could have expected a five percentage point bump had they campaigned,¹²⁵ which, if not resulting in retention, at least would have made the outcome very close. Chief Justice Wallace Jefferson of the Texas Supreme Court also highlighted the importance of campaigning, saying, "Retention judges need to think about campaigning. . . . This may not sound popular, but sometimes you need to tell your side of the story. I think retention is good because it forces judges to get out there to explain themselves and to be held accountable."¹²⁶

2. Illinois 2010

A counterexample to the 2010 Iowa retention election may be found in the 2010 retention contest of Chief Justice Kilbride of the Supreme Court of Illinois.¹²⁷ After Kilbride joined an Illinois Supreme Court ruling¹²⁸ that overturned limits on medical malpractice awards,¹²⁹ pro-business interests¹³⁰ organized a \$700,000 media campaign¹³¹ against Kilbride's retention.¹³² The campaign did not focus on Kilbride's role in the medical malpractice ruling, how-

123. Dave Price, *Judge Hanson Responds to Justice Retention Vote*, PRICE OF POLITICS, ETC. (Nov. 7, 2010, 8:15 PM), <http://whoipolitics.blogspot.com/2010/11/judge-hanson-responds-to-same-sex.html>.

124. Curriden, *supra* note 9, at 58.

125. *Id.* (quoting Albert Klumpp, Research Analyst at McDermott Will & Emery).

126. *Id.*

127. Recall that Illinois uses a hybrid system where justices are initially selected in partisan elections and then run in retention elections. *See supra* p. 6.

128. *Lebron v. Gottlieb Mem'l Hosp.*, 930 N.E.2d 895, 899, 917 (Ill. 2010).

129. Editorial, *How Should We Respond to the 2010 Judicial Elections?*, 94 JUDICATURE 102, 102 (2010); *see also* Monique Garcia, *State Supreme Court Justice Wins Retention Battle*, CHI. TRIB. (Nov. 2, 2010), http://articles.chicagotribune.com/2010-11-02/news/ct-elect-kilbride-20101102_1_jury-awards-retention-battle-constitutionality-of-state-law (describing pro-business group that sought to "dump a judge it sees as unwilling to stop large jury awards").

130. Aspin, *supra* note 8, at 230 (listing the U.S. Chamber of Commerce, the American Tort Reform Association, and the American Justice Partnership, a creation of the National Association of Manufacturers).

131. *Id.* at 227.

132. Roy A. Schotland, *Iowa's 2010 Judicial Election: Appropriate Accountability or Rampant Passion?*, 46 CT. REV. 118, 125 (2011).

2012] CAMPAIGNING BY JUDICIAL RETENTION CANDIDATES 257

ever—instead, Kilbride was portrayed as being soft on crime,¹³³ a nearly ubiquitous and often effective strategy when judges are up for election.¹³⁴ The opposition campaign featured “probably the most outrageous ad of the entire 2010 judicial election season . . . Dressed in orange jumpsuits, actors posing as convicted criminal[s] recounted the grisly details of their crimes, and then said that [Chief] Justice Thomas Kilbride had taken their side and voted against law enforcement and victims.”¹³⁵ In the end, however, the “soft on crime” strategy may have backfired because it may have caused voters to view the opposition effort as dishonest.¹³⁶

Unlike his Iowa peers, Kilbride opted to campaign and raised over \$2.7 million,¹³⁷ a sum greater than the \$2.2 million raised by all retention election candidates nationwide from 2000 to 2009 combined.¹³⁸ Kilbride’s campaign featured numerous “positive” television commercials and included endorsements from law enforcement.¹³⁹ Major plaintiffs’ law firms also supported

133. *Id.* at 230.

134. This strategy was “financed by groups focused solely on civil lawsuit awards,” including the American Tort Reform Association. SKAGGS ET AL., *supra* note 13, at 20; *see also* JudicialCampaignAds, *Vote No on Justice Kilbride*, YOUTUBE (Oct. 26, 2010), <http://www.youtube.com/watch?v=aPmGtxw2en8>. Kilbride’s opponents likely used this strategy because they feared that the medical malpractice issue would not resonate with voters. Generally, the American electorate has taken a very punitive (as opposed to rehabilitative) stance regarding criminals over the last four decades, and politicians who are “tough on crime” are rewarded. For a discussion on how emphasis on the issue of crime has affected American electoral politics, *see* Sara Sun Beale, *Still Tough on Crime? Prospects for Restorative Justice in the United States*, 2003 UTAH L. REV. 413, 414, 428 (2003).

135. SKAGGS ET AL., *supra* note 13, at 20; *see also* JudicialCampaignAds, *supra* note 134.

136. Aspin, *supra* note 8, at 230.

137. SKAGGS ET AL., *supra* note 13, at 5.

138. *Id.* at 20.

139. Aspin, *supra* note 8, at 230; *see also, e.g.*, FairCourtsPage, *Justice Kilbride Tough on Crime (Illinois 2010)*, YOUTUBE (Oct. 8, 2010), <http://www.youtube.com/watch?v=eTOUGFXgqS8> (featuring endorsements from law enforcement officers and a state attorney). Note that some of these commercials, however, similarly pandered to voters’ fear of criminals by portraying Kilbride as tough on crime. *See, e.g.*, Norman L. Reimer, *Fear Unleashed: Money, Power and the Threat to Judicial Independence (Inside NACDL)*, THE CHAMPION, Nov. 2010, *available at* <http://www.nacdl.org/champion.aspx?id=16254> (transcribing an ad that says “Justice Tom Kilbride—a strong advocate for the victims of crime—endorsed by our police and prosecutors. Tom Kilbride wrote the opinion that protected victims of sex crimes from their attackers and issued rulings to simplify the prosecution of sexual predators and domestic violence abusers. We need judges who stand up for victims—not criminals. For Fairness. For Victims. For Justice. Vote YES for Tom Kilbride.”); *see also* Schotland, *supra* note 132, at 125 (“Every judge’s campaign slogan,

Kilbride by contributing \$1.5 million to the Illinois Democratic Party.¹⁴⁰

Kilbride was retained with a 65.9% affirmative vote, enough to clear the requisite 60% threshold in Illinois.¹⁴¹ Rolloff was only 5.8%,¹⁴² far below the 34% historical national average.¹⁴³ The race was the most expensive retention election ever in the state, and the second most expensive ever nationwide.¹⁴⁴ While Kilbride did not apologize for raising such substantial sums to defend his seat on the bench, he also recognized the danger posed by the increasing politicization of the judiciary, saying, “If we are going to allow the courts to be politicized to this degree, where there’s more and more big-time money coming in, it’s going to ruin the court system and we might as well shut down the third branch.”¹⁴⁵

3. Importance, or Unimportance, of 2010

Given the events of 2010,¹⁴⁶ an essential question becomes whether those elections foreshadow a new reality for judicial elec-

in advertisements and on billboards, is some variation of ‘tough on crime.’ The liberal candidate is the one who advertises: ‘Tough but fair.’ Television campaigns have featured judges in their robes slamming shut a prison cell door.”) (quoting Hans A. Linde, Comment, *Elective Judges: Some Comparative Comments*, 61 S. CAL. L. REV. 1995, 2000 (1988)).

140. SKAGGS ET AL., *supra* note 13, at 20.

141. *See supra* note 25 and accompanying text.

142. Aspin, *supra* note 8, at 229.

143. *See supra* p. 248.

144. SKAGGS ET AL., *supra* note 13, at 20. Only the 1986 California retention election was costlier. *Id.*

145. Reimer, *supra* note 139.

146. Overall, there were 482 judges up for retention across 12 states in 2010. Aspin, *supra* note 8, at 224. Of those, eighteen judges in six states had opposition of some kind to their retention, with “opposition” consisting of a formal effort to prevent a judge’s retention or a recommendation against retention by an official judicial evaluating body. *See id.* at 226, 228–29. In this statistic, opposition does not include a recommendation against retention from a bar association. *Compare id.* at 224 fig.5, *with id.* at 228–29. Including the Iowa justices, six judges out of the eighteen were not retained. *Id.* at 228. Two Colorado district judges were not retained after allegations arose that they had made mistakes as prosecutors in 1999. *Id.* at 227. “The 1999 conviction was overturned in 2008, the defendant then won a \$10 million settlement, both judges were censured for not being more diligent, and the lead detective in the case was indicted in June of 2010 on eight counts of perjury related to the trial.” *Id.* The formal opposition against them was relatively modest, spending less than \$12,000. *Id.* Both judges received about 38% in affirmative votes. *Id.* at 228–29. The previously mentioned Judge Postma was not retained in Alaska. *See supra* pp. 248–49. Also in Alaska, “a late-breaking campaign during the final two weeks against Justice Dana Fabe failed, and she was retained.” *How Should We Respond to the 2010 Judicial Elections?*, *supra* note 129, at 102. Ten

tions or are instead merely a blip on the radar. In all likelihood, most retention election candidates will be retained without much fanfare,¹⁴⁷ and most judicial decisions will continue to be rendered with little regard for public opinion. Instead, the concern is that some judges facing retention elections will find themselves confronting the “crocodile in the bathtub” dilemma: whether to make the correct (in his or her professional opinion) but unpopular judicial ruling and potentially unleash the fury of oppositional forces, or to make the popular, but incorrect¹⁴⁸ ruling, and avoid public denunciation.¹⁴⁹

II. RECONSIDERATION OF RETENTION ELECTION ELECTIONEERING RULES

A. *Competing Goals Surrounding Retention Elections*

Before evaluating the propriety of judicial electioneering, it is helpful to take a step back and explore the values served by retention elections and the merit selection system. The primary aim of retention elections is to balance the competing values of judicial impartiality and judicial accountability. Indeed, retention elections were the result of a political compromise that attempted to reconcile the tension between these two concerns.¹⁵⁰

judges nationwide who were retained survived very close elections, and came within five percentage points of not being retained. Aspin, *supra* note 8, at 224.

147. *Cf.* Schotland, *supra* note 132, at 125 (“For any judge—even United States Supreme Court justices—‘hot-button’ issues are rare (even at the Supreme Court, only a minute fraction of their decisions). For most trial judges, such issues are non-existent.”).

148. No opinion is expressed or implied in this article regarding the correctness of the Iowa Supreme Court’s *Varnum v. Brien* decision or the Illinois Supreme Court’s *Lebron v. Gottlieb Memorial Hospital* decision.

149. There are two additional concerns resulting from the 2010 retention elections that are not explored in this Article. First, current or would-be judges, seeing the insecurity that accompanies being an elected judge, may choose to leave the bench or never seek it, respectively. *See* Schotland, *supra* note 132, at 127–28. Second, politicians may see retention elections as a way to rally the public, leading to campaigns with the only genuine purpose of furthering those politicians’ careers. *See id.* at 127.

150. *See supra* Part I.A.1.

1. Judicial Impartiality

Maintaining the impartiality of the judiciary is of paramount importance.¹⁵¹ As Supreme Court Justice Sandra Day O'Connor explained, "Our effectiveness as judges relies on the knowledge that we will not be subject to retaliation for our judicial acts."¹⁵² She went on:

We of course want judges to be impartial, in the sense of being free from any personal stake in the outcome of the cases to which they are assigned. But if judges are subject to regular elections they are likely to feel that they have at least some personal stake in the outcome of every publicized case. Elected judges cannot help being aware that if the public is not satisfied with the outcome of a particular case, it could hurt their reelection prospects.¹⁵³

As Justice Stevens has written, "[I]n litigation, issues of law or fact should not be determined by popular vote; it is the business of judges to be indifferent to unpopularity."¹⁵⁴ In addition to impartiality itself, the appearance thereof is important as well¹⁵⁵ so litigants

151. MODEL CODE OF JUDICIAL CONDUCT Preamble [1] (2010) ("An independent, fair and impartial judiciary is indispensable to our system of justice."). It is important to distinguish judicial impartiality from judicial independence. "The two concepts are often intertwined and interdependent, and when [separated] . . . independent decision making [refers to] judicial impartiality and judicial branch separateness [refers to] judicial independence." *Conversation, supra* note 6, at 339 (statement of Ruth McGregor, C.J.); *see also* Richard Briffault, *Judicial Campaign Codes After Republican Party of Minnesota v. White*, 153 U. PA. L. REV. 181, 198–99 (2004) ("Judicial impartiality refers to the constitutional imperative that judges treat all parties before them fairly and equally and decide cases according to the evidence and the law. . . . Judicial independence is linked to impartiality . . . [b]ut [] also implicates the separation of powers and the freedom of the courts from the other branches of government.").

152. *Conversation, supra* note 6, at 339 (statement of Sandra Day O'Connor, J.).

153. *Republican Party of Minn. v. White*, 536 U.S. 765, 788–89 (2002) (O'Connor, J., concurring) (emphasis added) (citing Julian N. Eule, *Crocodiles in the Bath tub: State Courts, Voter Initiatives and the Threat of Electoral Reprisal*, 65 U. COLO. L. REV. 733, 739 (1994)).

154. *White*, 536 U.S. at 798 (Stevens, J., dissenting).

155. *See Caperton v. A.T. Massey Coal Co.*, 129 S. Ct. 2252, 2266 (2009) ("A judge shall avoid impropriety and the appearance of impropriety.") (quoting ABA ANNOTATED MODEL CODE OF JUDICIAL CONDUCT Canon 2 (2004)) (internal quotation marks omitted); *see also* MODEL CODE OF JUDICIAL CONDUCT Canon 1 (2010) ("A judge shall uphold and promote the independence, integrity, and impartiality of the judiciary, and shall avoid impropriety and the appearance of impropriety.").

2012] CAMPAIGNING BY JUDICIAL RETENTION CANDIDATES 261

“can feel they are treated fairly.”¹⁵⁶ Impartiality is the cornerstone upon which the judicial branch is built.

The appearance of impartiality also holds constitutional significance: even if a judge is truly impartial, the appearance of bias may be so intolerable as to violate due process. In *Caperton v. A.T. Massey Coal Co.*, a West Virginia jury entered a \$50 million judgment against Massey Coal.¹⁵⁷ The timing of the case was such that one of the seats on the West Virginia Supreme Court would be up for election before the appeal would be heard. Don Blankenship, the CEO of Massey Coal, knew that the winner of the election would hear the appeal¹⁵⁸ and spent more than \$3 million to support West Virginia Supreme Court candidate Brent Benjamin—\$1 million more than the amount spent by the campaign committees of Benjamin and his opponent combined.¹⁵⁹ Benjamin was elected,¹⁶⁰ and the \$50 million verdict was reversed by a 3-2 vote, with Justice Benjamin siding with the majority.¹⁶¹ The U.S. Supreme Court, while specifically declining to determine whether Justice Benjamin was in fact biased,¹⁶² held that the “probability of actual bias on the part of the judge . . . [was] too high to be constitutionally tolerable,” reversed the West Virginia Supreme Court,¹⁶³ and remanded with instructions for Justice Benjamin to recuse himself.¹⁶⁴ In *Caperton*, the Supreme Court makes clear that the appearance of impartiality is as crucial as actual impartiality.¹⁶⁵

156. *Conversation*, *supra* note 6, at 340 (statement of Ruth McGregor, C.J.).

157. *Caperton*, 129 S. Ct. at 2257. West Virginia does not have intermediate appellate courts, so all trial level appeals are made to the state supreme court. *The West Virginia Judicial System*, OFFICIAL WEBSITE FOR THE STATE OF WEST VIRGINIA, <http://www.state.wv.us/wvsca/wvsystem.htm> (last visited May 19, 2012).

158. *Caperton*, 129 S. Ct. at 2257 (“Knowing the Supreme Court of Appeals of West Virginia would consider the appeal in the case, Blankenship decided to support an attorney who sought to replace Justice McGraw.”).

159. *Id.*

160. *Id.* The U.S. Supreme Court declined to determine whether Blankenship’s expenditures were the cause of the electoral outcome, however. *See id.* at 2264.

161. *See Caperton v. A.T. Massey Coal Co.*, 679 S.E.2d 223 (W. Va. 2008), *rev’d*, 129 S. Ct. 2252 (2009).

162. *Caperton*, 129 S. Ct. at 2263.

163. *Id.* at 2257, 2267 (internal citation omitted) (internal quotation marks omitted). Interestingly, the West Virginia Supreme Court, with a different membership that did not include the recused Justice Benjamin, went on to again reverse the \$50 million verdict in a subsequent opinion. *See Caperton v. A.T. Massey Coal Co.*, 690 S.E.2d 322, 357 (W. Va. 2009).

164. *Caperton*, 129 S. Ct. at 2263–64, 2267.

165. *See also* *Republican Party of Minn. v. White*, 536 U.S. 765, 793 (2002) (Kennedy, J., concurring) (“Courts, in our system, elaborate principles of law in

2. Judicial Accountability

Judicial accountability prevents a judge from disregarding the law or otherwise damaging the integrity of the court. According to the American Judicature Society, “Independence and accountability are different sides of the same coin. Both are necessary to maintain that delicate balance which permits our system of justice to function effectively.”¹⁶⁶ There are many things for which voters might hold a judge accountable. Beyond the correctness of judicial rulings, a judge might also be evaluated based upon his or her management of the court’s financial resources, efficiency in disposing of cases,¹⁶⁷ ability to ensure disadvantaged litigants (e.g., non-En-

the course of resolving disputes. The power and the prerogative of a court to perform this function rest, in the end, upon the respect accorded to its judgments. The citizen’s respect for judgments depends in turn upon the issuing court’s absolute probity. Judicial integrity is, in consequence, a state interest of the highest order.”).

166. Editorial, *The Judicial Independence and Accountability Task Force*, 88 JUDICATURE 108, 121 (2004); see also Charles Gardner Geyh, *Rescuing Judicial Accountability from the Realm of Political Rhetoric*, 56 CASE W. RES. L. REV. 911, 911 (2006) (explaining that judicial independence is well-examined, but that judicial accountability is often neglected, much in the same way one knows that F.D.R. is on the heads side of a dime but may only know that “a torch and stuff” are on the tails side).

[J]udicial independence enables judges to follow the facts and law without fear or favor, so as to uphold the rule of law, preserve the separation of governmental powers, and promote due process. Given these objectives, one may fairly conclude that judges who are subject to intimidation from outsiders interested in the outcomes of cases the judges decide lack the independence necessary to follow the facts and law. At the same time, one may just as fairly conclude that judges who are so independent that they can disregard the law altogether without fear of reprisal likewise undermine the rule of law values that judicial independence is supposed to further.

Geyh, *supra* at 915–16.

167. Judges have intense pressures to resolve cases quickly; it is well documented that the number of cases on their dockets generally overwhelms judges. “‘The justice system’s funding has been decreasing in constant dollars for at least two decades,’ said David Boies, co-chairman of a commission formed by the American Bar Association to study court budget issues. ‘We are now at the point where funding failures are not merely causing inconvenience, annoyances and burdens; the current funding failures are resulting in the failure to deliver basic justice.’” John Schwartz, *Critics Say Budget Cuts for Courts Risk Rights*, N.Y. TIMES, Nov. 27, 2011, at A18, available at <http://www.nytimes.com/2011/11/27/us/budget-cuts-for-state-courts-risk-rights-critics-say.html>. Signs indicate this trend is not going to relent soon. See CONFERENCE OF STATE COURT ADM’RS, PRELIMINARY COSCA BUDGET SURVEY: SUMMARY (2011), available at http://www.americanbar.org/content/dam/aba/administrative/task_force/cosca_bdgtsrvy_maysummaryv3.authcheckdam.pdf (indicating state court systems are facing significant budget shortfalls, and have taken drastic steps to reduce costs, including furloughing judicial officers and staff, delaying filling judicial vacancies, and reducing the use of

2012] CAMPAIGNING BY JUDICIAL RETENTION CANDIDATES 263

glish speakers and pro se litigants) receive a fair hearing, mental competence to discharge his or her duties, and so on.¹⁶⁸ Criticizing or removing a judge for failing to adequately perform these administrative duties should be relatively uncontroversial.

In contrast, only in certain circumstances should a judge be removed for his or her *substantive* rulings.¹⁶⁹ As Justice Stevens explained in his *Republican Party of Minnesota v. White* dissent,

Informed criticism of court rulings, or of the professional or personal conduct of judges, should play an important role in maintaining judicial accountability. However, attacking courts and judges—not because they are wrong on the law or the facts of a case, but because the decision is considered wrong simply as a matter of political judgment—maligins one of the basic tenets of judicial independence—intellectual honesty and dedication to enforcement of the rule of law regardless of popular sentiment.¹⁷⁰

To expound on Justice Stevens' position, judicial errors might be divided into two categories: intentional deviations from the law and honest mistakes. Intentional deviations constitute a willful violation of the judicial oath and should be punished, as they represent a usurpation of the legislative and executive powers to make the law.¹⁷¹ On the other hand, a judge who holds a good faith belief that he or she is following the law and nevertheless makes a mistake should generally avoid reprisal.¹⁷² Note, however, that enough honest mistakes, particularly on questions for which there

retired judges); *Obama Getting Fewer Judges Confirmed Than Nixon*, CBSNEWS (Sept. 6, 2010, 2:53 PM), <http://www.cbsnews.com/stories/2010/09/06/politics/main6839525.shtml> ("Forty-seven [federal judicial] vacancies have been labeled emergencies by the judiciary because of heavy caseloads.").

168. See *Conversation*, *supra* note 6, at 343 (statement of Shirley Abrahamson, C.J.).

169. See *id.* at 341.

170. See *Republican Party of Minn. v. White*, 536 U.S. 765, 802–03 (2002) (Stevens, J., dissenting) (quoting Paul J. De Muniz, *Politicizing State Judicial Elections: A Threat to Judicial Independence*, 38 WILLAMETTE L. REV. 367, 387 (2002)).

171. Geyh, *supra* note 166, at 934–35.

172. *Id.*

To threaten or punish judges with loss of tenure, resources, or jurisdiction for honestly held but unacceptable views of the law, encourages judges to jettison their conceptions of what the law requires in favor of what they believe those in a position to punish them want to hear—which is antithetical to the rule of law values that customary independence is calculated to preserve. This is not to suggest that judges should be unaccountable for unacceptable decisions—decisions at the edge are subject to appellate review. They give rise to discussions in the media, which elicit reactions from voters, who petition the politi-

is an objectively correct answer,¹⁷³ raise questions of judicial competence and make removal a potentially appropriate response.

The primary problem presented by this dichotomy, however, is in identifying which mistakes are intentional and which are honest—an especially difficult enterprise which has in some sense been left to the voters.¹⁷⁴ However, there is no check on the voters regarding the appropriateness of their reasons in voting for or against a judge.¹⁷⁵ As a result, judicial “[e]lections can . . . lumber off, crushing judges who have done their best to follow the law.”¹⁷⁶ Consequently, it seems possible or even likely that certain issues (for example, abortion, same-sex marriage, and gun rights) could become litmus tests in future retention elections, significantly politicizing the process, much in the same way abortion has essentially become a litmus test for U.S. Supreme Court nominees.¹⁷⁷

The possibility that the electorate will retaliate against a judge for a good faith, honest ruling is the danger that springs from excessively emphasizing judicial accountability. However, accountability has its benefits and should not be cast aside entirely. Instead, this value must simply be balanced against the competing interest of impartiality.

3. Informed Electorate

A final interest to be considered in evaluating campaigning rules in retention elections is that of having an informed electorate. Given that citizens will be voting on the retention of their judges, it follows that citizens should be informed about the process of judicial retention elections. As noted above, voters often have little or

cal branches, which explore amendments to existing law that judges must consider anew.

Id. at 926–27.

173. For example, allowing a conviction for treason to stand based upon the testimony of only one witness and no confession by the accused would clearly be incorrect because the U.S. Constitution explicitly requires two witnesses to convict in the absence of a confession. U.S. CONST. art. III, § 3, cl. 1. This is an extreme example to show that not all judicial decisions can be swallowed up by the “subjective” designation that might be thrust upon them. Without getting into an extensive discussion on this topic, suffice it to say that the more routine the matter being decided and the lower the court, the more likely that a judicial question has a legally objective answer.

174. See Geyh, *supra* note 166, at 927, 932–35.

175. Schotland, *supra* note 132, at 124–25 (“One cannot deny that voters can use their opportunity to vote any way they wish.”).

176. Geyh, *supra* note 166, at 924.

177. See, e.g., Daryl J. Levinson, *Parchment and Politics: The Positive Puzzle of Constitutional Commitment*, 124 HARV. L. REV. 657, 744 (2011).

2012] CAMPAIGNING BY JUDICIAL RETENTION CANDIDATES 265

no knowledge regarding the retention election process once they get to the polls.¹⁷⁸ Rolloff in judicial elections is also a concern.¹⁷⁹ Beyond retention elections, increasing knowledge about the judicial branch might also promote confidence in government because the judicial branch is the “least understood branch of government.”¹⁸⁰

Some efforts have already been undertaken to address voter ignorance. In several states, bar associations, state-sponsored commissions, or both have produced judicial performance evaluations to better inform the public regarding the candidates on whom they are voting.¹⁸¹ Whether these guides have been effective is unclear. State commissions almost always recommend that judges be retained and the voters usually follow suit.¹⁸² However, not all voters simply follow the recommendation to retain or reject. Instead, voters in some states, such as Alaska, Arizona, and Colorado, have tended to vote to retain judges with higher evaluation scores while voting against judges with lower evaluation scores, even though the evaluating body recommended that those lower-scoring judges be retained.¹⁸³ But these voting trends were not universal, and in other states, such as Missouri and Utah, were nonexistent.¹⁸⁴

While there is some correlation between recommendations against retention and lower affirmative vote totals, the difference is usually not sufficient to keep a judge from being retained.¹⁸⁵ For example, in 2010, a voting guide was prepared by eleven bar associations in Chicago, summarizing the recommendations of the as-

178. *See supra* pp. 247–48.

179. *See supra* p. 248.

180. *Conversation, supra* note 6, at 341–42 (statement of Shirley Abrahamson, C.J.); *see also* ERIC LANE & MEG BARNETTE, BRENNAN CTR. FOR JUSTICE, A REPORT CARD ON NEW YORK’S CIVIC LITERACY 14 (2011) (explaining that only 55% of those polled in a civic literacy exam could correctly identify the judiciary as the government branch with the power to deem laws unconstitutional, whereas 66% and 60% of respondents could identify the roles of the executive and legislative branches, respectively).

181. *See Judicial Campaigns and Elections: Voter Guides*, AM. JUDICATURE SOC’Y, http://www.judicialselection.us/judicial_selection/campaigns_and_elections/voter_guides.cfm?state (last visited Apr. 3, 2012).

182. Aspin, *supra* note 8, at 222. Although, some research indicates that voters may differentiate between judges who are highly recommended and judges who are also recommended, but not as highly. *Id.* at 222 n.8.

183. Larry T. Aspin, *Retention Elections and Evaluations: A Response to Current Trends in Contested Judicial Elections?*, in *FUTURE TRENDS IN STATE COURTS 2007* 104, 107 (Carol R. Flango et al. eds., 2007), *available at* http://ncsonline.org/WC/Publications/KIS_Ctfutu_Trends07.pdf.

184. *Id.*

185. *See Aspin, supra* note 8, at 223.

sociations.¹⁸⁶ Thirteen of the 65 retention election candidates were not recommended by at least one association, and one candidate failed to receive even one association recommendation. Nonetheless, all 65 judges were retained, including the one not recommended by any bar association.¹⁸⁷

Regardless of the efficacy of voter guides, encouraging judges to meet the electorate in the midst of a retention election may be worthwhile because it provides an additional opportunity for judges to explain the judicial role to the public and for the public to better understand its role in the process.

B. *Current Rules Governing Retention Elections*

1. Judicial Codes of Conduct

Electioneering by judicial candidates is regulated by a judicial code of conduct. The state codes are often arranged into about four to seven canons propounding general principles that guide judicial behavior.¹⁸⁸ Each canon has a number of more specific rules accompanied by commentary on the rules, which is usually derived from either the American Bar Association Model Code or the decisions of state judicial advisory committees.

All but one of the state judicial codes of conduct are based upon some version of the ABA Model Code of Judicial Conduct.¹⁸⁹ Nonetheless, there is great variety among the state analogs. Because the Model Code has changed so much over the years,¹⁹⁰ not all states have kept up with the ABA's changes. Additionally, many states have intentionally deviated from the ABA Model Code.

2. Three Types of Retention Election Campaign Rules

All retention election candidates are permitted to campaign in at least some circumstances; no state prohibits a retention election candidate from campaigning entirely. The particular rules in place

186. *Id.*

187. *See id.* at 223–24. The universally not-recommended judge received a 64.2% affirmative vote, which was sufficient to clear the requisite 60% threshold. *Id.* at 224.

188. *E.g.*, MODEL CODE OF JUDICIAL CONDUCT Canon 1 (2010) (“A judge shall uphold and promote the independence, integrity, and impartiality of the judiciary, and shall avoid impropriety and the appearance of impropriety.”).

189. Brent Dorner, Comment, *2007 Model Code of Judicial Conduct: Are the Speech Restrictions Necessary?*, 33 J. LEGAL PROF. 341, 342 (2009). Montana is the sole exception, though it has rules governing judicial conduct that resemble the Model Code. *Id.* at 342 n.13; *see also* Tom Lininger, *On Dworkin and Borkin*, 105 MICH. L. REV. 1315, 1323 (2007).

190. *See infra* Part II.B.3.

2012] CAMPAIGNING BY JUDICIAL RETENTION CANDIDATES 267

vary by state. There are essentially three different types of rule sets for campaigning in retention elections: the Fixed Time Rule, the Candidacy Rule, and the Active Opposition Rule. Each rule set creates a different time window during which judges are permitted to fundraise and campaign, though some states have separate rules for the timing of fundraising and campaigning (activities collectively referred to as “electioneering” in this Article). This section describes the content of the three rule sets.

Fixed Time Rule

The Fixed Time Rule permits a judge to begin electioneering a certain number of days before the election. The amount of time varies by jurisdiction, ranging from 180 days to 2 years.¹⁹¹ The current ABA Model Code of Judicial Conduct¹⁹² adopts the Fixed Time Rule, though it does not recommend a period of time.¹⁹³ In spite of the ABA’s endorsement, the Fixed Time Rule is the least commonly used rule by states with judicial retention elections, and is found in only three of the nineteen jurisdictions.¹⁹⁴

191. KAN. CODE OF JUDICIAL CONDUCT R. 4.2(B)(1)–(3) (2009), *available at* <http://www.kscourts.org/kansas-courts/supreme-court/orders/2009/2009SC006.pdf> (one year); OKLA. CODE OF JUDICIAL CONDUCT R. 4.2(B)(1)–(3) (2011) (codified at OKLA. STAT. tit. 5, ch.1, app. 4), *available at* <http://www.okbar.org/public/judges/codeOfJudicialConduct.pdf> (180 days); MD. CODE OF JUDICIAL CONDUCT R. 4.1(c)(2)(B)(i) (2010) (two years).

192. MODEL CODE OF JUDICIAL CONDUCT (2011), *available at* http://www.americanbar.org/content/dam/aba/administrative/professional_responsibility/2010_mjcj_final_for_website.authcheckdam.pdf.

193. *See id.* R. 4.2. For a full discussion of the history of the ABA model rule, *see infra* Part II.B.3.

194. Kansas and Oklahoma use the rule for all judicial retention elections. KAN. CODE OF JUDICIAL CONDUCT R. 4.2(B)(1)–(3) (2009), *available at* <http://www.kscourts.org/kansas-courts/supreme-court/orders/2009/2009SC006.pdf>; OKLA. CODE OF JUDICIAL CONDUCT R. 4.2 (2011), *available at* <http://www.okbar.org/public/judges/codeOfJudicialConduct.pdf>. First-time retention election candidates in Maryland can campaign at any time under the Candidacy Rule, *see infra* Part II.B.2.b (describing the Candidacy Rule), while judges who have been retained at least once campaign under the Fixed Time Rule. MD. CODE OF JUDICIAL CONDUCT R. 4.1, 4.4(c) (2010). Regarding Oklahoma, note that one of the comments to the state’s relatively new rule specifies that candidates cannot seek or use endorsements of any kind unless there is active opposition. OKLA. CODE OF JUDICIAL CONDUCT, R. 4.2 Comment [4] (2011). In examining the language of the current rule itself, this comment is likely the result of poor editing rather than an actual rule. First, an extraneous “and” appears after the last portion of the rule. *Id.* R. 4.2(B)(3) (“A candidate for elective judicial office may . . . seek, accept, or use endorsements from any person or organization other than a partisan political organization; and.”). Second, the rule’s predecessor did include an active opposition component. *In re Amendments to the Code of Judicial Conduct*, 1997 OK 79 (1997)

Candidacy Rule

The Candidacy Rule permits a judge to begin electioneering once he or she becomes a “candidate” for reelection. Though this rule appears simple on its face, many jurisdictions only vaguely define when a judge becomes a candidate or else allow a judge to determine when his or her candidacy starts by simply announcing his or her candidacy.¹⁹⁵ This could be years in advance of the actual election.¹⁹⁶ For example, in Illinois, a judge can become a candidate “as soon as he or she makes a public announcement of candidacy,”¹⁹⁷ and, in Iowa, a judge becomes a candidate “as soon as he or she declares or files as a candidate with the election or appointment authority.”¹⁹⁸ The Candidacy Rule is used in eight jurisdictions.¹⁹⁹

(mentioning an active opposition requirement in Canon 5D). Finally, the Scope of the Code states that “[c]omments neither add to nor subtract from the binding obligations set forth in the Rules,” OKLA. CODE OF JUDICIAL CONDUCT Scope (2011). Therefore, it seems that no active opposition requirement of any kind applies to any provision in Oklahoma.

195. See also CHARLES G. GEYH & W. WILLIAM HODES, REPORTERS’ NOTES TO THE MODEL CODE OF JUDICIAL CONDUCT 97, 97 (2009) (indicating one becomes a candidate “largely by self-designation”).

196. *Id.*

197. ILL. CODE OF JUDICIAL CONDUCT Terminology (2007) (codified at ILL. S. CT. R. 67) (2007), available at <http://www2.illinois.gov/jib/Documents/Codeof-Conduct.pdf> (defining “Candidate”).

198. IOWA CODE OF JUDICIAL CONDUCT Terminology (2010) (codified at Iowa Ct. R. ch. 51) (defining “Judicial candidate”).

199. The eight jurisdictions are Arizona, California, Illinois, Iowa, New Mexico, Pennsylvania, South Dakota, and Tennessee. In Illinois, Pennsylvania, and Tennessee, declared candidates may *campaign* at any time, but can only solicit funding in a Fixed Time Rule fashion. See ARIZ. CODE OF JUDICIAL CONDUCT Terminology (2009), available at http://www.azcourts.gov/portals/37/NewCode/2009Code_Internet_Maste_5-03-10.pdf (“Judicial candidate” means any person, including a sitting judge, who is seeking selection for or retention in judicial office by election or appointment.”); Ariz. Judicial Ethics Advisory Comm., Advisory Op. 05-03 (2005), available at http://www.azcourts.gov/Portals/37/ethics_opinions/2005/05-03.pdf (stating that judges running for retention are “subject to the same ethical standards and election laws as all other candidates”); CAL. CONST. art. VI, § 16(d) (1) (describing method of retaining judges); CAL. CODE OF JUDICIAL ETHICS Canon 5 (2011), available at http://www.courts.ca.gov/documents/ca_code_judicial_ethics.pdf (failing to mention any special requirements for judges facing retention elections); ILL. CODE OF JUDICIAL CONDUCT Terminology (codified at Ill. Ct. R. 67) (2009) (defining “candidate”); ILL. CODE OF JUDICIAL CONDUCT CANON 7B(1)(a)(i)–(iii) (allowing judicial candidates to campaign); CODE OF JUDICIAL CONDUCT Canon 7B(2) (prohibiting campaign committees from soliciting funds earlier than one year before the election); IOWA CODE OF JUDICIAL CONDUCT Terminology (defining “judicial candidate”); IOWA CODE OF JUDICIAL CONDUCT Canon 4(B)(1)–(3) (codified at Iowa Ct. R. ch. 51:4.2(B)(1)–(3) (2010)) (permitting re-

2012] CAMPAIGNING BY JUDICIAL RETENTION CANDIDATES 269

Although it may seem unlikely that a candidate would electioneer for an extended period of time, the possibility is not a trifling concern. The 2007 Model Code Revision Committee took the possibility very seriously and “often used as a discussion hypothetical a judge elected to a ten-year term who immediately announced plans to run for reelection.”²⁰⁰

Active Opposition Rule

The Active Opposition Rule permits a judge to campaign only in response to active opposition to his or her retention.²⁰¹ Only one

tion election candidates to campaign); N.M. CODE OF JUDICIAL CONDUCT R. 21-001(A) (defining “candidate”); N.M. CODE OF JUDICIAL CONDUCT R. 21-700(B) (permitting candidates to campaign); PA. CONFERENCE OF STATE TRIAL JUDGES, RETENTION MANUAL 30 (June 2011), *available at* <http://ethics.pacourts.us/RetentionManualFinal.pdf> (allowing a judge to “publicly declare his or her candidacy for retention at any time, thereby becoming a candidate for retention”); JUDICIAL ETHICS COMM. OF THE PA. CONFERENCE ON STATE TRIAL JUDGES, DIGEST OF INFORMAL OPINIONS – OPINION 1/11b/99, *available at* <http://ethics.pacourts.us/digests.htm> (“The time for campaigning is not the same as the time for soliciting under Canon 7B. The Canons do not set forth the time when campaigning can begin. On July 21, 1999 the Judicial Ethics Committee rescinded its Formal Opinion 90-1 which had held that the time when a candidate for judicial office was permitted to begin to campaign was the same as the time when the candidate’s campaign committee was permitted to begin to solicit funds under Canon 7B (2) of the Code of Judicial Conduct.”); PA. CODE OF JUDICIAL CONDUCT CANON 7(B)(2) (2012) (prohibiting soliciting funds earlier than “thirty days prior to the first day for filing nominating petitions or the last day for filing a declaration of intention to seek reelection on a retention basis”); PA. CONST. art. V, § 15(b) (specifying the last day for filing a declaration of candidacy for retention election is “the first Monday of January of the year preceding the year in which his term of office expires”); 42 PA. CONS. STAT. ANN. § 3153(a) (West 2011) (same); S.D. CODE OF JUDICIAL CONDUCT Terminology (2011) (codified at S.D. CODIFIED LAWS app. to ch. 16-2 (2011)) (defining “candidate”); S.D. CODE OF JUDICIAL CONDUCT Canon 5(C)(1)(a)(i)–(iv) (permitting candidates to campaign); TENN. CODE OF JUDICIAL CONDUCT Terminology (2011) (defining “candidate” as including judges facing retention elections); TENN. CODE OF JUDICIAL CONDUCT Canon 5C(1)(b)(i)–(iii) (permitting candidates to campaign); TENN. CODE OF JUDICIAL CONDUCT Canon 5C(2)(a) (prohibiting acceptance of *solicited* funds earlier than 180 days before the election); Tenn. Judicial Ethics Comm., Opinion 01-01 (2001), *available at* <http://www.tncourts.gov/sites/default/files/01-01.pdf> (permitting acceptance of *unsolicited* funds at any time).

200. GEYH & HODES, *supra* note 195, at 106.

201. The Active Opposition Rule was the creation of the 1972 ABA Model Code of Judicial Conduct and was retained through 1990. “The 1972 version of the Model Code of Judicial Conduct permitted a campaign by a judge who ‘has drawn active opposition.’” Utah Ethics Advisory Committee, Informal Op. 00-5 (Aug. 31, 2000), *available at* http://www.utcourts.gov/resources/ethadv/ethics_opinions/2000/00-5.htm; *see also* MODEL CODE OF JUDICIAL CONDUCT Canon 5C (1990) (provision on judges and candidates subject to public election does not include the

of the eight jurisdictions employing this rule requires official certification that there is active opposition.²⁰²

The Florida Code of Judicial Conduct commentary states that “active opposition is difficult to define but is intended to include any form of organized public opposition or an unfavorable vote on

Active Opposition Rule). For more background on the creation of the 1972 Model Code, see THODE, *supra* note 94, at 99.

202. Florida is the exception. In Florida, judges’ political activity is limited by statute. FLA. STAT. ANN. § 105.071 (West 2011). In order to campaign, the judge must submit a certification of active opposition, “specifying the nature thereof,” to the secretary of state. FLA. CODE OF JUDICIAL CONDUCT, Canon 7C(2) (2008). The eight states with the Active Opposition Rule are Alaska, Colorado, Florida, Indiana, Missouri, Nebraska, Utah, and Wyoming. *See* ALASKA CODE OF JUDICIAL CONDUCT Canon 5C(1) (as amended by Order Amending Code of Judicial Conduct Canon 5A(1), Canon 5C, and the Application Section Concerning a Judge’s Activities, 2011 Alaska Ct. Order 1762 (July 1, 2011)) (permitting limited campaign activities without active opposition); ALASKA CODE OF JUDICIAL CONDUCT Canon 5C(2) (permitting advertising in newspapers, on television, and other media, in addition to distributing campaign literature, in response to active opposition); COLO. CODE OF JUDICIAL CONDUCT Canon 4 R. 4.3(A) (2010) (codified at COLO. REV. STAT. ch. 24 app. (2011), *available at* <http://www.cobar.org/repository/JudicialConduct.pdf>) (stating that judges seeking retention “should” not engage in any campaign activity unless that judge faces “active opposition”); FLA. CODE OF JUDICIAL CONDUCT Canon 7C(2) (2008), *available at* http://www.floridasupremecourt.org/decisions/ethics/09-15-2008_Code_Judicial_Conduct.pdf (permitting limited campaign activities in the absence of active opposition); FLA. CODE OF JUDICIAL CONDUCT Canon 7C Commentary (“The term ‘limited campaign activities’ is not intended to permit the use of common forms of campaign advertisement which include, but are not limited to, billboards, bumperstickers, media commercials, newspaper advertisements, signs, etc. Informational brochures about the merit retention system, the law, the legal system or the administration of justice, and neutral, factual biographical sketches of the candidates do not violate this provision.”); FLA. CODE OF JUDICIAL CONDUCT Canon 7C(2) (requiring that judges submit a certification of active opposition “specifying the nature thereof” to the secretary of state in order to campaign “in any manner authorized by law”); IND. CODE OF JUDICIAL CONDUCT R. 4.2(D)(1)–(3) (2011), *available at* http://www.in.gov/judiciary/rules/jud_conduct/jud_conduct.pdf (permitting retention election candidates to campaign in response to active opposition); MO. CODE OF JUDICIAL CONDUCT Canon 5(B)(3) (2010) (permitting retention election candidates to campaign in response to active opposition); NEB. REV. CODE OF JUDICIAL CONDUCT § 5-304.2(B) (2011), *available at* <http://court.nol.org/rules/pdf/Ch5Art3.pdf> (permitting candidates to campaign in response to active opposition); UTAH CODE JUD. ADMIN. ch. 12, R. 4.2(B) (2012), *available at* <http://www.utcourts.gov/resources/rules/ucja/ch12/Canon.4.htm> (permitting a candidate to campaign in response to active opposition); WYO. CODE OF JUDICIAL CONDUCT Canon R. 4.2(B) (2011) (codified at WYO. CODE R. § 4.2 (LexisNexis 2011), *available at* http://www.courts.state.wy.us/CourtRules_Entities.aspx?RulesPages=JudicialConduct.xml (permitting a candidate to campaign in response to active opposition).

2012] CAMPAIGNING BY JUDICIAL RETENTION CANDIDATES 271

a bar poll.”²⁰³ Of course, other jurisdictions are free to adopt their own interpretations. Note, however, that the Active Opposition Rule does not necessarily prevent a judge from responding to all criticism of any kind. As the Utah Judicial Council explained,

[T]here may be other activities, short of operating a campaign, in which a judge could participate. . . . [F]or instance, a judge’s letter to the editor in response to a public letter to the editor would not constitute operating an election campaign. (Although a letter to the editor might implicate other Canons such as those involving the integrity of the judiciary, comment on pending cases, or exhibiting biases and prejudices.) It might therefore be possible for a judge to respond to public comments which do not rise to the level of active public opposition, without the response constituting the operating of an election campaign.²⁰⁴

Some jurisdictions allow candidates to begin fundraising before they are allowed to campaign, though other jurisdictions do not even permit the collection of contributions until active opposition has appeared. If judges are not permitted to fundraise prior to the emergence of active opposition, “judges up for retention can be vulnerable to last-minute attacks.”²⁰⁵

3. History of the ABA Model Rules

The existence of such a variety of rules in different states stems largely from the fact that the ABA Model Code has undergone significant changes over time. The ABA guidelines for judicial behavior were originally created in 1924 and received significant overhauls in 1972, 1990, and 2007.²⁰⁶ Minor amendments were

203. FLA. CODE OF JUDICIAL CONDUCT Canon 7C Commentary.

204. Utah Ethics Advisory Comm., Informal Op. 00-5 (Aug. 31 2000), *available at* http://www.utcourts.gov/resources/ethadv/ethics_opinions/2000/00-5.htm.

205. *How Should We Respond to the 2010 Judicial Elections?*, *supra* note 129, at 105. *See also infra* Part II.B.4 (discussing the merits of the Active Opposition Rule).

206. The original guidelines for judicial behavior were ratified by the ABA in 1924 as the “Canons of Judicial Ethics.” Benjamin B. Strawn, Note, *Do Judicial Ethics Canons Affect Perceptions of Judicial Impartiality?*, 88 B.U. L. REV. 781, 786 (2008). In 1972, these thirty-two canons were consolidated into seven canons contained in the newly-renamed “Model Code of Judicial Conduct.” *Id.* “By 1990, forty-seven states had adopted the 1972 Code or some variation of it.” *Id.* at 787. The Model Code was further consolidated into five canons in 1990 and then into four canons in 2007. *Id.*

made in the intervening years, with the most recent revisions having been made in 2010.²⁰⁷

The 1924 Canons of Judicial Ethics contained no timing recommendations,²⁰⁸ simply providing general guidelines for judicial office candidates.²⁰⁹ The 1971 draft of the Model Code included a Fixed Time Rule for fundraising²¹⁰ along with a prohibition on fundraising earlier than 90 days before the primary election,²¹¹ though each jurisdiction was encouraged to determine an appropriate time limit.²¹² The final version of the Model Code ratified by the ABA in 1972 instead created different rules for judicial candidates participating in contested elections and retention elections.²¹³ Candidates

207. ABA MODEL CODE OF JUDICIAL CONDUCT (2010), available at http://www.americanbar.org/groups/professional_responsibility/publications/model_code_of_judicial_conduct/model_code_of_judicial_conduct_table_of_contents.html.

208. The 1924 Canons of Judicial Ethics were considered “hortatory” rather than compulsory. See Strawn, *supra* note 206, at 786. But see THODE, *supra* note 94, at 43 (indicating that the “Committee consistently took the position that the Code should set enforceable mandatory standards” rather than “hortatory guide-lines”).

209. See CANONS OF JUDICIAL ETHICS CANONS 30 (1924), available at http://www.americanbar.org/content/dam/aba/migrated/cpr/pic/1924_canons.auth_checkdam.pdf (discussing “Candidacy for Office”). Interestingly, the 1924 Canons prohibited receiving campaign contributions from lawyers, see *id.* at Canon 32, but this practice was later approved in a 1941 ABA opinion, THODE, *supra* note 94, at 99; see also ABA Comm. on Prof'l Ethics & Grievances, Formal Op. 226 (1941). Note that because retention elections were not introduced in any state until 1934, separate rules for retention election candidates and other judicial candidates would probably not have been contemplated, though the American Judicature Society endorsed merit selection in 1920. See *supra* notes 43-44 and accompanying text.

210. When the term “fundraising” is used in this Article, this generally refers to fundraising by the candidate’s campaign committee rather than personal solicitations by the candidate, which are prohibited by the current model code. See MODEL CODE OF JUDICIAL CONDUCT R. 4.1(A)(8) (2010) (“[A] judge or a judicial candidate[] shall not . . . personally solicit[] or accept campaign contributions other than through a campaign committee . . .”). However, some courts have struck down prohibitions on personal solicitation as unconstitutional. GEYH & HODES, *supra* note 195, at 99.

211. The language suggests that fundraising for the general election can take place at any time after the 90-day threshold is reached for the primary election. See DRAFT MODEL CODE OF JUDICIAL CONDUCT Canon 7E (1971) (“A candidate’s committees may solicit funds for his campaign no earlier than [90] days before primary election and no later than [90] days after the last election in which he participates during the election year.”) (brackets in original).

212. *Id.* Canon 7E cmt. (1971).

213. The explanation for this distinction is that the Committee concluded that some aspects of merit system elections require special treatment In theory the merit system election removes a judge from politics and from the rigors of the campaign trail, but in a significant

2012] CAMPAIGNING BY JUDICIAL RETENTION CANDIDATES 273

in contested elections could only solicit funds under a Fixed Time Rule, while retention election candidates were governed by the newly created Active Opposition Rule.²¹⁴

This language was kept through 1990, when a Fixed Time Rule for the collection of campaign contributions²¹⁵ and a Candidacy Rule for campaigning were adopted.²¹⁶ This change was made, in part, because the canon containing the Active Opposition Rule was adopted less widely than other canons and was often ignored even when it was in place.²¹⁷ In the drafting stage, the revised rule originally called for different rules for different types of elections, but this approach proved to be too linguistically repetitive, eventually resulting in sets of rules that applied to all judicial candidates.²¹⁸ Because the committee “could find no basis for treating retention elections differently from other public elections,” a rule specifying that retention candidates could not engage in certain political activity was deleted.²¹⁹ The explanation for the removal of the active opposition requirement is surprisingly sparse; the consolidation of the rules for all judicial election candidates apparently simply “eliminated the need for” the Active Opposition Rule.²²⁰

In moving to a Fixed Time Rule, “the most controversial aspect . . . proved to be the suggested time period for election fundraising.”²²¹ The 1990 revised code settled on one year rather than the 90 days found in the 1972 code.²²² The committee extended the time for fundraising out of concern that candidates needed

number of instances the theory fails. . . . In thus authorizing a response analogous to self-defense, the *Code* allows a merit system candidate with active opposition to campaign under the same standard that is applicable to a candidate who is competing against another candidate for judicial office.

THODE, *supra* note 94, at 100.

214. See MODEL CODE OF JUDICIAL CONDUCT Canon 7B(2), (3) (1972) (“[A candidate] whose candidacy has drawn active opposition, may campaign in response thereto and may obtain publicly stated support and campaign funds in the manner provided in subsection B(2).”). See also *supra* note 201.

215. MODEL CODE OF JUDICIAL CONDUCT Canon 5C(2) (1990).

216. *Id.* at 5C(1)(b).

217. LISA L. MILORD, THE DEVELOPMENT OF THE ABA JUDICIAL CODE 46 (1992).

218. *Id.* at 47. The term “all judicial candidates” here includes all individuals seeking appointment to the judiciary, all judicial election candidates, and political activity by incumbent judges.

219. *Id.* at 53.

220. *Id.*

221. *Id.* at 54.

222. MODEL CODE OF JUDICIAL CONDUCT Canon 7B(2) (1972); MODEL CODE OF JUDICIAL CONDUCT Canon 5C(2) (1990).

R

more time to fundraise effectively.²²³ Rather than suggesting a specific period of time, the “Committee intended that each jurisdiction should adopt time limits for pre-election and post-election (if any) fund-raising that best suit[ed] local conditions.”²²⁴ The committee blessed post-election fundraising as a “necessary evil” in order to allow candidates to incur debt to respond to last-second attacks.²²⁵

As noted above, the current ABA Model Code²²⁶ adopts the Fixed Time Rule for both campaigning²²⁷ and fundraising,²²⁸ though it does not recommend a period of time:

Although the creation of this time period is not new, its use in this Rule to disconnect the status of *being* a judicial candidate from being permitted to engage in the *activities* of a candidate is an important feature of the reorganization of Canon 4. During its deliberations, the Commission was mindful of the need to establish a time period to ensure that a judge elected to a ten-year term could not immediately announce plans to run for reelection, establish a campaign committee, and raise campaign funds for almost ten full years. With the time period in place, the judge can continue to call himself or herself a candidate for ten years, but can raise campaign funds only after the time period has been satisfied, typically one year before the first primary.²²⁹

4. Evaluation of Electioneering Rules

In light of the 2010 elections and promises from various groups across the country to engage in active opposition to judges sitting for retention in 2012,²³⁰ electioneering rules and norms for

223. MILORD, *supra* note 217, at 54.

224. *Id.*

225. *Id.*

226. MODEL CODE OF JUDICIAL CONDUCT (2010), *available at* http://www.americanbar.org/content/dam/aba/administrative/professional_responsibility/2010_mcjc_final_for_website.authcheckdam.pdf.

227. *See id.* R. 4.2.

228. *Id.* R. 4.4(B)(2).

229. MODEL CODE OF JUDICIAL CONDUCT, Reporter’s Explanation of Changes, 65–66 (2007), *available at* http://www.americanbar.org/content/dam/aba/migrated/judicialethics/mcjc_2007.authcheckdam.pdf.

230. *See, e.g.*, SKAGGS ET AL., *supra* note 13, at 9 (“Bob Vander Plaats, a failed Iowa gubernatorial candidate who led the Vote No campaign, told his supporters, ‘We have ended 2010 by sending a strong message for freedom to the Iowa Supreme Court and to the entire nation that activist judges who seek to write their own law won’t be tolerated any longer.’”); Schotland, *supra* note 132, at 118 (“The 2010 Iowa judicial elections were, as former Arkansas Governor Mike Huckabee said soon after, of an ‘historic nature,’ likely ‘one that . . . will give legs to a larger

2012] CAMPAIGNING BY JUDICIAL RETENTION CANDIDATES 275

retention election candidates should be reevaluated.²³¹ This admonition is in no way a suggestion that all retention elections are in sudden danger of being politicized, resulting in a catastrophic destruction of judicial impartiality. Instead, this Article aims to initiate a dialogue regarding how judicial retention electioneering rules should be structured going forward.

Because under all three rule sets judges can always campaign in the presence of active opposition,²³² the question regarding what rule should be in place to govern the timing of retention election campaigns is really a question of what rule should govern when a judge does not yet face active opposition. This section explores this question and concludes that the Fixed Time Rule is the most desirable. The Fixed Time Rule best balances the competing values of judicial impartiality and accountability by providing a guaranteed window during which judges can prepare to defend themselves from opposition, and by preventing judges from electioneering perpetually, risking bias or fears of bias. Additionally, the Fixed Time Rule allows judges to engage in greater outreach to the citizenry to explain the courts' role in government.

"Thanks for voting to retain me! I hereby announce my candidacy . . ."

Allowing judges to campaign even when an election is not imminent, as permitted under the Candidacy Rule, unnecessarily opens the door to partiality and the appearance of partiality. Because many jurisdictions have such exploitable rules for becoming a retention candidate, an incumbent judge could become a "candidate" years in advance of the actual election,²³³ and could potentially shake down lawyers and parties for contributions as a regular matter. This risk is potentially ruinous to judicial impartiality while providing no obvious benefit; there is simply no legitimate reason for a judge to "campaign," if activity so far from the election could even fairly be characterized in that way, and fundraise ten years

movement over the next few years.'"). Current Iowa Supreme Court Chief Justice Mark Cady expressed in February 2012 that it is "hard for (him) to tell" if politics have shifted away from attacks on the court. O. Kay Henderson, *Chief Justice "Very Concerned" About November's Retention Election*, RADIO IOWA (Feb. 24, 2012), <http://www.radioiowa.com/2012/02/24/chief-justice-very-concerned-about-novembers-retention-election> (internal quotation marks omitted).

231. *Cf. How Should We Respond to the 2010 Judicial Elections?*, *supra* note 129, at 105 ("The states that are more restrictive ought to at least discuss the dangers these [Active Opposition Rule] restrictions can create.").

232. Assuming the Fixed Time Rule is sufficiently generous.

233. *See supra* Part II.B.2.

before an election. The ABA was correct in its 2007 assessment that such boundless electioneering should not be permitted.

As undesirable as the Candidacy Rule might be, however, one advantage is that it almost certainly passes constitutional muster. In *Republican Party of Minnesota v. White*,²³⁴ the Supreme Court considered constitutional limitations on judicial electioneering rules, striking down the “announce clause” in Minnesota’s Judicial Code of Conduct, which prohibited judicial candidates “from announcing their views on disputed legal and political issues.”²³⁵ Similarly, rules limiting the time during which a judicial candidate can campaign (that is, the Fixed Time and Active Opposition Rules) could be subject to challenge under *White*,²³⁶ given the uncertainty of the breadth of the case. One view holds that, after *White*, “efforts to preserve potent constraints on judicial campaign speech are overwhelmingly doomed to failure,”²³⁷ while the opposite camp takes solace in the majority’s rather cryptic admonition that “we neither assert nor imply that the First Amendment requires campaigns for judicial office to sound the same as those for legislative office.”²³⁸

The more persuasive interpretation of *White* is that the case should not be read as a prohibition against any restrictions upon judicial candidates. *White* certainly did not hold that states must allow judges to perpetually campaign. While drawing an appropriate line in the shadow of *White*’s ambiguities may be difficult, a rule that prohibits campaigning by judges when the next election is still several years away hardly seems unreasonable, particularly in light of the fact that *White* was concerned only with conduct less than two years away from an election.²³⁹

234. 536 U.S. 765 (2002).

235. *White*, 536 U.S. at 768, 787–88.

236. *Cf.* Briffault, *supra* note 151, at 223–28 (discussing questionable constitutionality of bans on personal contribution solicitations by judges). Because numerous articles have been written on *White* and its implications for other canons of judicial behavior, the nuances of the opinion will not be rehashed here.

237. Nat Stern, *The Looming Collapse of Restrictions on Judicial Campaign Speech*, 38 SETON HALL L. REV. 63, 64 (2008).

238. *Republican Party of Minn. v. White*, 536 U.S. 765, 783 (2002); *cf.* Briffault, *supra* note 151, at 186 (“[T]he special nature of the judicial function can justify restrictions on campaign conduct that would not be constitutional in the nonjudicial setting.”).

239. Gregory Wersal, the *White* plaintiff, ran for supreme court justice in 1996 and 1998. *White*, 536 at 768–69. He announced his candidacy for the November 1998 general election in January 1997, a little less than two years before the election. *Republican Party of Minn. v. Kelly*, 63 F. Supp. 2d 967, 973 (D. Minn. 1999). “In February 1998 [eight months before the election], Wersal sought an advisory

2012] CAMPAIGNING BY JUDICIAL RETENTION CANDIDATES 277

Fixed Time Rule vs. Active Opposition Rule

Having dismissed the Candidacy Rule, this Article now turns to the Fixed Time and Active Opposition Rules. Because retention elections are so rarely contested and uncontested judges are almost universally retained, in those jurisdictions employing the Fixed Time Rule, it is difficult to foresee many judges campaigning when they do not face any real opposition—whether it reaches the level of “active opposition” or not. However, the Fixed Time Rule has the benefit of permitting judges to both fundraise and campaign as insurance against potential opposition. During the Fixed Time window, judges might raise campaign funds and prepare campaign literature and media buys, but wait for active opposition to appear before actually campaigning.

The problem with the Active Opposition Rule is that judges cannot respond to late-breaking opposition. Knowing that the judge cannot campaign or even fundraise until the opposition emerges, some calculated opposition efforts will be timed to exploit this weakness. Indeed, precisely this strategy was utilized in 1996 against both Tennessee Supreme Court Justice Penny White and Nebraska Supreme Court Justice David Lanphier, both of whom were defeated in their retention bids after opposition groups launched their attacks only two months before the elections.²⁴⁰ More recently, an unsuccessful campaign opposing the retention of Alaska Justice Dana Fabe in 2010 emerged a mere two weeks before election day.²⁴¹ In light of this risk, some candidates in jurisdictions that permit preemptive fundraising have raised money as insurance against a possible attack that never materialized.²⁴² Thus while the Active Opposition Rule furthers the laudable goal of discouraging

opinion from the Office of Professional Responsibility” regarding whether it would enforce the relevant judicial canon provisions against him. *Id.*

240. Tarr, *supra* note 3, at 613–14. A similar strategy was also used against Kansas Supreme Court Justice Carol Beier in 2010, but was unsuccessful. Sears, *supra* note 34, at 877.

241. *See supra* note 146.

242. *See, e.g.*, Roy A. Schotland, *To the Endangered Species List, Add: Nonpartisan Judicial Elections*, 39 WILLAMETTE L. REV. 1397, 1407–08 n.40 (2003) (listing numerous preemptive fundraising efforts by judges); Tarr, *supra* note 3, at 614 (“[F]earful of opposition by anti-abortion groups that ultimately did not materialize, California Chief Justice Ronald George and Justice Ming Chin raised \$886,936 and \$710,139, respectively, for their retention elections in 1998.”). The preemptive fundraising was presumably permissible because there was either an exception to the pure Active Opposition Rule or the jurisdiction operated under either the Fixed Time or Candidacy Rules, which naturally permit a candidate to preemptively fundraise against feared opposition.

R

R

unnecessary electioneering activities by judges, it undesirably leaves judges open to being caught off guard by last minute attacks.

To some, the Fixed Time Rule might be considered worse than the Active Opposition Rule because the Fixed Time Rule allows judges to fundraise and campaign as insurance, injecting a monetary element into a system that might otherwise lack it and possibly damaging judicial impartiality in the process. This concern is not unfounded; there are real dangers that spring from monetary donations being given to a judicial candidate; the public should diligently ensure that judges are not surrendering their impartiality by being swayed by big donors rather than the law.

However, some element of politics may be unavoidable if *any* election system is employed. Additionally, many jurisdictions prohibit judges from personally soliciting funds,²⁴³ somewhat reducing fears of improper influence. Judges who have reservations about preemptively fundraising are free to refrain from doing so under the Fixed Time Rule, and indeed, could essentially proceed as if the Active Opposition Rule were in place if they so desired.

Moreover, there are strong arguments that judicial campaigning in retention elections promotes an informed electorate. Because of a lack of understanding regarding the judiciary generally, “judges in many states have been encouraged to meet with the public more ‘to talk about the role of the judge.’”²⁴⁴ Limited campaigning—in a prescribed pre-election window—is a way to further this goal. With the Fixed Time Rule, this interest in voter education can be furthered even in the absence of active opposition.

Ultimately, the potential benefits of the Active Opposition Rule are somewhat speculative, and are outweighed by the Fixed Time Rule’s capacity for mitigating the risks posed by last-second attacks and promoting public knowledge about the judiciary. Fixed Time Rule campaigning in the absence of active opposition *protects* the impartiality of the judiciary because it permits judges to shield themselves from political blitzes that would be more effective under a pure Active Opposition Rule, which leaves judges vulnerable by requiring them to react rather than being prepared in advance. If judges can only electioneer in response to opposition, their impartiality may be compromised out of fear that they could not effectively campaign if they needed to do so to retain their seats.

243. Briffault, *supra* note 151, at 223. *But see id.* at 224–28 (discussing possible unconstitutionality of such provisions).

244. *Conversation*, *supra* note 6, at 341 (statement of Shirley Abrahamson, C.J.).

2012] CAMPAIGNING BY JUDICIAL RETENTION CANDIDATES 279

The Fixed Time and Active Opposition Rules are not, however, mutually exclusive. For example, the Active Opposition Rule could be structured such that while campaigning is not permitted, preparatory activities—such as collecting contributions, buying airtime, and printing literature—could be allowed in a Fixed Time Rule manner. In other words, candidates could make preparations during a Fixed Time window but could not begin the actual campaign until active opposition appears. Perhaps judges subject to such a hybrid rule could get the best of both worlds: judges' campaign committees would still collect contributions, but if no campaign is ultimately waged, the potential corrupting influence of these contributions is likely diminished.

A different hybrid rule could permit judges to campaign both in a Fixed Time window prior to the election and before that window if active opposition appears. Because this hybrid takes the precision of the Fixed Time Rule and opens it up to abuse due to the discretionary nature of the Active Opposition Rule, requiring certification of active opposition may be appropriate if this hybrid rule is adopted. Otherwise, the active opposition exception potentially swallows the Fixed Time Rule. On the whole, such a hybrid rule is probably unnecessarily complicated, and a simple Fixed Time Rule is preferable to both hybrids, particularly if the time allotted for fixed time campaigning is sufficiently generous.

Regardless of which particular approach is ultimately accepted, a clear time limit for electioneering before a retention election is necessary. Rather than suggesting a time duration here, this Article follows the guidance of the ABA and leaves this decision up to the individual states.²⁴⁵ Different political conditions and different courts might require different Fixed Time windows. Note that judges compiling war chests for future elections can be thwarted; some states require that excess campaign funds for judicial candidates be surrendered to the state, returned to contributors, or donated to a charitable organization a short time after the election.²⁴⁶

245. *See supra* notes 226–28 and accompanying text.

246. *See, e.g.*, N.M. CODE OF JUDICIAL CONDUCT R. 21-800(E) (2004) (“A candidate for judicial office in either a partisan or retention election who has unused campaign funds remaining after election, and after all expenses of the campaign and election have been paid, shall refund the remaining funds pro rata to the campaign contributors, or donate the funds to a charitable organization, or to the State of New Mexico, as the candidate may choose, within thirty (30) days after the date the election results are certified.”).

The Fixed Time Rule best serves the values of retention elections. By being able to prepare for opposition ahead of time, judges cannot be caught off guard by last-minute attacks. By eliminating this weakness, judges will have more confidence in their ability to win a retention battle and will consequently be more willing to make unpopular, but correct, decisions. In this way, the rule promotes judicial impartiality but leaves the accountability aspect of retention elections undisturbed. Should judges choose to campaign in the absence of opposition, doing so would serve the goal of educating the citizenry about the nature of retention elections and the judicial branch generally. While a few judges might engage in “unbridled” campaigning that would not fulfill these values, experience dictates that the risk is small and insufficient to undermine the probable benefits of this rule.

III. SHOULD JUDGES FACING ACTIVE OPPOSITION CAMPAIGN?

Many see campaigning by judges as anathema to the judicial role. Nonjudicial candidates are permitted, and even encouraged, to engage in a wide range of activities, such as meeting with constituents and promising to change the law.²⁴⁷ The activities of judicial candidates, on the other hand, are much more limited.²⁴⁸ In spite of this, some judicial campaign ads seem to toe the line of propriety, most often because they explain how the judge will be tough on crime; this tactic is routinely decried by the defense bar.²⁴⁹ Judges taking money from potential litigants and their attorneys is also widely frowned upon.²⁵⁰ Yet a judge facing retention who refuses to campaign is more likely to be rejected by voters than one who does campaign.

247. See Schotland, *supra* note 132, at 126 n.32 (indicating that nonjudicial candidates can use open or private meetings, make promises, cultivate and reward sponsorship, participate in diverse multi-member bodies, build up patronage through constituent work, and fundraise).

248. *But see* Republican Party of Minn. v. White, 536 U.S. 765 (2002).

249. See, e.g., Keith Swisher, *Pro-Prosecution Judges: “Tough on Crime,” Soft on Strategy, Ripe for Disqualification*, 52 ARIZ. L. REV. 317, 358 (2010) (explaining that criminal defendants and their attorneys “could have a genuine concern that they will not be facing a fair and impartial tribunal”).

250. See, e.g., ADAM SKAGGS, BRENNAN CTR. FOR JUSTICE, *BUYING JUSTICE: THE IMPACT OF CITIZENS UNITED ON JUDICIAL ELECTIONS* 4–7 (2010) (citing polls of citizens, business leaders, trial lawyers, and judges all indicating a dislike of contributions to judicial campaigns).

2012] CAMPAIGNING BY JUDICIAL RETENTION CANDIDATES 281

Because it is more important for a judge in a retention election to be able to maintain his or her seat on the bench and therefore not subject his or her rulings to the will of the public, retention election candidates should campaign if they believe doing so is necessary. This conclusion is somewhat paradoxical, since fundraising and campaigning may sometimes infringe upon judicial impartiality and the appearance thereof. Nevertheless, campaigning allows a judge to better protect the rulings that he or she has made and in this way protect judicial impartiality generally.

As an initial matter, note that the participants on either side of a retention election battle are uniquely positioned relative to the participants in other judicial elections. In contested non-retention elections (meaning those where opposing candidates vie for the same position), the incumbent and challengers are all bound by the same rules.²⁵¹ In contrast, a judge facing a retention election is bound by a code of judicial conduct, which restricts some of his or her actions. The judge's opposition, meanwhile, has no such restrictions and is likely bound by few rules at all.²⁵² As the history of American politics has shown, some individuals will press the limits of propriety.²⁵³ Though the sitting judge is more restricted in his or

251. See MODEL CODE OF JUDICIAL CONDUCT R. 4.1–4.2 (2010).

252. Cf. Briffault, *supra* note 151, at 196.

253. The Iowa opposition's suggestion that other freedoms would be at risk should the justices be retained comes to mind. See *supra* Part I.B.1. At least some of these suggestions strain credulity. This is not to imply that anything the opposition movement did in Iowa, or any other campaign, was illegal. Instead, there is a wide range of tactics that are legal, but, at least in this author's opinion, are underhanded. As a more recent example, the American Crossroads super PAC, which only can make independent expenditures and is forbidden from making contributions to candidates, has recently explored the possibility of "coordinating" with candidates insofar as the candidates "would be consulted on the advertisement script and would then appear in the advertisement." American Crossroads, FEC Advisory Opinion Request 2011-23, at 3 (2011). Coordinated expenditures are deemed contributions by law. 2 U.S.C. § 441a(a)(7)(B)(i) (2006). The idea behind the American Crossroads request was that, because of the wording of the Federal Election Commission (FEC) regulation defining "coordinated communications," FEC Coordinated Expenditures Rule, 11 C.F.R. § 109.21 (2011), these proposed actions would not fall within the technicalities of the rule and therefore be permissible. See generally American Crossroads, *supra*. This tactic had already been utilized by a state Democratic committee at the time of American Crossroads' request. See *id.* at 2. The FEC ultimately deadlocked 3-3 on the request, leaving the permissibility of the tactic an open question. FEC, Certification in the Matter of American Crossroads, AO 2011-23 (Dec. 5, 2011), available at <http://saos.nictusa.com/aodocs/1189803.pdf>. Permitting this tactic would essentially destroy limits on contributions to candidates. Letter from J. Adam Skaggs & David Earley, Brennan Ctr. for Justice, to Anthony Herman, Gen. Counsel, Fed. Election Comm'n (Nov. 14, 2011), available at http://brennan.3cdn.net/3373da0d1d6f0197db_pum6bnoc

her campaigning, he or she still has the incumbency advantage to offset this edge held by the opposition.

The 2010 Iowa justices' decision not to campaign was certainly well-intentioned. Rather than risk further politicizing the Iowa Supreme Court, the justices stayed above the fray and let the people make their own decision, based upon the information available to them. Justice Streit would later say, "We didn't think Iowa voters would ever swallow Vander Plaats's story, and I don't have that confidence anymore today."²⁵⁴

Beyond a more general politicization of the judiciary, the Iowa justices may have feared the appearance of partiality that can stem from judges accepting campaign contributions.²⁵⁵ This fear is not unfounded. Almost half of surveyed state judges agreed that campaign donations influence judicial decisions, according to a 2001 poll.²⁵⁶ The Conference of Chief Justices, which represents 57 chief justices from every state and U.S. territory, wrote in a 2009 brief to the U.S. Supreme Court: "As judicial election campaigns become costlier and more politicized, public confidence in the fairness and integrity of the nation's elected judges may be imperiled."²⁵⁷

Fears of judicial partiality stemming from campaign contributions may be incongruous with a post-*Caperton* and *Citizens United* world, however. As a result of *Citizens United* and its progeny, nota-

1.pdf (advising the FEC to reject the position taken by American Crossroads in FEC Advisory Opinion Request 2011-23). Of course, this behavior must be expected from the most skilled players of the political "game." See generally DAVID SIRLIN, *PLAYING TO WIN: BECOMING THE CHAMPION* 18 (2005) (explaining that exploiting the weaknesses of the written rules, without violating those rules, and refusing to adhere to unwritten rules of "honor" are essential to success in the realm of high-level competitive game playing). Importantly, Sirlin notes that

[e]xploring extreme 'corner cases' of a game is what high-level play is about.

Exploring extreme situations in life can easily be socially unacceptable, morally wrong, and illegal. [In c]ompetitive games . . . the end (winning) justifies the means (as long as it's through moves the game defines as legal). Real life requires civic virtues like kindness, understanding, justice, and mercy.

Id. at 1.

254. Caldwell, *supra* note 111, at 19.

255. See *supra* p. 254.

256. SAMPLE ET AL., *supra* note 92, at 12. As one Ohio Supreme Court justice put it, "I never felt so much like a hooker down by the bus station in any race I've ever been in as I did in a judicial race." Adam Liptak & Jane Roberts, *Tilting the Scales?: The Ohio Experience; Campaign Cash Mirror's a High Court's Rulings*, N.Y. TIMES (Oct. 1, 2006), <http://query.nytimes.com/gst/fullpage.html?res=9A06E7D81730F932A35753C1A9609C8B63>.

257. SAMPLE ET AL., *supra* note 92, at 9, 12 (quoting Brief of the Conference of Chief Justices as Amicus Curiae in Support of Neither Party at 4, *Caperton v. A.T. Massey Coal Co.*, 129 S. Ct. 2252 (2009) (No. 08-22), 2009 WL 45973, at *4).

2012] CAMPAIGNING BY JUDICIAL RETENTION CANDIDATES 283

bly *SpeechNow.org v. FEC*,²⁵⁸ individuals, organizations, and corporations can now give unlimited amounts to “independent expenditure only committees,” more commonly referred to as “super PACs,” which in turn can make unlimited independent expenditures promoting or opposing candidates.²⁵⁹ Consequently, the decision of the judge to campaign or not may be largely irrelevant; if someone wants to promote a judge’s retention candidacy, that person can simply run his or her own independent campaign to support the candidate, whether the judge likes it or not. While direct campaign contributions can still be made, these are subject to statutory limits in most states.²⁶⁰ As of December 2011, super PACs are already having a major impact in the 2012 presidential election,²⁶¹ suggesting that they may also influence future judicial elections.

Indeed, the Iowa justices and the then-candidate in *Caperton* all received independent support unconnected with the candidates. *Caperton*, however, held that “there are objective standards that require recusal when the probability of actual bias on the part of the judge or decisionmaker is too high to be constitutionally tolerable” and that the \$3 million Blankenship spent in the election was suffi-

258. 599 F.3d 686 (D.C. Cir. 2010).

259. See Richard Briffault, *Super PACs*, 96 MINN. L. REV. 1629, 1631–32, 1647–56 (2012). An independent expenditure, by definition, is not coordinated with a candidate. Coordinating with a candidate causes the expenditure to be converted into an in-kind contribution, subject to contribution limits. Because of this lack of coordination, the Supreme Court held in *Citizens United* “that independent expenditures, including those made by corporations, do not give rise to corruption or the appearance of corruption.” *Citizens United v. FEC*, 130 S. Ct. 876, 909 (2010).

260. See, e.g., SKAGGS, *supra* note 250, at 15 n.2. See also STATE LIMITS ON CONTRIBUTIONS TO CANDIDATES 2011–2012 ELECTION CYCLE, NAT’L CONFERENCE OF STATE LEGISLATURES, http://www.ncsl.org/Portals/1/documents/legismgt/Limits_to_Candidates_2011-2012.pdf (last updated Sept. 30, 2011).

261. See, e.g., T.W. Farnam & Dan Eggen, *Super PACs Herald a Punishing Campaign Season*, WASH. POST (Dec. 8, 2011), http://www.washingtonpost.com/politics/super-pacs-herald-a-punishing-campaign-season/2011/12/08/gIQAEiHPgO_story.html; Maggie Haberman & Kenneth P. Vogel, *Adelson Discusses \$20 Million Check to Pro-Newt Gingrich Group, Denies Commitment*, POLITICO (Dec. 15, 2011, 1:49 PM), <http://www.politico.com/news/stories/1211/70501.html> (discussing possibility that casino magnate would direct \$20 million to super PACs supporting Newt Gingrich); *Who’s Financing the ‘Super PACs’*, N.Y. TIMES, <http://www.nytimes.com/interactive/2012/01/31/us/politics/super-pac-donors.html> (last updated May 7, 2012). The current contribution limit to a presidential candidate is \$5,000 annually—\$2,500 for the primary election and \$2,500 for the general election. See 2 U.S.C.A. § 431 (West 2012); *Contribution Limits 2011-12*, FEC, <http://www.fec.gov/pages/brochures/contriblimits.shtml> (last visited Dec. 16, 2011).

cient to require the justice to recuse himself.²⁶² Hence, while the independent expenditures themselves were permissible in both cases on First Amendment grounds, the independent expenditures in *Caperton* required recusal of the justice on due process grounds. It was in this way that Justice Kennedy—who authored both *Caperton* and *Citizens United* on behalf of five-member majorities—distinguished *Caperton* in *Citizens United*.²⁶³ The end result is that one is free to support a judicial candidate through independent expenditures as much as one wants. If, however, one goes over some indeterminate threshold of support, the judge is required to recuse him or herself from deciding particular cases.

While one might initially conclude that the Iowa justices valued remaining apolitical over winning, it was not that choosing to campaign led to certain victory and choosing not to campaign led to certain defeat. Instead, from the justices' perspective, campaigning entailed costs that they were unwilling to bear. If the justices viewed campaigning as an absolute evil, rather than an undesirable means to the desirable end of winning, then campaigning was simply never a possibility, whatever the cost. The effort and strain involved in campaigning might also have been viewed as burdensome and not worth the likely, though not certain, increase in affirmative votes.²⁶⁴ This author speculates that the justices perceived campaigning as an absolute evil and therefore decided they would rather face a greater risk of not being retained than taint their honor by violating a perceived absolute prohibition against campaigning.²⁶⁵

However noble their intentions might have been, the justices were not facing novices. The opposition consisted of “veterans of the previous battles against same-sex marriage in other states [who] had become adept at distilling years of history and legal scholarship into 30-second sound bites.”²⁶⁶ The justices needed all the help they could get if they were going to win. As mentioned earlier, failing to campaign may have actually been the difference in Iowa in

262. *Caperton v. A.T. Massey Coal Co.*, 129 S. Ct. 2252, 2257, 2263–64 (2009) (internal citation omitted) (internal quotation marks omitted).

263. See *Citizens United*, 130 S. Ct. at 910 (“*Caperton*’s holding was limited to the rule that the judge must be recused, not that the litigant’s political speech could be banned.”); see also Richard L. Hasen, *Citizens United and the Illusion of Coherence*, 109 MICH. L. REV. 581, 611–15 (2011) (discussing the inconsistency between *Citizens United* and *Caperton*).

264. The strain of campaigning on a judge might be even higher relative to a nonjudicial candidate due to the special restrictions with which the judge must comply. See *supra* Part III.

265. See *supra* Part I.B.1.

266. Cicoski, *supra* note 109, at 19.

2012] CAMPAIGNING BY JUDICIAL RETENTION CANDIDATES 285

2010.²⁶⁷ While the justices initially accepted their defeats with grace, Justice Streit later expressed that he was “still angry—about the lack of support he and his fellow justices received from groups that promised it [and] about the bind he and his fellow justices were put in.”²⁶⁸ Justice Streit believes in hindsight that the justices should have campaigned.²⁶⁹

Though the justices may have felt confident that they did the right thing at the time, both in their judicial decisions, most notably legalizing same-sex marriage in Iowa, and in their decision not to campaign, this is cold comfort to those in Iowa who would benefit from unpopular decisions. As Wisconsin Chief Justice Shirley Abrahamson explains, “[J]udicial independence is not for the judges, and it’s not for the lawyers; it’s for the people who come to court, and it’s for everyone else who doesn’t come to court but whose life is affected by what happens in court.”²⁷⁰

The Iowa justices should have campaigned in response to the opposition.²⁷¹ By failing to campaign and defend their seats, the justices partially ceded their judicial impartiality. Instead of ensuring that their own interpretation of the Iowa Constitution remained in place, they increased the risk that the views of the opposition would become governing law. Of course, this is a bit of an oversimplification, as it is not the opposition that gets to choose the judge’s successor, but the selection commission and the governor. However, a governor who knows why a judge was not retained would be likely to give the opposition what it wants. Therefore, the Iowa justices’ decision not to campaign may have endangered the *Varnum* opinion.

That being said, such a change in the membership of a court does not necessarily mean that the unpopular decision will be overruled. Perhaps the replacements will believe that the case was rightly decided or should remain undisturbed on stare decisis grounds.²⁷² However, knowing that the crocodile in the bathtub at-

267. Curriden, *supra* note 9, at 58.

268. Caldwell, *supra* note 111, at 19.

269. *See id.*

270. *Conversation*, *supra* note 6, at 341 (statement of Shirley Abrahamson, C.J.).

271. *See, e.g.*, Schotland, *supra* note 132, at 124 (explaining that “most knowledgeable observers have faulted” the justices for not campaigning, though Schotland himself does not).

272. For example, the Iowa Supreme Court has explained that “we do not overturn our precedents lightly and will not do so absent a showing the prior decision was clearly erroneous. Stare decisis does not prevent the court from correcting past judicial announcements that were plainly wrong.” *Iowa v. Derby*, 800

tacked once for a particular move, the replacement judge may reasonably believe the crocodile might attack again for the same “offense.”²⁷³ Indeed, the same-sex marriage crocodile would have two feet out of the bathtub before the replacement justices even arrived. To continue shaving after having seen what the crocodile did to one’s predecessor would be quite a feat indeed.

Ultimately, judges should make their decisions to the best of their professional ability and defend them by campaigning whenever opposition appears. “While . . . ‘good’ judges will ignore the political consequences of unpopular decisions, these ‘good’ judges will be the very ones most likely to be removed.”²⁷⁴ By the time a sitting judge is faced with the question of whether to campaign in response to active opposition, the crocodile is already out of the bathtub. It is too late to step out of the bathroom; the only choice is whether to fight back. Certainly there is something to be said for attempting to preserve the judiciary’s impartiality by not campaigning. But this value is largely unattainable once a major opposition appears—the election is already highly politicized. If the judge is rejected, his or her replacement will be viewed as the triumph of this politicized opposition movement. If the judge is retained, the nonacceptance of the political opposition becomes a political affirmation of the judge.

One need only look to Chief Justice Kilbride in the 2010 Illinois retention elections for a counterexample to the Iowa justices. Kilbride raised a staggering \$2.7 million to combat the \$700,000 campaign waged against his retention. Rather than allow the opposition effort to unseat him, he decided to fight back.²⁷⁵ It is certainly possible that Kilbride could have won without campaigning, particularly because the true motives behind the opposition effort were uncovered before the election and were received negatively.²⁷⁶ In essence, Justice Kilbride correctly valued winning highly and discounted the detriments of campaigning by judges. But it is certainly possible that Kilbride could have won without campaigning; particularly because the true motives behind the opposition effort were uncovered before the election and were received negatively.

N.W.2d 52, 59 (Iowa 2011) (internal citations omitted) (internal quotation marks omitted).

273. As Albert Einstein, and later Justice Scalia, put it, “Insanity . . . is doing the same thing over and over again, but expecting different results.” Sykes v. United States, 131 S. Ct. 2267, 2284 (Scalia, J., dissenting).

274. Uelmen, *supra* note 82, at 2072.

275. See *supra* notes 137–141 and accompanying text.

276. See *supra* notes 137–141 and accompanying text.

2012] CAMPAIGNING BY JUDICIAL RETENTION CANDIDATES 287

Finally, the mere existence of active judicial campaigns does not mean that judges will descend into the mud with their opposition. Judges can, and overwhelmingly do, run “predominantly positive, traditionally themed advertisements.”²⁷⁷ Candidates in nonpartisan elections ran zero attack ads in 2010.²⁷⁸ Most judges who campaign do so in a manner that is respectful of the judiciary’s role. The negative ads are largely being distributed by outside groups that the judge cannot control, regardless of whether he or she campaigns.²⁷⁹

CONCLUSION

An impartial judiciary is essential to this country’s system of law, but judges should not be completely unaccountable. The Fixed Time Rule best balances these competing interests by allowing a judge to make correct but unpopular rulings because it gives the judge a fair chance to defend his or her seat without allowing the judge to engage in a perpetual campaign. Even in the absence of opposition, judges should be permitted to electioneer a certain amount of time before a retention election to preemptively guard against would-be attackers and to inform the electorate regarding the retention election process as well as the judicial role. Thus states with retention elections should adopt the Fixed Time Rule to govern judicial electioneering.

Further, encouraging judges facing opposition to campaign makes the best of a bad situation and protects the judges’ impartiality and the rule of law. A judge facing retention does not relinquish his or her impartiality by campaigning, but instead embraces it by increasing the chance that he or she will remain on the bench and be able to defend the court’s precedents. A failure to campaign invites the rule of the masses rather than the rule of law. When faced

277. SKAGGS ET AL., *supra* note 13, at 13.

278. *Id.* at 18. However, such positive ads by the candidates do not necessarily lead to a collegial court; some judges are literally at each other’s throats. In June 2011, it was reported that Wisconsin Supreme Court Justice David Prosser, who had recently won a contested election to retain his seat, and Justice Ann Walsh Bradley were involved in a physical altercation where Bradley may have charged at Prosser and Prosser may have choked Bradley. See Crocker Stephenson, Cary Spivak & Patrick Marley, *Justices’ Feud Gets Physical*, JSONLINE (June 25, 2011), <http://www.jsonline.com/news/statepolitics/124546064.html>. No charges were filed over the alleged altercation. Andrew Harris, *Wisconsin Judges in Alleged Fracas Won’t Be Charged, DA Says*, BLOOMBERG (Aug. 25, 2011, 6:07 PM), <http://www.bloomberg.com/news/2011-08-25/wisconsin-judges-in-alleged-assault-won-t-face-charges-prosecutor-says.html>.

279. See *supra* Part III.

288 NYU ANNUAL SURVEY OF AMERICAN LAW [Vol. 68:239

with the bathtub crocodile lumbering toward him, the judge should fight back with his razor rather than be devoured by the beast out of fear that his blade might be dulled.

SCALING THE PATENT SYSTEM

CHRISTINA MULLIGAN & TIMOTHY B. LEE*

ABSTRACT

Why do firms in some industries ignore patents when developing new products? This paper posits a simple answer to this long-puzzling question: firms ignore patents because they are unable to discover the patents their activities might infringe. The costs of finding relevant patents, which we call “discovery costs,” are prohibitively high.

Not all industries face high patent discovery costs. Chemical patents are “indexable,” meaning that relevant patents can be efficiently retrieved by chemical formula. As a result, discovery costs in the chemical and pharmaceutical industries are low, and inadvertent infringement by firms in these industries is rare. But many other patent categories are not indexable, and in some cases that makes avoiding infringement practically impossible. In software, for example, patent clearance by all firms would require many times more hours of legal research than all patent lawyers in the United States can bill in a year. The result has been an explosion of patent litigation.

This paper attacks two core premises of patent law—that parties are always able to respect each other’s patent rights, and that firms should be punished for infringement even if they could not have avoided it. It concludes with several suggestions for how to change the patent system to alleviate the problems created by non-indexable patents.

Introduction	290	R
I. Scale and Big-O Notation	292	R

* Christina Mulligan is a Postdoctoral Associate and Lecturer in Law at the Information Society Project at Yale Law School, and beginning August 2013, an Assistant Professor of Law at the University of Georgia School of Law. Timothy B. Lee is a computer scientist and journalist. He is an adjunct scholar at the Cato Institute and covers tech policy for *Ars Technica*. The authors wish to thank Bryan Choi, James Grimmelman, Daniel Hemel, Ben Klemens, Tom Lowenthal, Evan McClanahan, Kevin E. Park, Amanda Rohn, Wendy Seltzer, Samson Vermont, Carly Weinreb, Aaron Williamson, and the attendants of Patent Conference 2 and IP Scholars Conference 2012 for their helpful comments on earlier drafts of this paper.

II. Analyzing the Patent System’s Scalability with Big-O		
Notation.....	294	R
A. Simple Model of the Patent System	294	R
B. Indexing Lowers Discovery Costs.....	295	R
C. Example: Chemical Patents are Indexable	297	R
D. Example: Software Patents Are Disorganized	297	R
E. Example: Discovery Costs for Corkscrews Are		
Low	305	R
III. Disorganized Patents Have Led to A Litigation		
Explosion	307	R
IV. Policy Suggestions: Beyond Patent Quality	309	R
A. Subject Matter Restriction	310	R
B. Independent Invention	313	R
C. Limiting Injunctions and Multiplied Damages for		
Patent Infringement	314	R
Conclusion	317	R

INTRODUCTION

In 1945, Friedrich Hayek wrote a famous essay called “The Use of Knowledge in Society.”¹ Responding to advocates for greater central planning of economic activity, Hayek pointed out that their theories assumed that knowledge about economic circumstances could be taken as “given” to economic decision-makers, an assumption he argued was unreasonable.² Knowledge about the state of the economy—about what resources exist and what goods and services consumers demand—is dispersed among millions of people. Gathering the information together in one place would be impractical, and even if it were done, no economic decision-maker could possibly absorb it all.³

The tendency to implicitly assume that economic actors are omniscient is a common pitfall of theoretical social science. By definition, the theorist knows everything there is to be known about the stylized model he has invented. Theorists often implicitly assume that economic actors automatically have the information they need to make decisions.⁴ Indeed, such assumptions may be essential to building a tractable model of the world. But the failure to ponder the feasibility of acquiring and using information can lead to flawed conclusions.

1. Friedrich A. Hayek, *The Use of Knowledge in Society*, 35 AM. ECON. REV. 519 (1945).

2. *Id.* at 519.

3. *Id.* at 524.

4. *See id.* at 519.

The contemporary patent debate suffers from this blind spot. Each patent is a demand that the world refrain from practicing a claimed art without the patent holder's permission. Potential infringers can only comply with this demand if they are aware of the patent's existence. On a blackboard or in the pages of a law review article, it is easy to assume that everyone knows about every patent.

But the real world is not so simple. To avoid infringement, a firm must expend resources to learn about potentially relevant patents. Typically this means hiring patent lawyers to conduct patent searches, which may or may not be affordable or effective.⁵ In this paper, we'll call the costs of such information-gathering activities the patent system's "discovery costs."⁶ One criterion for a well-functioning patent system—or any system of property-like rights—is that discovery costs be low enough to make it economically feasible for firms to obtain the information they need to comply with the law.⁷

5. For example, a recent Supreme Court case presumed that "simply asking an attorney to examine a product and compare it to the data base of existing patents is not a dependable way to see if a product is likely to infringe a patent." Stephen M. McJohn, *Top Tens in 2011: Patent & Trademark Cases*, 10 NW. J. TECH. & INTEL. PROP. 313, 317 (2012) (discussing *Global-Tech Appliances, Inc. v. SEB S.A.*, 131 S.Ct. 2060 (2011)). See also *Global-Tech Appliances*, 131 S.Ct. at 2064 (describing how a patent attorney failed to locate a deep fryer patent that his client's invention infringed).

6. Discovery costs are one of the types of transaction costs identified by Ronald Coase in his seminal essay, *The Problem of Social Cost*. Coase explained:

In order to carry out a market transaction *it is necessary to discover who it is that one wishes to deal with*, to inform people that one wishes to deal and on what terms, to conduct negotiations leading up to a bargain, to draw up the contract, to undertake the inspection needed to make sure that the terms of the contract are being observed, and so on. These operations are often extremely costly, sufficiently costly at any rate to prevent many transactions that would be carried out in a world in which the pricing system worked without cost.

R. H. Coase, *The Problem of Social Cost*, 3 J.L. & ECON. 1, 15 (1960) (emphasis added); see also Paul J. Heald, *Optimal Remedies for Patent Infringement: A Transactional Model*, 45 HOUS. L. REV. 1165, 1167–68 (2008) (explaining Coase's description of transaction costs "in the context of contracting for inventions"). Discovery costs as used in this article are notably not costs associated with the discovery period before a trial.

7. Clarisa Long has argued that the patent system does not "raise the information costs of searching and avoiding [most patented goods] unduly." Clarisa Long, *Information Costs in Patent and Copyright*, 90 VA. L. REV. 465, 524 (2004). See *id.* at 532–33 ("[Specialized and knowledgeable practitioners] will be able to draw on their preexisting knowledge of goods and technologies in the relevant field and as a result search costs will be lower than if they were not knowledgeable. . . . Requiring a small set of people to search exhaustively is not as socially expensive [in patent law as searches by many are in copyright]."); see also *id.* at 503 ("We would expect legal rules to force disclosure . . . , and . . . increase duties of avoidance, when the class of goods is small . . . , when the goods affect fewer observers, when

Thinking explicitly about discovery costs is a powerful tool for understanding the dysfunctions of the patent system. As we will see, discovery costs are relatively low in pharmaceuticals and other chemical industries.⁸ As a consequence, the patent system serves these industries relatively well. In contrast, discovery costs in the software industry are so high that most firms do not even try to avoid infringement.⁹ Unsurprisingly, software is a major contributor to the recent spike in patent litigation.

We will argue that this disparity can be explained by the fact that pharmaceutical inventions can be organized by chemical formula, while no analogous organizational scheme exists for software inventions.

I. SCALE AND BIG-O NOTATION

Our subject is what software engineers call “scalability,” the ability of a system to perform well as the “problem size” increases.¹⁰ A common experience for companies that build online services is to have a system that worked flawlessly with a limited number of test users grind to a halt when it is released to the public and used by millions of people.¹¹ Often, a system’s bottlenecks only become apparent when it is used at its full capacity.

He didn’t put it in these terms, but Hayek was essentially arguing that central planning doesn’t scale. Centralized economic decision-making can work in a small tribe whose chief knows every tribe member personally. But in a modern economy with millions of households, a single, central decision-maker would become a bottleneck, causing the entire system to grind to a halt. So modern societies have developed other mechanisms, such as the price sys-

those observers have greater tolerance for incurring costs of understanding the good, and when the disclosed information is objective and readily verifiable. This is indeed what we see with the patent form.”). This paper disagrees with Long’s view insofar as this paper argues that the costs of discovering whether an independently-created invention has already been patented are actually unduly high.

8. See *infra* Part III.C.

9. See *infra* Part III.D.

10. See generally CAL HENDERSON, BUILDING SCALABLE WEB SITES: BUILDING, SCALING, AND OPTIMIZING THE NEXT GENERATION OF WEB APPLICATIONS, ch. 9 (2006).

11. For example, Twitter has experienced repeated service outages as its user base has grown. See, e.g., Joe Tacopino, *World Cup causes Twitter outages, more ‘fail whales’ to come*, N.Y. DAILY NEWS, June 17, 2010, <http://www.nydailynews.com/news/money/world-cup-twitter-outages-fail-whales-article-1.180774>.

tem, to coordinate economic decisions in a decentralized, scalable fashion.

Scalability is also an important issue for the patent system. In an island nation of, say, 100,000 people, the patent system's discovery costs would not be a cause for concern. There might be a few patents granted each week, and it would be reasonable to simply expect every firm to read every patent. In contrast, the U.S. Patent & Trademark Office issues several *thousand* patents per week.¹² Clearly, it is not possible to expect every American firm to read and understand every issued patent. So a scalable patent system needs to offer firms efficient mechanisms to sort through those thousands of patents to find the ones that are relevant to them.

To think rigorously about whether patent searching is scalable, we're going to borrow the standard notation computer scientists use to talk about scalability, called "Big-O" notation.¹³ Big-O notation is a way to succinctly summarize how quickly a function grows relative to its input. To illustrate this concept, we'll use the example of a hypothetical chess tournament. Imagine you are planning a chess tournament with n players. You have only one chess set, so games have to be played in sequence. You are deciding between two tournament styles: a round-robin tournament in which every player plays one game with every other player, or a single-elimination tournament in which players are paired off and the loser of each game is dropped from the tournament.

If we assume that each game lasts one hour, and there are no breaks, the round-robin tournament will take $n(n-1)/2$ hours for n players.¹⁴ If n is large, the tournament will be intolerably long. For example, if $n=100$ the tournament will take more than six months!

Now consider the single-elimination tournament. If we again assume that each game lasts one hour, then the entire tournament will take $n-1$ hours. This is much more manageable; with $n=100$ players the tournament will take about four days.

12. U.S. Patent Statistics, Calendar Years 1963–2010, UNITED STATES PATENT & TRADEMARK OFFICE, http://www.uspto.gov/web/offices/ac/ido/oeip/taf/us_stat.pdf (noting that 219,614 utility patents were granted in 2010, with an average of 4223 utility patents granted per week).

13. See Paul E. Black, *big-O notation*, in *DICTIONARY OF ALGORITHMS AND DATA STRUCTURES* [ONLINE] (Paul E. Black, ed. 2010), <http://www.nist.gov/dads/HTML/bigOnotation.html> (last updated August 31, 2012); JON KLEINBERG & EVA TARDOS, *ALGORITHM DESIGN* 36–37, 47–56 (1st ed. 2006).

14. If each of n players played every other contestant, the tournament would take $n(n-1)$ hours. However, this would result in each pair playing each other twice, because after player A challenged player B, player B would then challenge player A. So we divide by two to prevent duplicate games.

To find the “Big-O” class of a function like $n(n-1)/2$, one expands the function to $1/2n^2 - 1/2n$ and drops everything other than the fastest-growing exponent, leaving n^2 .¹⁵ We can then say that the round robin tournament takes $O(n^2)$ time to complete, given n participants. In contrast, the single-elimination tournament takes just $O(n)$ time to complete.

This means that round-robin tournaments scale poorly compared to single-elimination tournaments. Doubling the number of players roughly doubles the length of the single-elimination tournament, while it increases the length of the round-robin tournament by a factor of four.¹⁶ For large tournaments, this effect will dwarf other considerations in choosing between the two options.

II.

ANALYZING THE PATENT SYSTEM'S SCALABILITY WITH BIG-O NOTATION

A. *Simple Model of the Patent System*

Now we use Big-O notation to evaluate the scalability of the patent system. We start by constructing a stylized model of the patent-eligible widget industry. Assume there are n firms producing widgets and each firm produces just one type of widget and holds just one patent. We will relax these assumptions later, but the simplicity of this model makes it a good starting point.

Each firm's widget may infringe multiple competitors' patents, and firms spend resources to learn which patents they must license or invent around. How large are these discovery costs? Suppose that the only known way to find the patents related to a particular widget is to examine all widget patents, and that lawyers can always correctly determine whether a widget violates a patent simply by

15. Factors besides the fastest-growing exponent are dropped because their effect on the length of time it takes to solve a problem becomes mathematically insignificant as the problem size increases.

16. We can demonstrate this by plugging p players and $2p$ players into the equations for determining the length of a single elimination and round robin tournament: $n-1$ and $1/2n^2 - 1/2n$, respectively. P players could finish a single elimination tournament in $p-1$ hours and could finish a round robin tournament in $1/2p^2 - 1/2p$ hours. To determine how long it would take $2p$ players to finish a tournament, we replace “ n ” in the equation with “ $2p$ ”. Twice as many players would take almost twice as long to finish a single elimination tournament as p players— $2p-1$ hours compared to $p-1$ hours. In contrast, a round robin tournament of $2p$ players would take $1/2(2p)^2 - 1/2(2p)$ hours, which multiplies out to $2p^2 - p$ hours. Whereas the round robin tournament of p players took almost about $1/2p^2$ hours, the tournament of $2p$ players takes almost $2p^2$ hours—quadrupling the time even though the players only doubled.

reading the patent claims.¹⁷ Suppose also that it takes a patent lawyer one hour to determine whether a given widget-related product infringes a given widget patent. Then each of the n widget firms will need to pay a patent lawyer to examine the patents held by each of the $n-1$ other widget firms. That's $n-1$ hours of work for each firm, leading to an industry-wide discovery cost of $n(n-1)$, or $n^2 - n$, billable hours. In other words, the discovery costs of the patent system are $O(n)$ for each firm, and $O(n^2)$ across the entire widget industry.

This means that the patent system scales poorly for widgets.¹⁸ If the number of widget firms (and, with it, the number of patents) doubles, the industry's total legal bills increase by a factor of four. In an industry with many firms, the patent system's discovery costs would be a large burden; they could even dwarf some firms' revenues altogether. For example, in a widget industry in which 30,000 firms had one patent apiece and could review one patent per hour, each firm would need to hire around fifteen full-time patent attorneys, resulting an industry-wide total discovery cost of almost a billion billable hours.¹⁹ That's a lot, even for deep-pocketed widget companies.

B. Indexing Lowers Discovery Costs

The above analysis assumes that every firm must examine every patent in some detail, but this is not practical in a world where there are hundreds of thousands of valid patents. Ideally, firms in real industries would have ways to quickly find the small number of

17. In the real world, lawyers frequently cannot state with certainty whether a given activity actually infringes a particular patent. In this paper, we will largely set this issue to the side and assume counterfactually that lawyers can always determine whether a particular activity infringes a particular patent in a reasonable amount of time. For further reading on the challenges of claim construction and determining the scope of patents, see Christopher A. Cotropia, *Patent Claim Interpretation and Information Costs*, 9 LEWIS & CLARK L. REV. 57 (2005); Christopher A. Cotropia, *Patent Claim Interpretation Methodologies and their Claim Scope Paradigms*, 47 WM. & MARY L. REV. 49 (2005); Jeanne Fromer, *Claiming Intellectual Property*, 76 U. CHI. L. REV. 719 (2009); Michael Risch, *The Failure of Public Notice in Patent Prosecution*, 21 HARV. J. L. & TECH. 179 (2007).

18. Running time can get much worse for computer scientists, who sometimes must solve problems that can only be solved in non-polynomial time, such as $O(k^n)$ for some constant k and number of inputs n .

19. Each of the 30,000 firms would need to hire a patent attorney to examine 29,999 other patents, which takes 1 hour per patent, so attorneys would spend a total of $30,000 * 29,999 * 1 \text{ hour} = 899,970,000$ hours. Assuming a typical attorney bills 2000 hours of work per year, each firm would need just shy of 15 attorneys to examine 29,999 patents.

patents that relate to its own products and ignore the rest. The ability to do this depends on a good system of organization.

We can see the power of organization in the system of real property. Counties do not store real property records in a random order. Rather, they place them in a predictable order based on their geographic location. Filing them in a predictable order allows rapid record retrieval in the same way that alphabetization allows rapid lookup of words in a dictionary. In Big-O terms, retrieving one item from a well-organized collection is roughly an $O(1)$, or constant-time, operation.²⁰ If a clerk needed to retrieve each record once during the year, he would only spend $O(n)$ time sorting through filing cabinets, where n is the total number of records.

By way of contrast, we can imagine a county with an incompetent records clerk who placed property records in filing cabinets at random. In this county, the property system would have the same scaling problem as widget patents: as the number of parcels increased, looking up who owned any given parcel would become more and more time-consuming. The only way to find a particular record would be to examine every record, one at a time. That means retrieving all records related to a particular parcel would be an $O(n)$ operation. If a clerk needed to retrieve each record once during the year, he would spend $O(n^2)$ time sorting through filing cabinets. In a county with many parcels, the system would be completely unmanageable.

The ability to organize claims to real property such that they can be quickly retrieved depends on the existence of a standardized and predictable representation such as geographic coordinates. If a group of items has such a representation, we call that group “indexable,” because the representation makes it possible to build an efficient index of the items. Whether a set is indexable or not depends on its inherent properties. Dictionary words are indexable because they can be organized alphabetically. Real property claims are indexable because they can be organized by their geographic location.

20. Constant-time operations, or operations that take $O(1)$ time, take the same amount of time to complete, regardless of the problem size. Depending on the details of the filing method, lookup times might be a slow-growing function like $O(\log_2 n)$. The difference between $O(1)$ and $O(\log_2 n)$ is not large, and we’re going to pretend that they are the same to simplify the presentation.

C. Example: Chemical Patents are Indexable

Are patents indexable? This is a difficult question to answer in general, since the answer varies by technology class. But at least one category of patents—chemical patents—is already being indexed.

The Food and Drug Administration produces a publication called *Approved Drug Products with Therapeutic Equivalence Evaluations*.²¹ This publication, colloquially known as the “Orange Book,”²² allows people to look up pharmaceutical patents based on the chemical formula of the active ingredient. A German organization, FIZ Karlsruhe, offers an electronic database called STN, which allows researchers to pull up all patents and other literature on particular molecules.²³

Chemical formulas allow efficient retrieval of chemical patents, just as geographic coordinates allow efficient retrieval of real estate records. In other words, chemical patents are indexable. That means that finding a patent based on its chemical formula is approximately an $O(1)$ operation, just as it is for real property records.

Recall that doubling the number of widget patents doubled every widget firm’s discovery costs, since each firm was forced to look at twice as many patents to weed out the irrelevant ones. In contrast, doubling the number of chemical patents does not increase chemical firms’ discovery costs because a database like STN can quickly filter out irrelevant patents, no matter how many of them are in the database. So the patent system scales well for chemicals.

D. Example: Software Patents Are Disorganized

Unfortunately, few, if any, non-chemical patents seem to be indexable. To be sure, there are searchable databases that include non-chemical patents.²⁴ But few, if any, non-chemical categories of patentable inventions have a standardized and comprehensive scheme for classifying patentable subject matter. That means that there may not be a faster way to find all patents relevant to a partic-

21. U.S. DEPT. OF HEALTH & HUMAN SERVS. ET AL., APPROVED DRUG PRODUCTS WITH THERAPEUTIC EQUIVALENCE EVALUATIONS (32d ed. 2012) [hereinafter “Orange Book”], available at <http://www.fda.gov/downloads/Drugs/DevelopmentApprovalProcess/UCM071436.pdf>.

22. *Id.* at iv (describing the origins of the nickname “Orange Book”).

23. See STN INTERNATIONAL, <http://www.stn-international.de/index.php?id=123> (last visited Dec. 10, 2011).

24. See, e.g., DELPHION, <http://www.delphion.com> (last visited Jan. 18, 2012) (cited in Jeanne C. Fromer, *Patent Disclosure*, 94 IOWA L. REV. 539, 585 n.208 (2009)); PATBASE, <http://www.patbase.com> (last visited Jan. 18, 2012).

ular product than to simply examine all of the patents in a particular technology class, one at a time.

We will focus on software patents, which provide a particularly good illustration of the problem.

As we have already noted, it would be completely impracticable for a firm to read every patent. So patent lawyers use a variety of methods to guess at what patents they should look at. These methods include searching by keyword, patent classification,²⁵ inventor or patent assignee, and searching for patents that cite to and are cited in similar patent applications.²⁶

The keyword search is crucial, but searching by keyword hardly approaches the speed and certainty of searching an indexable system. A firm producing a new word processor might search for patents containing phrases like “word processor,” “page layout,” “printing,” and so forth. But in the absence of a precise, standardized scheme for classifying software inventions, patent applicants are free to use any terms they like—or even make up new ones—to describe their software inventions. The scope of a patent’s claims will not always be obvious from a patent’s title or abstract.²⁷ And a single software patent can claim multiple applications that are only loosely connected to each other.²⁸

One particularly illustrative example of the limits of keyword searching is U.S. Patent No. 4,528,643—a “System for Reproducing Information in Material Objects at a Point of Sale Location.”²⁹ The

25. The patent classification system tries to categorize patented inventions as particular types. The Patent and Trademark Office assigns classification numbers based only on the patent’s claims, rather than the entire application. *See* U.S. PATENT & TRADEMARK OFFICE, U.S. DEPARTMENT OF COMMERCE HANDBOOK OF CLASSIFICATION 9 (2005), available at <http://www.uspto.gov/web/offices/opc/documents/handbook.pdf> (cited in Fromer, *supra* note 24, at 586 n.212). John R. Allison and Mark A. Lemley have discovered “numerous instances of what seem to be wrong or arbitrary classification decisions.” John R. Allison & Mark A. Lemley, *The Growing Complexity of the United States Patent System*, 82 B.U. L. REV. 77, 92 (2002) (cited in Fromer, *supra*, at 586). They note, “This problem is compounded by the fact that some patents involve more than one type of technology, so that classifying them into only a single category will necessarily mischaracterize them to some extent.” *Id.* at 92 n.49.

26. In addition to the USPTO website, numerous private companies provide searchable databases of patents and related information. *See, e.g.*, DELPHION, *supra* note 24; PATBASE, *supra* note 24.

27. The purpose of a patent abstract is to merely “enable the United States Patent and Trademark Office and the public generally to determine quickly from a cursory inspection the nature and gist of the technical disclosure.” 37 C.F.R. § 1.72(b) (2012).

28. *See* 37 C.F.R. § 1.75(b) (2012).

29. U.S. Patent No. 4,528,643, at [54] (filed Jan. 10, 1983).

invention that gave rise to the '643 patent was "a kiosk or vending machine to be used in retail locations for producing digital music tapes or other digital reproductions."³⁰ The patent owner attempted to enforce the patent against several software and publishing companies, claiming that the sale and transfer of software and documents over the Internet infringed the patent.³¹ The '643 patent was filed in 1983, long before "e-commerce" existed; indeed, the web page and browser were not created until 1990.³² Nonetheless, litigation over the meaning of the '643 patent took over seven years to complete.³³

The '643 patent illustrates the difficulty with relying on keywords to search for patents. The "Background of the Invention" section of the patent discusses "retail outlets (point of sale locations)" and their difficulty deciding which recordings to stock.³⁴ The patent makes no mention of the Internet or personal computers. An attorney trying to determine *ex ante* whether the process of selling and transferring software over the Internet had been patented would be unlikely to discover the '643 patent by conducting a keyword search.³⁵

30. See JAMES BESSEN & MICHAEL J. MEURER, PATENT FAILURE 194 (2008). Readers who find themselves confused as to what invention this patent describes are in good company. The exact scope of the '643 patent is not clear, even after years of litigation. The original application likely referred to a machine that could burn custom music recordings so the store would not have to keep them in stock.

31. See *Interactive Gift Exp., Inc., v. Compuserve Inc.*, 256 F.3d 1323, 1329 (Fed. Cir. 2001). *Catalina Marketing International, Inc. v. Coolsavings.com, Inc.*, 289 F.3d 801, 806–07 (Fed. Cir. 2002), presented a similar fact pattern.

32. See TIM BERNERS-LEE, WEAVING THE WEB 28–30 (1999).

33. See Order of Dismissal, *Interactive Gift Exp., Inc. v. Compuserve Inc.*, No. 95-6871 (S.D.N.Y. Jan. 21, 2003), ECF No. 224; Complaint, *Interactive Gift Exp. Inc., v. Compuserve Inc.*, No. 95-6871 (S.D.N.Y. July 25, 1995), ECF No. 1. The Federal Circuit had held the term "point of sale location" could include one's home. See *Interactive Gift Exp.*, 256 F.3d at 1335. However, it also held that the "material objects" referenced in the patent did not include a buyer's personal computer because the objects had to be offered for sale, removed from the device that wrote information to them after purchase, and intended for use on a device other than that which wrote information to them. *Id.* at 1338. Nonetheless, litigation following the Federal Circuit decision dragged on for one and a half years, concluding with voluntary dismissals and numerous settlement agreements. See Docket, *Interactive Gift Exp. Inc., v. Compuserve Inc.*, No. 95-6871 (S.D.N.Y. Nov. 7, 2001), ECF Nos. 184–86, 189, 201, 220–24 (dismissing parties, often after having reached a settlement).

34. U.S. Patent No. 4,528,643 col. 1 ll 6–8; col. 2 ll 13–15 (filed Jan. 10, 1983).

35. At first blush, one might suppose this problem could sometimes be averted by the reverse doctrine of equivalents. "The reverse doctrine of equivalents is an equitable doctrine designed to prevent unwarranted extension of the claims beyond a fair scope of the patentee's invention." *Roche Palo Alto LLC v. Apotex*,

This problem is exacerbated by the complexity of software products. Real software products often contain thousands—sometimes even millions—of lines of code.³⁶ Given that a handful of lines of code can constitute a patent-eligible invention, the number of potentially patentable inventions in a 100,000-line computer program can be very large. For example, the patent on raising a pop-up browser window when one attempts to leave a webpage³⁷ can be infringed by writing a mere three lines.³⁸

Hence, it is extremely difficult to anticipate all of the different aspects of a particular computer program that might be regarded as patent-eligible subject matter. It is even more difficult to anticipate all of the terms patent applicants could use to describe those various patentable concepts. The effectiveness of keyword searching is further undermined by the doctrine of equivalents, which holds that a patent can cover subject matter “equivalent” to its claims even if it does not fall within their literal scope.³⁹ This means that in

Inc., 531 F.3d 1372, 1377 (Fed. Cir. 2008) (internal quotation marks omitted). “[W]here a device is so far changed in principle from a patented article that it performs the same or similar function in a substantially different way, but nevertheless falls within the literal words of the claim, the [reverse] doctrine of equivalents may be used to restrict the claim and defeat the patentee’s action for infringement.” *Graver Tank & Mfg. Co. v. Linde Air Prods. Co.*, 339 U.S. 605, 608–09 (1950) (citation omitted). However, the reverse doctrine of equivalents is “all but defunct.” *See Long, supra* note 7, at 519 (describing how the United States Court of Appeals for the Federal Circuit “proudly declare[d] that it has struck down every successful assertion of the reverse doctrine of equivalents” (citing *Tate Access Floors, Inc. v. Interface Architectural Res.*, 279 F.3d 1357, 1368 (Fed. Cir. 2002))).

36. *See, e.g.*, MICROSOFT, *A History of Windows*, <http://windows.microsoft.com/en-US/windows/history> (last visited July 1, 2012) (noting that Windows XP was compiled from 45 million lines of code); Stewart Brand, *The Physicist*, WIRE MAG. (Sept. 1995), available at <http://www.wired.com/wired/archive/3.09/myhrvold.html> (interview with Nathan Myhrvold, estimating Microsoft Word consisted of two million lines of code in 1995).

37. U.S. Patent No. 6,389,458 (filed Oct. 30, 1998).

38. *See* BEN KLEMENS, MATH YOU CAN’T USE 1–2 (2006). The three lines of code necessary to create a pop up window when one attempts to leave a webpage in JavaScript are:

```
function onExit() {  
    popup = window.open (“pop.html”, “Don’t go!”) ;  
    popup.focus();}
```

Id. at 2.

39. A product or process may infringe a patent under the doctrine of equivalents if it performs “substantially the same function in substantially the same way to obtain the same result” as the patented invention. *Graver Tank Mfg. Co.*, 339 U.S. at 608 (quoting *Sanitary Refrigerator Co. v. Winters*, 280 U.S. 30, 42 (1929)); *see also* *Warner Jenkinson Co. v. Hilton Davis Chem. Co.*, 520 U.S. 17, 40

practice, a keyword-based patent search will either only find a fraction of the relevant patents, or produce so many results that it would be of little help to the searcher.⁴⁰

Other search strategies, such as searching by inventor, assignee, or citations in related patents, are no more promising. If a firm knows of an existing patent similar to its product, these methods may be useful for finding closely-related patents. But as we have seen, many different aspects of a software product may be patent-eligible, and there is no reason to think that all the patents relevant to a particular product will be linked together by citations, common inventors, or other similarities.

It is theoretically possible that future improvements in artificial intelligence will allow the creation of a search engine for software patents as powerful as conventional chemical patent databases. This search engine would have to be sophisticated enough to analyze a real-world machine or process, make a comprehensive list of characteristics that could constitute patent-eligible subject matter, produce a list of all possible terms that could be used to describe this subject matter, and find all patents that use these terms in a way that indicates possible infringement. But that technology doesn't

(1997) (explaining that the essential inquiry under the doctrine of equivalents is whether “the accused product or process contain[s] elements identical or equivalent to each claimed element of the patented invention”). John R. Allison and Mark A. Lemley argue, “[T]he doctrine of equivalents was . . . near death by the late 1990s . . . [and] district courts are more likely to reject doctrine of equivalents claims today than ever before.” John R. Allison & Mark A. Lemley, *The (Unnoticed) Demise of the Doctrine of Equivalents*, 59 STAN. L. REV. 955, 958 (2007); see also Lee Petherbridge, *On the Decline of the Doctrine of Equivalents*, 31 CARDOZO L. REV. 1371 (2010); David L. Schwartz, *Explaining the Demise of the Doctrine of Equivalents*, 26 BERKELEY TECH. L.J. 1157 (2011). Nonetheless, Samson Vermont points out that “one of every four or five cases in which a patentee wins a judgment of infringement is . . . a judgment of infringement under the doctrine of equivalents [Doctrine of equivalents] scope is litigated frequently. One of every two decisions on infringement is a decision on [doctrine of equivalents] infringement.” Samson Vermont, *Taming the Doctrine of Equivalents in Light of Patent Failure*, 16 J. INTELL. PROP. L. 83, 85 (2008).

40. Although keyword searches will find many relevant patents, finding only some relevant patents will not insulate an inventor from lawsuits or create incentives to license or design around found patents. If you can only find 50% of the patents on which your invention might infringe, there is little value in licensing or designing around those patents because you can still be sued by the owners of the other 50% of patents you did not find. Patent searching is not necessarily like searching for legal cases where the cases are similar and related to each other, and where, after a point, finding each new case produces diminishing returns. The first and last patent you find are equally likely to bring an accidental infringer economic ruin.

exist right now, and we are skeptical it will exist any time soon. That means the only reliable way to find all—not merely some—patents infringed by a particular software product is to have a human being look at all patents in software-related technology classes. So the discovery costs of software patents for a single firm is roughly $O(n)$ in the number of software patents, not $O(1)$ as with chemical patents.

The fact that the average firm has patent discovery costs that are $O(n)$ in the number of patents is not a problem if the number of firms and patents are both small. But the more firms and patents there are, the larger the discovery costs will be.

Once again, the software industry is a good example of an industry where the patent system works poorly. The number of firms producing patentable software is massive—much larger than the number of firms in the software industry as it is conventionally defined.⁴¹ Almost every medium and large American firm has an information technology (“IT”) department that performs backups, runs file and mail servers, runs the firm’s website, and so forth.⁴² IT professionals routinely create software to automate such tasks, and this software is potentially patent-eligible. Many firms also develop custom software to automate common business processes, and some of it is quite complex.⁴³ Hence, most medium and large American firms (as well as many non-profits, universities, and other organizations) are in the software industry as far as patent law is concerned.⁴⁴

And as a consequence, many kinds of firms are the targets of software patent lawsuits.⁴⁵ One complaint charged the Green Bay Packers, Caterpillar, Peapod, OfficeMax, and Kraft Foods with in-

41. See KLEMENS, *supra* note 38, at 92.

42. *Id.* at 93.

43. *Id.*

44. *Id.* at 4–5 (“[A] patent on a piece of code is a restriction . . . on the information technology department of every company in America, not to mention every person who writes macros to facilitate his or her work . . .”).

45. Even frivolous allegations of patent infringement can be very expensive to dispel. In 2005, the average cost of an opinion letter assessing the validity of a patent and whether an accused party infringed was about \$24,000. AM. INTELL. PROP. LAW ASS’N, REPORT OF THE ECONOMIC SURVEY 102 (2005) (*cited in* Matthew Sag & Kurt Rohde, *Patent Reform & Differential Impact*, 8 MINN. J.L. SCI. & TECH. 1, 10 n.41 (2007)). Getting a patent invalidated in court costs on average \$650,000. *Id.* at 108 (*cited in* Sag & Rohde, *supra*, at 10 n.44). It is thus often economically rational to pay high licensing fees for invalid patents that one did not even infringe rather than have to participate in a lawsuit. “This is the real perversity of the current patent system: rational actors will pay licensing fees for patents they strongly suspect are either invalid, or simply do not apply to them, because each of the alternatives is worse.” Sag & Rohde, *supra*, at 11.

fringing a patent for a “Remote Query Communications System”⁴⁶—specifically for having JPEG images on their websites.⁴⁷ Another plaintiff filed lawsuits against firms such as J. Crew and Linens ‘N Things for infringing its “Information Processing Methodology” patents⁴⁸ by transmitting data that customers entered on the defendants’ websites.⁴⁹ Other firms facing allegations of software patent infringement include McDonalds, Barnes & Noble, Jamba Juice, Aeropostale, 7-Eleven, and Oprah Winfrey’s Harpo Productions.⁵⁰

Not only do firms outside of the conventional software industry frequently produce potentially infringing software, they are also granted the lion’s share of software patents. James Bessen has

46. U.S. Patent No. 5,253,341 (filed Apr. 11, 1991).

47. Amended Complaint at 3, *Global Patent Holdings, LLC v. Green Bay Packers, Inc.*, No. 00-4623 (N.D. Ill. Oct. 23, 2007), ECF No. 50. An attempt to enforce the ‘341 patent in 2000 led to the invalidation of all sixteen claims, as well as to the addition of a seventeenth claim. *See* Motion to Reinstate at 1, *Techsearch LLC v. Internet Entm’t Grp.*, No. 00-4623 (N.D. Ill. Sept. 10, 2007), ECF No. 41; *see also id.* at Ex. A (*Ex parte* Global Patent Holdings, LLC, Appeal No. 2006-0698, Reexamination Control No. 90/005,742 (Dec. 26, 2006)). The amended complaint, *supra*, charged the Green Bay Packers and other parties with violating the new seventeenth claim, however an initial re-examination of the seventeenth claim found it invalid on nineteen grounds. *See* Initial Office Action in *Ex Parte* Reexamination of U.S. Patent No. 5,253, 341, at 3-5 (July 22, 2008), *available at* <http://www.scribd.com/doc/4328073/jpg-patent-reexam>. The case was dismissed without prejudice pending the reexamination on March 4, 2009. *See* Minute Entry, *Global Patent Holdings LLC v. Green Bay Packers, Inc.*, No. 00-4623 (N.D. 2009), ECF No. 154; *Agreed Motion to Dismiss, Global Patent Holdings, LLC v. Green Bay Packers, Inc.*, No. 00-4623, 2009 WL 3059752 (N.D. Ill. 2009), E.C.F. No. 153.

48. U.S. Patent No. 7,184,162 (filed Apr. 15, 2005); U.S. Patent No. 7,075,673 (filed Nov. 6, 2003); U.S. Patent No. 6,683,697 (filed Dec. 9, 1999).

49. *See* Complaint at ¶¶ 14, 17, 20, *Eon-Net L.P. v. J. Crew Inc.*, No. 07-10488 (S.D.N.Y. Nov. 20, 2007), ECF No. 1; Amended Complaint at ¶¶ 12, 15, *Eon-Net, L.P. v. Linens ‘N Things, Inc.*, No. 06-315 (D.N.J. Aug. 2, 2006), ECF No. 15. The cases were dismissed or settled before the court rendered a final judgment. *See* Stipulation and Order of Dismissal with Prejudice, *Eon-Net, L.P. v. J. Crew Inc.*, No. 07-10488 (S.D.N.Y. Aug. 4, 2008), ECF No. 11; Stipulation of Dismissal, *Eon-Net, L.P. v. Linens ‘N Things, Inc.*, No. 06-315 (D.N.J. Jan. 5, 2007), ECF No. 23; Stipulation of Dismissal, *Eon-Net, L.P. v. Linens ‘N Things, Inc.*, No. 06-315 (D.N.J. Dec. 22, 2006), ECF No. 22.

50. *See* Complaint, *Card Activation Techs., Inc. v. 7-Eleven Inc.*, No. 10-4984 (N.D. Ill. Aug. 9, 2010), ECF No. 1; Complaint, *Illinois Computer Research, LLC v. Harpo Productions, Inc.*, No. 08-7322 (N.D. Ill. Dec. 23, 2008), ECF No. 1; Complaint, *Card Activation Techs., Inc. v. Barnes & Noble, Inc.*, No. 07-1230 (N.D. Ill. Mar. 2, 2007), ECF No. 1 (also naming Jamba Juice Co. and Aeropostale Inc. as defendants); Complaint, *Card Activation Techs., Inc. v. McDonald’s Corp.*, No. 06-5578 (N.D. Ill. Oct. 13, 2006), ECF No. 1.

found that 83% of software patents granted to public firms in 2006 went to firms outside of the conventional software industry.⁵¹

In our discussion of the widget industry, we stipulated that each firm would hold exactly one patent and concluded that the industry-wide discovery costs of widget patents are $O(n^2)$. Obviously, real industries aren't like that. Some firms have many patents and others have none at all. So it's more precise to say that the patent system's discovery costs in non-indexable industries are $O(n_f n_p)$, where n_f is the number of firms and n_p is the number of patents. The widget industry in our example is a special case where $n = n_f = n_p$, so that total discovery costs are $O(n^2)$.

In the software industry, $n_f \neq n_p$, but n_f and n_p are both large. As we have seen, most medium and large firms produce patent-eligible software. There are roughly 635,000 firms in the United States with twenty or more employees.⁵² While not all of these firms produce software, many of the 1.7 million firms with five to nineteen employees do, so we'll estimate the number of firms that create software, n_f , to be 600,000 firms. And n_p , the number of software patents issued, is around 40,000 in a typical year (and growing).⁵³ That means that there are around twenty-four billion new patent-firm pairs each year that could produce accidental infringement. Even if a patent lawyer only needed to look at a patent for ten minutes, on average, to determine whether *any* part of a particular firm's software infringed it,⁵⁴ it would require roughly two million patent attorneys, working full-time, to compare every firm's prod-

51. James Bessen, *A Generation of Software Patents*, 18 B.U. J. SCI. TECH. L. 241, 256 (2012) (showing only 17.2% of software patents granted to public firms were granted to firms in the computer services and software industries).

52. According to the U.S. Census Bureau, there were 635,000 businesses with twenty or more employees in 2008, and 1.7 million firms with five to nineteen employees. See *Statistics About Business Size (including Small Business) from the Census Bureau*, CENSUS BUREAU HOMEPAGE, <http://www.census.gov/econ/smallbus.html> (last visited Dec. 9, 2011).

53. See Bessen, *supra* note 51, at 253.

54. Ten minutes is an unrealistically low amount of time. Patentable software can be written in only a few lines, see *supra* note 27, and many software programs consist of millions of lines of code. See, e.g., MICROSOFT *supra* note 36 (noting that Windows XP was compiled from 45 million lines of code); Brand, *supra* note 36 (interview with Nathan Myhrvold, estimating Microsoft Word consisted of two million lines of code in 1995). It is plainly beyond human capacity for an attorney to be able to hold in his or her mind everything that a large software program does, let alone to compare it to the content of a patent in a matter of minutes.

ucts with every patent issued in a given year.⁵⁵ At a rate of \$100 per hour, that would cost \$400 billion. For comparison, the software industry was valued at \$225.5 billion in 2010.⁵⁶

Obviously, \$400 billion is a highly speculative figure. But the exact number doesn't matter because there are only around 40,000 registered patent attorneys and agents in the United States.⁵⁷ Even if the entire patent bar worked full-time on patent clearance for software firms, there wouldn't be nearly enough lawyers to go around.

E. Example: Discovery Costs for Corkscrews Are Low Because There Are Few Corkscrew Patents

We have argued that discovery costs are low for the pharmaceutical industry because chemical patents are indexable by chemical formula. But even non-indexable patent classes can have modest discovery costs if n_p , the number of potentially-relevant patents, is small enough.

For example, consider corkscrews. A search of corkscrew-related technology classes⁵⁸ reveals that 301 utility patents were awarded between 1992 and 2011. Just five of these were issued in 2011.

We have argued that it would be impossible for anyone to read and understand the hundreds of thousands of existing software patents, or even to keep up with the hundreds of software patents the patent office issues each week. But it would only take a few weeks to read and understand the 301 utility patents related to corkscrews.⁵⁹ Given that only about fifteen patents are issued in corkscrew-related technology classes in a typical year, it would be fairly easy for an

55. The math behind this is straightforward: 40,000 patents*600,000 firms*(10 minutes per patent-firm pair)/(2,000 hours of work per attorney *60 minutes per hour)=2 million attorneys.

56. *Software: Global Industry Guide 2010*, DATA MONITOR RESEARCH STORE (2012), available at http://www.datamonitor.com/store/Product/software_global_industry_guide_2010?productid=4F026C5C-EBCC-4193-AD27-77260196E7F5.

57. See *Patent Attorney/Agent Search*, UNITED STATES PATENT AND TRADEMARK OFFICE, <https://oedci.uspto.gov/OEDCI/query.jsp> (search for "US" in the country field) (last visited Dec. 11, 2011).

58. We used patent classes 81/3.2, 81/3.4, 81/3.7, 81/3.9, 81/3.29, 81/3.36, 81/3.37, 81/3.45, 81/3.48. See e.g., Full-Text and Image Database, UNITED STATES PATENT & TRADEMARK OFFICE (current through Oct. 16, 2012), <http://patft1.uspto.gov/netahtml/PTO/search-bool.html>.

59. For example, if we assume that it takes an hour, on average, to understand a corkscrew-related patent, then it would take approximately eight forty-hour work-weeks to familiarize oneself with all 301 corkscrew-related utility patents.

attorney who specialized in corkscrew patents to keep abreast of new patents.

In March 2012, Amazon.com listed 4,551 corkscrews for sale in its “home and kitchen” section, made by 737 manufacturers.⁶⁰ Not every firm manufacturing corkscrews is listed on Amazon.com, but it’s reasonable to assume a majority of the commercially significant ones are. So we’ll estimate that n_f , the number of firms in the corkscrew industry, is no more than 1,500.

We have already estimated that $n_p=301$ for corkscrews. If we again assume that each firm-patent comparison takes ten minutes, then it would take approximately 75,250 hours to conduct patent clearance for all corkscrews currently on the market.⁶¹ That would require the services of approximately forty patent attorneys working full-time for a year.

And of course, this process wouldn’t need to be repeated every year. In a typical year, firms would only need to clear their new products and compare their existing products with newly-issued patents. So in a typical year, the corkscrew industry would require the services of significantly fewer than forty full-time patent attorneys.

Moreover, the fact that there are few enough patents that a single person could read and understand all of them means that the vetting process is likely to be considerably more efficient for corkscrews than for software. A patent is much harder to understand the first time it’s read than on the second, fifth, or twentieth encounter. There are so many software patents that no one could possibly read more than a small fraction of them, so attorneys doing software patent clearance spend most of their time reading patents for the first time. In contrast, attorneys that specialize in corkscrews would be looking at the same 301 patents over and over again. Their familiarity with the corkscrew patents would allow them to quickly identify which were relevant to a particular client’s products.

Hence, the fact that a patent class is non-indexable does not necessarily mean that discovery costs will be prohibitively high. If

60. AMAZON.COM, <http://www.amazon.com/> (last visited March 2012) (search for “corkscrews” in the category “Home and Kitchen”; click on “see more” hyperlink under “brands.”).

61. $301 * 1,500 = 451,500$ comparisons. If each takes ten minutes, this will take $451,500 / 6 = 75,250$ hours of work. Obviously, 10 minutes per comparison is a rough estimate, but one that suffices to illustrate the point that corkscrew patent discovery costs are several orders of magnitude smaller than software patent discovery costs.

n_p , the total number of patents in that technology class, is small, then the “brute force” approach of examining every patent may be feasible.

III. DISORGANIZED PATENTS HAVE LED TO A LITIGATION EXPLOSION

In practice, firms don’t—and can’t—spend whatever it takes to avoid infringement. Rather, they spend only as much money on patent searches as they believe will “pay off” in lower future litigation and licensing costs.⁶² In industries with low discovery costs, a rational firm is likely to spend enough money to find all patents relevant to its products. Inadvertent infringement in these industries is rare. On the other hand, in industries where discovery costs are high, the rational firm might not even try to avoid infringement, because a dollar spent on patent searches will produce much less than a dollar in savings due to reduced litigation.

Unsurprisingly, the software industry—and, indeed, the larger IT industry of which it is a part—is in the latter category. In a widely-cited paper, Mark Lemley documents the widespread IT industry practice of ignoring patents and tries to explain why IT firms behave as they do.⁶³ He suggests several explanations: patent negotiations take a long time, patent holders may not be willing to offer reasonable terms, many patents turn out to be invalid, and the number of patents a given firm must license would be large.⁶⁴ These are all plausible explanations, but there is a more important and fundamental one: firms have no cost-effective way of obtaining a complete list of relevant patents in the first place. Licensing the few they know about provides no protection against the many others they have not yet discovered.

Lemley notes that the pharmaceutical industry is one of the few industries that does not ignore patents, and he attributes the difference to the fact that the FDA forces patent holders to disclose relevant patents.⁶⁵ But such disclosures can only be compiled into a useful form because chemical patents are indexable. Without chemical formula as an organizational scheme, it would not be possible

62. See Rebecca S. Eisenberg, *Patent Costs and Unlicensed Use of Patented Inventions*, 78 U. CHI. L. REV. 53, 55 (2011) (“Information costs and transaction costs may dwarf potential gains to users from identifying and clearing rights . . .”).

63. Mark A. Lemley, *Ignoring Patents*, 2008 MICH. ST. L. REV. 19, 21 (2008).

64. *Id.* at 25–29.

65. *Id.* at 29–30.

to create a publication like the Orange Book that allows rapid retrieval of patents relevant to a particular real-world product.

The empirical evidence on litigation costs is consistent with our hypothesis about the importance of discovery costs. We should expect industries with high discovery costs to have high rates of inadvertent infringement and, as a consequence, high rates of litigation. In their 2008 book *Patent Failure*, James Bessen and Mike Meurer used stock market event studies to estimate the total costs of patent litigation for various industries during the 1980s and 1990s. They found that litigation costs for chemical patents are much lower than the profits from these patents.⁶⁶ This happy state of affairs can be explained by the low discovery costs of chemical patents—litigation is rare because infringement is rare.

Bessen and Meurer found litigation costs were much higher for non-chemical patents. From 1984 until 1994, the costs of litigation over non-chemical patents were roughly equal to the profits from those patents.⁶⁷ And from 1994 to until the end of their study period in 1999, the costs of litigation over non-chemical patents increased dramatically.⁶⁸

Why was there a spike starting in the mid-1990s? Between 1989 and 1998, courts made it dramatically easier to obtain patents on software and business methods.⁶⁹ Bessen and Meurer found that these patents contributed a disproportionate share of patent litigation. Software patents were more than twice as likely to be involved in patent litigation as other kinds of patents.⁷⁰ The closely-related category of business method patents was nearly seven times as likely to be involved in litigation.⁷¹ This is not surprising. We have already seen that software patents have particularly high discovery costs, and business method patents have high discovery costs for similar reasons.

66. See BESSEN & MEURER, *supra* note 30, at 139 fig.6.5.

67. *Id.* at 138–39.

68. *Id.*

69. The Federal Circuit began upholding patents on software in *In re Iwahashi*, 888 F.2d 1370, 1374 (Fed. Cir. 1989), and *In re Alappat*, 33 F.3d 1526, 1545 (Fed. Cir. 1994) (en banc). In 1998, the Federal Circuit eliminated the common-law ban on business method patents in *State Street Bank & Trust Co. v. Signature Fin. Grp., Inc.*, 149 F.3d 1368, 1375 (Fed. Cir. 1998). For a more general discussion of the expansion of patentable subject matter to cover software, see ROBERT PATRICK MERGES & JOHN FITZGERALD DUFFY, *PATENT LAW & POLICY* 151–55 (4th ed. 2007).

70. BESSEN & MEURER, *supra* note 30, at 22, 153 fig.7.2 .

71. *Id.*

IV. POLICY SUGGESTIONS: BEYOND PATENT QUALITY

We have shown that for technology classes with high discovery costs, the patent system is unlikely to ever work like a traditional property system. Rampant infringement is inevitable because firms have no way to discover which patents they are infringing. Because of this, firms have little incentive to even attempt to clear patent rights before introducing new products into the market.

No one would claim this state of affairs is ideal, but is this system the best we can do?

We don't think so. The patent system is supposed to be a mechanism for promoting the progress of the useful arts by transferring resources from the users of technologies to their inventors. But in industries where discovery costs are high, it does so in an erratic, wasteful, and unjust fashion. The system resembles a lottery more than a system of property rights—an unlucky minority of infringers is the target of ruinous lawsuits, and only a minority of patent owners “win” by catching infringers who pay up.

Not only is this unfair to the targets of these lawsuits, but it creates a generalized disincentive to innovate. Developing new products comes with the risk of incurring crippling liability and having one's business enjoined—precisely the opposite of the effect patents are supposed to have. And, of course, many of the resources consumed by the patent system flow not to inventors, but to pay patent attorneys and cover the patent system's other deadweight costs.

Many observers have argued that patent law reforms should focus on increasing patent “quality.”⁷² They usually mean that more patents should be invalidated for obviousness or non-novelty and that patents should have narrower and clearer claims. These are worthwhile goals to be sure. But they do not directly address the discovery cost problem.

There is no reason to think obvious or non-novel patents cost more to discover than “high quality” patents. So reforms that invalidate “low quality” patents reduce discovery costs only because they

72. See, e.g., U.S. PATENT & TRADEMARK OFFICE, 2010–2015 STRATEGIC PLAN 9–25 (2010), available at http://www.uspto.gov/about/stratplan/USPTO_2010-2015_Strategic_Plan.pdf (describing initiatives to improve patent examination timelines and patent quality); ADAM B. JAFFE & JOSH LERNER, INNOVATION AND ITS DISCONTENTS 171, 178–81 (2004) (discussing ways to improve patent quality); John R. Allison & Starling D. Hunter, *On the Feasibility of Improving Patent Quality One Technology at a Time: The Case of Business Methods*, 21 BERKELEY TECH. L.J. 729 (2006).

reduce the total number of patents being issued. That is valuable, of course, but really dramatic changes would be needed to invalidate enough patents to bring discovery costs under control in large, non-indexable industries like software.

Nor are high discovery costs primarily due to unclear patent boundaries. Unclear boundaries do raise discovery costs, since they require patent attorneys to spend more time examining each patent. But recall that we began our analysis with the counterfactual assumption that widget patents have perfectly clear boundaries. So discovery cost problems will crop up even where patent boundaries are crystal clear. Problems related to unclear patent boundaries exist in addition to the basic discovery cost problems that are the focus of this article.

There are several strategies the government could take to reduce discovery costs, ranging from making small changes to radically restructuring the patent system. Here we look at a few categories of policy changes that would lessen discovery costs to various degrees.

A. *Subject Matter Restriction*

The preceding analysis suggests that the patent system will work well when there is a clear and comprehensive way to index patents—reducing the entire industry's patent discovery time to $O(n_f)$ rather than $O(n_f n_p)$. And the patent system will also work better when n_f or n_p is very small—i.e., in industries with a small number of firms or inventions.

Together, these criteria arrange technologies on a spectrum. At one end of the spectrum are pharmaceuticals, an industry that is highly concentrated, has relatively few inventions, and can use chemical formulas to organize its patents. At the other end of the spectrum is software, an industry that is highly decentralized, produces many patents, and has no standardized classification system.

Notice that the factors that make software patents work poorly are characteristics of software itself and the software *industry*, not the patent system. This suggests that it is probably not possible to make the patent system work well for software. The basic problem is that there is a massive number of firms producing potentially-infringing software, a massive number of software patents, and no systematic way to organize them all. That is, discovery costs grow as $O(n_f n_p)$, and n_f and n_p are both large numbers. Changes to patent law probably can not make software inventions indexable, and policymakers certainly should not try to reduce the number of firms.

The fact that some categories of patentable subject matter have much higher discovery costs than others suggests an obvious policy response: exclude subject matter with high discovery costs from patent eligibility.⁷³ The strongest version of this reform would be to exclude all non-indexable technology classes from patent eligibility.⁷⁴ Or policymakers might exclude only those non-indexable categories for which litigation costs are highest.

The problem of high discovery costs provides a strong rationale for courts' traditional prohibition on patents on abstract ideas.⁷⁵ The more abstract an invention is, the more different par-

73. Excluding inventions with high discovery costs from patentability will not necessarily have deleterious effects. There are many arguments that software patents are not necessary to incentivize software development and that copyright, a sui generis regime, or no protection at all would be sufficient to encourage new innovation in software. These arguments are beyond the scope of this paper, but for more discussion, see generally KLEMENS, *supra* note 38, at 17–23; Pamela Samuelson, et al., *A Manifesto Concerning the Legal Protection of Computer Programs*, 94 COLUM. L. REV. 2308, 2404 (arguing for a sui generis intellectual property regime to protect software); Pamela Samuelson, *Benson Revisited: The Case Against Patent Protection for Algorithms and Other Computer Program-Related Inventions*, 39 EMORY L.J. 1025, 1148–53 (1990); Timothy B. Lee, *Patently Absurd - Copyright Law Can Meet the Needs of Software Developers*, NAT'L REV. (Oct. 3, 2011), available at <http://www.cato.org/publications/commentary/patently-absurd-copyright-law-can-meet-needs-software-developers>; Timothy B. Lee, *The Case against Literary (and Software) Patents*, TECHKNOWLEDGE (August 28, 2009), available at <http://www.cato.org/publications/techknowledge/case-against-literary-software-patents>; Wendy Seltzer, *Software Patents and/or Software Development* (Har. Univ. – Berman Ctr. for Internet & Soc. TPRC 2011, Sept. 24, 2011), available at <http://ssrn.com/abstract=1985780> (arguing software patents retard software development).

74. The patent system is often described as a “bargain with the public in which the inventor gives information about the invention in exchange for an exclusive right.” Mark A. Lemley, *The Myth of the Sole Inventor*, 110 MICH. L. REV. 709 745 (2012) (citing *Eldred v. Ashcroft*, 537 U.S. 186, 225 (2003) (referring to a patent as a “quid pro quo” for disclosure)); see also *Aronson v. Quick Point Pencil Co.*, 440 U.S. 257, 262 (1979) (identifying the fact that the patent system “promotes disclosure of inventions” as one of its key functions); Fromer, *supra* note 24, at 542; Lisa Larrimore Ouellette, *Do Patents Disclose Useful Information?*, 25 HARV. J.L. & TECH. 531, 532 (2012). If non-indexable patents are, in practice, undiscoverable, then the justification behind this “bargain” is significantly diminished because even fully disclosed patents would be relatively obscured among the many other nonindexable patents.

75. See *Diamond v. Diehr*, 450 U.S. 175, 185 (1981) (“Excluded from such patent protection are laws of nature, natural phenomena, and abstract ideas.”) (citations omitted); *Diamond v. Chakrabarty*, 447 U.S. 303, 309 (1980) (“The laws of nature, physical phenomena, and abstract ideas have been held not patentable.”) (citations omitted); *Parker v. Flook*, 437 U.S. 584, 589 (1978) (“Phenomena of nature, though just discovered, mental processes, and abstract intellectual concepts are not patentable”) (quoting *Gottschalk v. Benson*, 409 U.S. 63, 67 (1972)).

ties are likely to use it for different purposes, and the more flexibility parties will have to describe it.⁷⁶ All of these factors mean that more abstract patents will produce particularly high discovery costs and, as a consequence, particularly high rates of inadvertent infringement and litigation.

At one time, software and business method patents were considered too abstract to constitute patentable subject matter, but the United States Court of Appeals for the Federal Circuit took a more permissive stance towards software and business method patents during the 1990s.⁷⁷ The result has been an unprecedented explosion of litigation related to software and business method patents.⁷⁸ The Federal Circuit's de facto legalization of software and business method patents was a mistake and should be reversed.

76. For example, the relationship between energy and mass ($e=mc^2$) has applications for the study of radioactivity, space travel, nuclear energy, and the composition of the universe. See Peter Tyson, *The Legacy of E=MC²*, NOVA (Oct. 11, 2005), <http://www.pbs.org/wgbh/nova/physics/legacy-of-e-equals-mc2.html>. A patent application purporting to cover $e=mc^2$ could be described in terms of any of these applications.

77. Before the 1990s, courts generally held that software was not patentable. In *Gottschalk v. Benson*, 409 U.S. 63 (1972), the Supreme Court invalidated a process patent on software for converting signals from binary-coded decimal into binary, emphasizing that “[p]henomena of nature, . . . mental processes, and abstract intellectual concepts [we]re not patentable.” *Id.* at 67. The Court found that granting the patent “in practical effect would [grant] a patent on the [conversion] algorithm itself,” *id.* at 72, and indicated that patents on software programs were beyond the scope of the patent statute, absent legislative change. See *id.* at 72–73.

Following *Benson*, patent drafters attempted to redraft abstract process claims into claims for making a new machine, in the hopes of concealing any resemblance of their claims to the process claims at issue in *Benson*. See ROBERT PATRICK MERGES & JOHN FITZGERALD DUFFY, *PATENT LAW AND POLICY* 151–53 (4th ed. 2007). This tactic succeeded when the Federal Circuit upheld machine claims for software in *In re Iwahashi*, 888 F.2d 1370, 1374 (Fed. Cir. 1989), and *In re Alappat*, 33 F.3d 1526 (Fed. Cir. 1994) (*en banc*). The Federal Circuit in *Alappat* acknowledged that “many, or arguably even all, of the means elements recited in [the claim] represent circuitry elements that perform mathematical calculations.” 33 F.3d at 1544. Nonetheless, the majority concluded, “This [claim] is not a disembodied mathematical concept . . . but rather a specific machine to produce a useful, concrete, and tangible result.” *Id.*

In 1998, the Federal Circuit decided *State Street Bank & Trust Co. v. Signature Fin. Grp., Inc.*, 149 F.3d 1368, 1375 (Fed. Cir. 1998) (eliminating the common-law ban on business method patents).

78. In 2008, software patents were more than twice as likely to be litigated as other patents, and business method patents were nearly seven times more likely to be litigated than other patents. BESSEN & MEURER, *supra* note 30, at 22, 153 fig.7.2. In the late 1990s, software patents accounted for 38% of the cost of patent litigation for public firms. *Id.* at 22.

An enumeration of other technology classes with high discovery costs is beyond the scope of this paper, but would be a worthwhile subject for future study.

One alternative to exclusions based on subject matter would be to vary the application fee for patents based on the estimated discovery costs of each technology class. The patent office would take into account the indexability of the technology classes and the number of firms practicing in it. A party applying for a chemical patent would thus have to pay a very low fee, whereas the fee for patents in software-related technology classes would be drastically increased.⁷⁹ This would decrease n_p over time by discouraging applications for patents in areas with high discovery costs.⁸⁰

B. Independent Invention

Another powerful reform would be to create an independent invention defense to patent infringement.⁸¹ Ninety to ninety-eight percent of modern patent lawsuits are filed against independent

79. In 2001, Mark A. Lemley estimated that the cost of prosecuting a patent was between \$10,000 and \$30,000. See Mark A. Lemley, *Rational Ignorance at the Patent Office*, 95 Nw. U. L. REV. 1495, 1498 (2001). A change in patent application fees would have to be significant in relation to the already high costs of hiring attorneys to draft and prosecute the patent.

80. Peter S. Menell and Michael J. Meurer similarly suggest that the cost of evaluating patent applications “should be borne by the applicants and should be tailored to the costs of examining particular applications (or at least classes of applications).” Peter S. Menell & Michael J. Meurer, *Notice Failure and Notice Externalities* 35 (Boston Univ. Sch. of Law, Law and Econ. Research Paper No. 11-58, Boston Univ. Sch. of Law, Pub. Law Research Paper No. 11-58, Stanford Law and Econ. Olin Working Paper No. 418, UC Berkeley Pub. Law Research Paper No. 1973171, Feb. 16, 2012), available at <http://ssrn.com/abstract=1973171>.

81. Other authors have written about the possibility of creating an independent invention defense to patent infringement. See generally, e.g., John S. Leibovitz, *Inventing a Nonexclusive Patent System*, 111 YALE L.J. 2251 (2002) (suggesting that it would be economically efficient for independent inventors to receive independent patents on the same invention); Mark A. Lemley, *Should Patent Infringement Require Proof of Copying?*, 105 MICH. L. REV. 1525 (2007); Oskar Liivak, *Rethinking the Concept of Exclusion in Patent Law*, 98 GEO. L.J. 1643 (2010); Stephen M. Maurer & Suzanne Scotchmer, *The Independent Invention Defence in Intellectual Property*, 69 ECONOMICA 535 (2002), available at <http://www.dklevine.com/archive/scotchmer-independent-invention.pdf> (arguing that independent creation defense would preserve incentives to invent, but also permit more efficient use of inventions); Elisabetta Ottoz & Franco Cugno, *The Independent Invention Defence in a Cournot Duopoly Model*, ECON. BULL., June 20, 2004, at 1–7, available at <http://www.economicsbulletin.com/2004/volume12/EB-04L10005A.pdf>; Samson Vermont, *Independent Invention as a Defense to Patent Infringement*, 105 MICH. L. REV. 475, 480 (2007).

inventors, not copiers.⁸² Independent inventors develop and commercialize inventions without copying from existing, patented inventions. Patent owners who have patented the same invention and who identify an independent inventor can exact licensing fees or even stop the invention's use by the independently-inventing party entirely.⁸³ This system hurts independent inventors and the public by forcing independent inventors to negotiate licensing fees to keep their existing products on the market—or by removing otherwise-successful products from the market.

There is an extensive literature on proposed independent invention defenses,⁸⁴ and a full consideration of the arguments for and against such a defense is beyond the scope of this paper. Nonetheless, we note that in principle, an independent invention defense could reduce discovery costs to zero. With an independent invention defense on the books, patent holders would still have the security of knowing that a competitor could not *copy* their work, but independent parties who happened to create something that infringed a patent would not be liable to the patent holders. Inventors would also have no obligation to search for patents they potentially were infringing because so long as they were not copying another's work, their inventions would be safe from patent lawsuits.

C. *Limiting Injunctions and Multiplied Damages for Patent Infringement*

A final reform would be to limit patent remedies for infringement of non-indexable patents to actual damages, rather than per-

82. Lemley, *supra* note 74, at 713 (citing Christopher A. Cotropia & Mark A. Lemley, *Copying in Patent Law*, 87 N.C. L. REV. 1421, 1424 (2009)). Notably, the prevalence of independent invention acts as evidence that the patent system is frequently unnecessary to spur invention. Independent inventors often develop and commercialize the very same inventions others have patented. The fact that a later party developed and commercialized a patented invention independently indicates that it was not necessary to award the original patentee a patent in order for society to benefit from the invention. See Lemley, *supra* note 81, at 1527 (citing Vermont, *supra* note 81).

83. See 35 U.S.C. § 283 (2006); see also *eBay Inc. v. MercExchange, L.L.C.*, 547 U.S. 388 (2006) (describing test for injunctive relief under the patent act).

84. See Leibovitz, *supra* note 81; Lemley, *supra* note 81; Liivak, *supra* note 81; Long, *supra* note 7, at 525–33; Maurer & Scotchmer, *supra* note 81; Roger Milgrim, *An Independent Invention Defense to Patent Infringement: The Academy Talking to Itself: Should Anyone Listen?*, 90 J. PAT. & TRADEMARK. OFF. SOC'Y 295 (2008); Ottoz & Cugno, *supra* note 81; Vermont, *supra* note 81.

mitting injunctive relief and multiplied damage awards.⁸⁵ In other words, remedies for patent infringement could be limited to those instantiating “liability rules” rather than “property rules.”

“Property rules” are designed to prevent parties from using another’s asset without permission.⁸⁶ In the case of patent law, its property rule system “include[s] injunctions and supercompensatory damages that would make [patent infringement] less attractive than bargaining to a consensual price with the [patent] owner.”⁸⁷ In contrast, “liability rules” such as lost profits or a royalty, theoretically allow parties to infringe on another’s patent “as long as . . . officially determined damages are paid. The level of the damages is set to compensate the owner,” rather than punish the infringer.⁸⁸

Generally speaking, property-rule remedies are considered beneficial when transaction costs between property holders and those who want to acquire property are low.⁸⁹ As Stewart Sterk explains:

Because property rules require all potential users of a resource to buy rights from th[e resource] owner, property rules enable the owner to accumulate information about potential bidders and the values those bidders attach to those rights. As a result, property rules enable resource owners to channel those resources to the bidders who value them most—promoting efficient use of resources.⁹⁰

85. Currently, patent infringers may be enjoined from future infringement and made to pay damages in the form of a reasonable royalty or lost profits. See 35 U.S.C. § 283 (“[C]ourts. . . may grant injunctions in accordance with the principles of equity to prevent the violation of any right secured by patent, on such terms as the court deems reasonable.”); 35 U.S.C. § 284 (2006) (Courts may award “damages adequate to compensate for the infringement, but in no event less than a reasonable royalty for the use made of the invention by the infringer, together with interest and costs as fixed by the court. . . . [T]he court may increase the damages up to three times the amount found or assessed.”).

86. Guido Calabresi & A. Douglas Melamed, *Property Rules, Liability Rules, and Inalienability: One View of the Cathedral*, 85 HARV. L. REV. 1089, 1092 (1972); Henry Smith, *Property and Property Rules*, 79 N.Y.U. L. REV. 1719, 1720 (2004).

87. Smith, *supra* note 86, at 1720.

88. *Id.*

89. Stewart E. Sterk, *Property Rules, Liability Rules, and Uncertainty About Property Rights*, 106 MICH. L. REV. 1285, 1290 (2008) (citing Guido Calabresi & A. Douglas Melamed, *Property Rules, Liability Rules, and Inalienability: One View of the Cathedral*, 85 HARV. L. REV. 1089, 1106–10, 1118 (1972)) (explaining that Guido Calabresi and A. Douglas Melamed observed that “property rules are efficient in cases of low transaction costs, while liability rules are preferable in cases of high transaction costs”).

90. *Id.* at 1295 (footnote omitted).

But property rules fail to promote the efficient use of resources when transaction costs are high. The combination of very high transaction costs (e.g., the costs of locating a patent holder to negotiate a license with) and punitive legal penalties (e.g., an injunction or multiplied damages for infringement) can prevent beneficial uses of property and waste resources by making property use very costly.⁹¹

Permitting property-rule remedies in patent law is harmful because the discovery costs of locating relevant patents render transaction costs too high for many non-indexable patents.⁹² It is not merely costly for potential infringers of non-indexable patents to locate the patents they might infringe—it is completely impracticable. As a result, firms are faced with a *disincentive* to develop new products because of the liability that could result if those products infringed others' patents. Firms cannot determine *ex ante* what is infringing, but an injunction *ex post* could be crippling.

Once a product or process has been designed in a way that incorporates a patented invention, redesigning the product might require shutdown for retooling. In addition, especially when the patented invention is a small component in the design of a complex product or process, a redesign around the patented invention may take substantial effort⁹³

In this case, a patent owner may then exact enormous licensing fees from an accidental infringer that the infringer would not have agreed to if the infringed patent had been identified in the product-development stage. Eliminating property-rule remedies—specifically eliminating injunctions and multiplied damages—would lessen the disincentives to producers created by the high discovery costs of the patent system.⁹⁴ This proposal does not constitute a

91. *See id.* at 1290.

92. *See id.* at 1296 (“Only if potential resource users know that use of the resource would intrude on someone else’s property right, and can readily identify the owner of that right, will they approach the owner to act as an information clearing-house.”); *id.* at 1304 (“[C]ompared to a liability-rule regime, a property-rule regime creates excessive incentives to search [to determine the scope of one’s legal rights] even when the search costs are high, [and] the probability of encroachment [on another’s right] is relatively low”); *see also id.* at 1311 (“Because the consequences of using without search are so draconian [in a property-rule regime], the user will often be willing to undertake an expensive search even when the probability of liability is very low.”).

93. *Id.* at 1333.

94. For a more detailed discussion in favor of awarding non-punitive royalties, see Brian J. Love, *The Misuse of Reasonable Royalty Damages as a Patent Infringement Deterrent*, 74 MO. L. REV. 909 (2009).

complete solution: litigation costs are still very high, and potential infringers and patent holders will still dispute the existence of liability and the size of damages. But it would be a step in the right direction.

A more modest proposal would combine this proposal and the previous one: limit injunctions and multiplied damages to cases where the plaintiff can demonstrate that actual copying took place. Under this approach, independent inventors would still be subject to liability for infringement, but their products could not be enjoined and their damages would be limited to a reasonable royalty rate determined by a judge. In contrast, a party caught copying another's invention would be subject to harsher remedies, including injunctions and heightened damages.

CONCLUSION

The patent system is supposed to promote the progress of science and the useful arts, but in some industries it seems to be doing just the opposite. The sheer number of patents and firms, and the lack of an effective organizational scheme for patents, can mean that patent clearance is practically impossible. In software, for example, patent clearance would require the services of many more patent attorneys than exist in the United States. In short, the patent system doesn't scale.

It's a fundamental problem that inventions in certain industries are not indexable, and incremental changes to the patent rules, such as beefing up the novelty and obviousness requirements, are not going to fix the problem. Only dramatic reforms—such as excluding industries with high discovery costs from patent protection, establishing an independent invention defense, or eliminating injunctions—can return the patent system to its proper role of promoting innovation.

DECODING FIRST AMENDMENT COVERAGE OF COMPUTER SOURCE CODE IN THE AGE OF YOUTUBE, FACEBOOK, AND THE ARAB SPRING

*JORGE R. ROIG**

ABSTRACT

Computer source code is the lifeblood of the Internet. It is also the brick and mortar of cyberspace. As such, it has been argued that the degree of control that a government can wield over code can be a powerful tool for controlling new technologies. With the advent and proliferation in the Internet of social networking media and platforms for the publication and sharing of user-generated content, the ability of individuals across the world to communicate with each other has reached truly revolutionary dimensions.

The influence of Facebook in the popular revolutions of the Arab Spring has been well documented. The use of YouTube in the 2008 U.S. presidential campaign has also left its indelible mark on the political landscape. New platforms have allowed millions of individuals to unleash their artistic and creative potentials. Tools like Google Earth have expanded the ability of entire populations to learn about their surroundings, the world at large, and their places

* Assistant Professor of Law, Charleston School of Law, Charleston, South Carolina; Juris Doctor, University of California at Berkeley, Boalt Hall School of Law, 2000; Bachelor of Arts with Honors in Economics, Harvard University, 1997. I would like to thank Professor Mark A. Lemley for serving as my advisor as I drafted the first version of this article and providing me with many insightful comments and suggestions. I would also like to thank Professor Sheila B. Scheuerman for her selfless and outstanding stewardship as I navigated the publication process, and for her valuable comments and revisions to my draft. Likewise, I wish to thank Professor Miller W. Shealy Jr. for his extensive comments and revisions. I would also like to acknowledge the help of my research assistants, Aaron K. Heath and Nicholas A. Shalosky, both of whom helped me in the revision of the final versions of the article. I am also deeply grateful for the constant support and helpful hand of the entire Charleston School of Law Faculty and Staff. This article would never have been finished without the help and understanding of my wonderful wife, Joanne A. Tomasini Muñiz, who always allows me to bounce my ideas off her privileged legal mind, is always willing to serve as my favorite devil's advocate, manages to keep me honest, and spent countless hours revising and editing the final version of my work. A very special mention, finally, goes to my darling daughter, Anya, who patiently put up with me as I failed to pay nearly enough attention during long hours of working on this article.

in it. The combination of smartphones and Twitter has created new tactics for protests and redefined the way in which individuals assemble to petition their government for a redress of grievances.

The time has come to reconsider the issue of whether computer source code is “speech” for First Amendment purposes and how the government can regulate it in a manner consistent with First Amendment values. This article proposes a three-step framework for analyzing questions of First Amendment coverage consistent with Supreme Court doctrine. In applying this framework to computer source code, this article also explores the relation between the different values that have been ascribed to the First Amendment, discusses some insights regarding the speech-conduct distinction, and considers the extent of First Amendment coverage in general.

Introduction	321	R
I. First Amendment Principles and the Problem of Source Code	327	R
A. Decoding “Speech”	327	R
B. Coverage vs. Protection	328	R
II. A Proposed Three-Part Test for First Amendment Coverage	331	R
A. Step 1: Is the Activity Communicative Enough? ..	331	R
1. The <i>Spence-Hurley</i> Test and the Per Se Rule for “Pure Speech”	331	R
2. Applying the <i>Spence-Hurley</i> Test to Source Code	338	R
B. Step 2: Is the Activity Central to the Development of a Medium for the Communication of Ideas? ..	341	R
1. Of Movie Projectors, Printing Presses, and Newspaper Racks	341	R
2. Source Code is Central to the Development of the Internet	343	R
C. Step 3: Does the Activity Promote First Amendment Values?	347	R
1. Theories of First Amendment Values	347	R
a. Truth	350	R
b. Democracy	352	R
c. Autonomy	354	R
d. Community	355	R
e. Prioritizing Values	356	R
2. First Amendment Values in Source Code	359	R
a. Truth	359	R

2012]	FIRST AMENDMENT COVERAGE OF SOURCE CODE	321	
	b. Democracy	369	R
	c. Autonomy	382	R
	d. Community.....	389	R
	Conclusion	395	R

INTRODUCTION

Over a decade ago, Dean Robert Post argued that “First Amendment coverage is triggered by those forms of social interaction that realize First Amendment values.”¹ Yet in the context of computer source code, he observed that “[d]igital First Amendment media, like the Internet, are so new and have such labile patterns of social interaction, that it seems to me enormously difficult to acquire reliable normative or descriptive traction on the relevant questions,” and that “it will be necessary to pursue this line of inquiry.”² In the years since, commentators have indeed made mention of theories of First Amendment values when considering the question of whether computer source code should be considered speech.³ However, a formal and extensive inquiry has been lacking, particularly since the advent and explosion of Web 2.0.⁴

1. Robert Post, *Encryption Source Code and the First Amendment*, 15 BERKELEY TECH. L.J. 713, 716 (2000) [hereinafter Post, *Encryption*] (citing Robert Post, *Recuperating First Amendment Doctrine*, 47 STAN. L. REV. 1249, 1254–55 (1995) [hereinafter Post, *Recuperating*]).

2. *Id.* at 723.

3. See, e.g., David McGowan, *From Social Friction to Social Meaning: What Expressive Uses of Code Tell Us About Free Speech*, 64 OHIO ST. L.J. 1515, 1515 (2003). Additionally, some scholars have debated the Internet’s influence on, inter alia, First Amendment values, democracy, republicanism, and politics, albeit in much more general terms. See, e.g., MATTHEW HINDMAN, THE MYTH OF DIGITAL DEMOCRACY 4 (2009); CASS R. SUNSTEIN, REPUBLIC.COM 2.0 5 (2007).

4. There is “a huge amount of disagreement about just what Web 2.0 means, with some people decrying it as a meaningless marketing buzzword, and others accepting it as the new conventional wisdom.” Tim O’Reilly, *What is Web 2.0*, O’REILLY (Sept. 30, 2005), <http://oreilly.com/web2/archive/what-is-web-20.html>. An oft-cited attempt at defining the concept, however, proposes a set of principal features of Web 2.0: (1) the use of the web as a platform; (2) the recognition of the power of harnessing collective intelligence; (3) competent database management; (4) the end of the software release cycle by delivering software as a service, not as a product; (5) the support of lightweight programming models; (6) writing software above the level of a single device; and (7) providing rich user experiences. *Id.* O’Reilly has also summarized what he believes to be “the core competencies of Web 2.0 companies” as follows: (1) services, not packaged software, with cost-effective scalability; (2) control over unique, hard-to-recreate data sources that get richer as more people use them; (3) trusting users as co-developers and harnessing collective intelligence; (4) leveraging the long tail through customer self-service;

The question of First Amendment coverage of computer source code was a hot topic in both academic debate and litigation around the turn of the millennium.⁵ However, the hubbub about source code as speech seems to have died down considerably after, roughly, 2004. The principal cases dealing with the issue of whether source code is covered by the First Amendment arose from challenges to regulations concerning the export and publication of encryption software.⁶ However, the most explosive controversies surrounding the topic were defused by the federal government's amendment of these regulations.⁷ By amending the regulations, the federal government largely mooted the relevant causes of action.⁸

(5) software above the level of a single device; and (6) lightweight user interfaces, development models, and business models. *Id.*

5. *See infra* notes 8–12.

6. *See* *Junger v. Daley (Junger I)*, 8 F. Supp. 2d 708, 711 (N.D. Ohio 1998), *rev'd*, 209 F.3d 481 (6th Cir. 2000); *Bernstein v. U.S. Dep't of State (Bernstein I)*, 922 F. Supp. 1426, 1428 (N.D. Cal. 1996), *aff'd*, *Bernstein v. U.S. Dep't of Justice*, 176 F.3d 1132 (9th Cir.), *withdrawn & reh'g granted*, 192 F.3d 1308 (9th Cir. 1999); *Karn v. U.S. Dep't of State*, 925 F. Supp. 1, 3 (D.D.C. 1996).

7. *See* Ryan Christopher Fox, Comment, *Old Law and New Technology: The Problem of Computer Source Code and the First Amendment*, 49 UCLA L. REV. 871, 887–88 (2002).

On January 14, 2000, however, the state of the law changed. The Encryption Administration Regulations were revised to allow U.S. companies to “have new opportunities to sell their products in the global marketplace.” Among other changes, the revisions decontrol encryption software up to and including sixty-four bits, and allow unrestricted encryption source code to be released without review, provided that the code is not “subject to an express agreement for payment of a licensing fee or royalty.” The regulations were presumably modified this way in order to support the “open source” approach to software development. The revised regulations also provided for a number of other allowances that eased review of exports in other situations.

Id. (quoting Revisions to Encryption Items, 65 Fed. Reg. 2492, 2492, 2497, 2499 (Jan. 14, 2000) (codified at 15 C.F.R. pts. 734, 740, 742, 770, 772, 774)).

8. *See, e.g.*, *Bernstein v. U.S. Dep't of Commerce (Bernstein IV)*, No. C 95-0582 MHP, 2004 WL 838163, at *2 n.2 (N.D. Cal. Apr. 19, 2004) (“In January 2000, defendants added 14 C.F.R. section 740.13(e) to the Federal Register, which allows the DOC to exempt ‘publicly available’ encryption source code from license requirements. Plaintiff amended his complaint in January 2002, alleging that the changed regulations still amounted to a prior restraint under the First Amendment. The defendants brought a motion for summary judgment on the amended complaint on the grounds that he lacked the requisite standing, which this court granted on July 28, 2003.”); *Karn v. U.S. Dep't of State*, No. 95-CV-01812, Docket No. 79 (D.D.C. Mar. 3, 2000) (order dismissing the complaint as moot). Similarly, most of the academic literature on the subject was concerned with that same set of litigation. *See, e.g.*, Dan L. Burk, *Patenting Speech*, 79 TEX. L. REV. 99, 101 (2000); Fox, *supra* note 7, at 888; Steven E. Halpern, *Harmonizing the Convergence of Medium, Expression, and Functionality: A Study of the Speech Interest in Computer Software*, 14

In the years since, however, the issues at the heart of this debate have hardly abated. For one thing, the several federal courts that broached the subject were not of one mind in their conclusions—or even in their methods of analysis.⁹ In fact, there was and still remains great disagreement in the academic community regarding the appropriate approach and answer to the question of First Amendment coverage of computer source code.¹⁰

HARV. J.L. & TECH. 139, 140 n.3, 175–177 (2000); Orin S. KERT, *Are We Overprotecting Code? Thoughts on First-Generation Internet Law*, 57 WASH. & LEE L. REV. 1287, 1289 (2000); Post, *Encryption*, *supra* note 1, at 713; Lee Tien, *Publishing Software as a Speech Act*, 15 BERKELEY TECH. L.J. 629, 631–32 (2000); John P. Collins, Jr., Note, *Speaking in Code*, 106 YALE L.J. 2691, 2691 (1997); Norman A. Crain, Comment, Bernstein, Karn, and Junger: *Constitutional Challenges to Cryptographic Regulations*, 50 ALA. L. REV. 869, 870 (1999); Seth Hanson, Note, *Bernstein v. United States Dep't of Justice: A Cryptic Interpretation of Speech*, 2000 B.Y.U. L. REV. 663, 664 (2000); Katherine A. Moerke, Note, *Free Speech to a Machine? Encryption Software Source Code Is Not Constitutionally Protected "Speech" Under the First Amendment*, 84 MINN. L. REV. 1007, 1007–08 (2000); Yvonne C. Ocrant, Comment, *A Constitutional Challenge to Encryption Export Regulations: Software Is Speechless*, 48 DEPAUL L. REV. 503, 504 (1998). Hence, with the cases went the articles.

9. Compare *Junger v. Daley (Junger II)*, 209 F.3d 481, 485 (6th Cir. 2000) (holding source code to be an expressive means of exchange and therefore protected by the First Amendment), *Bernstein v. U.S. Dep't of Justice (Bernstein II)*, 176 F.3d 1132, 1147 (9th Cir.), *withdrawn & reh'g granted*, 192 F.3d 1308 (9th Cir. 1999) (holding that the encryption export controls violated the First Amendment as applied), *Universal City Studios, Inc. v. Reimerdes*, 111 F. Supp. 2d 294, 327 (S.D.N.Y. 2000) (“As computer code—whether source or object—is a means of expressing ideas, the First Amendment must be considered before its dissemination may be prohibited or regulated.”), *aff'd sub nom.*, *Universal City Studios, Inc. v. Corley*, 273 F.3d 429, 449 (2d Cir. 2001) (holding that “computer code, and computer programs constructed from code can merit First Amendment protection”), and *DVD Copy Control Ass'n, Inc. v. Bunner*, 75 P.3d 1, 10 (Cal. 2003) (holding that computer code and computer programs constructed from code are covered by the First Amendment), *with Universal City Studios v. Reimerdes*, 82 F. Supp. 2d 211, 222 (S.D.N.Y. 2000) (stating that source code at issue is not covered by the First Amendment because its “expressive aspect appears to be minimal when compared to its functional component”), and *Karn*, 925 F. Supp. at 8–13 (holding that the regulations in question did not violate freedom of speech).

10. Compare Brian F. Fitzgerald, *Software as Discourse: The Power of Intellectual Property in Digital Architecture*, 18 CARDOZO ARTS & ENT. L.J. 337, 383–85 (2000), *Fox*, *supra* note 7, at 907–08 (citing Post, *Encryption*, *supra* note 1) (noting that some argue for “an increased protection status” for computer source code), Halpern, *supra* note 8, Liam Séamus O'Melinn, *The New Software Jurisprudence and the Faltering First Amendment*, 6 VAND. J. ENT. L. & PRAC. 310, 310 (2004) (claiming that courts are failing to shield source code from regulation as the First Amendment should require), Post, *Encryption*, *supra* note 1, at 717, Tien, *supra* note 8 *passim* (arguing for increased protection by considering works of software as “speech acts”), and Crain, *supra* note 8, at 870 (arguing that encryption regulation should be found unconstitutional under the First Amendment), *with Burk*, *supra* note 8,

Additionally, and possibly more critically, along with the relevant developments in new technologies and related social practices, the questions framed concerning the First Amendment coverage of source code have multiplied.¹¹ Today we live in an age of video sharing, viral Internet memes, YouTube presidential debates, social networking, tweeting, blogging, smartphone-enabled protests, and Facebook-fueled popular overthrows of decades-long regimes. The development over the past decade of new code, corresponding Internet architectures, and resulting social practices makes it essential that we reengage in the discussion that showed promise roughly ten years ago. This article aims to do so.

Of course, the government's concern and appetite for regulation of these new technologies and social practices is as robust as ever. Some have even observed a "focus on prohibiting or restricting code itself as a dangerous tool rather than relying on laws against the undesirable activity that the code facilitates."¹² One issue currently before Congress illustrates the urgency, novelty and complexity of these questions:

at 101–02 (seeing the long-term implications of treating software as speech to be "troublesome"), Kerr, *supra* note 8, at 1291 (suggesting that the Sixth Circuit Court's holding in *Junger II*, 209 F.3d at 485, might be overprotective of code in the First Amendment context), Collins, *supra* note 8, at 2696 (finding "no First Amendment right to speak in cryptographic computer source code"), Hanson, *supra* note 8, at 693 (arguing an expressive/functional test effectively balances the preservation of social order and individual liberty interest in free speech), Moerke, *supra* note 8, at 1048 (arguing that while "source code itself is not speech under the First Amendment," the encryption software may be entitled to First Amendment protection as an activity that provides for speech), and Ocrant, *supra* note 8, at 538–47 (arguing that cryptographic software alternatively is not speech, not protected speech, or at the most, speech afforded limited protection).

11. See, e.g., Tim Wu, Op-Ed, *Free Speech for Computers?*, N.Y. TIMES (June 19, 2012), <http://www.nytimes.com/2012/06/20/opinion/free-speech-for-computers.html> (describing the questions of whether computers speak and whether their speech should be covered by the First Amendment as ones that have "become [] real issue[s] with important consequences").

12. Fox, *supra* note 7, at 874. In his comment, Fox further remarked that, "while judges are beginning to understand the creative and social uses of computer source code, they are all the while hesitant to give the idea too much latitude." *Id.* at 894. In this sense, we should bear in mind that the Ninth Circuit Court's three-judge panel opinion in *Bernstein II*, 176 F.3d 1132, which represents one of the most robust arguments for First Amendment coverage of source code, was withdrawn, pending a rehearing en banc that never materialized, and is no longer good law. *Bernstein II*, 176 F.3d at 1132, *withdrawn & reh'g granted*, 192 F.3d 1308 (9th Cir. 1999). The rehearing en banc never occurred due to the amendment of the encryption regulations at issue in the case, which deprived the plaintiff of standing. See *supra* note 8; *Bernstein IV*, 2004 WL 838163, at *2 n.2.

The FBI believes that the historic shift in communication from telephones to the Internet has made it far more difficult for agents to wiretap Americans suspected of illegal activities, which it refers to as the “Going Dark” problem. Its solution: a proposed law that would require Internet companies including Apple, Microsoft, Facebook, Yahoo, and Google, to build in back doors for government surveillance.¹³

In an effort to enhance its surveillance and law enforcement capabilities, the FBI is currently lobbying Congress for a new Internet wiretapping statute.¹⁴ Wiretapping is not a new concept, of course, and neither is the idea of law enforcement agencies seeking new ways of expanding their surveillance capabilities. But what makes this situation particularly interesting, and dangerous, is the fact that the FBI is seeking to force private entities and individuals to build into their code architecture the ability for the FBI to eavesdrop on users’ communications.¹⁵ The consequent lack of transparency and accountability must give us pause.

Many questions of constitutional significance are apparent from this scenario.¹⁶ But one area of particular concern is the extent to which such a policy would infringe upon the First Amendment. On one hand, such surveillance could cause citizenry’s speech to be unconstitutionally chilled. On the other hand, forcing entities to write an FBI backdoor into their source code amounts to compelled speech, a possible violation of their free speech rights.

However, before we can even begin to discuss these issues, we must determine whether the regulation of computer source code implicates First Amendment guarantees at all. This article seeks an answer to this threshold question.

This article will argue that a three-step approach to First Amendment coverage best embodies the thrust of First Amendment doctrine. First, a court must consider whether a particular activity is communicative enough to be considered “speech” for First

13. Declan McCullagh, *FBI ‘Looking at’ Law Making Web Sites Wiretap-Ready, Director Says*, CNET NEWS (May 18, 2012, 1:17 PM PDT) [hereinafter McCullagh, *FBI ‘Look at’*], http://news.cnet.com/8301-1009_3-57437391-83/fbi-looking-at-law-making-web-sites-wiretap-ready-director-says/. See also Declan McCullagh, *FBI: We Need Wiretap-Ready Web Sites - Now*, CNET NEWS (May 4, 2012, 9:24 AM PDT), http://news.cnet.com/8301-1009_3-57428067-83/fbi-we-need-wiretap-ready-web-sites-now/.

14. See McCullagh, *FBI ‘Look at’*, *supra* note 13.

15. *Id.*

16. For example: How should the Fourth Amendment’s prohibition on unreasonable searches limit the FBI’s ability to do what it proposes? What type of Due Process problems, both procedural and substantive, does such a regulation raise?

Amendment purposes. At this level of analysis, “pure speech” and so-called “expressive conduct” are distinguished from conduct that is not sufficiently communicative to trigger the First Amendment.

At the second stage, activities that did not pass the first stage are reconsidered. Activities and devices that facilitate the development of a medium for the expression of ideas, though not themselves “expressive,” trigger First Amendment coverage as readily as traditional speech.

Passing one of these two steps, however, does not guarantee First Amendment coverage. If a particular activity is found to be either (1) communicative or (2) central to the development of a medium for the expression of ideas, then the court must engage in a comprehensive analysis of First Amendment values in order to determine the extent of coverage that such activity will merit. It is at this stage that specific kinds of speech—such as obscenity, fighting words, or commercial speech—are defined as deserving only limited First Amendment coverage. Specifically, the analysis will consider four central values of free speech: truth, democracy, autonomy, and community. These values serve both to justify and limit coverage of certain types of speech.

In summation, only activities that (1) pass *either* of the first two stages in the three-part analysis, *and also* (2) further First Amendment values under the third stage of the analysis are speech covered by the First Amendment.

Source code, as a general category of activities, passes all three parts of this test. Under the first stage, not only is source code “expressive conduct,” but it should actually be considered “pure speech.” Yet even if source code were not to be deemed communicative, its regulation would nonetheless trigger First Amendment coverage under the second stage: source code both promotes communication and is crucial to the development of another recognized medium for the expression of ideas, the Internet. Finally, under the third stage, source code not only promotes the three core First Amendment values of truth, democracy, and autonomy, it does so without threatening to destroy—and while in fact promoting—the community that the First Amendment serves. In an age when source code is essential to the spread of political speech and thought on a global scale, the recognition that code is covered by the First Amendment is necessary to further First Amendment values themselves.

I.
FIRST AMENDMENT PRINCIPLES AND THE
PROBLEM OF SOURCE CODE

A. *Decoding "Speech"*

The First Amendment guarantees our freedom of speech.¹⁷ It therefore follows that First Amendment protection can only be applied to "speech" as the term of art has been defined by the courts.¹⁸ But in order to ascertain whether an action constitutes speech, one must first understand the nature of the action itself. This section begins this process by clarifying what we mean when we talk about source code.

Computer source code is the text of a computer program written in a high-level programming language that can be read and understood by humans, but which can also be easily translated into computer-executable object code through the use of a program called a compiler.¹⁹ Thus source code has the distinguishing characteristic of being both comprehensible to humans and readily translatable into a form that can be fed into a computer. When source code is compiled and run on a computer, the machine will perform the tasks that have been encoded into the algorithms embodied in the source code. This gives source code a distinctly functional nature. In fact, some people characterize source code as a machine itself because "any function that can be implemented in software can be implemented equally well in hardware."²⁰ This means that a particular set of functions described and implemented by a piece of source code could also be hardwired into the hardware of a computer to produce the same effect.²¹

In spite of its functional characteristics, the fact remains that source code is a "language" that can be written, read, and understood by humans. Many people write and read in computer languages such as C, C++, Fortran, COBOL, Python, Perl, and Java. This means that these individuals can communicate ideas to each

17. See U.S. CONST., amend. I.

18. See Post, *Encryption*, *supra* note 1, at 715. *But see* R.A.V. v. City of St. Paul, 505 U.S. 377, 383–84 (1992) (explaining that even specific categories of expression traditionally thought to reside outside the auspices of the First Amendment are not "entirely invisible to the Constitution").

19. See *Bernstein v. U.S. Dep't of Justice (Bernstein II)*, 176 F.3d 1132, 1140 (9th Cir.), *withdrawn*, 192 F.3d 1308 (9th Cir. 1999).

20. Burk, *supra* note 8, at 119 (citing Pamela Samuelson et al., *A Manifesto Concerning the Legal Protection of Computer Programs*, 94 COLUM. L. REV. 2308, 2319 (1994)).

21. *Id.* (citing Samuelson, *supra* note 20, at 2320 n.34.).

other through the use of a programming language. In the everyday sense, a person who knows a particular high-level computer language can express herself in source code.

Still, not everything that might be included in the dictionary definition of the word “speech” receives full First Amendment coverage and protection. Just to name a few examples, the government routinely regulates contracts made through the use of language,²² prohibits the publication of obscene materials,²³ and punishes the solicitation of crimes.²⁴ Thus source code, like any other language or form of speech, may receive full, partial, or no First Amendment coverage. The next section will begin to explore these differing levels of coverage and what it means to be *protected* by the First Amendment.

B. Coverage vs. Protection

For purposes of First Amendment analysis, the courts have tried to distinguish between fully covered speech, speech that receives limited coverage, and non-speech. Unfortunately, the courts have not allocated different types of activities across these categories in a completely sound or consistent manner. The development of coherent doctrine will thus require an analysis of the values underlying the First Amendment itself. However, before sorting different types of activities into these categories, it is important to understand how these categories function within First Amendment doctrine.

First off, it is crucial to distinguish “coverage” from “protection”: the fact that the First Amendment might *cover* certain activity does not necessarily mean that such activity is *protected* by the First Amendment. If the First Amendment “covers” certain conduct that the government seeks to regulate, “the constitutionality of the conduct’s regulation must be determined by reference to First Amendment doctrine and analysis.”²⁵ If, on the other hand, a particular activity is not covered by the First Amendment, courts need not consult First Amendment doctrine to determine the constitutionality of its regulation. Thus the secondary question of First Amendment protection only arises if the initial question of coverage has been answered affirmatively. To say that an activity is “protected” by the First Amendment from government regulation means *first* that

22. See *R.A.V.*, 505 U.S. at 420 (citing Frederick Schauer, *Categories and the First Amendment: A Play in Three Acts*, 34 VAND. L. REV. 265, 270 (1981)).

23. See *Miller v. California*, 413 U.S. 15, 23 (1973).

24. See *United States v. Williams*, 553 U.S. 285, 297–98 (2008).

25. Post, *Encryption*, *supra* note 1, at 714.

the activity is covered by the First Amendment and *second* that the regulation attempted by the government is unconstitutional under First Amendment doctrine.²⁶

First Amendment coverage can also be a matter of degree: it need not cover all activities to the same extent. For example, pure political speech in a public forum—as traditionally embodied by the proverbial soapbox orator—receives full coverage and triggers strict scrutiny.²⁷ On the other hand, other types of communicative conduct—such as commercial speech,²⁸ speech of a sexual nature,²⁹ speech on non-public forums,³⁰ or what the Court has termed “expressive conduct”³¹—may receive limited (and differing) levels of First Amendment coverage and trigger only intermediate or even lesser levels of scrutiny.³²

26. *Id.*

27. *See, e.g.* Citizens United v. Fed. Election Comm’n, 130 S. Ct. 876, 898 (2010) (“Laws that burden political speech are ‘subject to strict scrutiny’” (quoting Fed. Election Comm’n v. Wis. Right To Life, Inc., 551 U.S. 449, 464 (2007))); Burson v. Freeman, 504 U.S. 191, 197 n.3 (1992) (“[C]ontent-based regulation of political speech in a public forum is valid only if it can survive strict scrutiny.”); F.T.C. v. Superior Court Trial Lawyers Ass’n, 493 U.S. 411, 448 (1990) (“Like soapbox oratory in the streets and parks, political boycotts are a traditional means of ‘communicating thoughts between citizens’ and ‘discussing public questions.’ Any restrictions on such boycotts must be scrutinized with special care in light of their historic importance as a mode of expression.”) (citations omitted); Boos v. Barry, 485 U.S. 312, 321 (1988) (“Our cases indicate that as a *content-based* restriction on *political speech* in a *public forum*, [the Act in question] must be subjected to the most exacting scrutiny.”).

28. *See* Central Hudson Gas & Elec. Corp. v. Public Serv. Comm’n of N.Y., 447 U.S. 557, 562–63 (1980) (“The Constitution therefore accords a lesser protection to commercial speech than to other constitutionally guaranteed expression.”).

29. *See, e.g.*, Young v. Am. Mini Theatres, Inc., 427 U.S. 50, 70 (1976) (plurality opinion) (“[E]ven though we recognize that the First Amendment will not tolerate the total suppression of erotic materials that have some arguably artistic value, it is manifest that society’s interest in protecting this type of expression is of a wholly different, and lesser, magnitude than the interest in untrammelled political debate”); New York v. Ferber, 458 U.S. 747, 764 (1982) (child pornography); Miller, 413 U.S. at 23 (obscenity).

30. *See* Davenport v. Wash. Educ. Ass’n, 551 U.S. 177, 178–79 (2007) (“Thus, the government can make content-based distinctions when subsidizing speech, and can exclude speakers based on reasonable, viewpoint-neutral subject-matter grounds when permitting speech on government property that is a nonpublic forum.”) (citations omitted).

31. *See* Texas v. Johnson, 491 U.S. 397, 406 (1989) (“The government generally has a freer hand in restricting expressive conduct than it has in restricting the written or spoken word.”).

32. *See* Sorrell v. IMS Health Inc., 131 S. Ct. 2653, 2673–74 (2011) (“Thus, the First Amendment imposes tight constraints upon government efforts to restrict, *e.g.*, ‘core’ political speech, while imposing looser constraints when the govern-

As such, the level of constitutional scrutiny to be applied to a particular regulation is determined by the level of the targeted conduct's First Amendment coverage. If the regulated activity is not covered by the First Amendment at all, the courts should default to traditional rational basis analysis.³³ If the activity is covered as speech, the courts might still face a question of whether that activity, in the specific context, should receive full coverage—subjecting the challenged regulation to traditional strict scrutiny—or whether it should receive some level of limited coverage—subjecting the challenged regulation to some form of intermediate scrutiny.³⁴

Once coverage—and therefore the appropriate level of constitutional scrutiny—has been established, the question of First Amendment protection can be posed and answered. This inquiry involves the actual application of that scrutiny to the challenged regulation. At this stage, courts should consider the government interest being pursued through regulation and the fit between such ends and the means employed to achieve it.³⁵

To say that the First Amendment covers source code, then, does not mean that the government will not be able to regulate the behavior of computer programmers and users. It only means that

ment seeks to restrict, *e.g.*, commercial speech, the speech of its own employees, or the regulation-related speech of a firm subject to a traditional regulatory program.”); *Bd. of Educ. v. Grumet*, 512 U.S. 687, 718 (1994) (O'Connor, J., concurring in part and concurring in the judgment):

But the same constitutional principle may operate very differently in different contexts. We have, for instance, no one Free Speech Clause test. We have different tests for content-based speech restrictions, for content-neutral speech restrictions, for restrictions imposed by the government acting as employer, for restrictions in nonpublic fora, and so on. This simply reflects the necessary recognition that the interests relevant to the Free Speech Clause inquiry—personal liberty, an informed citizenry, government efficiency, public order, and so on—are present in different degrees in each context.

33. *See, e.g.*, *Madsen v. Women's Health Ctr., Inc.*, 512 U.S. 753, 791 (1994) (Scalia, J., concurring in the judgment in part and dissenting in part) (“[T]he ‘rational basis’ standard . . . is applied—under the Equal Protection Clause—to government regulation of nonspeech activities. . .”).

34. *See, e.g.*, *Central Hudson Gas & Elec. Corp. v. Public Serv. Comm'n of N.Y.*, 447 U.S. 557, 562-63 (1980) (applying an intermediate level of scrutiny to commercial speech).

35. *See, e.g.*, *Boos v. Barry*, 485 U.S. 312, 321 (1988) (applying strict scrutiny and requiring “the State to show that the ‘regulation is necessary to serve a compelling state interest and that it is narrowly drawn to achieve that end’” (quoting *Perry Educ. Ass'n v. Perry Local Educators' Ass'n*, 460 U.S. 37, 45 (1983))); *Central Hudson*, 447 U.S. at 564 (applying an intermediate level of scrutiny and requiring that the State “assert a substantial interest to be achieved by restrictions on commercial speech” and that “the regulatory technique . . . be in proportion to that interest”).

the First Amendment will limit the way in which such conduct can be regulated according to the values it embodies.³⁶

II. A PROPOSED THREE-PART TEST FOR FIRST AMENDMENT COVERAGE

A. *Step 1: Is the Activity Communicative Enough?*

1. The *Spence-Hurley* Test and the Per Se Rule for “Pure Speech”

The Supreme Court has addressed the issue of what constitutes speech for First Amendment purposes in a series of cases dealing with what has been termed “symbolic speech,”³⁷ “expressive conduct,”³⁸ or “the expression of an idea through activity.”³⁹ Perhaps the four most important cases in this series are *United States v. O’Brien*,⁴⁰ *Spence v. State of Washington*,⁴¹ *Texas v. Johnson*,⁴² and *Hurley v. Irish-American Gay, Lesbian and Bisexual Group of Boston*.⁴³ In all of these cases someone was engaged in an activity that did not include the oral or written word, but nonetheless claimed to be expressing an idea. These cases drew a doctrinal distinction between this kind of “expressive conduct” and what the Supreme Court considers to be “pure speech.”⁴⁴

According to the traditional interpretation of this Supreme Court doctrine, the oral or written word is “pure speech” and is automatically entitled to First Amendment coverage.⁴⁵ Under this interpretation, as Judge Patel stated in her *Bernstein I* decision, *John-*

36. It is important to bear in mind that this article will limit its analysis to the question of First Amendment coverage. A much more extensive analysis would be required to go into the issues of First Amendment protection and the constitutionality of specific regulations such as the encryption regulations at issue in the *Bernstein*, *Junger*, and *Kam* litigations. These questions are left open for further research.

37. *United States v. O’Brien*, 391 U.S. 367, 376 (1968).

38. *Texas v. Johnson*, 491 U.S. 397, 406 (1989).

39. *Spence v. Washington*, 418 U.S. 405, 411 (1974).

40. 391 U.S. 367.

41. 418 U.S. 405.

42. 491 U.S. 397.

43. 515 U.S. 557 (1995).

44. *Cf.* Harry Kalven, Jr., *The Concept of the Public Forum: Cox v. Louisiana*, 1965 SUP. CT. REV. 1, 22–23 (referring to these two kinds of activities as “speech plus” and “speech pure,” respectively.). *See also* Post, *Recuperating*, *supra* note 1, at 1257 (citing Kalven, *supra*, at 22–23).

45. *See Tinker v. Des Moines Indep. Cmty. Sch. Dist.*, 393 U.S. 503, 505–06 (1969) (“‘[P]ure speech’ . . . , we have repeatedly held, is entitled to comprehensive protection under the First Amendment.” (citation omitted)).

son and *Spence* “strongly imply that a court need only assess the expressiveness of conduct in the absence of ‘the spoken or written word.’”⁴⁶

But the Court’s statements and implications, as traditionally interpreted to mean that the condition of being written or oral gives words automatic and full coverage, cannot be right. There are many instances in which the government regulates the oral and written word without triggering the full force of the First Amendment. Some of these instances deal with the regulation of particular kinds of speech that have been defined as receiving very limited First Amendment coverage,⁴⁷ such as obscenities⁴⁸ and fighting words.⁴⁹ But oral and written communications are continuously regulated in contexts that do not seem to fit such handy and constrained categories. For example, language used during the commission of a crime like solicitation⁵⁰ and language used in commercial transactions that falls short of commercial speech⁵¹ are both routinely regulated. The Supreme Court has recently reaffirmed that “it has never been deemed an abridgment of freedom of speech or press to make a course of conduct illegal merely because the conduct was in part initiated, evidenced, or carried out by means of language, either spoken, written, or printed.”⁵² That words are spoken, it seems, is no guarantee of coverage.

In any case, even if these forms of “not fully covered” oral and written speech could be subject to specific definitions, there should be some sort of underlying doctrinal framework that explains why these particular types of speech, and not others, are to be granted only the most limited form of First Amendment coverage. An underlying framework of this sort could *also* help us in determining which types of “expressive conduct” are worthy of First Amendment coverage. At bottom, all oral and written words are a kind of “expressive conduct” or “symbolic speech.” The oral word is nothing more than the exercise of certain muscles in our throat that vibrate

46. *Bernstein v. U.S. Dep’t of State (Bernstein I)*, 922 F. Supp. 1426, 1434 (N.D. Cal. 1996) (quoting *Texas v. Johnson*, 491 U.S. 397, 404 (1989)), *aff’d*, *Bernstein v. U.S. Dep’t of Justice*, 176 F.3d 1132 (9th Cir.), *withdrawn & reh’g granted*, 192 F.3d 1308 (9th Cir. 1999).

47. *See, e.g., R.A.V. v. City of St. Paul*, 505 U.S. 377, 420 (1992).

48. *Miller v. California*, 413 U.S. 15, 23 (1973).

49. *Chaplinsky v. New Hampshire*, 315 U.S. 568, 571–72 (1942).

50. *United States v. Williams*, 553 U.S. 285, 297–98 (2008).

51. *R.A.V.*, 505 U.S. at 420 (citing *Schauer*, *supra* note 22, at 270).

52. *Rumsfeld v. Forum for Academic & Institutional Rights, Inc.*, 547 U.S. 47, 62 (2006) (quoting *Giboney v. Empire Storage & Ice Co.*, 336 U.S. 490, 502 (1949)) (internal quotation marks omitted).

to make certain sounds which others can hear.⁵³ This is not, in principle, that different from a person waving his hand, a group walking down the streets of Boston,⁵⁴ or someone performing any other kind of physical activity. The written word is nothing more than a series of symbols inscribed in some tangible medium, not unlike the adherence of a peace symbol on a flag.⁵⁵ By the same token, as the Court acknowledges in its “expressive conduct” decisions, activities that are not the oral or written word can be just as expressive as speaking or writing. Hence, in *Johnson*, the Court extended the full coverage of the First Amendment to the activity of flag burning because it recognized that such an act could be as expressive as soapbox oration.⁵⁶

What makes the oral and written word intuitively different from other activities is the existence of a specific set of social conventions that make the sounds and symbols that we use in speaking and writing especially expressive.⁵⁷ As the Court explains in *Spence*: “[T]he context may give meaning to the symbol.”⁵⁸ When we speak or write in English, or in any other language for that matter, we bring to the table a whole set of historical and social axioms and contexts which enable us to communicate effectively and efficiently with others who speak the same language and recognize and use the same set of conventions to decode our messages.

The formalistic distinction that the Court draws between pure and symbolic speech is, therefore, an illusory one. At the very least, it is not an objective or clear-cut distinction, conveniently lingering in the state of nature for us to grasp and apply with ease. The Court’s line of reasoning, however, is the correct one. Part of the analysis necessary to determine whether an activity is speech for First Amendment purposes consists of deciding whether the activity has any communicative value. That is, we must assess whether there are enough social conventions in place such that others can understand the specific activity as conveying some kind of message. The Court developed in *Spence* the following test to determine whether some form of symbolic speech merits First Amendment scrutiny: it

53. See *Bernstein v. U.S. Dep’t of State (Bernstein I)*, 922 F. Supp. 1426, 1435 (N.D. Cal. 1996), *aff’d*, *Bernstein v. U.S. Dep’t of Justice*, 176 F.3d 1132 (9th Cir.), *withdrawn & reh’g granted*, 192 F.3d 1308 (9th Cir. 1999).

54. See *Hurley v. Irish-American Gay, Lesbian & Bisexual Grp. of Bos.*, 515 U.S. 557 (1995).

55. See *Spence v. Washington*, 418 U.S. 405, 410 (1974).

56. *Texas v. Johnson*, 491 U.S. 397, 406 (1989).

57. See Post, *Recuperating*, *supra* note 1, at 1257.

58. 418 U.S. at 410 (citing *Tinker v. Des Moines Indep. Cmty. Sch. Dist.*, 393 U.S. 503 (1969)).

must have “an intent to convey a particularized message” and “in the surrounding circumstances the likelihood [must be] great that the message would be understood by those who viewed it.”⁵⁹

The *Spence* test, however, was modified in *Hurley*.⁶⁰ In *Hurley*, the Court made clear that a particularized message is not required: “[A] narrow, succinctly articulable message is not a condition of constitutional protection, which if confined to expressions conveying a ‘particularized message’ would never reach the unquestionably shielded painting of Jackson Pollock, music of Arnold Schoenberg, or Jabberwocky verse of Lewis Carroll.”⁶¹ This would, of course, seem to suggest that more than just communicative values are ascribed to the First Amendment and that some of those values might in some cases outweigh the communicative ones.

Thus *Spence* and its progeny establish that an activity is communicative enough to be considered speech under the First Amendment when, in a particular social context, sufficient conventions exist such that the communication of ideas between people is possible, even if not overwhelmingly probable or specifically intended. Such an activity’s communicative nature makes it equivalent to pure speech, potentially activating First Amendment scrutiny of some kind. Meanwhile, with respect to pure speech, the Court has made a *per se* determination that the oral and written word will always pass the *Spence-Hurley* test: for all oral or written communication, the requisite social contexts exist in the form of an established language.

However, the Court has said that “[t]he government generally has a freer hand in restricting expressive conduct than it has in restricting the written or spoken word.”⁶² This would suggest that the distinction between symbolic speech and pure speech is more than just the establishment of a *per se* rule for the written and oral word. The Court seems to suggest that the coverage extended to symbolic speech is going to be less than that extended to pure speech—that there is some substantive difference between the treatments accorded to pure speech versus symbolic speech. How-

59. *Id.* at 410–11.

60. *Hurley v. Irish-American Gay, Lesbian & Bisexual Grp. of Bos.*, 515 U.S. 557, 569 (1995).

61. *Id.* (citation omitted) (citing *Spence*, 418 U.S. at 411).

62. *Texas v. Johnson*, 491 U.S. 397, 406 (1989).

ever, the Court's actual analysis of First Amendment issues shows this not to be the case.⁶³

First, the Court gives the same treatment to the communicative part of symbolic speech and to the communicative part of pure speech. Content-based regulations will be subject to strict scrutiny regardless of whether they target the oral or written word, or whether they target the communicative aspects of some other form of symbolic speech.⁶⁴ For example, in *Johnson*, after stating that the government has a freer hand in regulating expressive conduct, the Court clarifies that “[a] law directed at the communicative nature of conduct must, *like a law directed at speech itself*, be justified by the substantial showing of need that the First Amendment requires.”⁶⁵ The Court then elaborates that “[i]t is, in short, not simply the verbal or nonverbal nature of the expression, but the governmental interest at stake, that helps to determine whether a restriction on that expression is valid.”⁶⁶ When the government tries to regulate the communicative aspects of symbolic speech, strict scrutiny applies just as if it were pure speech.

Second, there is also no difference between the treatment of the non-communicative elements of symbolic speech and the treatment of the conduct-like elements of pure speech.⁶⁷ The Court has consistently held that intermediate scrutiny applies whenever the government regulates in a content-neutral fashion the non-communicative aspects of any activity.⁶⁸ In *O'Brien*, the Court established that a more lenient standard applies when the government regulates the non-communicative part of symbolic speech and “the governmental interest is unrelated to the suppression of free

63. See James M. McGoldrick, Jr., *Symbolic Speech: A Message from Mind to Mind*, 61 OKLA. L. REV. 1, 6 (2008) (“[R]efut[ing] the common claim that the government has ‘a freer hand’ in regulating symbolic speech than pure speech.”).

64. *Id.* at 25 (“If something is speech, then the level of protection will depend on whether the law is content-based or content-neutral, not the speech itself and not whether it is pure speech or symbolic speech.”).

65. *Johnson*, 491 U.S. at 406 (emphasis added) (quoting *Cnty. for Creative Non-Violence v. Watt*, 703 F.2d 586, 622–23 (1983) (Scalia, J., dissenting), *rev'd sub nom.* *Clark v. Cnty. for Creative Non-Violence*, 468 U.S. 288, 293 (1984)) (internal quotation marks omitted).

66. *Id.* at 406–07.

67. McGoldrick, Jr., *supra* note 63, at 25 (“Even with regard to content-neutral regulations of symbolic speech, the *Johnson* claim that courts have a ‘freer hand’ in regulating symbolic speech was in error.”).

68. *Id.* at 31 (“The intermediate test—whether the *O'Brien* test or the essentially interchangeable time, place, and manner test—allows for the careful balancing of the competing interests at stake.”).

expression.”⁶⁹ Similarly, pure speech has its more lenient counterpart in the “time, place, or manner restrictions.” Under *Clark v. Community for Creative Non-Violence*:

*Expression, whether oral or written or symbolized by conduct, is subject to reasonable time, place, or manner restrictions. We have often noted that restrictions of this kind are valid provided that they are justified without reference to the content of the regulated speech, that they are narrowly tailored to serve a significant governmental interest, and that they leave open ample alternative channels for communication of the information.*⁷⁰

The Court has also held that “*O’Brien’s* test . . . ‘is little, if any, different from the standard applied to time, place, or manner restrictions.’”⁷¹ After all, what are time, place, or manner restrictions if not limitations on the non-communicative aspects of pure speech?

By equating the two tests, the Court is implicitly recognizing that both pure and symbolic speech have speech and non-speech characteristics,⁷² and that, therefore, both components of pure and symbolic speech should be assessed under similar frameworks. Thus the distinction between symbolic and pure speech is illusory and, for the purposes of applying First Amendment scrutiny, unnecessary.

In fact, the Supreme Court has recently all but admitted as much. In *Rumsfeld v. Forum for Academic and Institutional Rights, Inc.* (“*FAIR*”), the Court adamantly reminded us that “‘words can in some circumstances violate laws directed not against speech but against conduct.’”⁷³ The Court is quite aware that a distinction between the speech-like and conduct-like aspects of an activity needs to be made in some situations involving the oral or written word itself. Nevertheless the Supreme Court insisted in *FAIR* on maintaining the formal distinction between “speech” and “the expressive nature of the conduct”: “Having rejected the view that the [regulation] impermissibly regulates *speech*, we must still consider whether

69. *United States v. O’Brien*, 391 U.S. 367, 377 (1968).

70. *Clark v. Cmty. for Creative Non-Violence*, 468 U.S. 288, 293 (1984) (emphasis added).

71. *Johnson*, 491 U.S. at 407 (citing *Clark*, 468 U.S. at 298). See also McGoldrick, Jr., *supra* note 63, at 30 (“[T]here is no real difference between the *O’Brien* and *Clark* tests.”).

72. As Professor Kalven stated: “I would suggest that all speech is necessarily ‘speech plus.’ If it is oral, it is noise and may interrupt someone else; if it is written, it may be litter.” Kalven, *supra* note 44, at 23.

73. 547 U.S. 47, 62 (2006) (quoting *R.A.V. v. City of St. Paul*, 505 U.S. 377, 389 (1992)).

the expressive nature of the *conduct* regulated by the statute brings that conduct within the First Amendment's protection."⁷⁴ But on the other hand, in *City of Erie v. Pap's A.M.* a plurality of the Justices interchangeably cited to both *O'Brien* and *Clark* when applying intermediate scrutiny to what it deemed to be a content-neutral regulation of expressive conduct, namely nude erotic dancing.⁷⁵ Once again, regardless of what the Court might say about a formalistic distinction between "pure speech" and "expressive conduct," the tests it applies to the communicative and the non-communicative aspects of both, respectively, are the same.

The Court has simply given us a test—the *Spence-Hurley* test—for deciding when a particular activity is communicative enough to be considered speech for purposes of First Amendment coverage: when, in a particular social context, sufficient conventions exist such that the communication of ideas between people is *possible*, even if not overwhelmingly probable, or even specifically intended. Furthermore, the Court has created a *per se* rule that exempts the oral and written word from this test. However, the level of coverage to which an activity is entitled once it has been deemed communicative enough under the *Spence-Hurley* test does not actually, and should not, depend on a formal distinction between "pure speech" and "expressive conduct."

Dean Post has suggested that what the Court does is extend First Amendment coverage to activities that constitute recognized media for the communication of ideas: "The very concept of a medium presupposes that constitutionally protected expression does not inhere in abstract and disembodied acts of communication of the kind envisioned by *Spence*, but is instead always conveyed through social and material forms of interaction."⁷⁶ The *Spence-Hurley* test, then, provides the tools to determine whether a particular activity is communicative enough to constitute such a recognized medium of expression for First Amendment purposes.⁷⁷

74. *Id.* at 65 (emphasis in original).

75. See 529 U.S. 277, 279 (2000) (plurality opinion).

76. Post, *Recuperating*, *supra* note 1, at 1257.

77. It should be noted, however, that for an activity to receive full coverage under the First Amendment, merely passing the *Spence-Hurley* test does not suffice. The activity must also further First Amendment values without destroying the community that the First Amendment intends to protect. This part of the Court's doctrine excludes other types of activities that, although communicative enough to pass the *Spence-Hurley* test, are not fully covered by the First Amendment. Moreover, passing the *Spence-Hurley* test is not a *sine qua non* requirement for First Amendment coverage, either. Something can be so central to the development of a recognized medium for the communication of ideas that it triggers First Amend-

2. Applying the *Spence-Hurley* Test to Source Code

“[A] programming language is simply a formal language.”⁷⁸ Source code, like any language, uses a predetermined set of conventions to convey messages comprehensible to others who know and understand the rules of the language. Source code text is simply a set of symbols written down on a tangible medium that, within a specified context, can be understood by others. In this sense, source code is very much like the written word that automatically passes the *Spence-Hurley* test.

In fact, as source code is a generic term for a series of established formal languages, it should be considered “pure speech” and deemed to be per se communicative and treated as speech by the First Amendment. In *Bernstein I*, the court agreed and did not even apply the *Spence-Hurley* test to source code because it considered it to be “speech” and “language.”⁷⁹ Similarly, in his article entitled *Publishing Software as a Speech Act*, Lee Tien has argued that “programming languages are languages for First Amendment purposes and source code is, as a doctrinal matter, pure speech.”⁸⁰

Tien, however, advocates for a narrow reading of the *Spence-Hurley* test.⁸¹ In applying speech act theory to the question of First Amendment coverage, Tien explains: “[T]he critical question for coverage purposes is whether the act at issue is *an act of communication*.”⁸² Pursuant to this understanding, “[b]oth the *intent* and social context aspects are necessary to transform an utterance into a speech act.”⁸³

Yet the purpose of the communication, be it in source code or English, is immaterial to First Amendment analysis. Though the Court in *Spence* spoke of “intent to convey a particularized message,”⁸⁴ the actual reception of the message by others does not enter into this first-step analysis. This is the main teaching of *Hurley*: *Hurley* betrays an unwillingness by the Court to delve into the specific subjective intentions of a Jackson Pollock, an Arnold Schoen-

ment scrutiny, even if the activity being regulated is not communicative enough to pass the *Spence-Hurley* test itself.

78. Sebastian Zimmeck, *Patent Eligibility of Programming Languages and Tools*, 13 TUL. J. TECH. & INTELL. PROP. 133, 143 (2010).

79. *Bernstein v. U.S. Dep’t of State (Bernstein I)*, 922 F. Supp. 1426, 1435 (N.D. Cal. 1996), *aff’d*, *Bernstein v. U.S. Dep’t of Justice*, 176 F.3d 1132 (9th Cir.), *withdrawn & reh’g granted*, 192 F.3d 1308 (9th Cir. 1999).

80. Tien, *supra* note 8, at 681.

81. *Id.* at 637.

82. *Id.* (emphasis added).

83. *Id.* (emphasis added).

84. *Spence v. Washington*, 418 U.S. 405, 410–11 (1974).

berg, or a Lewis Carroll when they choose to produce works of art that, at least on the surface, appear to be crafted so as to confuse or hide their meaning from their audience.⁸⁵ The hermetic artist, like the explicit orator, finds shelter under the First Amendment.

Tien's error lies in his overestimation of the importance of communicative values. By positing intentional communication as the sine qua non factor of First Amendment coverage analysis,⁸⁶ Tien's approach becomes both over and under-inclusive: a secret diary meant not to be read would not be covered, while highly communicative, and constitutionally unprotected,⁸⁷ "fighting words" would be covered.⁸⁸ Furthermore, such an approach would not accommodate different levels of First Amendment coverage for different types of communicative acts, which would also be inconsistent with current doctrine.⁸⁹

What makes an activity communicative is its *potential* for communication; this is why even the often perplexing music of Schoenberg passes the *Spence-Hurley* test.⁹⁰ So the fact that source code is mainly written to convey messages to computers instead of people is irrelevant at this stage of the analysis. The pertinent question is whether the activity can communicate a message, not what *type* of message it communicates or *to whom* the message is communicated. Thus communication directed solely to an inanimate diary, which the author intends never to be read by another human being, is still deserving of First Amendment coverage under *Spence-Hurley*. Similarly, an unsuccessful parade should pass the *Spence-Hurley* test just as easily as a successful one. The *Spence-Hurley* test only concerns itself with setting *a threshold probability* that a message will be listened to and understood, and not with the existence of an actual audience. Any concerns over the type of message communicated or its

85. *Hurley v. Irish-American Gay, Lesbian & Bisexual Grp. of Bos.*, 515 U.S. 557, 569 (1995).

86. Tien, *supra* note 8, at 637.

87. *Chaplinsky v. New Hampshire*, 315 U.S. 568, 571–72 (1942).

88. As mentioned above, implicit in *Hurley* is the Court's acknowledgment that other First Amendment values must be considered when determining whether an activity is covered by the First Amendment. These other values are the subject of the second and third steps in the proposed analysis set out in this article and explain these examples of over and under-inclusiveness: the diarist is engaging in an autonomous act of self-expression and maybe even a personal search for truth, so his conduct would be covered, while the utterance of fighting words (and the ensuing acts of aggression) would be subject to severely limited coverage because it threatens to destroy the community served by the First Amendment without substantially furthering any other First Amendment values.

89. *See supra* notes 33–36.

90. *Hurley*, 515 U.S. at 569.

audience, and whether any such communications further First Amendment values enough to merit coverage, are to be resolved in the third step of the analysis proposed in this article.

Under this framework, there is no doubt that source code passes the *Spence-Hurley* test. Many people can write and understand source code. As such, source code can be drafted with “an intent to convey a particularized message” and “the likelihood [is] great that the message would be understood.”⁹¹ The fact that the language used to convey such messages is not English, or that others might not understand it, is irrelevant.⁹² This is why the court held in *Bernstein I* that there is “no meaningful difference between computer language, particularly high-level languages as defined above, and German or French. All participate in a complex system of understood meanings within specific communities.”⁹³

Furthermore, the fact that a computer can understand source code does not figure into the *Spence-Hurley* inquiry. It is irrational to suggest that because a computer can be designed and constructed to understand and execute commands in English, the whole English language ceases to be covered by the First Amendment.⁹⁴ This observation applies equally to high-level computer languages.

91. *Spence v. Washington*, 418 U.S. 405, 410–11 (1974).

92. *Yniguez v. Arizonans for Official English*, 69 F.3d 920, 933–34 (9th Cir. 1995) (en banc), *vacated as moot sub nom.*, *Arizonans for Official English v. Arizona*, 520 U.S. 43 (1997).

93. *Bernstein v. U.S. Dep’t of State (Bernstein I)*, 922 F. Supp. 1426, 1435 (N.D. Cal. 1996), *aff’d*, *Bernstein v. U.S. Dep’t of Justice*, 176 F.3d 1132 (9th Cir.), *withdrawn & reh’g granted*, 192 F.3d 1308 (9th Cir. 1999).

94. This hypothetical, of course, is no longer just in the realm of science fiction. Apple’s inclusion of Siri in the iPhone 4S has brought to the mainstream the practice of ordering a computer to do things in natural language. See *Apple - iPhone 4S - Ask Siri to help you get things done*, APPLE, <http://www.apple.com/ios/siri/> (last visited Nov. 11, 2012). Siri understands commands in English, French, German, and Japanese. *Apple - Siri - Frequently Asked Questions*, APPLE, <http://www.apple.com/ios/siri/siri-faq/> (last visited Nov. 11, 2012). Of course, Siri is not alone. The world of computers responsive to natural language commands is one of the many areas of technological endeavor that are currently experiencing considerable growth. See Natasha Singer, *The Human Voice, as Game Changer*, N.Y. TIMES (Mar. 31, 2012), <http://www.nytimes.com/2012/04/01/technology/nuance-communications-wants-a-world-of-voice-recognition.html> (discussing a series of recent developments and ongoing projects in voice operated machines). Furthermore, the existence of Siri and other computers responsive to natural language commands cannot condemn the English language (or any other language for that matter) to the netherworlds of First Amendment invisibility. The court in *Bernstein II* was already wise to this logic years before the advent of Siri: “The fact that computers will soon be able to respond directly to spoken commands, for example, should not confer on the government the unfettered power to impose prior restraints on speech in an effort to control its ‘functional’ aspects.” *Bernstein v. U.S. Dep’t of*

Lastly, it is important to note that source code might be a better medium for expressing ideas about computer science than traditional language.⁹⁵ “[P]rogramming languages avoid the difficulties that English has in describing algorithms and may stand as the only practical means of expressing certain algorithms that require precise articulation. Programming languages provide the best means for communicating highly technical ideas—such as mathematical concepts—within the community of computer scientists and programmers.”⁹⁶ This is why “[t]he First Amendment mandates that we presume that speakers, not the government, know best both what they want to say and how to say it.”⁹⁷ Programmers should, therefore, be able to choose to speak in code instead of English, as part of the exercise of their First Amendment rights.

B. *Step 2: Is the Activity Central to the Development of a Medium for the Communication of Ideas?*

1. Of Movie Projectors, Printing Presses, and Newspaper Racks

Even if source code is not deemed communicative enough under the *Spence-Hurley* test, its regulation still triggers First Amendment scrutiny. This is because “First Amendment coverage is not limited to speech acts. It extends to forms of interaction that realize First Amendment values.”⁹⁸ These forms of interaction are often designated as media for the communication of ideas.⁹⁹ In *Joseph Burstyn, Inc. v. Wilson*, for example, the Supreme Court held that “motion pictures are a significant medium for the communication of ideas.”¹⁰⁰

Most importantly for our present discussion, though, is the fact that the First Amendment extends its coverage over attempts to regulate activity, and even material things, that are central to the development of these media.¹⁰¹ Dean Post explains:

The genre of the cinema . . . encompasses far more than speech acts. It includes materials, like celluloid; functional ma-

Justice (*Bernstein II*), 176 F.3d 1132, 1142 (9th Cir.), *withdrawn & reh'g granted*, 192 F.3d 1308 (9th Cir. 1999).

95. See *Bernstein II*, 176 F.3d at 1141.

96. Tien, *supra* note 8, at 662–63 (footnote omitted) (citing DONALD KNUTH, *THE ART OF COMPUTER PROGRAMMING: FUNDAMENTAL ALGORITHMS* 5 (1st ed. 1968)).

97. *Riley v. Nat'l Fed'n of Blind of N.C., Inc.*, 487 U.S. 781, 790–91 (1988).

98. Post, *Encryption*, *supra* note 1, at 716.

99. *Id.*

100. 343 U.S. 495, 501 (1952); see also Post, *Encryption*, *supra* note 1, at 716 (quoting 343 U.S. at 501).

101. Post, *Encryption*, *supra* note 1, at 717.

chines, like projectors; buildings, like movie houses; social organizations, like studios; and so forth. If the state were to prohibit the use of projectors without a license, First Amendment coverage would undoubtedly be triggered. This is not because projectors constitute speech acts, but because they are integral to the forms of interaction that comprise the genre of the cinema.¹⁰²

First Amendment scrutiny might be triggered by these kinds of regulation even if they do not involve viewpoint discrimination:

An obvious instance might be a law that prohibits newsprint in order to save trees. Newsprint is a material necessary for the publication of most newspapers. Although a law proscribing newsprint would be viewpoint (and content) neutral, it would carry the potential for so significantly affecting the First Amendment medium of newspapers that we would certainly review it under First Amendment principles. We would want to assure ourselves that it would not compromise the constitutional value we attribute to newspapers.¹⁰³

Along these lines, the Sixth Circuit has held that “something as mundane as a newspaper rack might fall into the category of speech-facilitating devices,” and might trigger First Amendment scrutiny.¹⁰⁴ And the Supreme Court has held that the First Amendment prohibition on prior restraints extends to “expression or *conduct commonly associated with expression*.”¹⁰⁵

The Internet should present a perfectly analogous situation. As early as 1997, the Supreme Court held the Internet to be a medium for the communication of ideas, and that case law “provide[s] no basis for qualifying the level of First Amendment scrutiny that should be applied to [the Internet].”¹⁰⁶ Therefore, just as the newspaper’s protection extends to newsstands, any activity, mechanism, or object essential to the free use or development of the Internet as a medium for the communication of ideas will trigger First Amendment scrutiny. The following section will explore this analogy and

102. *Id.*

103. *Id.* at 721–22.

104. Burk, *supra* note 8, at 115 (citing Plain Dealer Publ’g Co. v. City of Lakewood, 794 F.2d 1139, 1143 (6th Cir. 1986) (“The right to distribute newspapers by means of newsracks is protected by the First Amendment . . .”), *aff’d sub nom.*, City of Lakewood v. Plain Dealer Publ’g Co., 486 U.S. 750 (1988)).

105. City of Lakewood v. Plain Dealer Publishing Co., 486 U.S. 750, 760 (1988) (emphasis added).

106. Reno v. ACLU, 521 U.S. 844, 870 (1997).

demonstrate how source code is the newspaper rack of the Internet era.

2. Source Code Is Central to the Development of the Internet

The code of the Internet is central to what cyberspace is. “The code of cyberspace—whether the Internet, or a net within the Internet—defines that space. It constitutes that space. And as with any constitution, it builds within itself a set of values and possibilities that governs life there.”¹⁰⁷ In this way, Professors Joel R. Reidenberg and Lawrence Lessig have convincingly argued that code is effectively a “*lex informatica*”:¹⁰⁸ a set of rules or laws that “defines what behavior is possible in cyberspace and what values cyberspace will uphold.”¹⁰⁹ In other words, the law of cyberspace *is* its source code.

Given this insight, it is now generally accepted that effective regulation of the Internet will happen through regulation of the code that constitutes it.¹¹⁰ “Laws would have their effect, if only indirectly, by inducing changes in the *lex [informatica]*.”¹¹¹ “Smart governments will instead regulate by regulating the *code* that regulates the behavior of people in cyberspace.”¹¹² They will regulate the code in such a way as “to assure that cyberspace is architected in a way to protect government’s interests.”¹¹³

If government can regulate the code, then government can require codewriters to build the standards that the government

107. Lawrence Lessig, The Charles Green Lecture, *Open Code and Open Societies: Values of Internet Governance*, 74 CHI.-KENT L. REV. 1405, 1408 (1999) [hereinafter Lessig, *Open Code*].

108. Joel R. Reidenberg, *Lex Informatica: The Formulation of Information Policy Rules Through Technology*, 76 TEX. L. REV. 553 (1998) [hereinafter Reidenberg, *Lex Informatica*]. See also Joel R. Reidenberg, *Governing Networks and Rule-Making in Cyberspace*, 45 EMORY L.J. 911, 929 (1996).

109. Lawrence Lessig, *The Limits in Open Code: Regulatory Standards and the Future of the Net*, 14 BERKELEY TECH. L.J. 759, 761 (1999) [hereinafter Lessig, *Limits*] (citing Reidenberg, *Lex Informatica*, *supra* note 109, at 568–73).

110. *Id.* at 762. But cf. Viktor Mayer-Schönberger, *Demystifying Lessig*, 2008 WIS. L. REV. 713, 713 (2008) (offering “a fundamental critique of Lessig’s core argument [that ‘code is law’]—and thus of core assumptions of cyberlaw scholarship”). Professor Mayer-Schönberger claims that the weaknesses he points out in Lessig’s argument “limit [] the capacity of his theory to adequately capture the full dynamic at play in free speech.” *Id.* at 746. In fact, his claim further supports this article’s conclusion that all the First Amendment values discussed should be taken into account in order to arrive at a coherent theory of First Amendment coverage. See *id.* Truth is not enough.

111. Lessig, *Limits*, *supra* note 109, at 762.

112. *Id.* at 763 (emphasis added).

113. *Id.*

needs into the code. The future of regulatory standards under this view, then, would simply be a future where the government tells codewriters how to architect their code so as to incorporate governmental regulatory standards.¹¹⁴

By regulating the code, the government can effectively reshape the medium for the communication of ideas that is constituted by that code. “Different code, different regulation, different worlds,”¹¹⁵ different medium for the communication of ideas.

Taking up this argument, Ryan Christopher Fox provides us with two concrete examples of how the regulation of source code has indirectly regulated conduct: DeCSS and encryption.

DeCSS . . . allows users to bypass security controls on DVDs. Because it allows for the copying of what is supposed to be uncopyable media, the DeCSS code has been attacked in the courts. Unlike the Napster litigation, though, in which legal claims were based on traditional copyright law and the code was only implicated by the facts of the case, *DeCSS was attacked under laws regulating the distribution of a specific class of computer code*. Another example of code that some might feel is dangerous is that used in software designed to encrypt data to prevent its being read by undesired individuals. Like DeCSS, *distribution of certain pieces of encryption software have been regulated by law that focuses specifically on computer code*, rather than on any illegal activities that might be performed using the code.¹¹⁶

Government attempts to regulate DeCSS and encryption “demonstrate a relatively new focus on prohibiting or restricting code itself as a dangerous tool rather than relying on laws against the undesirable activity that the code facilitates.”¹¹⁷

Professor Lessig argues that the Internet’s amenability to regulation will, in the first place, depend on how “open” the source code is.¹¹⁸ Open source code is code that is available for all to see, read, modify, and improve. “Open code is software in plain view. It is software that comes bundled with its source code as well as its object code.”¹¹⁹ As we have already discussed, programmers and computer scientists can read source code and quickly recognize any

114. *Id.* at 764.

115. *Id.* at 762.

116. Fox, *supra* note 7, at 874 (emphasis added) (footnotes omitted).

117. *Id.*

118. Lessig, *Limits*, *supra* note 110, at 764.

119. *Id.*

“controls” the code may hide.¹²⁰ Hence the more a code is open to review, the less amenable the code is to regulation.¹²¹ “[T]o the extent that code remains open, it is harder for government to regulate; to the extent it is closed, it is easier.”¹²²

Once source code is viewed as central to the development of a medium for the communication of ideas, its regulation necessarily implicates the First Amendment. Professor Lessig points out two important problems that arise from the government’s regulation of the Internet through code—problems with First Amendment implications and solutions.¹²³ First, regulation of code raises issues of over-inclusiveness.¹²⁴ This becomes apparent once we realize that the regulation of code tends to divorce the regulatory technique from the underlying purpose of the regulation. In doing so, regulating the code can allow the government to extend its grasp more than the protection of the values that it seeks to further through such regulation would otherwise require.¹²⁵ In other words, by separating the means from the ultimate end, additional activity is indirectly regulated through control of the code without regard for the balance of interests that direct regulation of such activity might otherwise consider.¹²⁶ Professor Lessig ably illustrates this phenomenon through two examples: (1) the criminalization of fair use through the anti-circumvention provisions of the Digital Millennium Copyright Act (which over-regulates conduct that would not otherwise be illegal under the underlying copyright law); and (2) a possible requirement of digital IDs as a way of zoning of speech in cyberspace.¹²⁷ Regulation of code as a means to regulate an underlying problem will therefore tend to be over-inclusive.¹²⁸ This will lead to the over-regulation of the use of the Internet, a medium for the communication of ideas, and thus an over-regulation of expression. By extending First Amendment coverage to source code, this result can be avoided. Strict scrutiny of such regulation would ensure that legislation is narrowly tailored.

120. *Id.* (“It is this code that allows a programmer to open an open source software project and see what makes it tick. By being able to see what makes it tick, open source software makes transparent any control that the code might carry.”).

121. *Id.*

122. *Id.* at 767.

123. Lawrence Lessig, *The Law of the Horse: What Cyberlaw Might Teach*, 113 HARV. L. REV. 501, 536 (1999) [hereinafter Lessig, *Law of the Horse*].

124. *Id.* at 536–37.

125. *Id.*

126. *Id.* at 537.

127. *Id.* at 537–39.

128. *Id.*

The second big problem with Internet regulation through code is that such regulation might not be transparent.¹²⁹ In general, we like government regulation to be as transparent as possible; we like to know how the government is controlling us and why. That way, if we do not like what the government is doing, we can vote it out of office. Since regulation of code can have effects on cyberspace that are indirect and hidden, it can threaten our democratic values. The regulation of code allows the government to “[e]nslave the code while telling the world that [it is] leaving the space free.”¹³⁰ Again, extending First Amendment coverage to source code can solve this problem. The application of strict scrutiny to the regulation of the underlying code would “‘smoke out’ illegitimate”¹³¹ governmental interests being furthered by such regulation. In applying strict scrutiny, courts would demand that the government justify the regulation of the code on the basis of the true underlying interests pursued and the relationship between the means employed and those interests.

At this juncture, an example of how the regulation of code can compromise First Amendment values might be helpful. The litigations in *Bernstein*, *Karn*, and *Junger* all involved the regulation of encryption source code.¹³² Though government regulation of encryption source code may not appear, at first glance, to directly limit covered speech, such a regulation would chill a great deal of First Amendment protected speech transmitted across the Internet. “Tien offers the valuable suggestion that encryption software might be conceived as providing the equivalent of envelopes which protect the privacy of underlying digital messages.”¹³³ A law that strips Internet speech of its privacy “would certainly merit First Amendment coverage.”¹³⁴

This is because the chill on participating in a First Amendment medium that comes from exposure is a well-recognized First Amendment interest. Laws prohibiting anonymous political leaflets have thus been struck down because of their potential impact on speakers.

129. *Id.* at 539.

130. Lessig, *Limits*, *supra* note 109, at 763.

131. *City of Richmond v. J.A. Croson Co.*, 488 U.S. 469, 493 (1989).

132. *See* cases cited *supra* note 9.

133. Post, *Encryption*, *supra* note 1, at 723 (citing Tien, *supra* note 8, at 672).

134. *Id.*

Encryption software is a way of preventing an analogous chill within digital media.¹³⁵

For this reason, the court in *Bernstein II* recognized in dicta that “the government’s efforts to retard progress in cryptography may implicate the Fourth Amendment, as well as the right to speak anonymously, the right against compelled speech, and the right to informational privacy.”¹³⁶

So even if source code is not considered communicative enough to constitute “speech” under the *Spence-Hurley* test, its regulation should still trigger First Amendment scrutiny. Code is at the very heart of the development of the Internet, a recognized medium for the communication of ideas. The regulation of source code, therefore, is nothing but an indirect way of regulating the Internet itself, a recognized medium for the communication of ideas. As such, its regulation can compromise the freedom of the communicative medium that it serves to create and shape. Consequently, even if source code is deemed not communicative itself under the *Spence-Hurley* test, it must still be treated to the same First Amendment coverage because source code is inextricably intertwined with a recognized medium for the communication of ideas.

However, our analysis cannot end here.

C. Step 3: Does the Activity Promote First Amendment Values?

1. Theories of First Amendment Values

The Supreme Court has delineated types of activities that, while passing steps one or two of this analysis, do not merit the full force of First Amendment coverage. For example, we know that certain uses of the written or oral word can trigger only a very limited level of First Amendment coverage when they are deemed to be “fighting words,”¹³⁷ obscene,¹³⁸ or criminal solicitations.¹³⁹ The question therefore becomes: how do we know which activities merit full First Amendment coverage?

135. *Id.* (citing *McIntyre v. Ohio Elections Comm’n*, 514 U.S. 334 (1995); *Talley v. California*, 362 U.S. 60, 60 (1960); *NAACP v. Alabama ex rel. Patterson*, 357 U.S. 449, 460–64 (1958)).

136. *Bernstein v. U.S. Dep’t of Justice (Bernstein II)*, 176 F.3d 1132, 1146 (9th Cir.) (citations omitted), *withdrawn & reh’g granted*, 192 F.3d 1308 (9th Cir. 1999).

137. *Chaplinsky v. New Hampshire*, 315 U.S. 568, 571–72 (1942).

138. *Miller v. California*, 413 U.S. 15, 23 (1973).

139. *United States v. Williams*, 553 U.S. 285, 297–98 (2008).

Tien proposes that solving this coverage problem does not require an overarching analysis of First Amendment values.¹⁴⁰ He believes that, “[w]hile many consider the Court’s coverage doctrine to be incoherent, it can be harmonized without appealing to a grand theoretical framework of First Amendment values.”¹⁴¹ Instead, Tien suggests that “coverage issues primarily raise practical problems about whether someone is speaking.”¹⁴² Still, Tien’s approach does not bypass the value question; rather, it provides a one-word answer: communication. Tien seems to argue that the Court’s coverage doctrine can be understood by reference to a single, albeit sophisticated, understanding of the communicative value of “speech acts.” “Under this theory, the critical question for coverage purposes is whether the act at issue is an act of communication.”¹⁴³

However, the Court’s opinions do not support this approach. Tien is correct that a central part of the coverage analysis involves determining an activity’s communicative value; that is what the *Spence-Hurley* test is designed to measure. However, this test fails to account for First Amendment coverage of simple objects related to media, such as movie projectors or printing presses, or the limited coverage given to highly communicative “fighting words” or commercial speech.

Furthermore, Tien’s approach fails to take into account First Amendment coverage of certain important subgenres of human activity. The main reason for this under-inclusiveness is his insistence upon the importance of “illocutionary intent”: under Tien’s approach, it is “the speaker’s intent to perform a speech act,” coupled with her intent that her act be understood by her audience, that “transforms an utterance act—like making noise—into a speech act.”¹⁴⁴ As Tien himself admits, “This approach would exclude, for example, the concept of found meaning, which bears no relation to the speaker’s intended utterance.”¹⁴⁵ But the First Amendment cannot be completely deaf to the concept of found meaning. The First

140. See Tien, *supra* note 8, at 636 (noting that the Supreme Court “hasn’t seemed to need such a theory to decide coverage issues”).

141. *Id.*

142. *Id.* at 637. It should be noted that Tien himself acknowledges that concentrating the coverage question upon a definition of “speaking” is a normative choice itself. *Id.* at 637 n.27. About this criticism, Tien explains: “[M]y constraints are based on ‘speech acts’ as normative social phenomena, not on a full-blown theory about First Amendment values. Put another way, I begin with communication and then freedom of speech, not the other way around.” *Id.*

143. *Id.* at 637.

144. *Id.* at 640.

145. *Id.* at 651.

Amendment does not only protect the speaker, but speech itself.¹⁴⁶ This is particularly relevant when we consider the values of truth and democracy, which may benefit from speech even when the speaker might not have intended it for anybody's ears. Think of the great benefit that has accrued to humankind from the publishing of Franz Kafka's final works and of Virgil's *Aeneid*, both of which were published against their authors' last wills and testaments.¹⁴⁷ For these reasons, the Court has given important notice to "the 'inherent worth of the speech' and 'its capacity for informing the public.'"¹⁴⁸ Similarly, the Court has completely disregarded the intent of those who may have substantially contributed economically to the production of certain speech.¹⁴⁹ Finally, Tien's position that "the relevant intent is the speaker's intent that the hearer understand the act as a speech act"¹⁵⁰ is also inconsistent with the Court's pronouncement in *Hurley* that the First Amendment reaches "the unquestionably shielded painting of Jackson Pollock, music of Arnold Schoenberg, or Jabberwocky verse of Lewis Carroll."¹⁵¹ Even though many people might not fully understand their underlying messages, and regardless of whether the cited artists had the intent to illuminate or obscure those messages, the Court has unequivocally found these works to be protected.

If we are to coherently explain the Court's coverage doctrine, more than communicative value must be found in the activities covered by the First Amendment. For the purposes of this discussion, three traditional theories of First Amendment values will be refer-

146. As Justice Scalia has so poignantly remarked: "But to return to, and summarize, my principal point, which is the conformity of today's opinion with the original meaning of the First Amendment. The Amendment is written in terms of 'speech,' not speakers." *Citizens United v. Fed. Election Comm'n*, 130 S. Ct. 876, 929 (2010) (Scalia, J., concurring). However, the First Amendment is not *only* concerned with speech. Instead, First Amendment doctrine must take into account the particularities of *both* the speaker *and* her speech.

147. Lior Jacob Strahilevitz, *The Right to Destroy*, 114 YALE L.J. 781, 830–31 (2005). See also M.A. Orthofer, *Weighing Words Over Last Wishes*, POETS & WRITERS (Nov. 1, 2003), http://www.pw.org/content/weighing_words_over_last_wishes. But see Strahilevitz, *supra*, at 831, 831 n.202 (noting that accounts of Virgil's final intentions for the *Aeneid* differ).

148. *Citizens United*, 130 S. Ct. at 929 (Scalia, J., concurring) (quoting *First Nat'l Bank of Bos. v. Bellotti*, 435 U.S. 765, 777 (1978)).

149. See *id.* at 905 ("The First Amendment protects the resulting speech, even if it was enabled by economic transactions with persons or entities who disagree with the speaker's ideas.") (majority opinion).

150. Tien, *supra* note 8, at 651.

151. *Hurley v. Irish-American Gay, Lesbian & Bisexual Grp. of Bos.*, 515 U.S. 557, 569 (1995).

enced: truth, democracy, and autonomy.¹⁵² The courts also frequently consider a fourth, sometimes competing, value: community. The exercise of free expression must, in some extreme and well-delineated cases, be checked by the need to protect our community from self-destruction, violence, and the annihilation of the channels of communication themselves.

The Court's seemingly inconsistent First Amendment coverage doctrine can only be explained by deciphering which values the Court promotes by extending coverage to some types of activities and not others.¹⁵³ Drawing heavily on the work of others but also adding some nuance, this article proposes that the Court has read into the First Amendment a very specific set of values that it believes are central to the First Amendment's intent and *raison d'être*: (1) truth; (2) democracy; (3) autonomy; and (4) community.¹⁵⁴

a. Truth

Justice Holmes introduced us, in his celebrated dissent in *Abrams v. United States*,¹⁵⁵ to one of the most important theories be-

152. This selection, of course, is not random. In fact, it has been profoundly influenced by both the Supreme Court's language in myriad cases and by Dean Post's previous work on this subject: "There are presently three major candidates for such values: (1) the creation of new knowledge; (2) individual autonomy; and (3) democratic self-government." Robert Post, *Participatory Democracy and Free Speech*, 97 VA. L. REV. 477, 478 (2011) [hereinafter Post, *Participatory Democracy*].

153. Dean Post has argued that "[i]t makes sense to conclude that the scope of the First Amendment extends only to those forms of speech (or regulation) that implicate constitutional values." *Id.*

154. The brutally normative question of exactly which values should make such a privileged list, of course, is fraught with peril and controversy. In fact, some commentators have suggested that no unifying theory of First Amendment values can easily be formed:

There have been many attempts to justify First Amendment coverage in terms of a free speech principle. Our "standard list of candidates" derives from the various theories of free speech: the pursuit of truth, self-governance, the "checking value" of free speech, and so on. But there seems to be no discernible general principle here, for these various theories don't cohere particularly well.

Tien, *supra* note 8, at 636 (citing Schauer, *supra* note 22, at 267-68; Post, *Recuperating*, *supra* note 1, at 1270). However, that a perfectly consistent unified theory of the values behind the First Amendment might be hard or even impossible to construct is not a good enough reason to stop trying to make an accurate approximation of what such a theoretical framework should look like. Judges will continue to make decisions that reflect their own subjective judgments about what the First Amendment is meant to protect. Denying this is at best naïve, and at worst disingenuous.

155. 250 U.S. 616, 624 (1919) (Holmes, J., dissenting).

hind First Amendment doctrine: the marketplace of ideas.¹⁵⁶ As Justice Holmes eloquently put it:

[W]hen men have realized that time has upset many fighting faiths, they may come to believe even more than they believe the very foundations of their own conduct that the ultimate good desired is better reached by free trade in ideas—that the best test of truth is the power of the thought to get itself accepted in the competition of the market, and that truth is the only ground upon which their wishes safely can be carried out. That at any rate is the theory of our Constitution. It is an experiment, as all life is an experiment.¹⁵⁷

Truth can best be obtained through an open and unencumbered dialogue in which all parties are free to express their opinions. “Factions should be checked by permitting them all to speak, and by entrusting the people to judge what is true and what is false.”¹⁵⁸ If something is true, rational people will be convinced, and that belief will be recognized and incorporated by society. “[The] theory of [the] market-place of ideas ‘is essentially the method of science,’ which seeks ‘progress through free and rational inquiry.’”¹⁵⁹ The Supreme Court has recently reaffirmed the idea that there is in our society an “‘open marketplace’ of ideas protected by the First Amendment,” and that “ideas ‘may compete’ in this marketplace ‘without government interference.’”¹⁶⁰

This theory, however, is only concerned with attaining truth, and is premised on a capitalist notion of information flow and of rational human behavior. It is, in many ways, the application of the scientific method to public discourse at large. It is central to the current social understanding of what the United States of America wants to be. Nonetheless, this theory serves truth as its principal master, or at least the search for truth in dialectical form. If we are to value other things, or if we are interested in a more cooperative, less competitive, search for truth—or if we are skeptical of the existence of any such singular Truth—¹⁶¹ we must consider other theo-

156. See also Post, *Participatory Democracy*, *supra* note 152, at 478.

157. *Abrams*, 250 U.S. at 630 (Holmes, J., dissenting).

158. *Citizens United v. Fed. Election Comm’n*, 130 S. Ct. 876, 907 (2010) (citation omitted).

159. Tien, *supra* note 8, at 664 n.145 (quoting Thomas Emerson, *Colonial Intentions and Current Realities of the First Amendment*, 125 U. PA. L. REV. 737, 741 (1977)).

160. *Citizens United*, 130 S. Ct. at 906 (quoting *New York State Bd. of Elections v. López Torres*, 552 U.S. 196, 208 (2008)).

161. See Post, *Participatory Democracy*, *supra* note 152, at 479 (emphasizing that truth cannot be the only value favored by the First Amendment, as “[t]he First

ries and values as well: values that are just as central to the spirit of the United States Constitution, and just as necessary for its long term survival and prosperity.

b. Democracy

Many commentators have emphasized the importance of democratic self-governance as an instructive value in First Amendment doctrine.¹⁶² However, democracy is itself served by the First Amendment through two models of self-governance: (1) the participatory model and (2) the Meiklejohnian model.¹⁶³

“The participatory model emphasizes the importance of preserving uncensored access to public discourse so that citizens can maintain the warranted sense that their government is responsive to them.”¹⁶⁴ As Dean Post has recently pointed out:

The value of democratic legitimation occurs, as Habermas and many others have theorized, specifically through processes of communication in the public sphere. It requires that citizens have access to the public sphere so that they can participate in the formation of public opinion, and it requires that governmental decision making be somehow rendered accountable to public opinion.¹⁶⁵

Hence, under this theory, the First Amendment should cover “those speech acts and media of communication that are socially regarded as necessary and proper means of participating in the formation of public opinion.”¹⁶⁶ If you let people engage in public discourse and get involved in the political life of their state, then they will feel as if they are a part of the body politic and will respect

Amendment recognizes no such thing as a ‘false’ idea” (quoting *Hustler Magazine v. Falwell*, 485 U.S. 46, 51 (1988)).

162. See, e.g., *id.* at 482; Vincent Blasi, *The Checking Value in First Amendment Theory*, 1977 AM. B. FOUND. RES. J. 521, 555; Robert Bork, *Neutral Principles and Some First Amendment Problems*, 47 IND. L.J. 1, 28 (1971).

163. See Post, *Participatory Democracy*, *supra* note 152, at 482; Robert Post, *The Constitutional Status of Commercial Speech*, 48 UCLA L. REV. 1, 5 (2000) [hereinafter Post, *Commercial Speech*].

164. Post, *Commercial Speech*, *supra* note 163, at 12.

165. Post, *Participatory Democracy*, *supra* note 152, at 482 (footnote omitted) (citing JÜRGEN HABERMAS, *BETWEEN FACTS AND NORMS: CONTRIBUTIONS TO A DISCOURSE THEORY OF LAW AND DEMOCRACY* app. I, at 472–77, 486–90 (William Rehg trans., Polity Press 1996) (1992)).

166. *Id.* at 483.

and feel comfortable with its actions.¹⁶⁷ This is a necessary condition for a healthy democracy.¹⁶⁸

On the other hand, Professor Alexander Meiklejohn has “famously argued that ‘the final aim’ of First Amendment freedom is to ensure the circulation of opinion and information necessary for ‘the voting of wise decisions.’”¹⁶⁹ Under this model, “What is essential is not that everyone shall speak, but that everything worth saying shall be said.”¹⁷⁰ People need as much information as possible so that when they vote they can make the best-informed, most intelligent, democratic decisions.¹⁷¹ Thus the First Amendment should primarily serve to guarantee access to information. “The right of citizens to inquire, to hear, to speak, and to use information to reach consensus is a precondition to enlightened self-government and a necessary means to protect it.”¹⁷² In this way, the First Amendment serves, as it does in the participatory model, to further the democratic values that the First Amendment was designed to protect.

However, an interest in promoting democratic values does not explain all of the Court’s First Amendment opinions. Even those who believe most ardently that democratic self-governance is the principal value furthered by the First Amendment recognize the need for additional considerations.¹⁷³ For example, the values of truth and democracy, taken together, still fail to account for cases such as *Stanley v. Georgia*, where the Supreme Court held that criminal prosecution for mere private possession of obscene materials was prohibited by the Constitution.¹⁷⁴ In this sense, it is important to remember that materials will only qualify as obscene if they lack

167. See *Citizens United v. Fed. Election Comm’n*, 130 S. Ct. 876, 898 (2010) (“Speech is an essential mechanism of democracy, for it is the means to hold officials accountable to the people.”).

168. Cf. *Buckley v. Valeo*, 424 U.S. 1, 14 (1976) (“Discussion of public issues and debate on the qualifications of candidates are integral to the operation of the system of government established by our Constitution.”).

169. Post, *Commercial Speech*, *supra* note 163, at 13 (quoting ALEXANDER MEIKLEJOHN, *POLITICAL FREEDOM: THE CONSTITUTIONAL POWERS OF THE PEOPLE* 26 (1960)).

170. MEIKLEJOHN, *supra* note 169, at 26.

171. Post, *Commercial Speech*, *supra* note 163, at 13.

172. *Citizens United*, 130 S. Ct. at 898.

173. See, e.g., Post, *Participatory Democracy*, *supra* note 152, at 488 (“I do not contend that the value of democratic self-governance can explain all First Amendment decisions.”).

174. 394 U.S. 557, 559 (1969). See also Post, *Participatory Democracy*, *supra* note 149, at 488 (“There are no doubt some decisions, like *Stanley v. Georgia*, that can be explained only by reference to the value of autonomy.”).

“serious literary, artistic, political, or scientific value.”¹⁷⁵ Hence any First Amendment coverage of obscene materials must be justified by values other than truth or democracy.

c. Autonomy

The First Amendment is also concerned with individual autonomy. The Supreme Court, as well as commentators, has made this point abundantly clear: “One fundamental concern of the First Amendment is to ‘protec[t] the individual’s interest in self-expression.’ Freedom of speech helps ‘make men free to develop their faculties,’ it respects their ‘dignity and choice,’ and it facilitates the value of ‘individual self-realization.’”¹⁷⁶ People must be allowed to express their own individuality. Human beings grow and learn about themselves and others through the exercise of their creativity. They must be free to try to reach their highest potential through their individual self-realization.¹⁷⁷ “This is often formulated as the constitutional value of autonomy, which is sometimes referred to as ‘self-fulfillment’ or ‘self-expression.’”¹⁷⁸ “There is no doubt that this form of liberal autonomy has deep roots in American constitutionalism, and it is clear that its influence can be detected in First Amendment doctrine.”¹⁷⁹

The First Amendment is at the forefront of the protections that allow us the freedom to create new worlds through the use of our creative expressions. Our society is premised on the idea that we are free to think whatever we want. “If the First Amendment means anything, it means that a State has no business telling a man, sitting alone in his own house, what books he may read or what films he may watch.”¹⁸⁰ Our thoughts cannot be controlled, because we are the only legitimate monarchs of our own minds. “Our whole constitutional heritage rebels at the thought of giving government the power to control men’s minds.”¹⁸¹ Expression is sufficiently close to

175. *Miller v. California*, 413 U.S. 15, 24 (1973).

176. *Citizens United*, 130 S. Ct. at 972 (Stevens, J., concurring) (citations omitted) (citing *Consol. Edison Co. of N.Y. v. Pub. Serv. Comm’n of N. Y.*, 447 U.S. 530, 534 n.2 (1980); *First Nat’l Bank of Bos. v. Bellotti*, 435 U.S. 765, 777 n.12 (1978); *Whitney v. California*, 274 U.S. 357, 375 (1927) (Brandeis, J., concurring); *Cohen v. California*, 403 U.S. 15, 24 (1971); Martin H. Redish, *The Value of Free Speech*, 130 U. PA. L. REV. 591, 594 (1982)).

177. See Tien, *supra* note 8, at 636 (stating that individual self-realization is the most general candidate for a free speech principle).

178. Post, *Participatory Democracy*, *supra* note 152, at 479.

179. *Id.*

180. *Stanley v. Georgia*, 394 U.S. 557, 565 (1969).

181. *Id.*

thought as to require the strictest protection of its free exercise if our autonomy is to be preserved. If we allow government to regulate what we can and cannot say, it would necessarily relinquish our control over what we can and cannot think. This type of mind control is unacceptable under the First Amendment.

But again, the value of autonomy, by itself, appears insufficient to explain First Amendment doctrine. A common attack on autonomy as a basis for First Amendment coverage is that all kinds of human activity, not just speech, can be seen as furthering self-realization.¹⁸² “[T]he value of autonomy extends not merely to the speech of persons but also to the actions of persons. This suggests that the value of autonomy is not unique to speech but instead extends to the full libertarian protection of personal action.”¹⁸³ As Professor Bork points out, “[One] cannot, on neutral grounds, choose to protect speech on this basis more than [one] protects any other claimed freedom.”¹⁸⁴ This argument, however, makes light of the fact that a court need not itself choose to protect speech over other human activities. The First Amendment already makes this choice. Self-realization through speech was specifically set apart by the Constitution for special protection. Thus the close relationship identified by the Supreme Court between speech, language, and thought must be seen as an important factor in this choice.

Another valid criticism to the autonomy rationale is that “there are many situations in which the autonomy of a speaker conflicts with the autonomy of an audience.”¹⁸⁵ Hence the value of autonomy does not help differentiate between certain types of speech that might receive limited coverage under the First Amendment, such as fighting words, defamation, or obscenity.¹⁸⁶ Any explanation of this limited coverage must rely on other First Amendment values, such as the final value proposed of community.

d. Community

The preceding theories of First Amendment values must be tempered by a reality often underestimated in the First Amendment context: the need to preserve and foster a sense of community. By allowing activities that further the aforementioned values of

182. See, e.g., Bork, *supra* note 162, at 25 (arguing that the autonomy rationale—the “development of individual faculties” rationale—fails to “distinguish speech from any other human activity”).

183. Post, *Participatory Democracy*, *supra* note 152, at 479.

184. Bork, *supra* note 162, at 25.

185. Post, *Participatory Democracy*, *supra* note 152, at 480.

186. *Id.*

truth, democracy, and autonomy, we often make it acceptable for people to express ideas that are in stark contrast with what the greater part of the community believes. This inevitably creates tension between different parts of a community. This tension must be kept within a certain limit so that the community itself is not completely destroyed.

More importantly, increased intra-community tensions might destroy the channels of free communication themselves. For example, Dean Post has remarked that, typically, “[f]reedom of thought is transmuted into new knowledge only when it is integrated into those forms of social practices” that “depend upon positive intellectual virtues like respect, reason, fairness, accuracy, integrity, honesty, logic, and civility.”¹⁸⁷ Conversely, though congruously, other commentators have argued that the First Amendment is integral to the development of a tolerant society.¹⁸⁸ As we analyze different types of activities and try to decide whether they should receive coverage under the First Amendment, we must keep in mind that the values of truth, democracy, and autonomy must be moderated by a keen awareness of the possible destructive effects that these activities might have on the community we are trying to protect and nurture.

e. Prioritizing Values

One final question must be asked regarding First Amendment values: Is one value more important than the others? This, of course, is a question that has generated quite a bit of controversy. Some commentators have argued that the value of democratic self-governance should be the only value taken into account.¹⁸⁹ Justice Oliver Wendell Holmes, of course, famously proclaimed the “marketplace of ideas” rationale.¹⁹⁰ Meanwhile, Justice Thurgood Marshall was known to emphasize the value of autonomy: “The First Amendment serves not only the needs of the polity but also those of the human spirit—a spirit that demands self-expression.”¹⁹¹ And Justice Louis Brandeis was quite dexterous at eloquently advocating

187. *Id.* at 478.

188. *See, e.g.*, LEE C. BOLLINGER, *THE TOLERANT SOCIETY: FREEDOM OF SPEECH AND EXTREMIST SPEECH IN AMERICA* 10 (1986); David A. Strauss, *Why Be Tolerant?*, 53 U. CHI. L. REV. 1485, 1485–86 (1986).

189. *See, e.g.*, Bork, *supra* note 162, at 27–28.

190. *Abrams v. United States*, 250 U.S. 616, 630 (1919) (Holmes, J., dissenting).

191. *Procunier v. Martínez*, 416 U.S. 396, 427 (1974) (Marshall, J., concurring).

for all three of these values at once.¹⁹² Finally, Dean Post has proposed arranging the several competing theories of the First Amendment according to a “lexical priority”:¹⁹³

The rules of the participatory theory will be imposed when required by that theory; the rules of the Meiklejohnian perspective will be imposed when required by that perspective and not incompatible with the participatory theory; the rules of autonomy theory will be imposed when required by that theory and not incompatible with the participatory and Meiklejohnian approaches; and so forth.¹⁹⁴

Quite recently, however, the Supreme Court warned of the dangers of addressing this question lightly.¹⁹⁵ In *U.S. v. Stevens*, the Court rejected the Government’s argument that Congress’s “categorical balancing of the value of the speech against its societal costs” should determine “[w]hether a given category of speech enjoys First Amendment protection.”¹⁹⁶ In writing for the Court, Justice Scalia characterized such “a free-floating test for First Amendment coverage” as “startling and dangerous.”¹⁹⁷ He further explained:

The First Amendment’s guarantee of free speech does not extend only to categories of speech that survive an ad hoc balancing of relative social costs and benefits. The First Amendment itself reflects a judgment by the American people that the benefits of its restrictions on the Government outweigh the costs. Our Constitution forecloses any attempt to revise that judgment simply on the basis that some speech is not worth it.¹⁹⁸

The Court, it should be noted, was reacting to the Government’s attempt to place such power to balance the costs and benefits of particular types of speech in the ever-changing hands of the political branches, Congress and the Executive. Doing so would amount to allowing the contours of the First Amendment to be rewritten through the legislative process as the whims of the electorate might blow. This, of course, would be antithetical to the power

192. See *Whitney v. California*, 274 U.S. 357, 375 (1927) (Brandeis, J., concurring).

193. Post, *Participatory Democracy*, *supra* note 152, at 489 (quoting Robert Post, *Reconciling Theory and Doctrine in First Amendment Jurisprudence*, 88 CAL. L. REV. 2353, 2373 (2000)).

194. *Id.*

195. *United States v. Stevens*, 130 S. Ct. 1577, 1585–86 (2010).

196. *Id.* at 1585.

197. *Id.*

198. *Id.*

of judicial review itself, as announced in *Marbury v. Madison*, and to the role of the Supreme Court as final arbiter and interpreter of the Constitution. “The Constitution is not a document ‘prescribing limits, and declaring that those limits may be passed at pleasure.’”¹⁹⁹ This article in no way argues for such an approach. Instead, this article points to what the Court itself recognized in *Stevens*: that *the Court* has to consider First Amendment values as it attempts to decipher and describe the limits imposed by *the Constitution*.²⁰⁰

In the end, nonetheless, it is important to heed the Supreme Court’s warning in *Stevens* that defining whole categories of speech out of First Amendment coverage must not be done “on the basis of a simple cost-benefit analysis.”²⁰¹ The dynamic nature of the underlying First Amendment values and their inherent interrelatedness suggest a need for open and unconstrained debate on the topic. The Constitution serves a multiplicity of masters: order, equality, autonomy, justice, and democracy, amongst others. There is no reason why the freedom of speech guaranteed by the First Amendment—a central tenet of our system—should be bound by rigid hierarchies as to the values it serves. In fact, we must be especially aware of the fact that the Court has recently been reluctant to create new categories of disfavored speech.²⁰² Consequently, as we analyze how source code furthers First Amendment values, we must err on the side of caution, fully conscious of the Court’s interpretation of the First Amendment as favoring more, rather than less, coverage of general classes of speech.

199. *Id.* (quoting *Marbury v. Madison*, 1 Cranch 137, 178, 2 L. Ed. 60 (1803)).

200. *United States v. Stevens*, 130 S. Ct. 1577, 1585–86 (2010) (“To be fair to the Government, its view did not emerge from a vacuum. As the Government correctly notes, this Court has often described historically unprotected categories of speech as being ‘of such slight social value as a step to truth that any benefit that may be derived from them is clearly outweighed by the social interest in order and morality.’ In *New York v. Ferber*, we noted that within these categories of unprotected speech, ‘the evil to be restricted so overwhelmingly outweighs the expressive interests, if any, at stake, that no process of case-by-case adjudication is required,’ because ‘the balance of competing interests is clearly struck’”) (citations omitted).

201. *Id.* at 1586.

202. *See, e.g., id.* (refusing to create a category of disfavored speech for depictions of animals being intentionally tortured and killed); *Brown v. Entm’t Merchants Ass’n*, 131 S. Ct. 2729 (2011) (refusing to create a category of disfavored speech for violent video games).

2. First Amendment Values in Source Code

a. Truth

Computer science is a science like any other: it is a search for mathematical truth through the use of the scientific method. Professor Lessig has made this point explicit by analogizing the story of Andrew Wiles's proof of Fermat's Last Theorem to the development of the Linux operating system.²⁰³ Through the use of the Internet and the open sharing of ideas and methods, both Andrew Wiles and Linus Torvalds were able to get help from countless people who improved, tested, and added to their ideas.

This is what the scientific method is all about. One person comes up with an idea on how to solve a problem. She implements and tests her idea. If her theory works out, she publishes her work and others replicate it. Yet scientists can only improve and extend the ideas of others if they are able to see and replicate them. This is why the free expression of ideas within the scientific world is so important.

"Academic freedom is 'a special concern of the First Amendment.'"²⁰⁴ "Teachers and students must always remain free to inquire, to study and to evaluate, to gain new maturity and understanding"²⁰⁵ "[T]he First Amendment protects scientific expression and debate just as it protects political and artistic expression."²⁰⁶ This is why "scientific seminars, discussions, and publications are covered by the First Amendment."²⁰⁷ "Authors routinely write books and articles in which they communicate procedures to each other. . . . [S]uch writings are unambiguously covered by the First Amendment."²⁰⁸

There is no reason to think that this would not also apply to computer science and source code. Computer source code, like mathematical equations or chemical formulas, is the description of

203. Lessig, *Open Code*, *supra* note 107, at 1417.

204. Tien, *supra* note 8, at 633 n.15 (quoting *Keyishian v. Bd. of Regents*, 385 U.S. 589, 603 (1967)).

205. *Sweezy v. New Hampshire*, 354 U.S. 234, 250 (1957).

206. *Bd. of Trs. of Leland Stanford Junior Univ. v. Sullivan*, 773 F. Supp. 472, 474 (D.D.C. 1991) (citing *Miller v. California*, 413 U.S. 15, 34 (1973)); *Roth v. United States*, 354 U.S. 476, 484 (1957); *United States v. U.S. Dist. Court for Cent. Dist. of Cal.*, 858 F.2d 534, 542 (9th Cir. 1988)).

207. Roy G. Spece, Jr. & Jennifer Weinzierl, *First Amendment Protection of Experimentation: A Critical Review and Tentative Synthesis/Reconstruction of the Literature*, 8 S. CAL. INTERDISC. L.J. 185, 187 n.4 (1998).

208. Post, *Encryption*, *supra* note 1, at 718.

ideas formulated by experts of a science.²⁰⁹ Consider the following description of the interchange of ideas that characterizes the computer science world:

In using and stating source code, programmers not only assert a particular procedure or set of procedures—they also participate in a scientific discourse about the asserted procedures. . . . Publishing an algorithm can create a discourse about those classes of algorithms and problems. . . . [T]he publication of computer programs—algorithms in source code form—contributes to the development of mathematics itself. Conversely, mathematical problems have stimulated various areas of computer science

*This discourse is central to the marketplace of ideas in computer science.*²¹⁰

Source code is both a participant and a good in the marketplace of ideas. The First Amendment should therefore cover the publication and exchange of source code in the academic context, as it furthers the First Amendment value of truth.²¹¹

Furthermore, the importance of computer code in the academic setting has recently garnered recognition across a wide array of fields. According to the *New York Times*, “Many professors of computer science say college graduates in every major should understand software fundamentals.”²¹² Similarly, Professor Jeannette M. Wing, head of the Computer Science Department at Carnegie Mellon University, argues that “[c]omputational thinking is a fundamental skill for everyone, not just for computer scientists.”²¹³ Professor Wing further claims that “[j]ust as the printing press facilitated the spread of the three Rs, what is appropriately incestuous about this vision is that *computing and computers facilitate the spread of computational thinking*.”²¹⁴ The spread of knowledge regarding com-

209. See *Bernstein v. U.S. Dep’t of Justice* (citations omitted), 176 F.3d 1132, 1141 (9th Cir.) (making the same comparison), *withdrawn & reh’g granted*, 192 F.3d 1308 (9th Cir. 1999).

210. Tien, *supra* note 8, at 664 (emphasis added).

211. *But see* Kerr, *supra* note 8, at 1291 (arguing that all sorts of things, including physical objects, express ideas about themselves, and that this is an insufficient reason to extend First Amendment coverage to such things, including source code).

212. Randall Stross, *Computer Science for the Rest of Us*, N.Y. TIMES (Mar. 31, 2012), <http://www.nytimes.com/2012/04/01/business/computer-science-for-non-majors-takes-many-forms.html>.

213. Jeannette M. Wing, *Computational Thinking*, COMM. ACM, Mar. 2006, at 33, 33.

214. *Id.* (emphasis added).

puter code will not only help in developing new knowledge about programming, but will also contribute to the development of general human knowledge.

It is particularly telling that the approach taken by some professors when teaching basic computer literacy to individuals concentrating in other fields has been strikingly multidisciplinary.²¹⁵ Consider one example:

At Wheaton College . . . a professor of computer science teaches “Computing for Poets.” The only prerequisite, according to the course syllabus, is “a love of the written (and digital) word.”

Professor LeBlanc has his students learn the basics of Python, another modern language used in the software industry. But this course is tied to two courses offered by the English department on J.R.R. Tolkien and Anglo-Saxon literature.²¹⁶

Similarly, Professor Wing argues that computational thinking will aid the development of knowledge in other fields because it teaches problem-solving techniques.²¹⁷ She also states that studying computer code can teach us all sorts of other things: “thinking recursively”; “parallel processing”; “recognizing both the virtues and the dangers of aliasing, or giving someone or something more than one name”; “judging a program not just for correctness and efficiency but for aesthetics, and a system’s design for simplicity and elegance”; “thinking in terms of prevention, protection, and recovery from worst-case scenarios through redundancy, damage containment, and error correction”; “planning, learning, and scheduling in the presence of uncertainty”; and “strategy for winning a game.”²¹⁸ Professor Wing also gives specific examples of how the interdisciplinary study of computer code has brought about exciting new developments in other fields of study such as statistics, biology, economics, chemistry, and physics.²¹⁹

The mistake too often made, Professor Wing argues, is “equat[ing] computer science with computer programming.”²²⁰ Computer science, at the end of the day, is “[a] way that humans, not computers, think.”²²¹ “Computational thinking is a way humans solve problems; it is not trying to get humans to think like com-

215. See Stross, *supra* note 212.

216. *Id.*

217. Wing, *supra* note 213, at 33.

218. *Id.* at 33–34.

219. See *id.*

220. *Id.* at 35.

221. *Id.* (emphasis removed).

puters.”²²² Similarly, computer code is the language in which *humans speak to other humans* about computer science. It is the language in which ideas about computational thinking are expressed. It is not just a way to get a machine to do something. Rather, it is a way of thinking about our reality and interacting with it in creative ways. Thus the ability to freely create and share source code is essential to the instruction, furtherance, and development of computational thinking.

It should be noted that the government conceded in the *Bernstein* litigation that the First Amendment covers source code printed in academic works.²²³ The government, however, argued in *Bernstein I* and *II*, as well as in other cases, that source code in electronic form is different because it can be directly fed into a computer.²²⁴ This distinction is immaterial to our discussion here. “For purposes of the First Amendment, the language in which books and articles are written is without importance.”²²⁵ It follows that the medium in which one publishes one’s ideas should also be immaterial to the question of First Amendment coverage. The values furthered by a particular activity are the same whether one publishes the ideas on paper or through the Internet in electronic form.²²⁶ “The decisive question for coverage is whether or not the publication of the source code forms part of a First Amendment medium, and this question can be affirmatively answered even if the publication of the source code is in electronic form.”²²⁷

Furthermore, the arguments advanced here do not only apply to professors and PhDs. The open source movement has shown that a large number of people are interested in engaging in computer programming as a hobby. By keeping their source code open and sharing it with others, open source programmers have been able to

222. *Id.*

223. *Bernstein v. U.S. Dep’t of State (Bernstein I)*, 922 F. Supp. 1426, 1434 (N.D. Cal. 1996), *aff’d*, *Bernstein v. U.S. Dep’t of Justice*, 176 F.3d 1132 (9th Cir.), *withdrawn & reh’g granted*, 192 F.3d 1308 (9th Cir. 1999).

224. *See, e.g., Bernstein v. U.S. Dep’t of Justice (Bernstein II)*, 176 F.3d 1132, 1141–42 (9th Cir.) (“[T]he government maintains that source code is different from other forms of expression . . . because it can be used to control directly the operation of a computer without conveying information to the user.”), *withdrawn*, 192 F.3d 1308 (9th Cir. 1999).

225. Post, *Encryption*, *supra* note 1, at 718.

226. *See id.* at 719 (“So long as the publication of . . . source code forms part of this public discourse and debate, it will be covered by the First Amendment, whether it is set forth in a printed article or in an online discussion.”).

227. *Id.*

develop better programs that continue to evolve and improve.²²⁸ This means that all sorts of people are contributing to the discourse of computer science by engaging in the free, open sharing of their source code with others.²²⁹

The open source model, furthermore, looks very much like a free marketplace of ideas where people give and take as they construct better ways of dealing with the problems they are tackling. Eric S. Raymond famously used the metaphor of the bazaar (a free marketplace) to describe the open source movement in his seminal piece entitled *The Cathedral and the Bazaar*.²³⁰ Raymond explained that the open source model incorporated the free exchange of ideas, much in the same way that goods are freely exchanged in a bazaar, as a more efficient and effective way of building code.²³¹ He drew a sharp contrast between this cooperative model and the isolationist, centralized, command and control model of cathedral construction.²³²

Similarly, Richard Stallman refers to the open source movement as the Free Software Movement: “[F]ree in the sense of ‘free speech,’ not in the sense of ‘free beer.’”²³³ Stallman goes on to emphasize the role of a free exchange of ideas in the development of open source code over any purely monetary or economic connotation of the word “free.”²³⁴

In this sense, Professor Lessig argues that when code is left open—when all of its modular components are subject to public tinkering— “[n]o rules say which way is right. Instead, the evolution of a market does that. The evolution of thousands of people trying their hand at improving a code, and thousands of people choosing

228. See generally Eric Steven Raymond, *The Cathedral and the Bazaar*, ERIC S. RAYMOND'S HOME PAGE (Aug. 2, 2002, 9:02 AM), <http://www.catb.org/~esr/writings/homesteading/cathedral-bazaar/cathedral-bazaar.ps>.

229. In fact, the uptrend in individuals' interest in learning programming languages has recently been noted in the national press. See, e.g., Jenna Wortham, *A Surge in Learning the Language of the Internet*, N.Y. TIMES (Mar. 27, 2012) <http://www.nytimes.com/2012/03/28/technology/for-an-edge-on-the-internet-computer-code-gains-a-following.html> (“The blooming interest in programming is part of a national trend of more people moving toward technical fields. According to the Computing Research Association, the number of students who enrolled in computer science degree programs rose 10 percent in 2010, the latest year for which figures are available.”).

230. See generally Raymond, *supra* note 228, at 3.

231. See *id.*

232. See *id.*

233. Lessig, *Open Code*, *supra* note 107, at 1406.

234. *Id.*

which improvement makes sense.”²³⁵ Source code, then, and open source code in particular, are prime examples of the “marketplace of ideas” theory of the First Amendment at work. When code is open, everyone can participate in its development. Through “running code, that by its power produces rough consensus,”²³⁶ an open marketplace is created, where everybody’s ideas can be pounded out and morphed into a better concept of the truth.²³⁷

But let us also consider how source code promotes truth through the development of new media for the communication of ideas. From this perspective, allowing government control and regulation of source code, without regard to First Amendment doctrine, could do a disservice to the acquisition of knowledge. This issue has become more dramatic in the past few years, as “publication of user-generated content (UGC) (also known as consumer-generated media) has exploded.”²³⁸

Perhaps one of the most significant examples of this issue is YouTube. YouTube catapulted itself to household name status after it “was purchased by Google for \$1.65 billion in stock in October 2006, a little more than a year after it started.”²³⁹ By mid-2008, ac-

235. *Id.* at 1415.

236. *Id.* at 1418.

237. But code need not be open or academic to further the value of truth and be worthy of coverage under the First Amendment. Source code is also having a dramatic effect on the marketplace of goods and services itself. A large number of people are interested in engaging in computer programming as a business. Every day our economy is becoming more and more intertwined with the development of computer code. This phenomenon is only natural, as technological growth is the engine of real economic growth. It was only recently that Apple overtook Exxon Mobil as the most valuable company in the world, in terms of market capitalization. James B. Stewart, *Confronting a Law of Limits*, N.Y. TIMES (Feb. 24, 2012), <http://www.nytimes.com/2012/02/25/business/apple-confronts-the-law-of-large-numbers-common-sense.html>. As of November 6, 2012, four of the top ten companies with the largest market capitalization in the world were computer-related enterprises (Apple, Microsoft, Google, and IBM). YCharts, *Market Cap Stock Rankings*, YCHARTS, http://ycharts.com/rankings/market_cap (last visited Nov. 7, 2012). According to one news article, “The thinking is that with so much business gravitating toward the Internet, it’s critical that today’s entrepreneurs learn the language of the computer—or at least enough that they won’t be left behind.” Colleen Debaise, *Do You Really Need to Code?*, ENTREPRENEUR.COM (Mar. 28, 2012), <http://www.entrepreneur.com/blog/223238>. Computer code, then, may also be reshaping the economic landscape itself, and becoming part of the vocabulary of entrepreneurs, financiers, venture capitalists, and the like. Government regulation of this code could compromise the development of new ideas through the investment in and development of new technologies.

238. Robert P. Latham et al., *Legal Implications of User-Generated Content: YouTube, MySpace, Facebook*, 20 INTELL. PROP. & TECH. L.J. 1 (2008).

239. *Id.*

ording to one estimate, “more than 65,000 videos [were being] uploaded to YouTube every day, and 100 million videos [were being] viewed daily.”²⁴⁰ But four years are a lifetime in Internet time and, by early 2012, YouTube’s popularity had mushroomed:

- Over 800 million unique users visit YouTube each month
- Over 4 billion hours of video are watched each month on YouTube
- 72 hours of video are uploaded to YouTube every minute
- 70% of YouTube traffic comes from outside the US
- YouTube is localized in 43 countries and across 60 languages
- In 2011, YouTube had more than 1 trillion views, or around 140 views for every person on Earth.²⁴¹

Furthermore, YouTube’s popularity is only compounded by its interrelation with social networking sites such as Facebook and Twitter: “500 years of YouTube video are watched every day on Facebook, and over 700 YouTube videos are shared on Twitter each minute.”²⁴² The resulting statistics are simply mind-blowing.²⁴³

240. *Id.*

241. *Statistics*, YOUTUBE, http://www.youtube.com/t/press_statistics (last visited Nov. 7, 2012).

242. *Id.*

243. *See id.* At this point, it is helpful to include a description of what YouTube does, how it works, and how its ease of use makes it uniquely convenient.

The mechanisms employed by YouTube are similar to those employed by its competitors: users can send digital video files from their computers or web-capable devices to the central YouTube system, where the video files are then converted into Adobe’s Flash Video format and stored on the YouTube servers. YouTube users may then visit the website and view the videos through their web browsers. This system allows users to create and share videos with friends and others without having to worry about issues such as the cost of, or limitations on, bandwidth.

One potential reason for YouTube’s immense popularity despite its numerous competitors is its ease of use. Once a user uploads his video to YouTube through the system described above, he is asked to give a brief description of the video and to apply certain keywords, or “tags.” From this point forward, the process is almost entirely automated. Barring any issues, the video will be available for viewing on YouTube in mere minutes.

It is just as easy to view the videos as it is to upload them. Users can visit the website and browse the YouTube library for whatever it is they seek. The video is then delivered quickly and efficiently to the viewer’s web browser. In addition, YouTube recommends featured or related videos for further viewing. Furthermore, videos may be “embedded” into other websites so that they may be shared outside of the YouTube website. The relative ease of use makes YouTube an attractive medium for sharing and viewing content.

Kevin C. Hormann, Comment, *The Death of the DMCA? How Viacom v. YouTube May Define the Future of Digital Content*, 46 HOUS. L. REV. 1345, 1353–54 (2009) (footnotes omitted). Consequently, it has recently been argued that “YouTube also

The ease and speed with which users can post, share, find, and view videos on YouTube is, of course, highly dependent on the computer code that runs the whole system. If you regulate the code underlying YouTube, you can therefore control the extensive user-generated content that is being constantly uploaded, viewed, and shared on it throughout the world. The sheer volume of content being shared is a testament to the capacity of government to hush an enormous amount of speech if it were to regulate the code that allows such expression. This alone should trigger our sensibilities concerning the need for First Amendment coverage of such code.

Let us consider some specific tangible examples of how government regulation could disserve the value of truth in this context. Users have posted a plethora of how-to videos on YouTube, teaching viewers how to perform all sorts of tasks. More formally, though, YouTube has created services called YouTube EDU and YouTube for Schools that allow users to use the site to engage in teaching activities by watching, posting, or sharing educational videos.²⁴⁴

It is not hard to imagine that a government might be inclined to limit the availability of some types of educational or how-to content to its citizens. For example, the federal government might seek to prohibit the posting of any tax advice on how to legally avoid certain tax liabilities so that it can maximize its tax revenues. Such a result could be implemented by requiring YouTube to include in the relevant code some algorithm that automatically limits access to any videos identified with certain keywords such as “IRS” or “tax advice.” The same could be done with regard to educational videos seeking to teach the theory of evolution to users. A state might determine that such educational goals go against its public policy and, again, seek to limit their availability to their citizens by regulating the underlying code. In such a scenario, the state might even benefit from regulating the code to limit such content when the user is located within the state’s territorial jurisdiction. Regulation of such speech would most likely be deemed content-based and subject to strict scrutiny under the First Amendment, yet these hypothetical

helps ‘promot[e] the Progress of Science’—the constitutional goal of copyright—by providing an exciting environment in which users mix their content with preexisting works—copyrighted or otherwise.” Peter K. Yu, *Digital Copyright and Confuzzling Rhetoric*, 13 VAND. J. ENT. & TECH. L. 881, 894 (2011) (footnote omitted) (quoting U.S. CONST. art. I, § 8, cl. 8).

244. See *YouTube EDU*, YOUTUBE, <http://www.youtube.com/education> (last visited Nov. 7, 2012); *YouTube for Schools*, YOUTUBE, <http://www.youtube.com/schools> (last visited Jan. 19, 2012).

regulations could evade such review if source code is deemed not to be covered by the First Amendment.²⁴⁵

Likewise, a government may seek to control indecent expression by regulating source code. The government, again, could force YouTube to include in the relevant code a mechanism for scanning videos to reject any images that might contain nudity. Such a process, however, would result in an overbroad limitation of expression under well-established First Amendment doctrine.²⁴⁶ Novel ideas could be censored, and the marketplace would be less rich as a result. That this result is accomplished by a change in computer code instead of a government censor looking at each video is merely a practical convenience that should not change the First Amendment scrutiny. Efficient censors are just as bad, and actually worse, than inefficient ones.

Another prime example of a cybermarket for ideas is the blogosphere. Professor Lessig argues that blogs offer a success story of what choice and transparency can do for freedom of speech.²⁴⁷ He sees them as facilitating the “vigorous exchange of ideas” that characterizes the “marketplace of ideas.”²⁴⁸ According to this perspective, “the blogosphere enables a wide spectrum of views to be presented through which an accepted construction of truth can emerge.”²⁴⁹ “Lessig equates a multitude of blogs with a plentitude of facts, views, and opinion, with choices available to the readers . . . transparent because their arguments are completely visible and linked simultaneously to counterarguments from other blogs.”²⁵⁰ To be sure, blogs have proliferated. As of 2007, “there

245. See *Consol. Edison Co. of N.Y. v. Pub. Serv. Comm’n of N.Y.*, 447 U.S. 530, 536 (1980) (strict scrutiny applied where speech is regulated based on content).

246. See *Reno v. ACLU*, 521 U.S. 844, 878–79 (1997) (finding statute restricting child pornography overbroad for its purpose and therefore unconstitutional).

247. See LAWRENCE LESSIG, *CODE: VERSION 2.0* 242 (2006) [hereinafter LESSIG, *CODE*].

248. *Id.*

249. Mayer-Schönberger, *supra* note 110, at 724 (citing LESSIG, *CODE*, *supra* note 236, at 244). *But cf. id.* at 731–34 (arguing that “Lessig’s conceptualization of . . . freedom of speech based on market and choice exposes severe weaknesses of the market mechanism”); SUNSTEIN, *supra* note 3, at 141–43 (arguing that the analogy between the blogosphere and the marketplace of ideas is imperfect).

250. Mayer-Schönberger, *supra* note 110, at 724–25 (citing LESSIG, *CODE*, *supra* note 236, at 236). *But cf. id.* at 733 (“[B]logs, may in the aggregate expose their readers to many different viewpoints, but that is of little value when trying to discover truth. . . . [B]logs are as susceptible to biases, fashions and fads as mainstream media is.” (footnote omitted)); SUNSTEIN, *supra* note 3, at 51 (claiming that bloggers routinely provide links to other sites only to “show how dangerous, or

[were] 55 million blogs, and over 40,000 new ones [were being] created each day, with a new one every 2.2 seconds.”²⁵¹ And “[y]ou can easily find blogs on countless subjects.”²⁵²

The enormous amount of speech generated and shared in the blogosphere should alert us to the high level of First Amendment value therein concentrated. So it should be troubling to conceive of a government regulatory scheme whereby all posts expressing certain viewpoints would be eradicated from the blogosphere through an ingenious filtering mechanism inserted in the code that runs the individual blogs. If the government were to prohibit outright the expression of certain opinions, it would certainly run into a First Amendment wall. Why should forcing hosting services to install code that accomplishes the same end be any different?

Of course, not all of the information shared by individuals on the Internet will be truthful. The awesome ability of social networking websites and blogs to spread falsity, half-truth, or over-simplified truth with lightning speed to millions of people across the world should not be underestimated.²⁵³ Then again, the whole point of the marketplace of ideas theory of First Amendment value is that the free exchange of information will serve to weed out the true from the false. As Justice Louis Brandeis famously said: “If there be time to expose through discussion the falsehood and fallacies, to avert the evil by the processes of education, *the remedy to be applied is more speech*, not enforced silence.”²⁵⁴

Source code furthers the First Amendment value of truth in a host of ways. The past few years have reinforced this point with exponential force. The social interactions that have recently arisen in the Internet could be subjected to severe government censorship if we fail to recognize that the First Amendment must cover source code.

how contemptible, competing views really are”); *id.* at 149 (“[A] plurality of cross-citations simply cast contempt on the views that are being cited! Only a quarter of cross-ideological posts involve genuine substantive discussion.”).

251. SUNSTEIN, *supra* note 3, at 138.

252. *Id.*

253. See, e.g., Noam Cohen, *A Video Campaign and the Power of Simplicity*, N.Y. TIMES, (Mar. 11, 2012), <http://www.nytimes.com/2012/03/12/business/media/kony-2012-video-illustrates-the-power-of-simplicity.html>; Jeremy W. Peters, *A Lie Races on Twitter Before Truth Can Boot Up*, N.Y. TIMES (Apr. 9, 2012), <http://www.nytimes.com/2012/04/10/us/politics/false-nikki-haley-twitter-report-spreads-fast.html>.

254. *Whitney v. California*, 274 U.S. 357, 377 (1927) (Brandeis, J., concurring) (emphasis added).

b. Democracy

Source code is the language of cyberspace: it defines and constitutes that space.²⁵⁵ Thus source code is also a *lex informatica*, a law that “defines what behavior is possible in cyberspace and what values cyberspace will uphold.”²⁵⁶ The Internet’s code architecture will be central in determining how the Internet community regulates itself and how the government will be able to regulate it.²⁵⁷

The First Amendment is, in part, designed to protect the ideals of democracy by providing for an open and free public discourse. Free participation in the source code discourse allows people to engage in the decision making process that constitutes the Internet community and shapes the constitution of cyberspace itself. Allowing people to freely read, write, publish, and distribute their own source code furthers the participatory model of democratic self-governance. The people’s ideas will be out there to be read, considered, and maybe incorporated into the architecture of the space. This is more than participation in an economic market; it is participation in a legislative process. If code is law, then writing code is legislating.

The Meiklejohnian theory of First Amendment values is similarly furthered by a free code. Since code is the law of the Internet, we need to have as much information as possible about the code. By letting people freely engage in the public discourse of code in

255. Lessig, *Open Code*, *supra* note 107, at 1408. This statement is reminiscent of Martin Heidegger’s and Ludwig Wittgenstein’s descriptions of our own “real” world as a reality constituted by language. See MARTIN HEIDEGGER, *POETRY, LANGUAGE, THOUGHT* 146 (Albert Hofstadter trans., Harper & Row 1975) (1971) (“Man acts as though *he* were the shaper and master of language, while in fact *language* remains the master of man.”); LUDWIG WITTGENSTEIN, *TRACTATUS LOGICO-PHILOSOPHICUS* §5.6 (C. K. Ogden trans., Routledge 1922) (1921) (“*The limits of my language* mean the limits of my world.”). This is in part why we understand freedom of speech to be so important. In contemporary society we have a basic understanding of how speech constitutes our reality. “Like everything metaphysical the harmony between thought and reality is to be found in the grammar of the language.” LUDWIG WITTGENSTEIN, *ZETTEL* § 55 (G.E.M. Anscombe & G.H. von Wright eds., G.E.M. Anscombe trans., Univ. of Cal. Press 1970) (1967). The close link between speech, language, and thought has not gone unnoticed by the Supreme Court. This is why, in the context of protecting freedom of speech, the Court has emphasized that “[o]ur whole constitutional heritage rebels at the thought of giving government the power to control men’s minds.” *Stanley v. Georgia*, 394 U.S. 557, 565 (1969).

256. Lessig, *Limits*, *supra* note 109, at 761.

257. Lawrence Lessig, *Keynote Address: Commons and Code*, 9 *FORDHAM INTELL. PROP. MEDIA & ENT. L.J.* 405, 410 (1999). See generally Lessig, *Open Code*, *supra* note 107; Lessig, *Limits*, *supra* note 109; Lessig, *Law of the Horse*, *supra* note 123.

the Internet, we are making available to everyone information on how others think the problems of cyberspace might better be solved. Furthermore, we are providing people with the tools to create their own corners of cyberspace in whichever way they want. The code itself, unencumbered by regulation, must form part of the public discourse.

Source code furthers the values of both models of democratic self-determination because the code is central to the development and regulation of cyberspace, a space in which we will be living a considerable part of our lives for years to come. We must therefore have as much information as possible about the code that determines how cyberspace is governed. Moreover, we need to translate the democratic values to which we adhere in our “real” space to cyberspace by allowing people to participate in the public discourse of that space in the language of that space: source code.

An important example of source code promoting democratic self-determination can be found in the proliferation of *open* code.²⁵⁸ Open code creates a discourse of code in which the participatory model of democratic self-determination is furthered through everyone’s participation in the development of the code. Yet at the same time, the Meiklejohnian model is also furthered by the virtual guarantee that, with everybody giving their input into the process, the necessary information for wise decision-making will be available to all. Implicitly, therefore, any restriction on open code will inhibit this free exchange and any resulting democratic value.

Similarly, Professor Jack M. Balkin has been seen as arguing that “the creative reuse and modification of preexisting materials help promote the development of a vibrant democratic culture, which in turn affects a country’s political future.”²⁵⁹ Regarding the ability to openly copy, sample, rip, change, and mix up ideas, and its effect on democratic self-governance, he states:

258. Professor Lessig argues that “we should look to the structure of our constitutional tradition, and extract from it the values that are constituted by it, and carry these values into the world of the Internet’s governance—whether the governance is through code, or the governance is through people.” Lessig, *Open Code*, *supra* note 107, at 1409. The balance to be struck between the open and the closed, the level of regulation that we want in cyberspace, should, therefore, be determined through an open and free public discourse, just like it is in real space. Hence, allowing the code to be as free as possible furthers the value of democratic self-governance.

259. Yu, *supra* note 243, at 896 (discussing 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, 45 (2004)).

A democratic culture is the culture of widespread “rip[ping], mix[ing], and burn[ing],” of nonexclusive appropriation, innovation, and combination. It is the culture of routing around and glomming on, the culture of annotation, innovation, and bricolage. Democratic culture . . . makes use of the instrumentalities of mass culture, but transforms them, individualizes them, and sends what it produces back into the cultural stream. In democratic culture, individuals are not mere consumers and recipients of mass culture but active appropriators.²⁶⁰

By ensuring that the discourse in source code is as free as possible, these positive effects can be extended *to the Internet as a whole*. This is not to say that all persons must participate in a culture of open code. Rather, extending First Amendment coverage to source code will create more incentives for people to freely participate in the discourse of code. This will make code more open and free, which should in turn further First Amendment values.

Moreover, source code can also have a direct effect on our ability to engage in public discourse regarding our conventional “real world” government. Any doubt as to the sheer power of new Internet and social networking technologies should have been dispelled by the recent Kony 2012 phenomenon.²⁶¹

With more than 100 million online views, “Kony 2012” became *the most viral video in history*. It drew global attention to the reviled leader of the Lord’s Resistance Army and his use of child soldiers to terrorize people in at least four countries over the years. *It also prompted Congressional resolutions* urging President Obama to bolster efforts in the region, where 100 U.S. military advisers were assigned last fall to help countries combat the Lord’s Resistance Army.²⁶²

Of course, the Kony 2012 campaign has not been free of criticism and detractors.²⁶³ Nevertheless, its demonstration of the

260. Balkin, *supra* note 259, at 45 (footnote omitted) (quoting LAWRENCE LESSIG, *THE FUTURE OF IDEAS: THE FATE OF THE COMMONS IN A CONNECTED WORLD* 9 (2001)).

261. See generally *Kony 2012*, INVISIBLE CHILDREN, <http://www.invisiblechildren.com/kony/> (last visited Nov. 7, 2012).

262. Jennifer Preston, *Sequel to ‘Kony 2012’ Video Addresses Critics and Outlines Call for Action*, N.Y. TIMES (Apr. 5, 2012, 7:59 PM), <http://thelede.blogs.nytimes.com/2012/04/05/sequel-to-kony-2012-video-addresses-critics-and-outlines-call-for-action/> (emphasis added).

263. *Id.*

power of new Internet media and networking tools to mobilize political will can hardly be denied.²⁶⁴

Similarly, what better example of the Internet changing the game of politics is there than the so-called “YouTube Election,” as the U.S. Presidential election of 2008 has come to be known?²⁶⁵ In fact, even before the official beginning of the 2008 election cycle, the influence that YouTube would have on politics became brutally evident to then Republican Senator George Allen. Mr. Allen was at the time seen as a major contender for the Republican Presidential nomination; however, his potential future presidential campaign unraveled before it even began when, still in the midst of his run for the Senate in the 2006 midterm elections, Mr. Allen “was caught on tape at a campaign event twice calling a college student of Indian descent a ‘macaca,’ an obscure racial slur.”²⁶⁶ The video was posted on YouTube and it quickly went viral.²⁶⁷ From there, “[i]t then bounced from the Web to the front page of The Washington Post to cable and network television news shows.”²⁶⁸ The whole fiasco was generally seen as costing Mr. Allen his presidential aspirations.²⁶⁹ He also eventually lost his Senate seat in the 2006 election

264. *Id.* Another recent example of a viral video directly influencing discussions of public policy, and allegedly causing riots in multiple locations around the world, is the trailer to the movie entitled *Innocence of Muslims*. Sam Bacile, *Muhammad Movie Trailer*, YOUTUBE (Jul. 2, 2012), <http://www.youtube.com/watch?v=qmo dVun16Q4>. See generally *The ‘Innocence of Muslims’ Riots* (Nakoula Basseley Nakoula), N.Y. TIMES, http://topics.nytimes.com/top/reference/timestopics/subjects/i/innocence_of_muslims_riots/ (last updated Nov. 26, 2012).

265. See, e.g., Beth Kowitz, *The YouTube Election*, CNNMONEY (Jul. 18, 2008), http://money.cnn.com/2008/07/18/magazines/fortune/kowitz_obamavideo.fortune/index.htm; Jessica Ramirez, *The Big Picture*, DAILY BEAST (Nov. 9, 2008), <http://www.thedailybeast.com/newsweek/2008/11/09/the-big-picture.html>; Rajini Vaidyanathan, *Top Hits of the YouTube Election*, BBC NEWS, http://news.bbc.co.uk/2/hi/americas/us_elections_2008/7699509.stm (last updated Oct. 30, 2008, 11:27 GMT); James Wolcott, *The YouTube Election*, VANITY FAIR, June 2007, at 96, 96, available at <http://www.vanityfair.com/ontheweb/features/2007/06/wolcott200706>. Curiously, the New York Times seemed to be ahead of the curve when it began using the term two years earlier, in 2006. Ryan Lizza, *The YouTube Election*, N.Y. TIMES (Aug. 20, 2006), <http://www.nytimes.com/2006/08/20/weekinreview/20lizza.html>.

266. Lizza, *supra* note 259.

267. *Id.*

268. *Id.*

269. See *id.*

to Democrat Jim Webb. As of November 8, 2012, the video had 672,027 total views on YouTube.²⁷⁰

But YouTube's influence on the presidential campaign was just beginning to rear its head. The campaigns themselves would soon get into the game. Consider this contemporaneous description of what had happened, on the campaigns' end, when all was said and done:

TubeMogul, a measurement service, estimates that just the videos that ran on Obama's YouTube channel alone were watched the equivalent of 14.5 million hours, with McCain's channel racking up about 488,152 hours. Had the Obama camp purchased the same amount of airtime on TV it would have cost them roughly \$46 million and the McCain camp \$1.5 million, according to an analysis on the TechPresident blog. On YouTube it was free. It was also priceless. A Pew Research Center report titled "Internet and Campaign 2008" found that 39 percent of voters watched campaign-related video online during the election cycle. That's higher than the percentage of voters who said they checked out candidate Web sites, political blogs or social-networking sites.²⁷¹

YouTube, then, can be "widely used as a political or fund-raising tool, as evident in the 2008 U.S. presidential election and other electoral campaigns. In April 2011, President Obama launched its re-election campaign bid in part through a YouTube video, 'It Begins with Us.'²⁷²

But perhaps the most dramatic effect that YouTube had during the election was the fact that it gave every individual with the time and interest the ability to participate in the political debate and to get his message out to millions of people across the nation and around the world. The 2008 U.S. presidential campaign gave us such viral phenomena as the Obama Girl (25,557,384 views as of November 8, 2012),²⁷³ the 1984 Apple commercial-inspired "Vote

270. Shekar Ramanuja Sidarth, *George Allen Introduces Macaca*, YOUTUBE (Aug. 15, 2006), <http://www.youtube.com/watch?v=r90z0PMnKwI> (last visited Nov. 8, 2012).

271. Ramirez, *supra* note 265 (citation omitted).

272. Yu, *supra* note 243, at 898 (footnote omitted) (citing MICHAEL STRANGELOVE, *WATCHING YOUTUBE: EXTRAORDINARY VIDEOS BY ORDINARY PEOPLE* 137-57 (2010); BarackObamadotcom, *Barack Obama 2012 Campaign Launch Video - "It Begins with Us,"* YOUTUBE (Apr. 3, 2011), <http://www.youtube.com/watch?v=fVZLvVFIFQ>).

273. barelypolitical, *Best of Obama Girl: Crush On Obama*, YOUTUBE (Jun. 13, 2007), <http://www.youtube.com/watch?v=wKsoXHYICqU> (last visited Nov. 8, 2012).

Different” video (6,351,034 views as of November 8, 2012),²⁷⁴ and will.i.am’s celebrity-laden “Yes We Can - Barack Obama Music Video” (24,353,527 views as of November 8, 2012).²⁷⁵ “When the election ended, all YouTube videos mentioning Senator Obama had received a total of 1.9 billion views compared with Sen. John McCain’s, which got 1.1 billion views.”²⁷⁶ Power to the people.

YouTube even managed to create a joint venture between its corporate interests, those of a more traditional media outlet, and those of the general public. The CNN-YouTube Presidential Debates allowed individuals to send questions in video format to be posed to the candidates directly. The experiment proved quite successful in terms of viewership.²⁷⁷

Allowing government regulation of the underlying code that runs YouTube could serve to undermine this newly found potential for robust public discourse. An unscrupulous government could use such regulation to undermine grass-roots movements for change and favor incumbents in their reelection campaigns.

On the other hand, it would be naïve to ignore the fact that these new tools serve more than the previously powerless; powerful factions and entities can also utilize the efficiencies of the Internet. “[T]he power of the Internet is two-way. In addition to all those who wished to speak and had no voice until now, there are those millions whom every lobbyist and advocate wished to reach but could not and now can.”²⁷⁸ This power could be harnessed by the government or special interests groups to give special treatment to those with the resources to engage in high-powered lobbying.

Such a result would be troubling indeed, particularly after the Supreme Court’s holding in *Citizens United*, which has given a free hand to corporate interests participating in political campaigns.²⁷⁹

274. Philip de Vellis, *Vote Different*, YOUTUBE (Mar. 5, 2007), <http://www.youtube.com/watch?v=6h3G-IMZxjo> (last visited Nov. 8, 2012).

275. WeCan08, *Yes We Can - Barack Obama Music Video*, YOUTUBE (Feb. 2, 2008), <http://www.youtube.com/watch?v=jjXyqcx-mYY> (last visited Nov. 8, 2012).

276. Ramirez, *supra* note 265.

277. Paul J. Gough, *CNN’s YouTube Debate Draws Impressive Numbers*, REUTERS NEWS (Jul. 25, 2007), <http://www.reuters.com/article/technologyNews/idUSN2425835220070725> (“Monday’s CNN-YouTube debate brought in pretty good numbers, delivering the highest viewership for a debate among adults 18-34 in cable news history.”).

278. N.J. Slabbert, *Orwell’s Ghost: How Teletechnology Is Reshaping Civil Society*, Keynote Address at The 2008 “You Tube” Election?: The Role and Influence of 21st-Century Media Symposium (Mar. 13, 2008), in 16 *COMMLAW CONSPPECTUS* 349, 355 (2008).

279. *Citizens United v. Fed. Election Comm’n*, 130 S. Ct. 876, 913 (2010).

“The inequality that comes from divergences in wealth is not, on the Court’s view, a proper subject for democratic control. According to the Court, campaign-finance restrictions cannot be justified by reference to equality at all.”²⁸⁰ Nonetheless, particularly in the post-*Citizens United* world, it would seem that the availability of low-cost alternatives to campaigning, such as YouTube, could serve as a fundamental equalizing tool which could help individuals and candidates with less resources level the playing field.

But YouTube is not the only new technology that has fundamentally altered the way in which individuals can participate in the democratic process. As previously mentioned, the blogosphere is an arena where individuals are finding a new vibrant space for the discussion of ideas. Professor Cass R. Sunstein has pointed out:

In recent years, the most highly rated political blogs—including Atrios, Instapundit, and the Daily Kos—have received over tens of thousands of visitors *each day*.

. . . Political blogs are a small percentage of the total [amount of blogs], but they are plentiful, and they seem to be having a real influence on people’s beliefs and judgments.²⁸¹

And even though some commentators, including Professor Sunstein, have pointed out some negative implications of the propagation of blogs and public discourse on the Internet,²⁸² it would be close to impossible to argue against the proposition that, *on balance*, the value of democratic self-governance is much better off with the blogosphere and the Internet than without them. After all, as Professor Sunstein admits, “No one doubts that the blogosphere enables interested readers to find an astounding range of opinions and facts.”²⁸³ Allowing the government to regulate code in a way that would create content-based restrictions in the blogosphere would certainly compromise the First Amendment value of democracy. It would do so in the same way that, as discussed previously, allowing content-based regulations of source code would compromise the value of truth.

Facebook and Twitter also provide good examples of how new developments in computer code have drastically altered the way in

280. SUNSTEIN, *supra* note 3, at 168.

281. *Id.* at 138.

282. *See id.* at 144–46 (arguing that blogs only provide a flawed contribution to deliberative democracy); *id.* at 44, 86–91 (explaining how cybercascades can be counterproductive to a well-functioning deliberative democracy); *id.* at 60–76 (expounding on how the Internet can lead to group polarization and fragmentation detrimental to the public discourse).

283. *Id.* at 139.

which individuals can interact with their government. The influence of these social networking media in the popular revolts in Egypt and Tunisia appears to be undeniable.²⁸⁴ Although the details of just how Facebook and Twitter were used by the participants in these uprisings have not been systematically compiled, it seems that the social networks were useful in at least two ways. As reported by the national press: “First, Facebook and elsewhere online is *where people saw and shared* horrifying videos and photographs of state brutality that inspired them to rebel. Second, these sites are where people found out the basic logistics of the protests—*where to go and when to show up.*”²⁸⁵

Facebook and other online tools, then, have played a central role in raising awareness both within and without the borders of individual countries regarding their problematic government regimes. This drive towards change is continuing even in countries, such as Saudi Arabia, where the people’s dissatisfaction need not boil over into full-fledged revolution, but rather pressures the government in the direction of peaceful reform.²⁸⁶ The positive influence in public discourse of these media from a Meiklejohnian perspective should be clear. These media allow the posting and sharing of an immense amount of information, *including first-hand visual and audio accounts* of government activities around the world. “[A]s shown in relation to the recent Japanese earthquake and political protests in the Middle East and North Africa, home videos shot by citizen journalists provide real-time audio and visual reports without the filtering of the mainstream press.”²⁸⁷ This level of ac-

284. See generally John Pollock, *Streetbook: How Egyptian and Tunisian youth hacked the Arab Spring*, MIT TECH. REV. (September/October 2011), <http://www.technologyreview.com/web/38379/>; Rebecca J. Rosen, *So, Was Facebook Responsible for the Arab Spring After All?*, ATLANTIC (Sep. 3 2011), <http://www.theatlantic.com/technology/archive/2011/09/so-was-facebook-responsible-for-the-arab-spring-after-all/244314>. But see Thomas L. Friedman, *Facebook Meets Brick-and-Mortar Politics*, N.Y. TIMES (June 9, 2012), <http://www.nytimes.com/2012/06/10/opinion/sunday/friedman-facebook-meets-brick-and-mortar-politics.html> (describing how the groups that were successful in utilizing Facebook and social networking media to stimulate and drive the Arab Spring movements have been stifled by traditional politics in their attempts to shape the ensuing regimes).

285. Rosen, *supra* note 284 (emphasis added).

286. See Robert F. Worth, *Twitter Gives Saudi Arabia a Revolution of Its Own*, N.Y. TIMES (Oct. 20, 2012), <http://www.nytimes.com/2012/10/21/world/middleeast/twitter-gives-saudi-arabia-a-revolution-of-its-own.html>.

287. Yu, *supra* note 243, at 897 (citing Jennifer Preston, *Volunteer Site with Harvard Roots Spreads Citizen Journalism’s Voice*, BOSTON GLOBE, Mar. 14, 2011, at 9 (describing the work of Global Voices, which “turned to Facebook, YouTube, and Twitter, where other bloggers and hundreds of ordinary people stepped into the

cess, all around the world, to primary sources, to personal accounts of historical events as they happen, is unprecedented and paradigm-shifting.²⁸⁸

Furthermore, Internet technologies greatly increase information-consumers' access to all types of data from around the world, or from their own corner of the world. Consider, for example, this account of how the availability of Google Earth might have influenced the revolt in Bahrain:

While Facebook has gotten all the face time in Egypt, Tunisia and Bahrain, don't forget Google Earth, which began roiling Bahraini politics in 2006. A big issue in Bahrain, particularly among Shiite men who want to get married and build homes, is the unequal distribution of land. On Nov. 27, 2006, on the eve of parliamentary elections in Bahrain, The Washington Post ran this report from there: "Mahmood, who lives in a house with his parents, four siblings and their children, said he became even more frustrated when he looked up Bahrain on Google Earth and saw vast tracts of empty land, while tens of thousands of mainly poor Shiites were squashed together in small, dense areas. 'We are 17 people crowded in one small house, like many people in the southern district,' he said. 'And you see on Google how many palaces there are and how the al-Khalifas [the Sunni ruling family] have the rest of the country to themselves.' Bahraini activists have encouraged people to take a look at the country on Google Earth, and they have set up a special user group whose members have access to more than 40 images of royal palaces."²⁸⁹

Internet tools like Google Earth can even help individuals learn more about their own surroundings and about their own lives in the context of a huge world just beyond their reach.

role of citizen journalist and shared their experiences, cellphone photos, and videos online"); Steve Sternberg, *The World to the Rescue*, USA TODAY, Apr. 12, 2011, at 1A ("Japan's disaster has spotlighted the critical role that social media websites such as Twitter, Facebook, Google, YouTube and Skype increasingly are playing in responses to crises around the world. They may have been designed largely for online socializing and fun, but such sites and others have empowered people caught up in crises and others wanting to help to share vivid, unfiltered images, audio and text reports before governments or more traditional media can do so.")).

288. See generally THOMAS KUHN, *THE STRUCTURE OF SCIENTIFIC REVOLUTIONS* (3d ed. 1996).

289. Thomas L. Friedman, *This Is Just the Start*, N.Y. TIMES (March 2, 2011), <http://www.nytimes.com/2011/03/02/opinion/02friedman.html>.

Of course, it would be simple for a government to require that a system such as Google Earth alter its source code to prohibit the public from accessing views of particular geographic locations. For all the reasons previously discussed, the First Amendment should cover such measures, even if they only regulate the underlying source code and not individual usage.²⁹⁰

Additionally, free usage of Internet media furthers the value of democratic self-governance as articulated in the participatory model. These media allow all types of individuals—including those who, in the past, did not have a voice in public discourse—to post and share their experiences and opinions regarding all manner of public policy issues. Individuals from all walks of life, as long as they can get their hands on a device connected to the Internet—a friend’s smartphone will suffice—can participate in a meaningful public dialogue and feel like they have an opportunity to make a difference. For example, “The ability to publicly disseminate . . . videos [shot by citizen journalists] has also empowered citizens against oppressive governments.”²⁹¹ This is a huge step forward for democracy.

Facebook and Twitter have also fundamentally altered the manner in which individuals participate in public discourse by combining their power with the mobility of smartphones and the precision of Global Positioning Satellite Systems (GPS). To understand this point, one need only consider the effectiveness of groups of protestors armed with nothing but an Internet- and GPS-enabled smartphone. In an article published in *Wired*, Bill Wasik explains that a way in which governments have traditionally dealt with gath-

290. It is worth remembering at this point that First Amendment coverage does not always translate into First Amendment protection. So, for example, if the U.S. Department of Defense were concerned with individuals accessing pictures of sensitive sites, such as nuclear missile silos or military bases, it could still make an argument that regulation of Google Earth would be justified in such cases by a compelling government interest in national security. The same would be true regarding Google Earth publication of pictures of active theaters of war around the world or right here at home.

291. Yu, *supra* note 243, at 897 (citing Jennifer Preston & Brian Stelter, *Cellphone Cameras Become World’s Eyes and Ears on Protests Across the Middle East*, N.Y. TIMES, Feb. 19, 2011, at A11 (“For some of the protesters facing Bahrain’s heavily armed security forces in and around Pearl Square in Manama, the most powerful weapon against shotguns and tear gas has been the tiny camera inside their cellphones. By uploading images of . . . violence in Manama, the capital, to Web sites like YouTube and yFrog, and then sharing them on Facebook and Twitter, the protesters upstaged government accounts and drew worldwide attention to their demands.”)).

ering crowds of protesters is by separating them.²⁹² In many cases, “in the pre-cell-phone era . . . overall numbers didn’t matter one bit if you could not keep physically connected.”²⁹³ Proximity in space sold at a premium. “Step out of the phalanx . . . and you might never find your fellows again; in the meantime, the opposing mob might find you alone.”²⁹⁴ So all the government had to do to defuse a protest was use the old, tried and true approach: divide and conquer. Facebook, Twitter, smartphones, and GPS have changed the game.

By allowing members of a protest to maintain close informational proximity, these new technologies eliminate the need for constant physical proximity. Protestors can always regroup and can do so almost immediately. All they need to do is set new GPS (and time) coordinates and share them through their smartphones, in real-time, and at a distance.²⁹⁵ “Today, . . . a crowd’s power is amplified by the fact that its members can never really get separated. A crowd that’s always connected can never really be dispersed. It’s always still out there.”²⁹⁶ The consequent empowerment of previously powerless groups is awesome. Previously unsophisticated constituencies can become organized in a hurry. According to Wasik, “What’s really revolutionary about all these gatherings—what remains both dangerous and magnificent about them—is the way they represent a disconnected group getting connected, a mega-underground casting off its invisibility to embody itself, formidably, in physical space.”²⁹⁷

292. See Bill Wasik, *Crowd Control*, WIREd, January 2012, at 76, 112.

293. *Id.*

294. *Id.*

295. To the government forces trying to contain them, these protestors must seem almost as perturbing and disconcerting as quantum nonlocality, quantum entanglement, and the phenomenon informally known as “spooky action at a distance” did to physicists when they were first theorized and observed.

296. Wasik, *supra* note 292, at 112.

297. *Id.* As this quote suggests, of course, the emergence of what has been termed “flash mobs,” *id.* at 80, can lead to problematic situations, such as highly efficient and hard to control rioting. When a peaceful protest turns into a violent riot, social media can become a rioter’s best weapon. Much in the same way that a telephone can be used to “put out a hit” on someone, these new technologies can be used in criminal enterprises. This problem, however, is one of First Amendment *protection* and the control of the particular use that some ill-intentioned individuals or groups might find for their smartphones. No one is arguing that telephone conversations be totally excluded from First Amendment *coverage*. Instead we should be vigilant that governments do not use such bad conduct to stifle the legitimate use of these technologies in ways that greatly further First Amendment values.

The possibility of government regulation of these new media is not mere mirage. For example, China's tug of war with Google over their censorship of search results has been widely covered by the national press.²⁹⁸ "Google's decision to team up with the Chinese government and provide Chinese-Google users search results that match the preferences of the Chinese political leadership" has also been widely criticized.²⁹⁹ These government censorship programs restricting citizens' access to Internet sources, of course, would be deemed unconstitutional prior restraints under First Amendment doctrine if they were attempted here by the federal or state governments. However, if source code were deemed not covered by the First Amendment, such censorship might avoid constitutional scrutiny if it was embedded in direct regulation of the underlying code.

Quite recently, "China started a sweeping crackdown of its vibrant social networking media . . . , detaining six people, closing 16 Web sites and shutting off the comment function for two gigantic microblog services."³⁰⁰ This crackdown resulted from "the political instability that has gripped China since one of its most charismatic politicians, Bo Xilai, lost his post in March."³⁰¹ That, in turn, "spurred rumors of a coup, which the government-run Xinhua news agency cited as the reason for the measures."³⁰² Similarly, in Egypt, "Virtually all internet access . . . [was] cut off . . . as the government battle[d] to contain the street protests that [eventually] topple[d] President Hosni Mubarak[']s" totalitarian regime.³⁰³

298. See, e.g., Miguel Helft & David Barboza, *Google Shuts China Site in Dispute Over Censorship*, N.Y. TIMES, (Mar. 22, 2010), <http://www.nytimes.com/2010/03/23/technology/23google.html> ("Just over two months after threatening to leave China because of censorship and intrusions from hackers, Google on Monday closed its Internet search service there and began directing users in that country to its uncensored search engine in Hong Kong."); Aaron Smith, *China Renews Google License, Ending Standoff*, CNNMONEY (July 9, 2010), http://money.cnn.com/2010/07/09/technology/google_china/index.htm ("Google said Friday that it has renewed its license with the Chinese government to continue operating in that country, ending a standoff over censorship.")

299. Mayer-Schönberger, *supra* note 110, at 725 (citing LESSIG, *CODE*, *supra* note 236, at 80).

300. Ian Johnson, *China Limits Online Discussion Over Rumors*, N.Y. TIMES (Mar. 31, 2012), <http://www.nytimes.com/2012/04/01/world/asia/china-shuts-down-web-sites-after-coup-rumors.html>.

301. *Id.*

302. *Id.*

303. Christopher Williams, *How Egypt Shut Down the Internet*, TELEGRAPH (Jan. 28, 2011), <http://www.telegraph.co.uk/news/worldnews/africaandindianocean/egypt/8288163/How-Egypt-shut-down-the-internet.html>.

Again, the possibility of the government here in the United States requiring that some form of “kill switch” be embedded in the source code of particular websites, or in the code architecture of the Internet at large, should certainly give us pause. Such a formidable source of government power must be subject to scrutiny under First Amendment doctrine.

It should be noted that it is not only repressive regimes in lands that sound far away from our “western sensibilities” that have taken or considered these types of government actions. Here in the United States a robust public debate about First Amendment concerns, among other things, recently brought to a screeching halt the advance of the Stop Online Piracy Act (SOPA),³⁰⁴ a bill introduced by U.S. Representative Lamar S. Smith,³⁰⁵ and its equivalent in the Senate, the PROTECT IP Act (PIPA).³⁰⁶ These bills sought to grant expanded powers to federal law enforcement agencies to combat online trafficking in copyrighted intellectual property and counterfeit goods.³⁰⁷ The bills’ opponents argued that some of their provisions threatened First Amendment liberties.³⁰⁸ These provisions allowed law enforcement to block entire Internet domains because of a single instance of infringement, eliminated the “safe harbor” provisions in the Digital Millennium Copyright Act granting immunity from liability to Internet sites, and required search engines to delete domain names from their search results.³⁰⁹ A more sophisticated attempt by Congress to achieve the same objectives of SOPA and PIPA, but couched in regulation of the computer code itself, would avoid constitutional scrutiny if it were determined that source code is not covered by the First Amendment.

Another warning sign of potentially troubling government action took place in the United Kingdom after the 2011 riots that

304. See Jonathan Weisman, *After an Online Firestorm, Congress Shelves Antipiracy Bills*, N.Y. TIMES (Jan. 20, 2012), <http://www.nytimes.com/2012/01/21/technology/senate-postpones-piracy-vote.html>; Julianne Pepitone, *SOPA and PIPA Postponed Indefinitely After Protests*, CNNMONEY (Jan. 20, 2012), http://money.cnn.com/2012/01/20/technology/SOPA_PIPA_postponed/index.htm.

305. H.R. 3261, 112th Cong. (2011).

306. S. 968, 112th Cong. (as introduced in Senate, May 12, 2011).

307. See generally H.R. 3261; S. 968.

308. See, e.g., Brendan Sasso, *Legal Expert Says Online Piracy Bill Is Unconstitutional*, HILL (Dec. 11, 2011, 07:05 AM ET), <http://thehill.com/blogs/hillicon-valley/technology/198575-legal-expert-online-piracy-bill-is-unconstitutional> (summarizing Professor Laurence Tribe’s memorandum to Congress arguing that the Stop Online Piracy Act (SOPA) would violate the First Amendment).

309. See, e.g., *id.*

spread through some of its main cities. As Bill Wasik reported, “In the aftermath of the UK riots, the proposals floating around Parliament sounded . . . intrusive.”³¹⁰ “Representatives of Facebook and Twitter were called in to discuss emergency plans to throttle their services. Research in Motion, the maker of BlackBerry, has promised (or so it has been reported) that it would halt BBM if riots happened again.”³¹¹ To quell the social unrest, Parliament was asking these *private* media providers to block the people’s speech in the government’s stead.

While such measures may very well be justified in some circumstances, we should be very concerned about the possibility of our government regulating source code to create broad, undifferentiated, and centralized control mechanisms to shut the population out of these new media. Such regulation would be a serious threat to our First Amendment freedoms.³¹² Hence any such attempt at regulating computer source code must be subject to First Amendment scrutiny.

As explained above, new technologies and new media in the Web 2.0 have created new, cheap, and highly effective ways for individual citizens to organize themselves collectively in ways that promote democratic self-governance. These new media have given new, more sensitive eyes and ears to those who were blind and deaf to the abuses of the rich and powerful. They have also given the powerless a way to have their voices heard, in many cases for the first time. But all of these tools for positive change are at the mercy of those who would control the underlying code that defines their availability, their shape, and their very existence. We must be vigilant and ensure that regulation of the source code that gives the Internet life is subjected to the purifying light of First Amendment scrutiny.

c. Autonomy

The First Amendment is also profoundly concerned with an individual’s autonomy, self-fulfillment, or self-realization.³¹³ In many ways, the open source movement embodies this First Amendment value. Programmers involved in the open source movement

310. Wasik, *supra* note 292, at 113.

311. *Id.*

312. It should also be pointed out that First Amendment concerns are not the only ones to be taken into account in this context. “Vital emergency personnel routinely rely on BBM and other smartphone services, so an outright shutdown might easily sacrifice more lives than it saves.” *Id.*

313. *See supra* note 200.

take their source code very personally. Many of them see writing code as their own kind of art form. Eric S. Raymond, one of the leading proponents of the open source movement, discusses how this applies to open source Linux programmers: “The ‘utility function’ Linux hackers are maximizing is not classically economic, but is the intangible of their own ego satisfaction and reputation among other hackers.”³¹⁴ These programmers are out to solve their own problems and to help others solve theirs because they *like* writing source code. They express themselves and their own individuality through their way of writing code. Why else would they spend countless hours of their lives writing programs and then give them away for free? They are not pure cyberspace altruists. They derive great satisfaction out of expressing themselves through their source code. This is why they attach their personal reputations to the code. *The code is their art.*

Furthermore, open source programmers express, in English, their intentions of making a political statement through their source code. By keeping their source code open they are saying something about what they think the new frontier of cyberspace should look like. An excerpt from the Free Software Foundation’s web site makes this clear:

The term “free software” is sometimes misunderstood—it has nothing to do with price. It is about freedom. Here, therefore, is the definition of free software.

A program is free software, for you, a particular user, if:

- You have the freedom to run the program as you wish, for any purpose.
- You have the freedom to modify the program to suit your needs. (To make this freedom effective in practice, you must have access to the source code, since making changes in a program without having the source code is exceedingly difficult.)
- You have the freedom to redistribute copies, either gratis or for a fee.
- You have the freedom to distribute modified versions of the program, so that the community can benefit from your improvements.³¹⁵

Open source programming is not just a business model, but an expression of autonomy. Programmers are making statements

314. Raymond, *supra* note 228, at 22.

315. Richard Stallman, *The GNU Project*, GNU OPERATING SYSTEM, <http://www.gnu.org/gnu/thegnuproject.html> (last updated Sep. 13, 2012).

through their source code about how they think their world should be.

The programmers' self-expression is embodied not only in their licensing polices, but also in the code itself. Imagine that Congress passed a law that required that all software include in its code certain technical protection measures to prevent people from copying it. Forcing open source programmers to write into their code these protections would undermine the statements about free software that they are trying to make. Such a regulation of the code itself would impinge on the programmers' autonomy: it would change the way in which programmers express their individuality, establish a reputation, and distinguish themselves from others. Source code is the way programmers expound their ideas about what the world should look like, and it thereby furthers the values of autonomy and self-realization that the First Amendment embodies.

On the other hand, the value of source code is not limited to furthering programmers' autonomy. This has become ever more evident with the advent of websites that allow for the posting and sharing of user-generated content, such as YouTube, Facebook, and Twitter.³¹⁶ Just as the amount of political speech posted on these media has ballooned, the quantity of *artistic and personal expression* shared by individuals on the Internet has also exploded.³¹⁷

In his discussion about the influence of new Internet technologies on the production of copyrighted works, Professor Peter K. Yu points out that “[t]he arrival of new digital technologies and social

316. See Debora Halbert, *Mass Culture and the Culture of the Masses: A Manifesto for User-Generated Rights*, 11 VAND. J. ENT. & TECH. L. 921, 923 (2009) (explaining the history and effect of user-generated content).

317. See Olufunmilayo B. Arewa, *YouTube, UGC, and Digital Music: Competing Business and Cultural Models in the Internet Age*, 104 NW. U. L. REV. 431, 432–33 (“On the creation side, estimates suggest that the number of UGC content creators will rise from 83 million in 2008 to 115 million by 2013. On the viewer side, UGC websites are increasingly becoming dominant locales for the consumption of content. YouTube, for example, had over 112 million U.S. viewers with 6.6 billion videos viewed in January 2010. Facebook also experienced explosive growth in the 2000s, with the number of active Facebook users growing to more than 400 million worldwide by 2010 User statistics and valuation figures for Facebook and other UGC websites attest to the potential future growth of UGC more generally.” (footnotes omitted)); Daniel Gervais, *The Tangled Web of UGC: Making Copyright Sense of User-Generated Content*, 11 VAND. J. ENT. & TECH. L. 841, 845–46 (2009) (“Hundreds of millions of Internet users are downloading, altering, mixing, uploading, and/or making available audio, video, and text content on personal web pages, social sites, or using peer-to-peer technology to allow others to access content on their computer.”).

networking platforms has opened the door for the public to actively participate in cultural production.”³¹⁸ Yu quotes Professor Balkin’s argument that democratic cultural participation is important for two reasons:

First, culture is a source of the self. Human beings are made out of culture. A democratic culture is valuable because it gives ordinary people a fair opportunity to participate in the creation and evolution of the processes of meaning-making that shape them and become part of them; a democratic culture is valuable because it gives ordinary people a say in the progress and development of the cultural forces that in turn produce them.

Second, participation in culture has a constitutive or performative value: When people are creative, when they make new things out of old things, when they become producers of their culture, they exercise and perform their freedom and become the sort of people who are free. That freedom is something more than just choosing which cultural products to purchase and consume; the freedom to create is an active engagement with the world.³¹⁹

Furthermore, the “viral” nature of the spread of popular user-generated content is redefining the boundaries of both celebrity and artistic expression.³²⁰ No longer does an aspiring musician need the good auspices of a big-name, corporate record company to make his work known. He need only record himself playing his music and post the video on YouTube. If people like it, they will distribute it to others. This phenomenon has already catapulted myriad unknown artists to international success or celebrity.³²¹

318. Yu, *supra* note 243, at 895.

319. *Id.* at 896 n.67 (quoting Balkin, *supra* note 259, at 35).

320. For an interesting discussion on some of the multiple arguments made by content owners, content creators, and users regarding the influence of new Internet technologies on the production of copyrighted works, see Yu, *supra* note 243.

321. See Whitney Baker, *Five Celebrities Who Got Their Start on YouTube*, PASTE MAG. (June 5, 2010, 7:00 AM), <http://www.pastemagazine.com/blogs/lists/2010/06/-celebrities-who-started-on-youtube.html>; Claudine Beaumont, *YouTube: Top 10 Celebrities*, TELEGRAPH (Nov. 26, 2008), <http://www.telegraph.co.uk/technology/3527671/YouTube-top-10-celebrities.html>; Amanda Greene, *10 YouTube Success Stories*, WOMAN’S DAY, <http://www.womansday.com/life/10-youtube-success-stories-108958> (last visited Nov. 12, 2012); *List of YouTube Personalities*, WIKIPEDIA, FREE ENCYCLOPEDIA, http://en.wikipedia.org/wiki/List_of_YouTube_personalities (last modified Nov. 9, 2012); Cheryl L. Slay, *MySpace or Whose Space?*, MD. B.J., Jan.–Feb. 2008, at 16 (“For example, former receptionist Brooke Brodack was signed to a deal with NBC’s Carson Daly show after her appearance on [YouTube], and sev-

Justin Bieber was discovered on YouTube!³²²

One type of user-generated content that has become quite prevalent in sites such as YouTube is the parody.

[M]any YouTube videos are themselves parodies or are posted in order to permit users to append satirical, informative, or critical commentary about the videos or their stars. Readers doubting this fact are invited to type into YouTube's search box the name of a famous politician or cultural icon in their community or nation.³²³

Parodies, of course, receive quite robust First Amendment protection.³²⁴ But parodies often make fun of other works; oftentimes *copyrighted* works. So it is only natural that copyright holders (and the big business intermediaries who publish and distribute such copyrighted works) would grow concerned about the unauthorized use of their works in user-generated content. Nevertheless, a great number of such parodic uses would be covered by the First Amendment or by the Copyright Act's own fair use defense.³²⁵

Nevertheless, the government could choose, under the constant pressure and lobbying from the movie and music industries, to attempt to regulate YouTube's underlying code and require that any user-generated content that made use of copyrighted works be automatically blocked from being shared on the site. The government could also choose to force the artists themselves to include certain code in their copyrighted digital works that would help sites like YouTube automatically recognize copyrighted material. Absent First Amendment coverage of source code, both of these types of regulation would escape constitutional scrutiny. This would result, on the one hand, in the trampling of parodists' First Amendment and fair use rights and, on the other hand, in a limitation to some artists' autonomy and potential choice to keep their works "open

eral YouTube celebrities have received record deals."); Bob Tedeschi, *New Hot Properties: YouTube Celebrities*, N.Y. TIMES (Feb. 26, 2007), <http://www.nytimes.com/2007/02/26/technology/26ecom.html>; William Wei, *Meet the YouTube Stars Making \$100,000 Plus per Year*, BUS. INSIDER (Aug. 19, 2010), <http://www.businessinsider.com/meet-the-richest-independent-youtube-stars-2010-8?op=1>.

322. Desiree Adib, *Pop Star Justin Bieber Is on the Brink of Superstardom*, ABC NEWS (Nov. 14, 2009), <http://abcnews.go.com/GMA/Weekend/teen-pop-star-justin-bieber-discovered-youtube/story?id=9068403>.

323. Hannibal Travis, *Opting Out of the Internet in the United States and the European Union: Copyright, Safe Harbors, and International Law*, 84 NOTRE DAME L. REV. 331, 399–400 (2008).

324. *Hustler Magazine, Inc. v. Falwell*, 485 U.S. 46, 57 (1988).

325. *See, e.g., id.*; *Campbell v. Acuff-Rose Music*, 510 U.S. 569, 579 (1994) (holding commercial parody protected under fair use doctrine).

sourced” and available for copying by others (much in the same way that government regulation could limit open source programmers’ autonomy).

A similar situation could occur with respect to user-generated content that includes a portion of a copyrighted work as part of a personal fair use of such work.³²⁶ A prime example of this arose in litigation concerning an individual who “uploaded to YouTube a film of her young children dancing in the kitchen to the song ‘Let’s Go Crazy’ by Prince.”³²⁷ The record company then sent a takedown notice to YouTube, allegedly at Prince’s behest, under the pertinent provisions of the Digital Millennium Copyright Act.³²⁸ The record company claimed that the posting of the video violated copyright law.³²⁹ “After YouTube removed the video, the plaintiff filed suit against [the record company] for issuing the takedown notice in bad faith because it did not consider whether the video would fall under the fair use doctrine.”³³⁰ Eventually, the district court in *Lenz* “held that before issuing a DMCA takedown notice, a content owner must consider whether the content falls into the fair use doctrine.”³³¹ “[I]n order for a copyright owner to proceed under the DMCA with ‘a good faith belief that use of the material in the manner complained of is not authorized by the copyright owner, its agent, or the law,’ the owner must evaluate whether the material makes fair use of the copyright.”³³² “The court held that a takedown notice is not sent in good faith unless the content owner first considers the fair use doctrine.”³³³

Social networking media have also given rise to the Internet meme. An Internet meme has been defined as “a humorous image, video, piece of text, etc. that is copied (often with slight variations) and spread rapidly by Internet users,”³³⁴ or “a concept that spreads

326. In fact, it is not hard to imagine similar situations arising in other fair use contexts, such as in an academic setting.

327. Hormann, *supra* note 243, at 1356 (citing *Lenz v. Universal Music Corp.*, 572 F. Supp. 2d 1150, 1151–52 (N.D. Cal. 2008)).

328. *Lenz*, 572 F. Supp. 2d at 1152–53.

329. *Id.* at 1152.

330. Hormann, *supra* note 243, at 1356 (citing *Lenz*, 572 F. Supp. 2d at 1153).

331. *Id.* (citing *Lenz*, 572 F. Supp. 2d at 1154).

332. *Lenz*, 572 F. Supp. 2d at 1154 (quoting 17 U.S.C. §512(c)(3)(A)(v) (2006)). See generally 17 U.S.C. §107 (2006) (describing the parameters of the fair use doctrine).

333. Hormann, *supra* note 243, at 1356 (citing *Lenz*, 572 F. Supp. 2d at 1154).

334. *meme*, OXFORD DICTIONARIES, <http://oxforddictionaries.com/definition/english/meme> (last visited Nov. 12, 2012).

via the Internet.”³³⁵ But how are Internet memes created? Someone must think them up in the first place, of course. And then that initial creator will set the Internet meme loose in cyberspace, where myriad users will make it their own, twisting and turning it, adding to and subtracting from it, until it becomes a part of our culture.

Creative reuse and modification of preexisting materials . . . are highly valuable to society. They ensure that “everyone—not just political, economic, or cultural elites—has a fair chance to participate in the production of culture, and in the development of the ideas and meanings that constitute them and the communities and subcommunities to which they belong.”³³⁶

The creation of an Internet meme enables an expression of autonomy for both the creator of the meme and the Internet user who spreads it through Internet culture. They surely must, at least in some cases, grow quite proud of their contribution to human culture.

But one need not receive international recognition to exercise one’s right to self-realization. The Internet also provides new tools for social interactions that may not be as grandiose as those described above, but that nevertheless provide spaces for individuals to express their autonomy. The blogosphere again appears to be tailored for such individual self-expression:

The power of the blogosphere has yet to be fully demonstrated and grasped. To say it is immense is a gross understatement. The blogosphere is *a universe including all the people in the world who want to be published*, who prior to the Internet would not have been able to get published easily if at all, and *who now can be published to their hearts’ content* as long as they can gain access to a computer. It may be unduly colorful to say that this is an informational counterpart of the unlocking of nuclear power, but it is not inaccurate.³³⁷

335. *Internet Meme*, WIKIPEDIA, FREE ENCYCLOPEDIA, http://en.wikipedia.org/wiki/Internet_meme (last modified Nov. 12, 2012). For a list of Internet memes, see *List of Internet Phenomena*, WIKIPEDIA, FREE ENCYCLOPEDIA, http://en.wikipedia.org/wiki/List_of_Internet_memes (last modified Nov. 7, 2012). The concept derives from the term “meme,” which was coined some decades ago by British scientist Richard Dawkins. See *Internet Meme*, *supra*. A meme is “an element of a culture or system of behavior that may be considered to be passed from one individual to another by nongenetic means, especially imitation.” *meme*, *supra* note 334.

336. Yu, *supra* note 243, at 896 (quoting Balkin, *supra* note 259, at 4).

337. Slabbert, *supra* note 278, at 355 (emphasis added).

But as a perfect platform for people of all walks of life to share their ideas and talents, blogs are hardly alone. Facebook is also a paradise of self-expression.

Facebook has become one of the most popular websites in the world. In 2009, Facebook became the second most-visited website on the Internet. Social networking sites like Facebook enable users to *create profiles about themselves that other users are able to view*. Users can communicate with one another by sending private messages or by posting public messages on the profiles of other users. Additionally, users can create or join groups that focus on particular or common interests, or create invitations for events, parties, and informal gatherings.

Facebook also permits users to *upload photographs of themselves* and others onto the site, and allows users to “tag,” or identify, people in the posted photos, which can then be accessed from the profile of a “tagged” user. Many Facebook users have hundreds of photos of themselves posted on the site. In addition, *users can post their current “status” to communicate plans, thoughts, or quips*. The statuses, along with all other recent activity undertaken by the user on Facebook, appear both in the user’s profile and in a “news feed” that all friends of that user see when they log into the site.³³⁸

The Facebook profile, status, and news feed have become a central tool for autonomous self-expression in our contemporary society, and they are all utterly dependent on code. Government regulation of that code could serve to frustrate individuals’ exercise of their autonomy.

d. Community

Finally, let us consider how source code can either build or destroy the community that the First Amendment seeks to protect. When considering its destructive potential, source code will be compared with other categories of speech to which the Supreme Court has extended only the most limited of First Amendment coverage and protections. But first, let us briefly consider how source code is central to the development of new technologies and media that actually strengthen the glue that ties our communities together.

338. Bryan Starrett, *Tinker’s Facebook Profile: A New Test For Protecting Student Cyber Speech*, 14 VA. J.L. & TECH 212, 215 (2009) (emphasis added) (footnotes omitted).

One social networking site not yet mentioned in this article is LinkedIn. “As of February 9, 2012, LinkedIn operates the world’s largest professional network on the Internet with more than 150 million members in over 200 countries and territories.”³³⁹ LinkedIn provides its users with the opportunity to establish professional connections with all manner of individuals around the world, helping them to establish a community of potential business partners, clients, employers, and employees. By allowing people to exponentially grow their professional networks, LinkedIn fosters economic efficiencies and interdependencies that bring the community together, making the interchange of both ideas and goods more likely. A tighter knit economic and professional community promotes a more stable, peaceful, profitable, and amicable social order.

Something similar can also be said of Facebook and Twitter, of course. It is widely recognized that “[s]ocial networking sites provide benefits for users of all ages and backgrounds.”³⁴⁰ For one thing, these sites “allow users the ability to reconnect with old friends and make new friends,” thereby engendering a wider sense of community.³⁴¹ “In addition, most social networking sites are ‘global,’ which provides for diverse relationships,” and can lead to the creation of a more tolerant society and a more global sense of community.³⁴²

Furthermore, other Internet media such as YouTube can inject “an important social element often missing from passive media, such as movies, television, music, and books.”³⁴³ After all, “[s]ocialization is one of the reasons why YouTube, Facebook, Twitter, and Tumblr have become wildly popular today.”³⁴⁴ It is only natural that “[g]iven the choice between watching an unfamiliar program put together professionally by an entertainment firm and

339. *About*, LINKEDIN, <http://press.linkedin.com/about> (last visited Nov. 12, 2012).

340. Shannon N. Sterritt, Comment, *Applying the Common-Law Cause of Action Negligent Enablement of Imposter Fraud to Social Networking Sites*, 41 SETON HALL L. REV. 1695, 1696 (2011).

341. *Id.* (citing Karen Goldberg Goff, *Social Networking Benefits Validated*, WASH. TIMES (Jan. 28, 2009), <http://www.washingtontimes.com/news/2009/jan/28/social-networking-benefits-validated>).

342. *Id.* (citing WHAT IS SOCIAL NETWORKING, <http://www.whatissocialnetworking.com/> (last updated July 18, 2012)).

343. Yu, *supra* note 243, at 898.

344. *Id.* (citing JEAN BURGESS & JOSHUA GREEN, *YOUTUBE: ONLINE VIDEO AND PARTICIPATORY CULTURE* 58–74 (2009) (discussing YouTube as a social network); STRANGELOVE, *supra* note 272, at 103–36 (discussing the YouTube community)).

a few short videos involving the user's friends goofing around, some users undoubtedly will select the latter."³⁴⁵ "Even if the homemade videos are of lower quality, the users' familiarity with the subject and their interest in what happens to their friends will make up for the difference."³⁴⁶ "Frankly, you would have to be dead inside not to find something emotionally or intellectually compelling on YouTube. After all, it is you, it is me, it is our neighbours, our families, our friends (and, all too often, our darn kids) who can be seen on YouTube."³⁴⁷

Professor Sunstein has argued that "some of the experiences made possible by modern technologies are *solidarity goods*, in the sense that their value goes up when and because many other people are enjoying or consuming them."³⁴⁸ He further explains that such experiences are desirable and important for three principal reasons: (1) they give individuals "simple enjoyment"; (2) they "provide a form of social glue" and create "common memories and experiences, and a sense of a common enterprise"; and (3) they allow "people who would otherwise see one another as quite unfamiliar" to "come instead to regard one another as fellow citizens with shared hopes, goals, and concerns."³⁴⁹ Professor Sunstein describes the stock of the ensuing "relationships of trust and reciprocity, in which people see their fellow citizens as potential allies, willing to help and deserving of help when help is needed," as a form of "social capital."³⁵⁰ Consider, in this context, this recent description of the creation of a collective identity in the electronic dance music scene:

In the past decade or so, though, despite all the ways that the Internet encourages music to nichify, the rise of social media has actually pushed electronic dance music in the opposite direction. Witnessing its sheer numbers, sensing its collective power, the dance scene has reunified, becoming more of a mass phenomenon—an undifferentiated subculture of millions. *It turns out that the thrill of collective identity, a moblike feeling of shared enormity, is far more exciting to fans than were their endless dives down rabbit holes of sonic purism.*³⁵¹

345. *Id.*

346. *Id.*

347. STRANGELOVE, *supra* note 272, at 3.

348. SUNSTEIN, *supra* note 3, at 102.

349. *Id.* at 103–04.

350. *Id.* at 104 (quoting ROBERT D. PUTNAM, *BOWLING ALONE: THE COLLAPSE AND REVIVAL OF AMERICAN COMMUNITY* 18–24 (2000)) (internal quotation marks omitted).

351. Wasik, *supra* note 292, at 83 (emphasis added).

It is precisely the sense of community that these new Internet technologies engender that has made them so popular. Thus social networking opportunities provided by the development of source code and the Internet also further the First Amendment value of community.

To determine whether source code can perform a negative function and destroy the community that the First Amendment seeks to protect one must answer the following question: how does source code compare with other types of activities that are communicative enough to pass the *Spence-Hurley* test but are, *because of their tendency to destroy the community*, nevertheless granted only very limited First Amendment coverage? The short answer is that source code is quite different.

First of all, we must understand more precisely why the Court has decided that some particular classes of speech only deserve the most limited First Amendment coverage.³⁵²

There are certain well-defined and narrowly limited classes of speech, the prevention and punishment of which have never been thought to raise any Constitutional problem. These include the lewd and obscene, the profane, the libelous, and the insulting or “fighting” words [S]uch utterances are no essential part of any exposition of ideas, and are of such slight social value as a step to truth that any benefit that may be derived from them is clearly outweighed by the social interest in order and morality.³⁵³

The value of community is important in defining some of these special types of disfavored speech. Let us now discuss these categories of disfavored speech and compare them to source code.

Obscenity was most recently defined by the Supreme Court in *Miller v. California*.³⁵⁴ Obscenity is “limited to works which, taken as a whole, appeal to the prurient interest in sex, which portray sexual conduct in a patently offensive way, and which, taken as a whole, do not have serious literary, artistic, political, or scientific value.”³⁵⁵

352. It should be noted that a potential for destruction of the community is not the only reason why some particular communicative activity will receive limited First Amendment coverage. For each type of activity, a comprehensive analysis of all the First Amendment values discussed in this article is required.

353. *Chaplinsky v. New Hampshire*, 315 U.S. 568, 571–72 (1942). *But see R.A.V. v. City of St. Paul*, 505 U.S. 377, 383–84 (1992) (footnote omitted) (explaining that even specific categories of expression traditionally thought to reside outside the auspices of the First Amendment are not “entirely invisible to the Constitution”).

354. 413 U.S. 15 (1973).

355. *Id.* at 24.

This is a very narrow definition. It has two especially salient features. The work must be “patently offensive,” and it must lack “serious literary, artistic, political, or scientific value.” We see that the Court is concerned with protecting the integrity of the community by allowing the prohibition of speech that might harm the sensibilities of its members. However, these values will only give way in the case where the speech is devoid of serious value. There is no reason to think that source code, *as a whole category of speech*, is either offensive in a way that would threaten to destroy our community, or devoid of any serious value. For one thing, almost all source code is scientifically valuable to computer scientists. Any rationale analogous to that which gives rise to the exclusion of obscenity is, therefore, inapplicable to source code as a whole.

“Fighting words” are disfavored under the First Amendment for basically the same reasons obscenity is. Fighting words are words that, as explained in *Chaplinsky*, are not an essential part of any exposition of ideas.³⁵⁶ Furthermore, they are words that threaten to destroy the basic community norms of respect towards the person.³⁵⁷ These words are not a part of the public discourse, *and they threaten to destroy this discourse itself*. The Court has defined fighting words as utterances that will imminently lead to violent confrontation, thereby *destroying the very channels of communication that the First Amendment seeks to protect*.³⁵⁸ There is no reason to include source code in this kind of category. Source code, as we discussed previously, furthers all the values embodied in the First Amendment and does not threaten to destroy the community in any specific or systematic way. If some source code can be used for malicious and destructive purposes, then that specific kind of source code might be excluded, just like some uses of the English language might be excluded. There is no good reason, however, to exclude source code as a whole category of disfavored speech.

Libel and defamation are disfavored for similar reasons. These are kinds of speech that harm innocent people without furthering any of the values previously discussed. It should be noted that under libel and defamation, only false statements of fact are relegated to limited coverage. And even then, specific intent requirements are necessary to exclude this speech from First Amendment coverage.³⁵⁹ Similar intent requirements are necessary to exclude speech that might constitute a crime from First Amendment cover-

356. *Chaplinsky v. New Hampshire*, 315 U.S. 568, 571–72 (1942).

357. *Id.* at 572.

358. *Id.*

359. *See New York Times v. Sullivan*, 376 U.S. 254, 262 (1964).

age.³⁶⁰ Such speech will be excluded only when “it is the very vehicle of the crime itself.”³⁶¹ “[W]hat these sorts of factual statements contribute to the general understanding of listeners is minimal, and the justifications for free speech that apply to speakers do not reach communications that are simply means to get a crime successfully committed.”³⁶² Some specific malicious uses of source code might amount to this kind of activity that would not be covered by the First Amendment, just like some uses of the English language will constitute the vehicle for a crime and will not be covered by the First Amendment. Source code as a whole, however, cannot be deemed defamatory or criminal in all cases.

Some of the categories of speech just discussed could further the value of autonomy. People might want to express themselves in obscene ways. And it is possible that some might claim that insults amounting to fighting words are an expression of their individuality. The Court, however, has determined that these values are outweighed by the possible harm to the community. In the extreme cases of obscenity and epithets, the normal preference for the protection of a free public discourse is suspended for the sake of the community. Some specific instances of source code might be determined to be as pernicious as the types of speech discussed in this section. *The same is true, however, of any other language.* This is why the courts have defined these narrow exceptions. There is no reason, however, to say that *all of source code* falls into any such category or to relegate it to the very limited First Amendment coverage the Supreme Court has so sparingly allotted.

Source code furthers all of the First Amendment values that have traditionally been identified by the Supreme Court as central to the constitutional protection of our freedom of speech.³⁶³ Furthermore, source code does not tend to destroy the community or

360. *Rice v. Paladin Enters., Inc.*, 128 F.3d 233, 248 (4th Cir. 1997).

361. *United States v. Varani*, 435 F.2d 758, 762 (6th Cir. 1970).

362. *Rice*, 128 F.3d at 244 (quoting KENT GREENAWALT, *SPEECH, CRIME, AND THE USES OF LANGUAGE* 85 (1989)).

363. *Accord* Balkin, *supra* note 259, at 45 (arguing that technological advances in communication can help democratize speech); Alex Colangelo & Alana Maurushat, *Exploring the Limits of Computer Code as a Protected Form of Expression: A Suggested Approach to Encryption, Computer Viruses, and Technological Protection Measures*, 51 *MCGILL L.J.* 47, 60 (2006) (“[T]he creation and dissemination of software code satisfy the principles of seeking truth, attaining individual self-fulfillment, and allowing for human-flourishing, all of which govern the freedom of expression guarantees in the [Canadian] *Charter* [of Rights and Freedoms].”); Fitzgerald, *supra* note 10, at 337 (citing the work of philosophers such as Foucault, Derrida, Baudrillard, and Heidegger to show how code can help influence public discourse and construct powerful communication tools).

channels of communication that the First Amendment seeks to protect in any systematic or all-encompassing way. For these reasons, source code is fundamentally different from the specific types of disfavored speech that the Supreme Court has excluded from full First Amendment coverage.

CONCLUSION

Computer source code is the lifeblood of the Internet. It is also the brick and mortar of cyberspace itself. As such, any control that a government can wield over code is a powerful tool for controlling the development of new technologies and idiosyncratic voices. With the advent and dramatic proliferation in the Internet of social networking media and platforms for the publication and sharing of user-generated content, the ability of individuals across the world to communicate and share ideas with each other has reached truly revolutionary dimensions.

As this article has argued, there are many reasons why First Amendment coverage should be extended to computer source code. Source code is sufficiently communicative under the *Spence-Hurley* test as either a kind of written word or as an activity that carries with it sufficient social conventions to convey messages understandable by others. Yet even if source code were considered non-speech for First Amendment purposes, its regulation would trigger First Amendment scrutiny because it is at the heart of the constitution of the Internet, a recognized medium for the communication of ideas.

Perhaps most importantly, source code furthers all the values embodied in the First Amendment without posing, as a category of speech, a threat to the community of First Amendment agents. None of the narrowly defined classes of speech that are excluded from First Amendment coverage are similar enough to source code to justify carving out a similar exception for computer code.

The awesome potential that computer source code has to empower individuals and groups all across the globe in their struggle for truth-seeking, democratic self-governance, self-realization, and community-building makes it deserving of full First Amendment coverage. Let the people speak in code.

THE BOUNDARY BETWEEN ZEALOUS ADVOCACY AND OBSTRUCTION OF JUSTICE AFTER SARBANES-OXLEY

GRETA FAILS

Introduction	397	R
I. The Legal Framework—Ethical Duties and the Sarbanes-Oxley Criminal Law Overlay	400	R
II. Key Elements and Ambiguities of Sections 1512(c) and 1519	406	R
A. Aguilar and the “Nexus Requirement”	406	R
B. Arthur Andersen and “Corruptly”	407	R
C. Aguilar, Arthur Andersen, and Sections 1512(c) and 1519	409	R
III. Lawyers Charged with Violations of the Sarbanes- Oxley Obstruction of Justice Provisions	413	R
A. The Prosecution of Lauren Stevens	416	R
B. The Broader Implications of <i>United States v.</i> <i>Stevens</i> for Lawyers	419	R
1. Overcriminalization	421	R
2. Overdeterrence	423	R
3. Prosecutorial Overreach	428	R
IV. Potential Solutions to the Problems Created by the Sarbanes-Oxley Obstruction of Justice Provisions	430	R
A. The Broad Reach of Section 1519 Must Be Constrained	431	R
1. Section 1519 Should Be Interpreted to Only Reach Conduct Endeavored with Dishonest Intent	431	R
2. Section 1519 Should Be Interpreted to Implicitly Contain a Nexus Requirement	433	R
B. Steps Lawyers and Courts Can Take to Avoid the Harsh Consequences of a Broad Interpretation of the Obstruction Provisions	434	R
Conclusion	439	R

INTRODUCTION

When does a lawyer cross the boundary between zealous advocacy and obstruction of justice? Traditionally this boundary was clear, with lawyers only prosecuted for obstruction for covering up

their own crimes or for taking their clients' cases too far, e.g., intimidating witnesses or destroying documents.¹ Lawyers had leeway to aggressively represent their clients up to this boundary pursuant to their ethical duty of zealous advocacy.² This landscape was significantly changed, however, when Congress enacted the Sarbanes-Oxley Act of 2002 and created two new obstruction of justice regulations.³ Because Congress implemented the Act in response to corporate scandals in which lawyers had played integral roles,⁴ it intended for the new obstruction provisions to bolster the arsenal of prosecutors, to ferret out obstructive conduct that previously was not policed, and to remove constraints imposed by the Supreme Court.⁵ The broad and ambiguous language of the statute, coupled with unsettled legal doctrine interpreting the scope of the new provisions, has created a gray area in which lawyers are vulnerable to prosecution for conduct that was previously only characterized as zealous advocacy.⁶

The breadth of the Sarbanes-Oxley obstruction provisions was tested in 2011 when prosecutors charged a lawyer with obstruction of justice for conduct that would have previously been characterized as permissible lawyering on behalf of her client. The Department of Justice prosecuted Lauren Stevens, an attorney for GlaxoSmithKline, based on her voluntary response to a government inquiry into her client's marketing practices.⁷ *United States v. Stevens* is noteworthy because it entails the prosecution of a lawyer who was neither out for personal gain nor using the legal system to her personal advantage; this was a case in which a lawyer was simply advocating on behalf of her client. Stevens would not have been prosecuted for this conduct under the traditional obstruction of justice scheme, but Sarbanes-Oxley created a gray area, and in their discretion prosecutors determined that Stevens fell on the wrong side of the line. Stevens was ultimately acquitted,⁸ but the mere fact that prosecutors brought this case is troubling. The threat of similar future prosecutions against lawyers, coupled with the harsh penal-

1. See *infra* notes 79–82 and accompanying text.

2. See *infra* notes 11–12 and accompanying text.

3. Sarbanes-Oxley Act of 2002, Pub. L. No. 107-204, 116 Stat. 745 (2002). The two new obstruction provisions are codified at 18 U.S.C. § 1512(c) (2006) and 18 U.S.C. § 1519 (2006).

4. See *infra* notes 20–26 and accompanying text.

5. See *infra* Part II.C.

6. See *infra* Part III.B.

7. *United States v. Stevens*, 771 F. Supp. 2d 556 (D. Md. 2011).

8. See Transcript of Oral Argument at 10, *United States v. Stevens*, 771 F. Supp. 2d 556 (D. Md. 2011) (No. 10-694).

ties Sarbanes-Oxley obstruction prosecutions entail,⁹ will lead lawyers to be more risk averse in the representation of their clients, to the detriment of client representation generally. If courts and prosecutors interpret these provisions broadly, three negative consequences will inevitably result. First, traditionally innocent conduct considered zealous advocacy will be punishable, resulting in overcriminalization of attorney conduct. Second, lawyers will be disincentivized to fight for their clients when their careers and livelihoods could be jeopardized by an obstruction prosecution, resulting in overdeterrence of legitimate advocacy. Third, prosecutors will push the boundaries of their discretion and charge lawyers for pretextual or coercive reasons, resulting in prosecutorial overreach. Courts should construe these Sarbanes-Oxley obstruction provisions narrowly so that lawyers may continue to advocate for their clients without fear of criminal prosecution.

Part I lays out the ethical and legal framework with which lawyers must comply. It describes how the Model Rules create a duty to zealously advocate on behalf of clients while external regulation qualifies this duty and delineates the boundaries of permissible advocacy. It also outlines the Sarbanes-Oxley obstruction of justice provisions at issue in this Note. Part II analyzes the key elements and ambiguities of the Sarbanes-Oxley obstruction provisions based on their text and on the major Supreme Court precedent interpreting other obstruction of justice statutes. Part III delves into prosecutions of lawyers for Sarbanes-Oxley obstruction violations, focusing specifically on the prosecution of Lauren Stevens. It then discusses the implications of Sarbanes-Oxley and *United States v. Stevens* for lawyers generally, explaining that a broad interpretation of the Sarbanes-Oxley provisions will lead to overcriminalization, overdeterrence, and prosecutorial overreach. Ultimately, the incentive structure created by harsh penalties and vague obstruction provisions will lead to risk-averse lawyer behavior and less effective client representation. Part IV concludes with a discussion of how courts should construe the Sarbanes-Oxley obstruction of justice provisions in the future and, alternatively, what lawyers and courts can do to stave off negative consequences if the provisions remain broadly applicable to their conduct.

9. See *infra* notes 30 and 32 and accompanying text.

I.
THE LEGAL FRAMEWORK—ETHICAL DUTIES AND
THE SARBANES-OXLEY CRIMINAL
LAW OVERLAY

The legal profession is both internally and externally regulated. Internally the profession is governed by ethical rules promulgated by the bar and the states' highest courts.¹⁰ These ethical rules set the boundaries of permissible lawyer behavior and describe the lawyer's role in the profession. Included within this role is "the lawyer's obligation zealously to protect and pursue a client's legitimate interests, within the bounds of the law, while maintaining a professional, courteous and civil attitude toward all persons involved in the legal system."¹¹ A lawyer's ethical role can thus be seen as client-centered;¹² he has an ethical *duty* not only to advocate for his client,

10. See Ted Schneyer, *How Things Have Changed: Contrasting the Regulatory Environments of the Canons and the Model Rules*, 2008 J. PROF. LAW. 161, 161–62 (2008) [hereinafter Schneyer, *How Things Have Changed*] ("In internal regulation (often called 'professional self-regulation'), the bar, in tandem with the states' highest courts, develops, interprets, and enforces practice norms. When the ABA drafts a model legal ethics code, bar associations or court-created agencies administer the disciplinary process, and bar or court-created committees render advisory ethics opinions, they engage in internal regulation."). While the ABA Model Rules of Professional Conduct are not binding, and many states' rules differ dramatically from the Model Rules, this Note will focus on the general precepts and several specific provisions from the Model Rules when discussing ethical duties and internal regulation of the legal profession.

11. MODEL RULES OF PROF'L CONDUCT Preamble [9] (2012). The "zealously" advocate language is only found in the Preamble of the Model Rules, and the "zeal" descriptor has gradually been phased out of state ethical rules across the country. See, e.g., Paul C. Saunders, *Whatever Happened To 'Zealous Advocacy'?*, 245 N.Y. L.J., no. 47, Mar. 11, 2011, at 1, 2 (discussing the disappearance of the language "zealous advocacy" from the New York Rules of Professional Conduct). The debate over the role of "zealous" advocacy in the legal profession is beyond the scope of this Note. To the extent that this Note discusses the existence of an ethical duty of zealous advocacy, that duty is derived both from the explicit mention of zeal in the Preamble of the Model Rules, and from as the general duties across all states' codes of professional conduct to maintain client confidences and to avoid conflicts of interest. The concept of "zealous" advocacy in this context is used more as a rhetorical device to represent the general ethical duties to advocate competently, loyally, and diligently in a client's best interests within the bounds of the law.

12. See, e.g., Susan D. Carle, *Power as a Factor in Lawyers' Ethical Deliberation*, 35 HOFSTRA L. REV. 115, 116 (2006) (discussing the client-centered and justice-centered conceptions of legal advocacy). See also MODEL CODE OF PROF'L RESPONSIBILITY Canon 7 (2007) ("A Lawyer Should Represent a Client Zealously Within the Bounds of the Law"). While the Model Rules describe additional roles for lawyers, the lawyer's ethical duty to his client is predominant. MODEL RULES OF PROF'L CON-

but to do so zealously. This duty, however, is not boundless. Lawyers may not forcefully advocate for their clients to the detriment of all other interests. Lawyers are internally constrained in their advocacy by competing ethical obligations,¹³ as well as externally constrained, namely in that they may not break the law to further the interests of their clients.¹⁴ A lawyer's duty to his client is thus a delicate balance between zealous advocacy and staying within the bounds of the law.

Lawyers are increasingly governed by external rules.¹⁵ General civil and criminal laws enacted at both the state and federal levels are being used to police lawyer conduct,¹⁶ and laws specifically targeting attorney conduct are also being implemented.¹⁷ While these external constraints ensure that lawyers are held to the same legal standards as other citizens and help to promote the rule of law within the legal industry, they also limit a lawyer's ability to advocate on behalf of his client. The relationship between internal and external regulation of lawyers can therefore be a rocky one, with the two

DUCT Preamble [1] (2012) (describing a lawyer's roles as "a representative of clients, an officer of the legal system and a public citizen having special responsibility for the quality of justice").

13. An example of internal constraints can be seen in the confidentiality rules. On the one hand, lawyers have a duty to safeguard all information relating to the representation of a client. *See* MODEL RULES OF PROF'L CONDUCT R. 1.6(a) (2012). On the other, they are required, in some circumstances, to reveal confidential information to further other interests. *See, e.g., id.* R. 4.1(b) (requiring a lawyer to disclose material confidential client information to third parties "when disclosure is necessary to avoid assisting a criminal or fraudulent act by a client").

14. This prohibition is found within the Model Rules themselves, as well as in external regulations. *See id.* Preamble [9] (describing a lawyer's duty to "zealously . . . protect and pursue a client's legitimate interests, *within the bounds of the law*") (emphasis added).

15. *See* Schneyer, *How Things Have Changed*, *supra* note 10, at 162 ("The external sector consists of statutes, regulations, and judicial doctrines that are primarily interpreted and enforced outside the realm of professional discipline. External law has been growing at an accelerating rate since 1970, both in the states and especially at the federal level."); Ted Schneyer, *An Interpretation of Recent Developments in the Regulation of Law Practice*, 30 OKLA. CITY U. L. REV. 559, 566-67 (2005) [hereinafter Schneyer, *An Interpretation of Recent Developments*] ("[F]ederal regulation has expanded markedly . . . appears to be accelerating and is likely to gain momentum with further globalization of the legal services market.").

16. *See, e.g.,* James M. Fischer, *External Control over the American Bar*, 19 GEO. J. LEGAL ETHICS 59, 97-98 (2006) (discussing cases in which consumer protection laws were applied to lawyers).

17. *See, e.g.,* Bankruptcy Abuse Prevention and Consumer Protection Act of 2005, Pub. L. No. 109-8, § 226, 119 Stat. 23, 45 (2005) (codified at 11 U.S.C. §§ 101 et seq. (2006)) (including lawyers in the regulated category of "debt relief agencies").

regulatory frameworks seeking to achieve different goals and potentially conceptualizing the role of the lawyer in fundamentally different ways.¹⁸

One external constraint that is becoming increasingly relevant to lawyers is Sarbanes-Oxley.¹⁹ Enacted in 2002, Sarbanes-Oxley was a response to the corporate scandals of Enron, Arthur Andersen, and WorldCom, among others, and it was an attempt by Congress to “legislate[] ethical behavior for both publicly traded companies and their auditor firms.”²⁰ The Act included provisions requiring more transparency within publicly traded corporations and accounting and auditing firms.²¹ The goal was to hold corporations responsible to their shareholders and the public and to prevent future corporate scandals by making it harder for fraud and misconduct to persist unnoticed.²²

Although Sarbanes-Oxley was particularly aimed at accountants and corporations, Congress did not overlook lawyer involve-

18. See Schneyer, *How Things Have Changed*, *supra* note 10, at 163 (“Since at least the mid-1970s, just before the ABA began to draft the Model Rules, the relationship between the internal and external sectors has been an uneasy one. . . . [T]he sectors are in considerable tension, so much so that Susan Koniak has argued that the ‘bar’s law’ (internal) and the ‘state’s law’ (external) are grounded in very different conceptions of the lawyer’s proper role.”) (citing Susan P. Koniak, *The Law Between the Bar and the State*, 70 N.C. L. REV. 1389, 1409–27 (1992)); Schneyer, *An Interpretation of Recent Developments*, *supra* note 15, at 566–67 (“Shifting the regulatory center of gravity toward Washington is significant largely because it entails a shift toward legislative and administrative regulation, which may feature a different balance of policy concerns than has prevailed in the states, where the supreme courts, acting in tandem with the ABA and state and local bar associations, have held sway for a century.”).

19. Sarbanes-Oxley Act of 2002, Pub. L. No. 107-204, 116 Stat. 745 (2006).

20. Howard Rockness & Joanne Rockness, *Legislated Ethics: From Enron to Sarbanes-Oxley, the Impact on Corporate America*, 57 J. BUS. ETHICS 31, 31 (2005). See also H.R. REP. NO. 107-414, at 19 (2002) (“The Committee’s hearings on the Enron matter, the collapse of Global Crossing LLC, and the operations of the Nation’s capital markets all indicated that reforms were necessary both for the regulators and the regulated.”). The stated purpose of the Act was “[t]o protect investors by improving the accuracy and reliability of corporate disclosures made pursuant to the securities laws, and for other purposes.” Sarbanes-Oxley Act of 2002, Pub. L. No. 107-204, 116 Stat. 745 (2006).

21. See, e.g., *id.* §§ 401–09 (requiring enhanced financial disclosures from public companies).

22. The Act “will protect investors by improving the accuracy and reliability of corporate disclosures made pursuant to the securities laws. The bill achieves this goal through increased supervision of accountants that audit public companies, strengthened corporate responsibility, increased transparency of corporate financial statements, and protections for employee access to retirement accounts.” H.R. REP. NO. 107-414, at 16.

ment in the scandals that precipitated the Act. The public was made aware of the problem of lawyers acquiescing in and even enabling their clients' misconduct and corporate fraud²³ and responded with outrage, demanding that the law hold lawyers comparably responsible for their wrongdoing. Congress answered by enacting Sarbanes-Oxley, which imposed two new regulations on lawyers. First, the Act clarified that lawyer conduct would be scrutinized on the same level as the actions of accountants and CEOs.²⁴ Second, it deputized corporate lawyers to police the conduct of their clients and their clients' constituents.²⁵ This new federal law thus acted as a constraint on lawyer behavior and it additionally defined a new role for lawyers outside of their traditional ethical obligations.²⁶

23. See, e.g., Editorial, *Enron and the Lawyers*, N.Y. TIMES, Jan. 28, 2002, available at <http://www.nytimes.com/2002/01/28/opinion/enron-and-the-lawyers.html> ("In the Enron scandal, the accounting industry has been the profession taking the most heat. Before the dust settles, however, it seems inevitable that more questions will be raised about the role that lawyers played in Enron's alleged misdeeds.").

24. See § 703(a) (codified at 15 U.S.C. § 7201 (2006)) ("The Commission shall conduct a study to determine . . . (1) the number of securities professionals, defined as public accountants, public accounting firms, investment bankers, investment advisers, brokers, dealers, *attorneys*, and other securities professionals practicing before the Commission (A) who have been found to have aided and abetted a violation of the Federal securities laws . . .") (emphasis added); see also Stephen M. Cutler, Dir. of Enforcement Div., SEC, Speech before UCLA School of Law: The Themes of Sarbanes-Oxley as Reflected in the Commission's Enforcement Program (Sept. 20, 2004), available at www.sec.gov/news/speech/spch0920045mc.htm ("Consistent with Sarbanes-Oxley's focus on the important role of lawyers as gatekeepers, we have stepped up our scrutiny of the role of lawyers in the corporate frauds we investigate.").

25. See § 307 (codified at 15 U.S.C. § 7245 (2006)) ("Not later than 180 days after the date of enactment of this Act, the Commission shall issue rules, in the public interest and for the protection of investors, setting forth minimum standards of professional conduct for attorneys appearing and practicing before the Commission in any way in the representation of issuers, including a rule (1) requiring an attorney to report evidence of a material violation of securities law or breach of fiduciary duty or similar violation by the company or any agent thereof, to the chief legal counsel or the chief executive officer of the company (or the equivalent thereof); and (2) if the counsel or officer does not appropriately respond to the evidence (adopting, as necessary, appropriate remedial measures or sanctions with respect to the violation), requiring the attorney to report the evidence to the audit committee of the board of directors of the issuer or to another committee of the board of directors comprised solely of directors not employed directly or indirectly by the issuer, or to the board of directors.").

26. The Sarbanes-Oxley Act only applies to lawyers "appearing and practicing" before the Securities and Exchange Commission. See *id.* This phrase is interpreted broadly and all lawyers doing work for publicly traded companies could come under the reach of these provisions. See Jeffrey W. Stempel, *The Sarbanes-*

This Note will focus on two sections of Sarbanes-Oxley that were enacted for general application but that have been increasingly, and worrisomely, used to police lawyer conduct.²⁷ These sections are the obstruction of justice provisions: 18 U.S.C. § 1512(c) and 18 U.S.C. § 1519.²⁸ Section 1512(c) makes it a crime for anyone to “corruptly” destroy or conceal documents with the intent to make them unavailable in an official proceeding, or to otherwise obstruct any official proceeding.²⁹ A violation of this provision is punishable by a fine and/or up to twenty years in prison.³⁰ Section 1519 makes it a crime to “knowingly” alter, destroy, conceal, or make a false entry in any document with the intent to obstruct a federal investigation or bankruptcy case, “or in relation to or contemplation of any such matter or case.”³¹ A violation of this section is similarly punishable by a fine and/or up to twenty years in prison.³²

These two provisions substantially altered the scope of the federal obstruction of justice scheme. Prior to 2002, the Department of Justice primarily prosecuted obstruction of justice crimes under 18 U.S.C. § 1503(a)³³ and 18 U.S.C. § 1512.³⁴ Both sections are still in force today, but their scope has been significantly cabined by Supreme Court decisions, as will be discussed below.³⁵ Congress en-

Oxley Act: Lawyer Professional Responsibility, and a Heightened Role for Business Lawyers, 11 NEV. LAW. 8, 13 (2003).

27. This Note will not focus on the reporting-up provisions specifically geared toward lawyers in Section 307 of the Sarbanes-Oxley Act. These provisions have been controversial and hotly debated, but they are beyond the scope of this Note.

28. These provisions are found in sections 1102 and 802 of the Sarbanes-Oxley Act, respectively. While section 307 of the Sarbanes-Oxley Act imposes specific requirements on lawyers “appearing and practicing” before the SEC, these obstruction of justice provisions are of general application and apply to anyone, including *all* lawyers.

29. 18 U.S.C. § 1512(c) (2006).

30. *Id.*

31. 18 U.S.C. § 1519 (2006).

32. *Id.*

33. 18 U.S.C. § 1503 (2006) (“Whoever corruptly, or by threats or force, or by any threatening letter or communication, endeavors to influence, intimidate, or impede any grand or petit juror, or officer in or of any court of the United States . . . or corruptly or by threats or force, or by any threatening letter or communication, influences, obstructs, or impedes, or endeavors to influence, obstruct, or impede, the due administration of justice, shall be punished as provided in subsection (b).”). Section 1503 is often referred to as the “omnibus” obstruction statute or “catchall” provision because of its general language and broad applicability. *See infra* note 44 and accompanying text.

34. 18 U.S.C. § 1512(b) (2006).

35. *See infra* Part II.

acted sections 1512(c) and 1519 in Sarbanes-Oxley to broaden the reach of the obstruction of justice provisions.³⁶ Previously section 1512 only authorized prosecution for anyone who corruptly persuaded *another person* to destroy or alter documents to make them unavailable in an official proceeding—the Sarbanes-Oxley addition of section 1512(c) criminalizes destruction of documents by persons *acting alone*.³⁷ Section 1512(c) also increases the penalty for a violation of this provision.³⁸ Whereas section 1512(c) could thus be considered a minor expansion of the previous obstruction statutes, section 1519 is a major departure from the previous scheme. Section 1519 significantly broadens the reach of the omnibus obstruction provision by making it clear that it not only applies to obstructive actions taken against formal, pending government proceedings, but that it also applies to actions taken against informal investigations and “in relation to or contemplation of” a future proceeding.³⁹ The changes to the obstruction of justice statutes made

36. S. REP. NO. 107-146, at 14 (2002) (“[T]he current laws regarding destruction of evidence are full of ambiguities and technical limitations that should be corrected. [These] provision[s] [are] meant to accomplish those ends.”). Congress, despite wanting to broaden the reach of the obstruction of justice provisions, maintained a carve-out for the lawful provision of legal services. See 18 U.S.C. § 1515(c) (2006) (the “Safe Harbor Provision”) (“This chapter does not prohibit or punish the providing of lawful, bona fide, legal representation services in connection with or anticipation of an official proceeding.”). As will be discussed more fully herein, this exception does little to protect lawyers accused of ambiguously unlawful conduct under the broadened obstruction of justice provisions.

37. Compare 18 U.S.C. § 1512(b) (2006) (“Whoever knowingly uses intimidation, threatens, or corruptly persuades *another person*, or attempts to do so, or engages in misleading conduct toward another person, with intent to . . . (2) cause or induce any person to . . . (A) withhold testimony, or withhold a record, document, or other object, from an official proceeding; (B) alter, destroy, mutilate, or conceal an object with intent to impair the object’s integrity or availability for use in an official proceeding; (C) evade legal process summoning that person to appear as a witness, or to produce a record, document, or other object, in an official proceeding . . . shall be fined under this title or imprisoned not more than *ten years*, or both.”) (emphasis added), with 18 U.S.C. § 1512(c) (2006) (“Whoever corruptly (1) alters, destroys, mutilates, or conceals a record, document, or other object, or attempts to do so, with the intent to impair the object’s integrity or availability for use in an official proceeding; or (2) otherwise obstructs, influences, or impedes any official proceeding, or attempts to do so, shall be fined under this title or imprisoned not more than *20 years*, or both.”) (emphasis added).

38. See 18 U.S.C. § 1512(b) (2006) (punishing a violation of the act with a fine and/or ten years in prison).

39. 18 U.S.C. § 1519 (2006) (“Whoever knowingly alters, destroys, mutilates, conceals, covers up, falsifies, or makes a false entry in any record, document, or tangible object with the intent to impede, obstruct, or influence the investigation or proper administration of *any matter* within the jurisdiction of any department or agency of the United States or any case filed under title 11, or *in relation to or*

by the Sarbanes-Oxley additions are troublesome and require further elaboration.

II. KEY ELEMENTS AND AMBIGUITIES OF SECTIONS 1512(C) AND 1519

The Supreme Court has not yet ruled on the Sarbanes-Oxley obstruction of justice provisions. It did, however, interpret the obstruction of justice framework prior to the Sarbanes-Oxley additions,⁴⁰ and it also interpreted the preexisting provisions post-Sarbanes-Oxley.⁴¹ These Supreme Court precedents define the boundaries of federal obstruction of justice prosecutions and lay the foundation for interpreting the new statutes.

A. *Aguilar and the “Nexus Requirement”*

In *United States v. Aguilar*, the Supreme Court interpreted section 1503 after a judge was convicted of endeavoring to obstruct justice by disclosing a wiretap and lying to FBI agents when questioned about it.⁴² The Court overturned the judge’s obstruction conviction because it held that making false statements to an investigating agent who may or may not testify before a grand jury did not constitute a violation of section 1503.⁴³ It focused on what lower courts had termed a “nexus requirement” implicit in section 1503 that “tended to place metes and bounds on the very broad language of the catchall provision.”⁴⁴ The nexus requirement mandated that the accused’s actions “have a relationship in time, causation, or logic with the judicial proceedings.”⁴⁵ The Court described

contemplation of any such matter or case, shall be fined under this title, imprisoned not more than 20 years, or both.”) (emphasis added); S. REP. NO. 107-146, 14–15 (2002) (“This statute is specifically meant not to include any technical requirement, which some courts have read into other obstruction of justice statutes, to tie the obstructive conduct to a pending or imminent proceeding or matter. It is also sufficient that the act is done ‘in contemplation’ of or in relation to a matter or investigation. It is also meant to do away with the distinctions, which some courts have read into obstruction statutes, between court proceedings, investigations, regulatory or administrative proceedings (whether formal or not), and less formal government inquiries, regardless of their title. Destroying or falsifying documents to obstruct any of these types of matters or investigations, which in fact are proved to be within the jurisdiction of any federal agency are covered by this statute.”).

40. See *United States v. Aguilar*, 515 U.S. 593 (1995).

41. See *Arthur Andersen LLP v. United States*, 544 U.S. 696 (2005).

42. *Aguilar*, 515 U.S. at 595.

43. *Id.* at 600.

44. *Id.* at 599.

45. *United States v. Aguilar*, 515 U.S. 593, 599 (1993).

how this meant that “the endeavor must have the ‘natural and probable effect’ of interfering with the due administration of justice.”⁴⁶ The heart of the requirement is that “if the defendant lacks knowledge that his actions are likely to affect the judicial proceeding, he lacks the requisite intent to obstruct.”⁴⁷ In the case of the judge, since he did not know that his lies would affect the grand jury proceeding, the judge could not have intended to obstruct the proceeding.

This nexus requirement is an obligation nowhere enumerated in the statute, created by courts out of concern that the statute would otherwise not provide sufficient warning about what constitutes illegal conduct.⁴⁸ If no nexus between the obstructive conduct and a judicial proceeding was required, a person could be held liable for obstruction for lying to his wife about his whereabouts after committing a crime if he knew about a pending investigation and his wife happened to be subsequently interviewed as part of the investigation.⁴⁹ The scope of what could potentially be considered obstructive conduct absent a nexus requirement would therefore be vast. The Court in *Aguilar* thus placed a significant restraint on the reach of section 1503 by holding that this nexus requirement—that the act must have the “natural and probable effect” of obstructing judicial or grand jury proceedings—would be strictly enforced.

B. *Arthur Andersen and “Corruptly”*

Ten years later, the Court was once again faced with interpreting an obstruction of justice provision. Sarbanes-Oxley had been enacted three years earlier, but neither of its obstruction provisions was at issue. The Court was instead asked to interpret the obstruction of justice statutes as they stood in 2000, specifically section 1512(b).⁵⁰ The corporate scandals that led to the passage of Sarbanes-Oxley’s new obstruction provisions, however, loomed in the background of the Court’s opinion; the petitioner was Arthur Andersen LLP, Enron’s auditor, and the firm had been convicted of violating section 1512(b) by instructing its employees to destroy

46. *Id.* (elaborating on this point that “the defendant’s actions need [not] be successful; an ‘endeavor’ suffices”).

47. *Id.*

48. *See Aguilar*, 515 U.S. at 600 (“We have traditionally exercised restraint in assessing the reach of a federal criminal statute . . . out of concern that ‘a fair warning should be given to the world in language that the common world will understand, of what the law intends to do if a certain line is passed.’”) (quoting *McBoyle v. United States*, 283 U.S. 25, 27 (1931)).

49. *See id.* at 602.

50. *Arthur Andersen LLP v. United States*, 544 U.S. 696, 698 n.1 (2005).

documents related to Enron's collapse.⁵¹ The Court was focused on resolving a circuit split over the meaning of section 1512(b), particularly "what it means to 'knowingly . . . corruptly persuade' another person 'with intent to . . . cause' that person to 'withhold' documents from, or 'alter' documents for use in, an 'official proceeding.'"⁵²

In a unanimous decision, the Court interpreted the terms "knowingly . . . corruptly persuade" in section 1512(b) to encompass "only persons conscious of wrongdoing."⁵³ It rejected the lower court's jury instructions because they charged the jury to convict even when the petitioner "honestly and sincerely believed that its conduct was lawful" and even when the petitioner's sole intent was to "impede" a government proceeding without any dishonest purpose (i.e., a lawyer instructing a client to assert his legitimate attorney-client privilege and not answer a question).⁵⁴ Furthermore, the Court picked up the *Aguilar* nexus requirement and held that a "'knowingly . . . corrupt persuader' [sic] cannot be someone who persuades others to shred documents under a document retention policy when he does not have in contemplation any particular official proceeding in which those documents might be material."⁵⁵ It specifically questioned whether a proceeding needed to be pending for the nexus requirement to be applicable and it determined that a proceeding "need not be pending or about to be instituted at the time of the offense,"⁵⁶ but that the defendant must have "foreseen" a potential future proceeding for the nexus requirement to be satisfied.⁵⁷

The *Arthur Andersen* Court therefore narrowed the reach of section 1512(b) in two important ways. First, it imported into section 1512(b) the nexus requirement it had found in *Aguilar* with section 1503. Second, and more importantly, it drew a sharp distinction between the mere intent to impede a proceeding and the dishonest intent to obstruct justice. The Court was clear that unless the reach of section 1512(b) was abruptly cut off at the line of dishonest intent, the provision would impermissibly sweep up innocent conduct

51. *Id.* at 698.

52. *Id.* at 702, 703.

53. *Id.* at 706.

54. *Id.*

55. *Id.* at 708.

56. *Arthur Anderson LLP v. United States*, 544 U.S. 696, 707–08 (2005) (quoting 18 U.S.C. § 1512(e)(1) (2006) (current version codified at 18 U.S.C. § 1512(f)(1) (2006)), which makes it clear that proceedings need not be pending for section 1512 violations to occur).

57. *Id.*

and would, in essence, be limitless.⁵⁸ Traditionally protected assertions of attorney-client privilege, marital privilege, and the right against self-incrimination, as well as compliance with valid document retention policies, could all be considered obstructive conduct under a broader reading of the intent requirement.⁵⁹ Because withholding testimony or documents from a judicial proceeding is not necessarily corrupt, a more malicious intent was necessary to trigger a violation of the obstruction provisions.⁶⁰ After *Aguilar* and *Arthur Andersen*, it is clear that in order to violate the pre-Sarbanes-Oxley obstruction provisions a person must (1) be conscious of wrongdoing, (2) have a dishonest intent, and (3) have foreseen a judicial proceeding which his actions could materially impact.

C. *Aguilar, Arthur Andersen, and Sections 1512(c) and 1519*

The requirement that a person must possess a dishonest intent to obstruct justice, embodied in the modifier “corruptly” in the *Arthur Andersen* decision, may not apply to both of the Sarbanes-Oxley additions. Section 1512(c) only slightly modifies the language of section 1512(b) from “knowingly . . . corruptly persuade” to “corruptly,”⁶¹ doing away with the knowledge requirement and making the statute applicable to individuals acting alone. This modification likely does little to affect the applicability of the *Arthur Andersen* analysis, and section 1512(c) has been interpreted to require the same awareness of wrongdoing and dishonest intent that section 1512(b) necessitates.⁶²

The construction of section 1519 is a different matter. Section 1519 does not contain the same “corruptly” language as the other

58. *See id.*

59. *See id.* at 703–04.

60. *See id.*

61. 18 U.S.C. § 1512(c) (2006).

62. *See, e.g.,* *United States v. McKibbins*, 656 F.3d 707, 711 (7th Cir. 2011) (“The intent element is important here because the word ‘corruptly’ is what ‘serves to separate criminal and innocent acts of obstruction.’ . . . Without a showing of a willful, corrupt *mens rea* that has a nexus to an official proceeding, the government cannot meet its burden.”) (citation omitted) (citing *Arthur Andersen*, 544 U.S. at 704–08). Furthermore, the Supreme Court is unlikely to require a lesser standard for individuals acting alone than it does for individuals persuading others to obstruct justice. *See* A. Michael Warnecke & George W. Morrison, *Responding to Allegations of Improper Corporate Conduct*, HAYNES & BOONE LLP (Oct. 31, 2005), <http://haynesboone.com/files/Publication/d9877913-2a6c-4681-84a1-62c5af91e8b2/Presentation/PublicationAttachment/54c0ab6e-ae0-49af-9059-968f9969f787/Warnecke%20-%20Responding%20to%20Allegations%20-%202011-28-06.pdf> (“Given the scope of SOX’s main new section 1519 . . . and preexisting law, section 1512(c) is of relatively little practical importance.”).

obstruction of justice provisions.⁶³ Instead, section 1519 states that a violation occurs when someone “knowingly” obstructs justice.⁶⁴ The conspicuous absence of the modifier “corruptly” in section 1519 was not unintentional; Congress specifically meant to extend the government’s ability to prosecute obstruction of justice crimes by enacting a broadly worded and widely applicable provision.⁶⁵ The legislative history shows that Congress did not intend for section 1519 to contain a dishonesty requirement; it intended that the mere intent to impede a proceeding was sufficient to constitute a violation of the section.⁶⁶ Congress thus sent a clear signal to the judiciary, both in the text and legislative history of section 1519, that it wanted this provision to be treated differently than previous obstruction of justice statutes. Even if a court were to overlook this expression of legislative intent when construing section 1519, it would still need to distinguish the Supreme Court’s parsing of language signaling the level of intent in obstruction statutes in *Arthur Andersen*, which rather straightforwardly appears to indicate that section 1519’s language elucidates a different intent requirement than section 1512(b).⁶⁷

The courts that have analyzed the level of intent necessary to find a section 1519 violation have failed to coalesce on one definitive interpretation of the provision. Some courts have firmly read

63. 18 U.S.C. § 1519 (2006).

64. *Id.*

65. See S. REP. NO. 107-146, at 14 (2002).

66. *Id.* (“Section 1519 is meant to apply broadly to any acts to destroy or fabricate physical evidence so long as they are done with the intent to obstruct, impede or influence the investigation or proper administration of any matter, and such matter is within the jurisdiction of an agency of the United States, or such acts done either in relation to or in contemplation of such a matter or investigation.”).

67. See Elkan Abramowitz & Barry A. Bohrer, *The ‘Andersen’ Decision: Its Effects on 18 USC §1519 and Attorneys*, 234 N.Y. L.J., July 5, 2005, at 3, 6 (“[A]lthough both sections have a requirement that the destruction be done ‘knowingly,’ §1519 does not require that it be done ‘corruptly.’ Section 1519 also specifically refers to ‘imped[ing]’ an investigation. Thus the government may argue that the language in *Andersen*—indicating that impeding an investigation is not sufficient grounds for a conviction under §1512—may be distinguishable under §1519. Courts that have any doubt about how to interpret ‘impede’ under §1519 may look to the court’s dissection of the term in *Andersen*: “‘impede’ has broader connotations than “subvert” or even “undermine” . . . and many of these connotations do not incorporate any “corrupt[ness]” at all.’ The court equated ‘to impede’ with “‘to interfere with or get in the way of the progress of” or “hold up” or “detract from.””). It might be possible to argue that section 1519 should be narrowed pursuant to the *Arthur Andersen* opinion based on the broad language in the case about the need to confine the obstruction statutes so that people know when they might be violating it. See *id.*

an *Arthur Andersen* “corruptly” requirement into the “knowingly” language of section 1519, requiring knowledge of conscious wrongdoing as opposed to the mere honest intent to impede a proceeding.⁶⁸ Many courts, however, have neglected to fully flesh out the requirement, simply stating that section 1519 requires intent to obstruct.⁶⁹ These precedents are vulnerable to the interpretation that the removal of “corruptly” from the text of section 1519 makes a difference in the level of intent required to find a violation. Too few courts have addressed this issue to comfortably say that section 1519 falls in line with the Supreme Court’s previous elucidation of obstruction of justice intent requirements, especially given Congress’ stark move in utilizing different language in section 1519. If section 1519 ultimately is not found to contain a dishonest intent requirement comparable to that in other obstruction of justice provisions, the reach of section 1519 will be vast and the fears announced in *Arthur Andersen* about innocent conduct being swept up under the statute will be realized.⁷⁰ Whether section 1519 includes an implicit nexus requirement thus becomes an extremely important question, as that is one final avenue by which to constrain the reach of the otherwise limitless provision.

68. See *United States v. Kernell*, 667 F.3d 746, 754 (6th Cir. 2012); *United States v. Stevens*, 771 F. Supp. 2d 556, 560–61 (D. Md. 2011); *United States v. Hayes*, No. 3:09-cr-397, 2010 U.S. Dist. LEXIS 67446, at *12–13 (M.D. Pa. July 7, 2010); *United States v. Nestor*, No. 3:09-CR-00397, 2010 U.S. Dist. LEXIS 67097, at *18 (M.D. Pa. July 6, 2010); *United States v. Moyer*, 726 F. Supp. 2d 498, 506 (M.D. Pa. 2010); *United States v. Fumo*, No. 06-319, 2009 U.S. Dist. LEXIS 51581, at *172 (E.D. Pa. June 17, 2009); *United States v. Russell*, 639 F. Supp. 2d 226, 240 (D. Conn. 2007). Recently, in *United States v. Maury*, the Third Circuit referenced section 1519 in a string cite discussing the dishonest intent requirement of another obstruction of justice provision, seemingly implying without elaborating that section 1519 contains a dishonest intent requirement. No. 09-2305, 2012 U.S. App. LEXIS 19474, at *87–88 (3d Cir. Sept. 17, 2012).

69. See *United States v. Yielding*, 657 F.3d 688, 711 (8th Cir. 2011); *United States v. Gray*, 642 F.3d 371, 376 (2d Cir. 2011); *United States v. Moore*, No. 10-213 Section “L”(2), 2011 U.S. Dist. LEXIS 55817, at * 3 (E.D. La. May 23, 2011), *aff’d*, No. 11-30877, 2013 WL 512342, at *6-7 (5th Cir. Feb. 11, 2013); *United States v. Kun Yun Jho*, 465 F. Supp. 2d 618, 636 (E.D. Tex. 2006), *rev’d on other grounds*, 534 F.3d 398 (5th Cir. 2008).

70. See *Arthur Andersen LLP v. United States*, 544 U.S. 696, 703–04 (2005) (discussing the invocation of the right against compelled self-incrimination, the right not to disclose marital confidences, the right not to turn over documents protected by attorney-client privilege, and compliance with valid document retention policies under ordinary circumstances as legitimate, innocent conduct that would be swept up under a broader interpretation of section 1512(b)).

Courts have interpreted section 1512(c) to contain a nexus requirement,⁷¹ but it is becoming increasingly settled that section 1519 does not contain such a requirement. In enacting section 1519, Congress provided that no proceeding need be pending for the section to be applicable,⁷² and it also specifically made clear that it did not intend for the provision to contain a nexus requirement.⁷³ By relying on the legislative history and language of the section, the Second, Third, Fifth, Sixth, and Eighth Circuits have determined that section 1519 does not contain a nexus requirement.⁷⁴ The Second and Third Circuits, however, did so by overturning district court opinions that had followed the *Aguilar-Arthur Andersen* line of precedent and read a nexus requirement into section 1519,⁷⁵ requiring that the obstructive conduct bear some rela-

71. See *United States v. Johnson*, 655 F.3d 594, 605 (7th Cir. 2011); *United States v. Friske*, 640 F.3d 1288, 1292 (11th Cir. 2011); *United States v. Phillips*, 583 F.3d 1261, 1264 (10th Cir. 2009); *United States v. Carson*, 560 F.3d 566, 584 (6th Cir. 2009); *United States v. Ortiz*, 220 F. App'x 13, 16 (2d Cir. 2007). As of the time of the *Friske* decision, no circuit had rejected the nexus requirement for section 1512(c). See *Friske*, 640 F.3d at 1292 n.4.

72. See 18 U.S.C. § 1519 (2006).

73. S. REP. NO. 107-146, at 14–15 (2002) (“This statute is specifically meant not to include any technical requirement, which some courts have read into other obstruction of justice statutes, to tie the obstructive conduct to a pending or imminent proceeding or matter.”). The Senate did express concern that section 1519 could be interpreted more broadly than intended, and specifically stated that it should not be construed to prosecute individuals complying with normal document retention policies. See *id.* at 27 (“In our view, section 1519 should be used to prosecute only those individuals who destroy evidence with the specific intent to impede or obstruct a pending or future criminal investigation, a formal administrative proceeding, or bankruptcy case. It should not cover the destruction of documents in the ordinary course of business, even where the individual may have reason to believe that the documents may tangentially relate to some future matter within the conceivable jurisdiction of an arm of the federal bureaucracy.”).

74. See *United States v. Moore*, No. 11-30877, 2013 WL 512342, at *7 (5th Cir. Feb. 11, 2013); *United States v. Moyer*, 674 F.3d 192, 209 (3d Cir. 2012); *United States v. Kernell*, 667 F.3d 746, 754–55 (6th Cir. 2012); *Yielding*, 657 F.3d at 712; *Gray*, 642 F.3d at 377–78 (“[I]n enacting § 1519, Congress rejected any requirement that the government prove a link between a defendant’s conduct and an imminent or pending official proceeding. The defendants therefore are incorrect in assuming that because the Supreme Court has required a nexus to an official proceeding for purposes of other obstruction statutes, the same nexus requirement must apply to prosecutions under § 1519. . . . By the plain terms of § 1519, knowledge of a pending federal investigation or proceeding is not an element of the obstruction crime.”).

75. See *United States v. Piekarsky*, No. 3:09-CR-396, 2010 U.S. Dist. LEXIS 81360, at *14 (M.D. Pa. Aug. 11, 2010); *United States v. Hayes*, No. 3:09-CR-397, 2010 U.S. Dist. LEXIS 67446, at *19 (M.D. Pa. July 7, 2010); *United States v. Nestor*, No. 3:09-CR-00397, 2010 U.S. Dist. LEXIS 67097, at *18 (M.D. Pa. July 6,

tion to a contemplated proceeding or investigation.⁷⁶ These earlier district court opinions looked past Congress' expression of intent in the legislative history of the section and saw strong similarities between section 1519 and *Arthur Andersen's* clarification that no proceeding need be pending for a cognizable section 1512 violation.⁷⁷ It thus appears that, despite earlier opinions emphasizing the importance of reading a nexus requirement into section 1519, a trend toward eliminating that requirement has developed. How other courts interpret this aspect of section 1519 going forward will be extremely consequential; if future courts follow the trend and find that section 1519 does not contain a nexus requirement, the category of conduct that could be considered obstructive will greatly expand, as was discussed in *Aguilar*.⁷⁸

Overall, while section 1512(c) falls neatly into the pre-Sarbanes-Oxley obstruction of justice scheme, section 1519 is a radical departure. Both in its text and legislative history, Congress made clear that it was doing something different with this provision, and that it was attempting to provide prosecutors with additional ammunition. While it might be seen as a good thing, in light of the Enron and Arthur Andersen scandals, that prosecutors are better equipped post-Sarbanes-Oxley to combat obstruction of justice with these new provisions, the breadth of section 1519 could have unintended collateral effects, policing conduct Congress did not intend to deter, if it is not constrained by established obstruction of justice precedents.

III. LAWYERS CHARGED WITH VIOLATIONS OF THE SARBANES-OXLEY OBSTRUCTION OF JUSTICE PROVISIONS

Obstruction of justice prosecutions brought against lawyers illustrate the problems with the potential breadth of section 1519.

2010); *United States v. Moyer*, 726 F. Supp. 2d 498, 506 (M.D. Pa. 2010); *United States v. Russell*, 639 F. Supp. 2d 226, 240 (D. Conn. 2007).

76. *See Hayes*, 2010 U.S. Dist. LEXIS 67446, at *11–12, 19 (“Without the requirement of a nexus, the public may not know that their actions are illegal because they would not be aware of the federal proceeding they were obstructing. . . . Therefore, § 1519 requires a nexus between the alleged obstruction and the matter within United States jurisdiction which the action is contemplated to obstruct.”).

77. These courts also believe that the policy goals justifying restraint in interpreting the reach of the pre-Sarbanes-Oxley obstruction provisions apply equally when interpreting section 1519. *See, e.g., id.* at *14.

78. *See supra* note 49 and accompanying text.

Although there are many straightforward prosecutions of clear misconduct on the part of lawyers, the Lauren Stevens prosecution pushes the line marking the boundary between criminal conduct and zealous advocacy further back to prosecute conduct previously only considered zealous advocacy.

The traditional obstruction prosecutions of lawyers arise out of three scenarios. First, lawyers are prosecuted for obstruction violations in the context of a large underlying criminal scheme in which the lawyers were intimately involved for their own personal gain and actively obstructed a proceeding to cover their tracks.⁷⁹ These prosecutions focus on lawyers' personal actions and their personal involvement in criminal schemes like fraud and money laundering, not necessarily their advocacy on behalf of their clients. Second, lawyers are prosecuted for obstruction violations when an underlying crime in which the lawyers were involved is difficult to prove and it is easier for prosecutors to meet their burden of proof on the obstruction charge.⁸⁰ Third, lawyers are prosecuted for obstruction of justice on its own, not connected to any underlying crimes, when they clearly leap over the line from zealous advocacy into criminal behavior.⁸¹ For example, lawyers who bribe or threaten witnesses

79. *See, e.g.*, *United States v. Blair*, 661 F.3d 755, 775 (4th Cir. 2011) (affirming the attorney's conviction for laundering drug proceeds he obtained from a client and reversing his § 1503 obstruction conviction due to the prosecution's failure to meet the nexus requirement); *United States v. Reich*, 479 F.3d 179, 192 (2d Cir. 2007) (affirming the lawyer's conviction for forging a judge's order during an arbitration proceeding against a brokerage firm he alleged had mishandled his account); *United States v. Ellis*, 419 F.3d 1189, 1191 (11th Cir. 2005) (vacating the lawyer's sentence enhancement for his guilty plea to having sex with his criminal defendant in a then-pending case); *United States v. Stoll*, No. 10-60194-CR-COHN/SELTZER, 2011 U.S. Dist. LEXIS 18906, at *6 (S.D. Fla. Feb. 16, 2011) (denying the lawyer's motion to dismiss the 34-count indictment arising out of a mortgage fraud conspiracy); *United States v. Maze*, No. 5:06-CR-155-S-JMH, 2007 U.S. Dist. LEXIS 6694, at *14 (E.D. Ky. Jan. 30, 2007) (denying the lawyer's motion to dismiss his indictment arising from his conspiracy to rig an election); *United States v. Kaplan*, No. 02 Cr. 883 (DAB), 2003 U.S. Dist. LEXIS 21825, at *3, *47-48 (S.D.N.Y. Dec. 5, 2003) (discussing the lawyer's eleven-count indictment arising out of a health care fraud scheme).

80. *See, e.g.*, *United States v. Crawford*, 60 F. App'x 520, 532-34 (6th Cir. 2003) (affirming the lawyer's obstruction conviction arising out of a failed attempt to convict him of possessing and distributing controlled substances).

81. *See, e.g.*, *United States v. Simels*, 654 F.3d 161, 167-68 (2d Cir. 2011) (affirming the lawyer's obstruction convictions arising out of his attempts to bribe and threaten potential witnesses against his client); *United States v. Mintmire*, 507 F.3d 1273, 1274-75 (11th Cir. 2007) (affirming the lawyer's obstruction convictions arising out of his attempts to obstruct grand jury and SEC investigations of his clients); *United States v. Kennon*, No. 3:08CR42, 2009 U.S. Dist. LEXIS 30801, at *15-16 (W.D.N.C. Mar. 24, 2009) (refusing to join three separate incidents of

are unmistakably vulnerable to prosecution. These three types of prosecutions are typical obstruction of justice cases, and they are also instances in which lawyers' ethical duties align with their legal obligations.⁸² Furthermore, these cases display both a clear nexus between the obstructive behavior and a judicial proceeding and evident dishonest intent, and thus do not raise any of the issues or ambiguities discussed above.

If these types of prosecutions were the only cases brought against lawyers under the obstruction provisions, the bar would have no reason to worry about the breadth of the Sarbanes-Oxley obstruction provisions. Truly reprehensible, dishonest conduct would be prosecuted and zealous advocacy would be allowed to flourish unscathed; prosecutors would be drawing a bright line between advocacy and criminality and lawyers would know when they were about to cross it. Not all prosecutors, however, have focused solely on these clear cases. Instead, at least one prosecutor recently went after a lawyer for more ambiguously inappropriate conduct.⁸³

alleged witness tampering and obstruction of justice committed by an attorney advocating for his clients in three different criminal cases); *United States v. Coren*, No. 07-CR-265 (ENV), 2008 U.S. Dist. LEXIS 71564, at *1-2 (E.D.N.Y. Aug. 29, 2008) (denying the lawyer's motion to dismiss the seventeen-count indictment arising out of the scheme he developed for his clients to defraud various government entities relating to construction contracts).

82. See MODEL RULES OF PROF'L CONDUCT R. 1.2(d) (2012) ("A lawyer shall not counsel a client to engage, or assist a client, in conduct that the lawyer knows is criminal or fraudulent . . ."); *id.* R. 3.3(b) ("A lawyer who represents a client in an adjudicative proceeding and who knows that a person intends to engage, is engaging or has engaged in criminal or fraudulent conduct related to the proceeding shall take reasonable remedial measures, including, if necessary, disclosure to the tribunal."); *id.* R. 3.4 ("A lawyer shall not: (a) unlawfully obstruct another party's access to evidence or unlawfully alter, destroy or conceal a document or other material having potential evidentiary value. A lawyer shall not counsel or assist another person to do any such act . . .").

83. See *United States v. Stevens*, 771 F. Supp. 2d 556 (D. Md. 2011). See generally Lisa Kern Griffin, *Criminal Lying, Prosecutorial Power, and Social Meaning*, 97 CALIF. L. REV. 1515, 1516 (2009) [hereinafter Griffin, *Criminal Lying*] ("Process offenses that arise during white collar investigations rather than from the commission of the crime itself have increasingly been the focus of federal prosecution."); Lisa Kern Griffin, *Wanting the Truth: Comparing Prosecutions of Investigative and Institutional Deception*, 7 INT'L COMMENT. ON EVIDENCE, no. 1, 2009 at 7 [hereinafter Griffin, *Wanting the Truth*] ("[B]ecause the criminal lying prohibitions can be stretched to cover very ordinary human behavior, and because lying is an everyday occurrence, there is an obvious gap between statutory over-deterrence and on-the-ground under-enforcement. Recent cases suggest that whether and when prosecutors choose to close that gap by prosecuting investigative lies has little to do with truth-seeking in the false statements context and more to do with the need for efficiency where those statements are pretexts for more serious but unprovable

When prosecutors charge lawyers with obstruction violations unconnected to any underlying criminal activity, related solely to the lawyers' advocacy for their clients, and for debatably ethical behavior, they enter a gray area where lawyers do not know whether their conduct is permissible or criminal. How courts construe the breadth of the Sarbanes-Oxley provisions in this gray area will have vast import for the regulation of the legal profession and for lawyers' incentives to zealously advocate for their clients.

A. *The Prosecution of Lauren Stevens*

Prosecutors delved into this gray area to police lawyer conduct in 2011 with the prosecution of Lauren Stevens. Stevens was Vice President and Associate General Counsel for GlaxoSmithKline at the time of her indictment, and she was accused of making false statements and obstructing a Food and Drug Administration investigation into her client's alleged illegal drug marketing.⁸⁴ She was not charged with involvement in any of her client's alleged criminal activity; she was prosecuted solely for purportedly overstepping the bounds of permissible lawyering in her voluntary response to a government inquiry. Stevens was targeted by prosecutors as "part of the government's long-promised crackdown on individual executives for their roles in pharmaceutical company cases"⁸⁵ and because of the "mounting complaints from consumer groups and Congress that companies are paying nine-figure fines as a cost of doing business while executives are almost never held accountable."⁸⁶ Prosecutors were thus using Stevens's indictment to send a warning signal to other pharmaceutical executives that they were not immune to prosecution for their involvement in their companies' wrongdoing.⁸⁷

crimes, with the assertion of authority where defendants are recalcitrant, and with the desire for apology where defendants have failed to take responsibility.").

84. See Duff Wilson, *Former Glaxo Lawyer Indicted*, N.Y. TIMES (Nov. 9, 2010, 2:25 PM), <http://prescriptions.blogs.nytimes.com/2010/11/09/former-glaxo-lawyer-indicted>.

85. *Id.*

86. *Id.* Interestingly, none of the executives at GlaxoSmithKline who were allegedly involved in the underlying criminal activity were indicted and the company had not been formally charged with any crimes at the time Stevens was prosecuted. See Christina Pazzanese, *DOJ lawsuit could dampen future role of in-house counsel*, NEW ENGLAND IN-HOUSE, Jan. 2011, <http://newenglandinhouse.com/2011/02/02/doj-lawsuit-could-dampen-future-role-of-in-house-counsel>.

87. See Duff Wilson, *Ex-Glaxo Executive Is Charged in Drug Fraud*, N.Y. TIMES (Nov. 9, 2010), <http://www.nytimes.com/2010/11/10/health/10glaxo.html> ("This is absolutely precedent-setting—this is really going to set people's hair on

The indictment⁸⁸ alleged that Stevens, through her advocacy, played an integral role in attempting to cover up her company's wrongdoing and obstruct the FDA's investigation into off-label marketing of the antidepressant Wellbutrin.⁸⁹ Specifically it accused her of violating both sections 1512 and 1519 by sending signed letters to the FDA in response to its investigatory requests "in which she made materially false statements and concealed and covered up documents and other evidence that showed the extent of K-Corp.'s promotion of W-Drug for unapproved uses."⁹⁰ She allegedly withheld slide presentations, handouts, and audio cassettes from the government that included potentially incriminating information and instead produced false, misleading, and incomplete information to the government in an attempt to obstruct its investigation.⁹¹

In an unprecedented move by the judge, the court granted Stevens's motion for a judgment of acquittal on May 10, 2011 at the close of the prosecution's case-in-chief.⁹² The court homed in on the boundary between zealous advocacy and criminal conduct and stated that this prosecution pushed the boundary too far. What Stevens did was simply zealous advocacy on behalf of her client,⁹³ and she "should never have been prosecuted and she should be permit-

fire . . . This is indicative of the F.D.A. and Justice strategy to go after the very top-ranking managing officials at regulated companies.").

88. Indictment, *United States v. Stevens*, 771 F. Supp. 2d 556 (D. Md. 2011) (No. 10-694). Ms. Stevens was first indicted on November 8, 2010. On March 23, 2011, the indictment was dismissed without prejudice after the court learned that the prosecutors misstated the law during the grand jury proceedings. *See United States v. Stevens*, 771 F. Supp. 2d 556 (D. Md. 2011). On April 14, 2011, Ms. Stevens was re-indicted with the same charges. *See Adam Ramirez, Feds Re-Indict Ex-Glaxo In House Lauren Stevens for Corporate Duty*, FINDLAW (Apr. 18, 2011, 5:49 AM), http://blogs.findlaw.com/in_house/2011/04/ex-glaxo-in-house-is-reindicted.html.

89. *See Wilson, supra* note 87.

90. Indictment, *supra* note 88, ¶ 25.

91. *Id.* ¶¶ 26–41, Counts One and Two.

92. Transcript of Oral Argument, *supra* note 8, at 8, 9 ("I conclude on the basis of the record before me that only with a jaundiced eye and with an inference of guilt that's inconsistent with the presumption of innocence could a reasonable jury ever convict this defendant. . . . In my seven and a half years as a jurist I have never granted one. There is, however, always a first. . . . I believe that it would be a miscarriage of justice to permit this case to go to the jury.").

93. *See Sue Reisinger, Crossing the Line: The Trial of GlaxoSmithKline Lawyer Lauren Stevens*, CORPORATE COUNSEL (June 23, 2011), <http://www.law.com/jsp/cc/PubArticleCC.jsp?id=1202497750428>; Alicia Mundy & Brent Kendall, *U.S. Rebuffed in Glaxo Misconduct Case*, WALL ST. J. (May 11, 2011), <http://online.wsj.com/article/SB10001424052748703730804576315101670843340.html>.

ted to resume her career.”⁹⁴ The court focused on two facts in coming to this conclusion. First, Stevens consulted outside counsel, who assured her that the responses she planned to make to the FDA’s requests were both legally and ethically permissible, before voluntarily responding to the FDA’s inquiry.⁹⁵ Because she relied on outside advice, the court held that she was acting in good faith, negating the *mens rea* required in both sections 1512 and 1519.⁹⁶ Additionally, Stevens’s conduct fit squarely into the Safe Harbor Provision of the obstruction of justice statutes.⁹⁷ Second, the court noted the strong negative consequences of finding Stevens liable for her conduct, specifically focusing on the “serious implications for the practice of law generated by this prosecution.”⁹⁸ It stated that “a lawyer should never fear prosecution because of advice that he or she has given to a client who consults him or her.”⁹⁹ Noting the “enormous potential for abuse in allowing prosecution of an

94. Transcript of Oral Argument, *supra* note 8, at 10.

95. *Id.* at 7 (“[T]he evidence in this case can only support one conclusion, and that is that the defendant sought and obtained the advice and counsel of numerous lawyers. She made full disclosure to them. Every decision that she made and every letter she wrote was done by a consensus.”).

96. *Id.* (“[E]ven if some of these statements were not literally true, it is clear that they were made in good faith which would negate the requisite element required for all six of the crimes charged in this case.”). In its order dismissing the first indictment on March 23, 2011, the court discussed the *mens rea* requirement for section 1519 actions. *See United States v. Stevens*, 771 F. Supp. 2d 556, 561 (D. Md. 2011). It determined that section 1519, despite omitting the term “corruptly,” still required the same dishonest intent as the other obstruction provisions. *Id.* The court reasoned that:

To hold otherwise would allow § 1519 to reach inherently innocent conduct, such as a lawyer’s instruction to his client to withhold documents the lawyer in good faith believes are privileged. Any other interpretation of § 1519 would ignore the admonition of the Supreme Court in *Arthur Andersen* that criminal liability ordinarily may only be imposed on those with consciousness of their wrongdoing.

Id. By construing section 1519 in this way, the court found that it was a specific-intent crime for which proof of good faith reliance on advice of counsel negated wrongful intent. *Id.* at 562.

97. The court stated, “As to Counts One and Two, the Safe Harbor Provision of Section 15(c) is an absolute bar. GlaxoSmithKline did not come to Ms. Stevens and say, assist us in committing a crime or fraud. It came to her for assistance in responding to a letter from the FDA. I conclude on the basis of this record that no reasonable juror could conclude otherwise beyond a reasonable doubt. . . . [T]he Safe Harbor Provision is designed specifically to protect an attorney who is acting in accordance with the obligation that every lawyer has to zealously represent his or her client and place their position in the most favorable possible light.” Transcript of Oral Argument, *supra* note 8, at 6.

98. *Id.* at 9.

99. *Id.*

attorney for the giving of legal advice,” the court avowed that its decision did not immunize lawyers from prosecution or mean that their conduct is unreviewable.¹⁰⁰ It instead carved out a specific role for the judiciary: “[T]he Court should be vigilant to permit the practice of law to be carried on, to be engaged in, and to allow lawyers to do their job of zealously representing the interests of their client. Anything that interferes with that is something that the court system should not countenance.”¹⁰¹ With this admonishment of prosecutorial overreach, the court acquitted Lauren Stevens.¹⁰²

B. The Broader Implications of United States v. Stevens for Lawyers

United States v. Stevens may be one of the first examples of prosecutors wielding the Sarbanes-Oxley obstruction provisions to police arguably legitimate lawyering, but it is most certainly not the last.¹⁰³ The vagueness of the provisions, the lack of substantial case law in this area, and the growing public discontent with corporate executives escaping legal action for their companies’ misdeeds¹⁰⁴ all point toward broader use of these obstruction provisions in the future. As the court intuited, the issues raised by *United States v. Stevens* have implications for lawyers that reach beyond the individual livelihood of Lauren Stevens. The fact that prosecutors brought this indictment in the first place (not only once, but twice) shows that they believe the Sarbanes-Oxley obstruction of justice provisions

100. *Id.* at 10.

101. *Id.*

102. Because of the procedural posture of the case, this disposition is unreviewable.

103. While the *Stevens* case could arguably be a one-off—an example of rogue prosecutorial overreach that is unlikely to be repeated, *see infra* note 105—the prosecution’s legitimate arguments, the instability of interpretations of section 1519, and the policy goals underlying Sarbanes-Oxley counsel that this may be the first in a line of cases in which prosecutors attempt to police lawyer conduct. *See also* Debra Cassens Weiss, *Acquitted In-House Lawyer Warns of the ‘Criminalization’ of Law Practice*, A.B.A. J. (Oct. 3, 2012), http://www.abajournal.com/news/article/acquitted_in-house_lawyer_warns_of_the_criminalization_of_law_practice/ (“I think the criminalization of the practice of law is here, and I don’t think it’s necessarily going away. . . . The government will continue to be aggressive in looking at in-house counsel.”) (quoting Lauren Stevens).

104. *See, e.g.*, Edward Wyatt, *Obama Urges Tougher Laws on Financial Fraud*, N.Y. TIMES, Jan. 24, 2012, http://www.nytimes.com/2012/01/25/business/obama-urges-tougher-laws-on-financial-fraud.html?_r=1&nl=todaysheadlines&emc=th24 (discussing how President Obama’s 2012 State of the Union Address set forth “proposals [which] seek to acknowledge the continuing frustration among many Americans—exemplified by the Occupy Wall Street movement—that few financial executives have been prosecuted for their actions leading up to the crisis”).

stretch far enough to cover this type of conduct.¹⁰⁵ This discretionary judgment, as well as the threat that prosecutors could target more attorney conduct traditionally considered advocacy in the future, are worrisome. Even though the court was quick to rebuke the prosecution's attempts to convict Stevens,¹⁰⁶ the opinion will not necessarily deter future prosecutors from bringing similar cases against other lawyers. First of all, it is a lone opinion from a Maryland district court. Second, the facts of the case were very favorable for the defense. Lauren Stevens was sixty years old at the time of the indictment and well-loved by her co-workers and peers.¹⁰⁷ She also relied on the advice of outside counsel who told her that her response to the FDA complied with the law.¹⁰⁸ Future cases with less sympathetic defendants who do not rely on outside legal advice and who do not have favorable facts on their side may not garner the same fortunate result.

If prosecutors continue to bring obstruction cases against lawyers for crossing the vague boundary between zealous advocacy and obstruction, the ramifications for individual lawyers will be immense. Those charged will shoulder the burden of defending years-long lawsuits, reputational harms, massive expense, and potential job loss. If convicted, or if they take a plea deal to avoid a drawn out suit, they will face the harsh penalties imposed by the Sarbanes-Oxley provisions: fines and/or up to twenty years in prison.¹⁰⁹ Additionally, if they are convicted or plead guilty, they will be disbarred, potentially permanently.¹¹⁰ Furthermore, a broad interpretation of

105. See Reisinger, *supra* note 93 (“[T]he stunned prosecutors privately complained that the jury would have found Stevens guilty had the judge let the trial continue.”). But see Sue Reisinger, *Why Didn't the Maryland U.S. Attorney Sign the Lauren Stevens Indictment?*, CORPORATE COUNSEL (June 20, 2011), http://www.law.com/corporatecounsel/PubArticleCC.jsp?id=1202497761188&Why_Didnt_the_Maryland_US_Attorney_Sign_the_Lauren_Stevens_Indictment&slreturn=20121020110158 (“Maryland U.S. attorney Rod Rosenstein refused to sign the Lauren Stevens indictments because he didn't think there was enough evidence to support the charges. . .”).

106. See Transcript of Oral Argument, *supra* note 8, at 6 (“The Subcommittee on Criminal Justice had received complaints of prosecutors harassing members of the defense bar, and that vigorously and zealously representing a client is no [sic] a basis for charging an offense under the Obstruction of Justice chapter.”).

107. See Wilson, *supra* note 84; Reisinger, *supra* note 93.

108. See Reisinger, *supra* note 93.

109. See 18 U.S.C. § 1512(c) (2006); 18 U.S.C. § 1519 (2006).

110. See, e.g., *In re Libby*, 945 A.2d 1169, 1169 (D.C. 2008) (disbarring a lawyer because an obstruction conviction per se involves moral turpitude, which mandates disbarment in the state); *In re Laudumiey*, 849 So. 2d 515, 524–25 (La. 2003) (permanently disbarring two lawyers after they pled guilty to obstruction of justice); *In re Finneran*, 919 N.E.2d 698, 703–07 (Mass. 2010) (finding that lawyer

section 1519 wielded against lawyers will have three major consequences on the legal profession as a whole: (1) it will result in overcriminalization of previously innocent conduct; (2) it will lead to overdeterrence of legitimate lawyering; and (3) it will open the door to prosecutorial overreach. Each of these effects will lead to impaired client representation, both on an individual and systemic basis, by creating an incentive structure that pits lawyers' interests in self-preservation against their clients' interests.

1. Overcriminalization

Overcriminalization will result if courts interpret section 1519 broadly. Faced with judicial constraints on the reach of its obstruction statutes, Congress enacted section 1519 using broad language and stated its intent that the provision be uninhibited by those prior legal rules.¹¹¹ Attempting to give prosecutors room to police more conduct, Congress failed to take into account the broader reasons behind the Supreme Court's actions. When the Supreme Court cabined the reach of the pre-Sarbanes-Oxley obstruction provisions in *Aguilar* and *Arthur Andersen*, it specifically did so to prevent overcriminalization of innocent conduct the legal system strives to promote.¹¹² By imposing a dishonest intent requirement, the Court drew a line between situations in which lawyers intended to impede a proceeding and those in which they dishonestly wished to obstruct it.¹¹³ The Court realized that intent to impede a proceeding could be found whenever a lawyer counseled his client to withhold documents under exercise of his attorney-client privilege or to refuse to answer a question because of a marital privilege or a Fifth Amendment right.¹¹⁴ If section 1519 is construed to require the mere intent to impede, this traditionally innocent conduct will be swept up as obstruction.¹¹⁵ Furthermore, the Supreme Court mandated a nexus requirement to avoid criminalizing otherwise innocent document retention policies and to ensure that the public

pleading guilty to obstruction of justice warranted disbarment); *Miss. Bar v. De-Laughter*, 38 So. 3d 631, 631 (Miss. 2010) (permanently disbarring a lawyer after he pled guilty to one count of obstruction under section 1512(c) because a felony conviction in the state warrants automatic and permanent disbarment); *In re Coren*, 905 N.Y.S.2d 62, 63 (N.Y. App. Div. 2010) (disbarring a lawyer after he pled guilty to a sixteen-count indictment, including section 1512(c), because a felony conviction in the state requires automatic disbarment).

111. *See supra* Part II.C.

112. *See supra* Parts II.A and II.B.

113. *See supra* notes 53–54 and accompanying text.

114. *See supra* note 59 and accompanying text.

115. *See supra* notes 53–54, 58–60 and accompanying text.

knows when its actions cross the line.¹¹⁶ If no nexus is required, normal destruction of documents in the ordinary course of business will be vulnerable to obstruction prosecutions and lawyers will not know when they will be indicted for destroying something.¹¹⁷ Additionally, otherwise innocent statements by lawyers could be policed as obstructive in the absence of a nexus requirement if those statements, although unconnected to any judicial proceeding when uttered, later become relevant to a judicial proceeding.¹¹⁸

Overcriminalization punishes the innocent, wastes precious judicial resources, and, more importantly, “dilutes the moral force of the criminal justice system.”¹¹⁹ Overcriminalization also spawns the two related consequences of overdeterrence and prosecutorial overreach. As lawyers realize that their previously innocent conduct is now exposed to prosecution and severe penalties, they will be disincentivized to advise their clients to do anything that could be interpreted to impede a proceeding.¹²⁰ This will not only impact clients on a case-by-case basis, but it will also damage the quality of legal representation across the profession. The valid policy goals behind privilege and other doctrines that would be swept up under the obstruction provisions would be undermined by a broad interpretation of section 1519, and the apprehensions expressed in *Agui-lar* and *Arthur Andersen* would be brought to life.¹²¹ Additionally, by expanding the boundaries of criminal law, overcriminalization cre-

116. See *supra* notes 48–49 and accompanying text.

117. In enacting the Sarbanes-Oxley obstruction provisions, Congress exempted document retention policies from the reach of the statute in its legislative history. See *supra* note 73. Congress did not similarly carve out exceptions for assertions of privilege.

118. See *supra* note 49 and accompanying text.

119. Erik Luna, *The Overcriminalization Phenomenon*, 54 AM. U. L. REV. 703, 726, 727 (2005).

120. See *infra* Part III.B.2.

121. See *supra* notes 48–49, 58–60 and accompanying text. Amicus briefs filed in the *Andersen* case also discussed the negative implications of broadly interpreting the obstruction provisions. See Abramowitz & Bohrer, *supra* note 67, at 6 (“[E]xpanding the definition of ‘corruptly persuades’ to encompass persuasion ‘having an improper purpose’ to impede the fact-finding of an official proceeding would ‘potentially criminalize’ a broad range of legitimate attorney advice and appropriately zealous advocacy. It would ‘chill zealous legal representation, create potential conflicts between counsel and client, and undermine faith in the privacy of attorney-client communications.’”) (quoting Brief for New York Council of Defense Lawyers as Amicus Curiae Supporting Petitioner at 2, *Arthur Andersen LLP v. United States*, 544 U.S. 696 (2005) (No. 04-368), 2005 WL 435901; Brief for National Association of Criminal Defense Lawyers as Amicus Curiae Supporting Petitioner, *Arthur Andersen LLP v. United States*, 544 U.S. 696 (2005) (No. 04-368), 2005 WL 435903).

ates greater scope for the discretionary enforcement of criminal laws, opening the door for abuse of that discretion by prosecutors.¹²²

2. Overdeterrence

The Sarbanes-Oxley obstruction provisions create a conflict of interest for lawyers that will lead to overdeterrence of legitimate lawyer behavior. Faced with the threat of criminal charges if they overstep the boundaries of zealous advocacy, lawyers are confronted with a dilemma. Take Lauren Stevens, for example. When the FDA began investigating her client's practices, she was forced to choose between voluntarily turning over potentially incriminating documents to the government, which could leave her client vulnerable to civil and criminal charges and herself to a malpractice claim, and maintaining her fidelity to her client and zealously advocating on its behalf, which would leave her personally vulnerable to criminal obstruction charges. Her allegiance was pulled in three different directions: to her client, to the public, and to herself.¹²³

122. See Luna, *supra* note 119, at 724–25.

123. Another example of a lawyer prosecuted under the Sarbanes-Oxley obstruction provisions for making the wrong choice (in prosecutors' eyes) between irreconcilable options is *United States v. Russell*, 639 F. Supp. 2d 226 (D. Conn. 2007). Philip Russell was charged with violating sections 1512(c) and 1519 after he destroyed a laptop containing child pornography that belonged to the choirmaster of his client church. *See id.* at 230. Russell had been retained by the church after it discovered the contents of the laptop, and he took possession of the laptop after the choirmaster resigned. *See* Evan T. Barr, 'Russell': Prosecuting Defense Counsel for Obstruction, 238 N.Y. L.J., Nov. 21, 2007, available at <http://www.steptoe.com/assets/attachments/3254.pdf>. The church made clear that it did not wish to press charges, and Russell had no reason to believe that the government was investigating the church or the choirmaster. *See id.* Believing there would be no future proceedings in which the evidence on the laptop would be necessary, Russell had several options for what to do with it: he could keep the laptop in his possession; he could return it to his client; or he could destroy it. If he kept the laptop, he could personally be prosecuted for possession of child pornography. *See* 18 U.S.C. § 2252 (2006). If he returned it to his client, he would leave the church vulnerable to the same prosecution. Because he reasoned that no future proceedings would be initiated that would ever need the laptop, and because destroying the laptop was the best way to protect himself and his client from child pornography prosecutions, Russell chose the third option. Unbeknownst to him, the government had already commenced an investigation into the choirmaster, and when it learned that Russell had destroyed key evidence in the investigation, it charged him with obstruction of justice. *See* Barr, *supra*. After unsuccessfully attempting to dismiss his indictment, *see* Russell, 639 F. Supp. 2d at 230, Russell pled guilty to a lesser charge of misprision of a felony. STEPHEN GILLERS, REGULATION OF LAWYERS: PROBLEMS OF LAW AND ETHICS 462 (8th ed. 2009). He was sentenced to community service and home confinement for six months, and he was suspended from practice for six

This divided duty is a direct result of the Sarbanes-Oxley criminal law overlay to lawyers' ethical duties. How is a lawyer supposed to choose which allegiance to honor? A traditional feature of the attorney-client relationship is undivided allegiance to the client.¹²⁴ But the "reporting up" obligations in Sarbanes-Oxley show that Congress intended to deputize lawyers¹²⁵ after the Enron scandal, requiring them to police their clients' conduct for the benefit of the public.¹²⁶ Congress also made a strong statement that it would hold lawyers personally accountable for their involvement in corporate fraud and misconduct.¹²⁷ By imposing these external regulations, Congress became directly involved in regulating lawyers and formulated additional ethical duties for them.¹²⁸ Congress legislated what zealous advocacy "within the bounds of the law" means in this context.¹²⁹ Because it created harsh penalties for going be-

months. *Id.* Cases like *Russell* and *Stevens* show the conflicting positions lawyers confront when charged with zealously advocating for their clients in the face of the Sarbanes-Oxley obstruction provisions.

124. For example, the duty of confidentiality and the conflict rules in the Model Rules place allegiance to the client above other interests. See, e.g., *id.* R. 1.8(a)-(d), (h), (i) (protecting the client's interests over the lawyer's interests when they conflict). While exceptions to these rules allow promotion of the public interest or the lawyer's views over the client's interests, the default position articulated in the Model Rules is allegiance to the client. See, e.g., *id.* R. 1.6(a)-(b) (generally prohibiting disclosure of client confidences and permitting disclosure to benefit the public interest only in very limited circumstances).

125. Congress intended to deputize at least those lawyers "appearing and practicing" before the SEC, which could be interpreted to include all lawyers doing work for public companies. See *supra* note 26.

126. See *supra* note 25 and accompanying text.

127. See *supra* note 24 and accompanying text.

128. After Sarbanes-Oxley, the Model Rules were amended to reflect this new reporting up obligation and to stave off any additional congressional attempt at legislating mandatory ethical requirements. See GILLERS, *supra* note 123, at 12, 583-84. Specifically, Rule 1.13 "strengthen[s] the reporting-up obligation. Although reporting up is not obligatory, it is now presumptively required 'unless the lawyer reasonably believes that it is not necessary in the best interest of the organization to do so.' Previously, reporting up was simply one option available to the lawyer. Of greater consequence, Rule 1.13 now contains its own exception to confidentiality. It permits, but does not require, reporting out if, after reporting up, 'the highest authority insists upon or fails to address in a timely and appropriate manner an action, or a refusal to act, that is clearly a violation of law,' and if, in addition, 'the lawyer reasonably believes that the violation is reasonably certain to result in substantial injury to the organization.'" *Id.* at 584. While the Sarbanes-Oxley Act only applies to lawyers "appearing and practicing" before the SEC, Rule 1.13 applies to "all lawyers for organizational clients, whatever the nature of the work." *Id.*

129. See *supra* notes 11-14 and accompanying text.

yond the boundaries of the law,¹³⁰ the weight of the duty has now tipped away from zealous advocacy toward strict compliance with the boundaries imposed by Sarbanes-Oxley.¹³¹

This shift toward compliance with law after Sarbanes-Oxley's imposition of new duties and harsh penalties is exacerbated because of the vagueness of the obstruction provisions. The ambiguity of these provisions and the lack of coherent precedent analyzing them add to the confusion surrounding where lawyers' allegiances lie and what conduct is permissible advocacy.¹³² When the laws are vague and could be interpreted broadly, lawyers will tend to be more risk averse in the hopes of avoiding potential criminal sanctions. Zealous advocacy for clients is sacrificed as lawyers must abandon previously permissible conduct occurring on the margins of obstruction of justice law to protect their own livelihoods.

The change imposed by Sarbanes-Oxley thus pits multiple interests against one another, resulting in a quintessential conflict of interest. In describing the "tempted lawyer problem" embodied in Model Rules 1.7 and 1.8, Stephen Gillers writes, "Since client trust is crucial in enabling lawyers to pursue their clients' goals and protect the clients' autonomy, the rules and law governing lawyers should forbid lawyers (absent client consent) ever to occupy a position in which they are tempted to betray their clients, without regard to whether any particular lawyer would actually succumb to

130. See *supra* notes 30, 32 and accompanying text.

131. Cf. Bruce A. Green, *Criminal Defense Lawyering at the Edge: A Look Back*, 36 *HOFSTRA L. REV.* 353, 395 (2007) ("Maybe criminal defense lawyers need to be cut some slack, in order to revivify an idea of 'zealous [] [advocacy] within the bounds of the law' that places as much emphasis on zealous advocacy as on the legal bounds.") (citation omitted). But see Marc I. Steinberg, *Lawyer Liability After Sarbanes-Oxley—Has the Landscape Changed?*, 3 *WYO. L. REV.* 371, 373 (2003) (arguing that Sarbanes-Oxley has not changed the regulatory landscape, that "the promulgation of SEC standards in this context will not greatly impact counsel's obligations under applicable state ethical rules as well as liability exposure under federal and state law," and that the incentive structure for lawyer behavior has not shifted).

132. See also Ilya O. Podolyako, *The Law of Unintended Consequences: A Critique of the Dilutive Effects and Efficiency Costs of Multilayer Regulation* 6 (Yale Law Sch. Student Scholarship Papers, Paper 91, 2009) ("[I]f a prosecutor charges a defendant with obstruction every time the latter violates a provision of Sarbanes-Oxley, the obstruction category subsumes any substantive prescriptions of the statute. Simultaneously, concepts of justice and the rule of law demand that the acts that are punished carry some logical similarity that allows a citizen to identify prohibited behavior and the accompanying punishment *ex ante*. When certain business decisions count as a breach of positive requirements while others fall into a catchall category, executives are bound to get confused.").

the temptation.”¹³³ Sarbanes-Oxley, however, incentivizes lawyers to “betray their clients” in order to protect their own livelihood.¹³⁴ Risk-averse attorneys faced with a Lauren Stevens-type dilemma will be incentivized to turn everything over to the government when they suspect their client may be engaged in a crime or fraud; they will turn in their clients to avoid the harsh Sarbanes-Oxley obstruction punishments and their collateral effects.¹³⁵ Even when lawyers may reasonably believe their conduct in refusing to disclose information is both ethical and lawful, or that their clients’ conduct is lawful, if there is a fear that the lawyers could be prosecuted they will be incentivized to turn any and all potentially material information over to the government.¹³⁶ This incentive is exactly what the

133. GILLERS, *supra* note 123, at 4.

134. See Dan Reidy & James Burnham, *Federal Criminal Investigations of Lawyers: Risks and Consequences*, JONES DAY, 19 (Mar. 17, 2011), <http://www.jonesday.com/files/Publication/049833b4-8375-439a-8c1d-f3f5a18ad60e/Presentation/PublicationAttachment/e9c84cfd-70cd-4ddc-9e82-9ef61c75457d/FederalCriminalInvest.pdf> (“If prosecutions of lawyers increase, no matter how justified each individual prosecution, lawyers will more often consider their personal exposure when giving advice or otherwise acting on a client’s behalf—a consideration that can be fundamentally at odds with their obligation to zealously advocate for their clients’ interests. And of course, the more lawyers fear punishment for their advocacy or for their advice, the more pressure there is for them to prioritize covering their backs over serving their clients.”).

135. While the law should arguably encourage truth-seeking, promote prosecution of guilty conduct, and incentivize lawyers to reveal evidence of their clients’ clear misdeeds, the real impact of the Sarbanes-Oxley Act will be on the boundaries. It will primarily harm those clients whose conduct is questionable, borderline unlawful, who would most benefit from the zealous protection of their attorneys. It is with those clients, those who are most in need of legal counsel, that lawyers will be most incentivized to be risk averse.

136. See Green, *supra* note 131, at 395 (“[B]oth [the blurriness of the legal and ethical lines and the fallibility of fact findings] encourage risk-averse lawyers to temper their advocacy, avoiding conduct that they believe is lawful and that may well be, but that may be construed differently by prosecutors, disciplinary authorities, and courts. In some regulatory areas, we have no concern about telling people to stay well back from the edge and about punishing them when they stray over it. But legal representation, particularly on behalf of a criminal defendant, is an area where over-deterrence comes at a cost.”); Thomas D. Morgan, *Comment on Lawyers as Gatekeepers*, 57 CASE W. RES. 375, 377–78 (2007) (“Most corporate wrongdoers do not wear signs saying ‘Criminal.’ The normative view [of the role of lawyers as gatekeepers] is sometimes expressed as a preference that we set up a kind screener group [sic] who will certify, or give a ‘Good Housekeeping seal,’ to disclosure documents. . . . [T]hat is virtually impossible to do unless one assumes perfect knowledge or sufficient imagination to anticipate where most of the hidden problems are. I think we can all look back at Enron, and say people should have known something was wrong. But there are many circumstances that are not remotely that clear.”).

Model Rules and internal ethical regulations attempt to avoid.¹³⁷ Additionally it is evidence of the conflicts that arise when external regulation is imposed on lawyers' ethical duties and different regulating bodies have conflicting visions about the role lawyers play.¹³⁸ This is also a client's worst nightmare: at the time when they need their lawyers most, their lawyers will not be able to help them.

The overdeterrence of legitimate lawyering resulting from the Sarbanes-Oxley provisions will have a significant deleterious effect on client representation. Once clients are aware that their lawyers may feel compelled to turn over potentially incriminating information to the government in order to avoid a personal prosecution for obstruction,¹³⁹ clients will be less likely to consult them about potentially incriminating matters.¹⁴⁰ The crux of the protection of the attorney-client relationship is that clients will feel safe and comfortable discussing sensitive matters with their lawyers; the moment clients fail to trust their lawyers, client representation suffers.¹⁴¹ Lawyers will be unable to adequately advocate for their clients if they are unaware of information material to the representation.

In addition, the purposes of Sarbanes-Oxley are not advanced by this risk averse behavior and subsequent client distrust. As clients pull away from their lawyers, lawyers can no longer perform an ade-

137. See MODEL RULES OF PROF'L CONDUCT R. 1.7 cmt. [10] (2012) ("The lawyer's own interests should not be permitted to have an adverse effect on representation of a client.").

138. See *supra* note 18 and accompanying text.

139. Lawyers will want to discuss their conflicting interests with their clients, both because it will help to protect their clients and keep them aware of future actions the lawyer may need to take and because lawyers have an ethical obligation to secure their clients' informed consent to their conflicted representation. See MODEL RULES OF PROF'L CONDUCT R. 1.7(b)(4) (2012); see also *id.* R. 1.7 cmt. [22] (discussing the nuances of obtaining sufficient informed consent to potential future conflicts in the representation).

140. See generally GILLERS, *supra* note 123, at 37 (discussing the policy goals behind the privilege and confidentiality rules and stating that "[t]hey will encourage the client to trust her lawyer and to be forthcoming with information (and sources of information) that the lawyer may need to represent her").

141. See MODEL RULES OF PROF'L CONDUCT R. 1.6 cmt. [2] (2012) ("A fundamental principle in the client-lawyer relationship is that, in the absence of the client's informed consent, the lawyer must not reveal information relating to the representation. . . . This contributes to the trust that is the hallmark of the client-lawyer relationship. The client is thereby encouraged to seek legal assistance and to communicate fully and frankly with the lawyer even as to embarrassing or legally damaging subject matter. The lawyer needs this information to represent the client effectively and, if necessary, to advise the client to refrain from wrongful conduct.").

quate gatekeeping function.¹⁴² When they are not included in major decisions or not privy to information about wrongdoing, they will be unable to identify legal problems for their clients or take any remedial action for the benefit of their clients and the public.¹⁴³ Sarbanes-Oxley meant to encourage the free-flow of information within corporations and with the public so that future corporate scandals could be discovered quickly and prevented sooner.¹⁴⁴ By imposing such harsh changes, the Act has instead driven information about misconduct further underground, incentivizing corporations to keep incriminating information buried deep within their recesses, far from their lawyers and the public.

3. Prosecutorial Overreach

When laws are vague and broad, they leave room for prosecutors to abuse their reach. Prosecutors can push the boundaries of the Sarbanes-Oxley obstruction provisions to police all kinds of lawyer conduct because the boundaries are blurry and malleable. Prosecutors may also be motivated to target lawyers for obstruction violations because they want to punish lawyers whom they perceive to be breaking the rules.¹⁴⁵ If lawyers will not adhere to the rules of

142. See Stephen Fraidin & Laura B. Mutterperl, *Advice for Lawyers: Navigating the New Realm of Federal Regulation of Legal Ethics*, 72 U. CIN. L. REV. 609, 613 (2003) (“A central premise underlying Section 307 and the SEC’s implementing rules is that lawyers are, and will continue to be, essential to corporate activity and to the interpretation and implementation of the laws relating to corporate governance and transactions.”); GILLERS, *supra* note 123, at 533 (“‘[D]eputizing’ [lawyers] to be the eyes and ears of government . . . will lead to less, not more compliance, because the company’s officers may then exclude lawyers from learning about questionable behavior for fear that the lawyer will then have a duty or authority to turn them in. As a consequence . . . the lawyers will never learn about the behavior in time to stop it.”).

143. See Geralyn M. Presti, *Current Ethical Issues for Securities Lawyers—A Comment on Humes*, 57 CASE W. RES. L. REV. 357, 363 (2007) (“I think the worst thing that could happen is that I would be excluded from sensitive corporate discussions, due to a fear that I might disclose privileged information to the SEC. It would be very distressing for me, and it would not be a good consequence for the client. It is critical that I am able to give the client appropriate legal advice and that the client can trust the privilege between us and give me complete information.”).

144. See *supra* notes 20–22 and accompanying text.

145. See Erin Murphy, *Manufacturing Crime: Process, Pretext, and Criminal Justice*, 97 GEO. L.J. 1435, 1446 (2009) (identifying “obstinacy” as a new motivation for prosecutors to charge obstruction of justice, reflecting “a strain of process crime prosecutions aimed at securing convictions against simply defiant or insubordinate individuals—not because their actions actually threaten the integrity of judicial processes or because they are otherwise difficult to convict—but solely because their acts constitute an affront to the formal dignity or authority of the state”); *cf.* Griffin, *Criminal Lying*, *supra* note 83, at 1533 (arguing that prosecutors use false

the game and cooperate with the government (fellow lawyers), prosecutors may wish to send the message that they will prosecute their colleagues for this affront. For example, Philip Russell, an attorney for a church, was slammed with a twenty-year obstruction violation for destroying child pornography found on a church computer before the government was able to use it to indict Russell's client's constituent.¹⁴⁶ Russell was not aware of the government's pending investigation into the constituent, and because his client had no intention of pressing charges, Russell did not believe he was doing anything wrong when he destroyed the images.¹⁴⁷ Commentators largely condemned his prosecution as retribution and prosecutorial overreach.¹⁴⁸

In addition, prosecutors are using obstruction charges pretextually as tools to negotiate deals in cases involving more substantive crimes.¹⁴⁹ An example of this can be seen in *United States v. Stevens*. Instead of truly trying to police her misconduct, Stevens's prosecution was also a ploy to encourage GlaxoSmithKline to settle with the FDA.¹⁵⁰ The vague Sarbanes-Oxley provisions enable this pretextual use of obstruction prosecutions, which will likely increase as the public demands more accountability from executives in the wake of corporate scandals.

These uses of the obstruction of justice provisions are undesirable. First, unconstrained prosecutorial discretion in this area could

statement charges to assert their authority over defendants and force them to apologize); Griffin, *Wanting the Truth*, *supra* note 83, at 7 ("The government's choice to exercise enforcement discretion as broadly as it does seems, in many false statement cases, to be a symbolic assertion of government power.").

146. See *supra* note 123.

147. See GILLERS, *supra* note 123, at 461-62.

148. See, e.g., *id.*

149. See Murphy, *supra* note 145, at 1450.

150. See Mundy & Kendall, *supra* note 93 ("Pharmaceutical companies have paid billions of dollars to settle various marketing-related charges with the government, but only a few executives have pleaded guilty to any crimes. . . . The government hasn't said whether the prosecution of Ms. Stevens was part of an effort to push Glaxo into a plea deal. It said in court documents in December that the Stevens case was part of an 'ongoing underlying health-care fraud investigation' looking at her and 'potential criminal activity by others.'"). Despite the fact that the government claimed it was not prosecuting Stevens with an eye toward incentivizing GlaxoSmithKline to settle with the FDA regarding its allegations of off-label marketing, this issue was on the minds of all those involved. Shortly after Stevens was acquitted, GlaxoSmithKline agreed to settle with the FDA. See Nate Raymond, *Glaxo Agrees 'In Principle' to Record \$3 Billion Settlement with U.S.*, CORPORATE COUNSEL (Nov. 7, 2011), http://www.law.com/corporatecounsel/PubArticleCC.jsp?id=1202524804094&Glaxo_Agrees_In_Principle_to_Record_3_Billion_Settlement_with_US.

lead to “prosecutorial abuse” and “procedural injustice,” making it difficult for lawyers to predict when their behavior crosses the line.¹⁵¹ Second, when laws are vague and prosecutors police conduct that most people perceive as legitimate, confidence in the rule of law is eroded.¹⁵² Frequently prosecuting this type of conduct might have counterproductive effects that “detract from a clear focus on the most salient cases of deception and obstruction [and] suggest that lying to the government is both a standard and an understandable response to investigative inquiries.”¹⁵³

All in all, the implications of courts’ broad interpretations of section 1519 and prosecutors’ willingness to push the reach of the Sarbanes-Oxley obstruction of justice provisions to police the zealous advocacy of lawyers, as evidenced by the recent prosecution of Lauren Stevens, are worrisome. *United States v. Stevens* was a warning signal to the bar, and the threat that more lawyers could be prosecuted for their lawyering still remains palpable. The fear of potential prosecution for conduct that was previously considered legitimate advocacy on behalf of their clients creates a divided loyalty between a lawyer’s personal livelihood and the representation of their clients, and incentivizes lawyers to look out for themselves. This incentive structure undermines both the goals of the ethics rules and Sarbanes-Oxley by compromising the sanctity of the attorney-client relationship and encouraging clients to withhold information from their lawyers. The present scheme, with its resulting overcriminalization, overdeterrence, and prosecutorial overreach, is undesirable, both for lawyers and their clients, and must be reevaluated lest the compelling policies underlying the legal profession’s internal ethical regulations be abandoned.

IV. POTENTIAL SOLUTIONS TO THE PROBLEMS CREATED BY THE SARBANES-OXLEY OBSTRUCTION OF JUSTICE PROVISIONS

The Sarbanes-Oxley obstruction of justice provisions tip the balance between zealous advocacy and compliance with law to promote strict legal compliance to the detriment of legitimate advocacy. This jostling of interests leads to the conflicts and undesirable

151. See Murphy, *supra* note 145, at 1498.

152. See Griffin, *Criminal Lying*, *supra* note 83, at 1550 (“There is . . . strong social scientific support for the general proposition that divergence between commonsense views of justice and the conduct of enforcers diminishes compliance.”).

153. Griffin, *Wanting the Truth*, *supra* note 83, at 12.

incentives discussed above,¹⁵⁴ and subjugates the established policy justifications underlying ethical rules¹⁵⁵ to promote the public interest goals of Sarbanes-Oxley.¹⁵⁶ Although it is not a crime to represent a client who may be engaged in criminal behavior, the Sarbanes-Oxley obstruction provisions instruct prosecutors to harshly scrutinize lawyers' advocacy on behalf of potentially blameworthy clients. Sarbanes-Oxley enables prosecutors to go further than simply prosecuting lawyers who are accused of joining in with their clients' misconduct, requiring prosecution of lawyers who are just doing their jobs.¹⁵⁷ Balance between zealous advocacy and "the bounds of the law" should be restored.¹⁵⁸ To do so, it will be necessary to narrowly interpret the Sarbanes-Oxley provisions so that they fall into the scheme elucidated in *Aguilar* and *Arthur Andersen*. Constraining the reach of section 1519 by imposing a nexus requirement and a dishonest intent requirement resolves all three of the problems discussed above. If courts fail to limit these statutes and allow them to be interpreted broadly, there are alternative steps lawyers and courts can take to avoid the severe consequences of overcriminalization, overdeterrence, and prosecutorial overreach in future obstruction prosecutions.

A. *The Broad Reach of Section 1519 Must Be Constrained*

1. Section 1519 Should Be Interpreted to Only Reach Conduct Endeavored with Dishonest Intent

Courts should analyze Section 1519 according to the boundaries set forth in *Arthur Andersen*. Although Congress may have intended to eschew this type of analysis by omitting the modifier

154. See *supra* Part III.B.

155. See *supra* notes 133, 137 and accompanying text.

156. See *supra* notes 142–44 and accompanying text.

157. See generally REPORT OF THE TASK FORCE ON THE LAWYER'S ROLE IN CORPORATE GOVERNANCE, N.Y. CITY BAR ASSOC. (2006), available at <http://www.nycbar.org>.

158. Some commentators perceive the Sarbanes-Oxley Act as a response to problems created by the promotion of zealous advocacy over other interests. See Carle, *supra* note 12, at 135 (describing how the Enron scandal "exacerbated the crisis of confidence about the effectiveness of current client-centered models of legal ethics regulation"). They argue that we need to shift our priorities and promote public interests over strict adherence to client interests. See Deborah L. Rhode, *Legal Ethics in an Adversary System: The Persistent Questions*, 34 HOFSTRA L. REV. 641, 649 (2006) (arguing that "the public has paid a substantial price for the ethic of undivided client allegiance" and that lawyers need to take into account the public effects of their client representation). While promotion of public interests in the wake of corporate scandals is laudable, it must not be advanced to the detriment of all reasonable client interests.

“corruptly” from the text of section 1519,¹⁵⁹ it could not have intended to criminalize the innocent conduct this revised language targets. In both *Aguilar* and *Arthur Andersen*, the Supreme Court recognized that the obstruction of justice provisions could be interpreted to criminalize previously innocent conduct, and in both cases the Court acted to narrow the scope of the obstruction provisions.¹⁶⁰ Courts should engage in a similar type of analysis with section 1519. They should recognize the negative implications of construing the section broadly and they should constrain it despite Congress’ use of different statutory language. Even though the Supreme Court specifically discussed the difference between “corruptly” and “intent to impede” in *Arthur Andersen*¹⁶¹ (a decision handed down after Congress passed section 1519, which employs the “intent to impede” language), the Court’s policy considerations for requiring a showing of dishonest intent should trump this textual analysis.¹⁶² Therefore, individuals should only be found to violate section 1519 if they possessed a dishonest intent to obstruct a proceeding.

The Sixth Circuit recently interpreted section 1519 in this way, and its argument is instructive. In *United States v. Kernell*, the court observed that no court has definitively held that the intent required for a section 1519 violation is anything less than the corrupt intent required in *Arthur Andersen*.¹⁶³ The court reasoned that despite the different language, section 1519 should be read to require the same corrupt intent as previous obstruction of justice provisions.¹⁶⁴ The same argument can be made when reading the legislative history of section 1519. Because Congress did not explicitly discuss a lesser intent requirement (it merely employed different language), section 1519 should be read to fit into the scheme of previous obstruction provisions. Congress could have intended to broaden the

159. See *supra* notes 64–66 and accompanying text.

160. See *supra* Parts II.A and II.B.

161. See *supra* notes 53–54 and accompanying text.

162. See Abramowitz & Bohrer, *supra* note 67, at 6 (“Although the Supreme Court reversed the Fifth Circuit and recognized the potentially chilling affects that a broad application of the obstruction laws may have on legitimate advocacy [in *Arthur Andersen*], courts faced with §1519 cases would be wise to keep these warnings in mind.”). Even though Sarbanes-Oxley was on the books at the time of the *Arthur Andersen* decision, the Court was not evaluating section 1519 in that case. While its specific textual analysis of section 1512(b) could potentially influence a future interpretation of section 1519, the Court’s policy discussion is broadly applicable to all of the obstruction provisions.

163. See *United States v. Kernell*, 667 F.3d 746, 754 (6th Cir. 2012).

164. See *id.*

reach of the obstruction of justice provisions in other ways, such as by eliminating the nexus requirement, and may not have intended to modify the intent requirement. Because Congress did not explicitly address a change in the intent requirement for a violation of the obstruction provisions, courts should not implicitly construe a modification.

This analysis of section 1519 to require dishonest intent will prevent overcriminalization. Innocent actors with an honest intent to impede a proceeding, such as a lawyer encouraging his client to exercise his attorney-client privilege to not speak to the government, will not be swept up into obstruction prosecutions. This analysis will also prevent future prosecutions like that of Lauren Stevens. By setting a clear boundary between innocent and dishonest conduct and drawing a bright line in this otherwise vague area, this construction of section 1519 will also prevent overdeterrence and prosecutorial overreach.

2. Section 1519 Should Be Interpreted to Implicitly Contain a Nexus Requirement

Section 1519 should be construed to contain a nexus requirement, as articulated in *Aguilar* and refined in *Arthur Andersen*. In *Arthur Andersen*, the Supreme Court made it clear that even though no proceeding need be pending for a violation of section 1512 to occur, a nexus requirement was still applicable.¹⁶⁵ Defendants need to have foreseen a potential future proceeding that their conduct could possibly obstruct to have violated the obstruction statutes.¹⁶⁶ Section 1519 should not be held to a different standard; it should be construed to only condemn those who endeavor to obstruct a foreseeable “contemplat[ed]” proceeding or investigation.¹⁶⁷ Interpreted in this way, innocent destruction or alteration of documents in the ordinary course of business will not be prosecuted unless, at the time of the destruction, it was reasonably foreseeable that the documents would be material to a potential future proceeding.¹⁶⁸

165. See *supra* notes 55–57 and accompanying text.

166. See *id.*

167. The language in section 1519 is similar to the gloss that section 1512(f)(1) puts on section 1512 violations. Compare 18 U.S.C. § 1519 (2006), with 18 U.S.C. § 1512(f)(1) (2006). See also Kyle R. Taylor, Note, *The Obstruction of Justice Nexus Requirement After Arthur Andersen and Sarbanes-Oxley*, 93 CORNELL L. REV. 401, 426–27 (2008) (arguing that the language of section 1519 “lends itself to application of the nexus requirement”).

168. This will prevent the prosecution of defendants who had an innocent intent at the time of destruction but whose conduct in hindsight appears questionable. See Barr, *supra* note 123 (“As a matter of discretion, the government should

This interpretation will also restrain the scope of potentially obstructive conduct and ensure that people like the lying husband in the hypothetical discussed in *Aguilar* are not prosecuted for their conduct unrelated to judicial proceedings.¹⁶⁹ This analysis will prevent future prosecutions like that of Philip Russell and solve all three of the problems discussed above. Furthermore, if section 1519 is not construed to contain a nexus requirement, it could be vulnerable to challenge as unconstitutionally vague, as discussed generally in *Aguilar* and *Arthur Andersen*.¹⁷⁰ The public needs to know when it could violate a criminal law, and without a nexus requirement there is no way for anyone to know when they have crossed the line.

Section 1519 more easily fits into the pre-Sarbanes-Oxley obstruction of justice framework's nexus requirement because of the nuances the Supreme Court added to the requirement in *Arthur Andersen*. It will be harder for courts to blend section 1519 into the intent requirement analysis as set forth in *Arthur Andersen* because of the different statutory language employed in section 1519—a difficulty heightened by the fact that *Arthur Andersen*, which discusses the implications of the use of different statutory language, was decided after section 1519 was passed.¹⁷¹ The negative consequences for client representation and corporate transparency, however, should counsel courts to use caution when articulating the breadth of this provision.

*B. Steps Lawyers and Courts Can Take to Avoid the Harsh
Consequences of a Broad Interpretation of
the Obstruction Provisions*

If courts disregard the potentially vast negative implications and broadly construe section 1519, lawyers will be vulnerable to prosecutions, like that of Lauren Stevens, for their advocacy on behalf of their clients. There are solutions beyond a narrow judicial

in the future simply limit these kinds of obstruction prosecutions to defendants whose actual knowledge of an investigation can be established. Otherwise, we may face the possibility of more individuals (and especially lawyers) being charged with impeding merely hypothetical proceedings, surely an odd application of the federal obstruction statutes.”).

169. See *supra* note 49 and accompanying text.

170. See *Arthur Andersen LLP v. United States*, 544 U.S. 696, 708 (2005); *United States v. Aguilar*, 515 U.S. 593, 600 (1995).

171. Ironically, this does not appear to be the way courts are interpreting section 1519. They are typically instead finding a dishonest intent requirement implied in the statute, but holding that there is no nexus requirement. See, e.g., *United States v. Kernell*, 667 F.3d 746, 754 (6th Cir. 2012).

construction of section 1519, however, to combat the overcriminalization, overdeterrence, and prosecutorial overreach problems that result from the persistent uncertainty of the provision.

First, to prevent overcriminalization, courts could create exceptions for legitimate instances of intent to impede a proceeding that have been traditionally protected. For example they should carve out reasonable, good faith assertions of attorney-client privilege and genuine document retention policies as protected.¹⁷² This could be accomplished by construing the Safe Harbor Provision¹⁷³ to encompass this conduct or by enunciating special exceptions. While creating these exceptions would alleviate some of the pressure on attorneys to be risk-averse in their client representation, the vagueness problems of the statute would remain and the legitimate lawyering previously considered innocent but not protected by an established legal principle like privilege would still be vulnerable to prosecution (i.e., situations like *United States v. Stevens*). This overcriminalization of zealous advocacy is unlikely to be remedied absent a narrow construction of section 1519.

Second, there are ways lawyers can prevent overdeterrence of reasonable zealous advocacy. Lawyers confronted with government investigations can consult outside counsel to confirm the legality of their actions.¹⁷⁴ This was one of the main reasons why Lauren Stevens was able to demonstrate her good faith and escape conviction.¹⁷⁵ If the lawyer can prove that he received all the advice necessary to make his decision, that he relied on the advice in good faith, that his reliance was reasonable, and that he acted on the advice, he may be able to escape conviction.¹⁷⁶ But there are several concerns with relying solely on this defense to escape prosecution.

172. Section 1519 likely would not survive a constitutional challenge if it did not create carve outs for privilege and Fifth Amendment rights.

173. See *supra* note 97.

174. This solution, however, is primarily geared toward aiding in-house counsel. In-house counsel are more likely to be viewed as obstructing prosecutions through their advocacy than other lawyers because of their unique position within their clients' organizations. Outside counsel are not prone to the same self-dealing to which in-house counsel are susceptible. Thus, while in-house counsel may be able to vindicate their legitimate lawyering by proving they relied on consultations with independent, unbiased outside counsel, outside counsel should not have to jump through these hoops. This is a questionable system though, which creates two tiers of burdens of production based on the type of attorney, when all attorneys, no matter their particular career path, are held to the same ethical standards.

175. See *supra* notes 95–96 and accompanying text.

176. See Douglas R. Richmond, *Advice of Counsel as a Defense to Sanctions and Professional Discipline*, 27 *Laws. Manual on Prof. Conduct (ABA/BNA)* 418, 2011 WL 2463403 (June 22, 2011).

First, depending on the situation, consulting outside counsel could entail an additional expense that clients may not be willing to shoulder. If the client refuses to allow his lawyer to consult outside counsel, the lawyer would once again face the dilemma discussed above of choosing between his client and his own livelihood. This solution also raises efficiency concerns by requiring additional layers of legal counsel potentially at multiple and repeated stages of client representation—when we require all lawyers to adhere to ethical standards, why should we demand that one lawyer consult another to ensure good faith compliance with law? Additionally, depending on how section 1519 is interpreted, the prosecution could argue that the defense is inapplicable. The advice of counsel defense only serves to negate the *mens rea* for specific intent crimes.¹⁷⁷ If section 1519 is interpreted to be a general intent crime, as the prosecution argued in *United States v. Stevens*, the advice of counsel defense will be unhelpful to defendant lawyers.¹⁷⁸ Despite the limitations of the advice of counsel defense, consulting outside counsel cannot hurt a lawyer's claims that he was acting in good faith and not seeking to obstruct justice.

177. See *United States v. Stevens*, 771 F. Supp. 2d 556, 560 (D. Md. 2011) (“Good faith reliance on the advice of counsel is only relevant to *specific* intent crimes because such reliance demonstrates a defendant's lack of the requisite intent to violate the law.”) (emphasis added).

178. See *id.* See also Elizabeth R. Sheyn, *Toward a Specific Intent Requirement in White Collar Crime Statutes: How the Patient Protection and Affordable Care Act of 2010 Sheds Light on the “General Intent Revolution”*, 64 FLA. L. REV. 449, 453 (2012) (“A relatively recent trend in the criminal law is the movement away from specific intent to general intent crimes, particularly with respect to white collar crimes.”); Reidy & Burnham, *supra* note 134, at 11–12 (“The government's interpretation means—in theory—that a lawyer can commit obstruction by withholding a document even if that lawyer was advised by outside counsel that the company did not need to produce it. The lawyer need only knowingly not produce the document, at least in part, because the document was harmful to the client. Further, the prosecution of GSK's former in-house counsel even hints at the possibility of the advice of outside counsel forming the basis of a conspiracy between in-house and outside counsel to deceive regulators. . . . This reasoning suggests that if GSK's outside counsel advised GSK's in-house lawyer that certain documents need not be produced, that advice not only failed to insulate GSK's in-house lawyer from liability, but may have exposed the outside lawyers to co-conspirator liability.”). Despite the arguments that could be made that section 1519 is a general intent crime, Congress expressed in the legislative history of the Sarbanes-Oxley Act that it wished for the provision to be a specific intent crime. S. REP. NO. 107–146, at 27 (2002) (“In our view, section 1519 should be used to prosecute only those individuals who destroy evidence with the *specific intent* to impede or obstruct a pending or future criminal investigation, a formal administrative proceeding, or bankruptcy case.”) (emphasis added).

Lawyers can also prevent overdeterrence through education. As more cases are decided in this area, lawyers will be presented with concrete examples of conduct that is permissible and conduct that is prosecutable. They can then strive to educate themselves and their peers about how to simultaneously promote their clients' interests while protecting their own livelihoods.¹⁷⁹ Of course, an inherent dilemma with relying on education to remedy overdeterrence is that the doctrine in this area may be vague and open-ended for quite some time. Additionally, if cases are decided against lawyers so as to constrain zealous advocacy, overdeterrence will increase.

A third way to prevent overdeterrence is to bolster the Safe Harbor Provision of the obstruction statute.¹⁸⁰ The Safe Harbor Provision could be interpreted to encompass all situations in which a lawyer assists his client in legal endeavors, thus protecting all legitimate zealous advocacy.¹⁸¹ The Safe Harbor Provision would not extend to protect legal services administered to help further a client's underlying crime or fraud. While at first glance a bright line rule, however, the Safe Harbor Provision will not necessarily alleviate the broad reach of obstruction prosecutions. As in *United States v. Stevens*, prosecutors could allege that the attorney's "zealous advocacy" was actually an attempt to assist the client in covering up the underlying criminal conduct (or even the lawyer's own misconduct).¹⁸² Reliance on this provision may not provide a complete solution, but it could help to prevent conviction in some cases, as it did in *United*

179. At the 2012 annual meeting of the Association of Corporation Counsel, Lauren Stevens described her experience and advised her fellow in-house counsel, "That is what you do—go back and defend your client zealously and don't back away because you are afraid of my experience." Sue Reisinger, *How Ex-GSK GC Lauren Stevens Fought the Law—and Won*, CORPORATE COUNSEL (Oct. 2, 2012), <http://www.law.com/corporatecounsel/PubArticleCC.jsp?id=1202573330716&the page=1>.

180. See *supra* note 97 and accompanying text. Recently, the Eastern District of North Carolina did just that when it dismissed an indictment against a lawyer for failing to allege an essential element of the charge, holding that, at least in the Fourth Circuit, "where an attorney or one under his direction is charged with an obstruction offense under § 1512(b)(2)(B) or § 1519, the government must expressly allege as an element of the offense in the indictment that the defendant in engaging in the conduct alleged was not 'providing lawful, bona fide, legal representation services in connection with or anticipation of an official proceeding.'" *United States v. Jackson*, No. 2:10-CR-8-FL, 2013 WL 782602, at *21 (E.D.N.C. Feb. 4, 2013) (quoting 18 U.S.C. § 1515(c) (2006)).

181. The court seemed to suggest a similar interpretation of the provision in *United States v. Stevens*. See *supra* note 97 and accompanying text.

182. See *supra* notes 88–91.

States v. Stevens.¹⁸³ This solution is less desirable than a blanket narrow interpretation of section 1519, however, because it would leave non-lawyers vulnerable to broad obstruction prosecutions.

Prosecutorial overreach can be constrained by instituting prosecutorial guidelines that clearly delineate the boundary between legitimate obstruction prosecutions and prosecutorial abuse. This can be done by providing articulable standards for construing the Sarbanes-Oxley obstruction provisions with which prosecutors must comply.¹⁸⁴ Instituting these guidelines could help to prevent prosecutors from believing they have free reign to wield the Sarbanes-Oxley provisions against lawyers in pretextual and coercive ways. While the obstruction doctrine remains broad and unclear, however, it will be difficult to elucidate clear guidelines or to determine when prosecutors have gone too far. This is another area in which the negative effects of Sarbanes-Oxley are unlikely to be resolved absent judicial intervention to cabin the reach of the Act.

Another way in which lawyers can combat prosecutorial overreach and avoid obstruction prosecutions is to cooperate fully with the government during investigations of their clients. Lawyers can be candid in their communications with the government and forthcoming about documents in their possession. For example, if Lauren Stevens had told the FDA that she had produced all documents relevant to the inquiry that she believed, after consulting outside counsel, she was legally obligated to produce, instead of categorically saying that the production was complete, the government may not have been so harsh on her.¹⁸⁵ This solution, however, is riddled with pitfalls. By indicating that he may not have disclosed all the information the government wants, does the lawyer open himself up to harsher sanctions? How far should the lawyer go to cooperate? Should he waive attorney-client privilege? Is he sacrificing zealous advocacy for his client the more he cooperates with the government? The fact that prosecutors will respond more favorably to lawyers who cooperate with them than to lawyers who put up a fight leads to bad incentives; knowing that they could receive more favorable treatment, lawyers will be incentivized to waive privilege

183. See *supra* note 97 and accompanying text.

184. See generally Fred C. Zacharias & Bruce A. Green, *Prosecutorial Neutrality*, 2004 WIS. L. REV. 837 (2004).

185. See Reidy & Burnham, *supra* note 134, at 9, 18 (“[C]ounsel responding to document requests from government agencies should use great care in their communications as to how they describe their production. . . . Every communication matters, and lawyers need to be careful about making overly broad or categorical statements to federal officials.”).

and turn documents over to the government that could prove to be detrimental to their clients.¹⁸⁶ Despite the pitfalls inherent in cooperation with the government, lawyers will be well-served if they recognize the advantages of forthright communication with the government and seek to strike a balance between cooperation and client protection.

As this discussion shows, these various potential solutions are rife with problems and not one is preferable to a narrow judicial construction of section 1519. The only way to truly protect zealous advocacy in the face of the Sarbanes-Oxley obstruction provisions is to cabin the reach of the statute. If courts decline to address the overcriminalization, overdeterrence, and prosecutorial overreach problems created by a broad interpretation of section 1519, lawyers should do all they can to ward off the government's scrutinizing gaze by adhering closely to ethical duties and seeking outside advice. They should try to avoid the temptation to sacrifice their clients' interests to promote their own until courts confine the reach of the obstruction provisions and remove the conflicts created by Sarbanes-Oxley.

CONCLUSION

The obstruction of justice provisions added by Sarbanes-Oxley are extremely broad and could have vast consequences for the ability of lawyers to zealously advocate for their clients. As the prosecution of Lauren Stevens shows, lawyers are vulnerable to prosecution

186. See Lonnie T. Brown, Jr., *Reconsidering the Corporate Attorney-Client Privilege: A Response to the Compelled-Voluntary Waiver Paradox*, 34 HOFSTRA L. REV. 897, 898 (2006) ("The Department of Justice . . . has adopted guidelines that seem to make waiver of the attorney-client privilege and work product protection a prerequisite for being deemed 'cooperative,' a significant designation that carries with it the prospect for more favorable penal treatment."); Robert C. Hockett, *Valuing the Waiver: The Real Beauty of Ex Ante Over Ex Post*, 57 CASE W. RES. L. REV. 381, 387 (2007) ("[I]n the cases we are talking about—namely, waiver that is undertaken in order to be deemed cooperative in the already commenced investigation of a crime alleged already to have been committed—one is seeking the benefit of lighter sentencing. This is an inherently coercive context. And so, unsurprisingly, corporate counsel often have remarked that even when waiver might be 'voluntary' in some metaphysical sense in these cases, it often does not feel that way in any motivational sense.") (emphasis omitted); William R. McLucas, Howard M. Shapiro & Julie J. Song, *The Decline of Attorney-Client Privilege in the Corporate Setting*, 96 J. CRIM. L. & CRIMINOLOGY 621, 622 (2006) ("The current trend has, at a minimum, eroded our traditional adversarial process and skewed the balance of power between government investigators and their corporate targets. . . . It forces corporate managers to think first of their own liability and not the broader good of the enterprise that should be—and once was—at the core of their professional lives.").

if they place a toe over the line separating zealous advocacy from obstruction of justice. Just where this line is and how bright it is will depend on how courts resolve the ambiguities regarding the *mens rea* and nexus requirements of section 1519. Courts should avoid the negative implications of creating a legal framework that incentivizes lawyers to betray their clients and construe section 1519 narrowly to fit within established Supreme Court obstruction of justice precedent. In doing so they will rectify the present imbalance between zealous advocacy and the “bounds of the law” and return the ethical ideal of undivided loyalty to clients back to where it is balanced with legal constraints and the public interest.

INTO THE PERFECT STORM: THE FAILURE OF TRUSTEE ACTIONS AGAINST THIRD PARTIES

NICK AXELROD

Introduction	441				R
I. The Modern Ponzi Scheme	446				R
A. Three Reasons to Target Third Parties	446				R
B. Available Causes of Action	448				R
II. Case Law and Formalistic Shortcomings	455				R
A. The Failure of Section 541	456				R
B. Section 544(a)(1) Actions	463				R
C. Conclusion	470				R
III. Fundamentals of Bankruptcy and Application to Creditor Actions Against Third Parties	471				R
A. Theoretical Justifications	471				R
1. The Theory of the Creditors' Bargain	471				R
2. The Traditional Account	474				R
3. Internalizing Costs and the Moralistic Account	476				R
B. When The Two Theories Agree: Ponzi Schemes .	477				R
1. Trustee Suits and the Creditors Bargain	477				R
2. Traditionalist Conceptions	483				R
3. Forum Shopping and Incentives	485				R
4. Conclusion	489				R
Conclusion	489				R

INTRODUCTION

In 2009, Ponzi schemes¹ collapsed in record numbers. From January of 2007 to June of 2009, they made up 3.5% of the federal

1. The courts define a Ponzi scheme as any scheme that uses fraudulently obtained funds to pay off previous investors, forestalling disclosure of the fraud. *Bayou Superfund, LLC v. WAM Long/Short Fund II (In re Bayou Grp., LLC)*, 362 B.R. 624, 633 (Bankr. S.D.N.Y. 2007) (rejecting a narrower definition of Ponzi scheme that required high returns and no legitimate business and collecting cases embracing the more general definition). *See also* *Danning v. Bozek (In re Bullion Reserve of N. Am.)*, 836 F.2d 1214, 1219 n.8 (9th Cir. 1988) (“[a] ‘Ponzi’ scheme is any sort of fraudulent arrangement that uses later acquired funds or products to pay off previous investors.”).

class action docket, up from just .2% throughout all of 2005–2006.² Bernard L. Madoff's multi-decade fraud was the largest of these Ponzi schemes. Over the course of 20 years, Madoff defrauded his investors of billions of dollars in cash, generating sixty-five billion dollars in paper losses.³ The victims included major institutional investors, individuals, and charities, including, among others, Nobel Laureate Elie Weisel and his Elie Weisel Foundation for Humanity.⁴ Though the Madoff fraud was the largest such scheme in history, it was by no means an isolated incident: R. Allen Stanford is alleged to have defrauded investors of more than \$7 billion⁵ and Mark Dreier another \$400 million dollars.⁶ The schemes themselves are not new. After all, Charles Ponzi perpetrated his eponymous fraud over 80 years ago. Modern Ponzi schemers, however, take advantage of global capital markets and sophisticated electronic trading systems to prolong their schemes for far longer than they previously could have.⁷ Yet this only delays an inevitable collapse. Eventually the music stops and the party ends. A rush of redemptions, like those tied to the global financial meltdown, often triggers the collapse.⁸ When the end does come, the most common result is a Chapter 7 liquidation under the Bankruptcy Code ("Code").

In a Chapter 7 liquidation of a Ponzi scheme, the Bankruptcy Trustee, a practitioner appointed by the court and overseen by the United States' Trustee's office, is charged with unwinding such a scheme. Courts, however, have interpreted the Code to leave the trustee with few prospects for recovery.⁹ First, the nature of the

2. Kenneth C. Johnston et al., *Ponzi Schemes and Litigation Risks: What Every Financial Services Company Should Know*, 14 N.C. BANKING INST. 29, 34 n.38 (2010).

3. See Christine Hurt, *Evil Has A New Name (And A New Narrative): Bernard Madoff*, 2009 MICH. ST. L. REV. 947, 952 (2009) (recording sixty-five billion dollars in paper losses and citing estimates that put actual losses at twenty billion).

4. See Stephanie Strom, *Elie Wiesel Levels Scorn at Madoff*, N.Y. TIMES, Feb. 26, 2009, <http://www.nytimes.com/2009/02/27/business/27madoff.html>.

5. Complaint, *S.E.C. v. Int'l Bank, Ltd.*, available at <http://www.sec.gov/litigation/complaints/2009/comp20901.pdf>.

6. See Benjamin Weiser, *Lawyer Pleads Guilty in \$400 Million Fraud*, N.Y. TIMES, May 12, 2009, at A23, available at <http://www.nytimes.com/2009/05/12/nyregion/12dreier.html>.

7. See, e.g., *Picard v. JPMorgan Chase & Co.*, 460 B.R. 84 (S.D.N.Y. 2011) (summarizing decades long relationship between Bernard Madoff and JPMorgan Chase); see also Clarence L. Pozza, Jr., et al., *A Review of Recent Investor Issues in the Madoff, Stanford and Forte Ponzi Scheme Cases*, 10 J. BUS. & SEC. L. 113, 115 (2010) (discussing claims against the banks and counterparties involved in various Ponzi Schemes).

8. Joshua Marcus & Jake Greenberg, *Ponzi Schemes: Washed Ashore by Recession's Low Tide, Reveal Controversial Issues*, 29 AM. BANKR. INST. J. 48 (2010).

9. See *infra* Part 2 and accompanying notes.

scheme means that the debtor itself is unlikely to have many assets.¹⁰ Second, although veil-piercing is an option, in recent cases like the Madoff Ponzi scheme, it has done little good.¹¹ Finally, a series of formalistic bars prevents the trustee from going after the deep-pocketed third parties that may have contributed to the fraud.¹² Commentators and judges alike have pointed out the inequities created by this “perfect storm” of code provisions.¹³ The trustee cannot sue on behalf of the estate because, in theory, he was a part of the fraud¹⁴ and often there are no transfers or preferences to avoid. At the same time, the trustee’s power under Section 544 has been interpreted by courts to preclude the trustee from suing as a hypothetical lien creditor.¹⁵ The result is that creditors who could not have contracted for protection are out of luck. Commentators have spilled much ink analyzing potential loopholes in the courts’ reasoning and others have brought a contortionist’s skill to statutory interpretation seeking an exception.¹⁶ The answer is more straightforward: the trustee’s powers should extend only as far as

10. Ponzi schemes collapse when there are insufficient new funds to cover redemptions, thus by definition they are already insolvent at the time of collapse. Compare Madoff’s 65 billion in liabilities and 200–300 million in unredeemed assets. Complaint at 5, 6, SEC v. Madoff, No. Civ. 08-10791, 2009 WL 980288 (S.D.N.Y. Apr. 10, 2009), 2008 WL 5197070, available at <http://www.sec.gov/litigation/complaints/2008/comp-madoff121108.pdf>; Pozza, Jr., et al., *supra* note 7, at 116.

11. The Madoff proceeding was substantively consolidated with his personal case. Consent Order Substantively Consolidating the Estate of Bernard L. Madoff into the SIPA Proceeding of Bernard L. Madoff Investment Securities LLC and Expressly Preserving All Rights, Claims and Powers of Both Estates, Sec. Investor Prot. Corp. v. Bernard L. Madoff Investment Securities LLC (*In re* Bernard L. Madoff), No. 09–11893 (Bankr. S.D.N.Y. June 9, 2009), available at <http://madoff.com/document/dockets/000525-252-order.pdf>. See generally 1 COLLIER ON BANKRUPTCY ¶ 105.08[1], at 105–23 (Alan N. Resnick & Henry J. Sommer eds., LexisNexis, 16th ed. 2008) (discussing the linkages in bankruptcy law between veil piercing and substantive consolidation); *Union Savings Bank v. Augie/Restivo Banking Co.*, 860 F.2d 515, 518 (2d Cir. 1988).

12. See *infra* Part 3.

13. See, e.g., Tanvir Alam, *Fraudulent Advisors Exploit Confusion in the Bankruptcy Code: How In Pari Delicto Has Been Perverted to Prevent Recovery for Innocent Creditors*, 77 AM. BANKR. L.J. 305, 315–17 (2003); Jeffrey Davis, *Ending the Nonsense: The In Pari Delicto Doctrine Has Nothing to Do with What Is § 541 Property of the Bankruptcy Estate*, 21 EMORY BANKR. DEV. J. 519 (2005); William McGrane, *The Erroneous Application of the Defense of In Pari Delicto to Bankruptcy Trustees*, 29 CAL. BANKR. J. 275 (2007); Samuel C. Wasserman, *Can the Trustee Recover? Imputation of Fraud to Bankruptcy Trustees in Suits Against Third-Party Service Providers*, 77 FORDHAM L. REV. 365 (2008).

14. See *infra* Part 3.A.ii.

15. See *infra* Part 3.A.ii.

16. See note 13 & accompanying text (collecting articles).

they can be justified by the doctrines that motivate bankruptcy law. Doing so would provide the trustee with precisely the powers she has been denied with far less complexity. This Note argues that the courts have misconstrued the role of the bankruptcy trustee by limiting her powers without any theoretical justification. Specifically it argues that an application of the basic, foundational principles of the Bankruptcy Code suggests that the trustee should have the power to bring actions against third parties on behalf of creditors. Practitioners and judges cannot be expected to remedy the situation alone without departing from a plain reading of the Code. Legislative action would therefore provide the most effective mechanism for aligning bankruptcy law with its theoretical justifications. In the absence of a willing legislature, however, the Bar and the judges who face this issue on a daily basis would do well to confront the policy implications of the issue head-on and interpret the Code to permit the trustee to bring actions against third parties.

There are two leading theoretical explanations for bankruptcy: the collectivist and traditional theories. The collectivist account focuses on maximizing recovery and justifies bankruptcy as a tool to reduce administrative costs and decrease strategic behavior by individual creditors.¹⁷ The traditional account is broader and considers the benefits of bankruptcy for the debtor and the creditors.¹⁸ In the collectivist account, the trustee's powers to maximize the creditors'

17. See generally THOMAS H. JACKSON, *THE LOGIC AND LIMITS OF BANKRUPTCY LAW* (1986); BARRY E. ADLER, *BANKRUPTCY AND RISK ALLOCATION*, 77 *CORNELL L. REV.* 439 (1992); BARRY E. ADLER, *FINANCIAL AND POLITICAL THEORIES OF AMERICAN CORPORATE BANKRUPTCY*, 45 *STAN. L. REV.* 311 (1993); DOUGLAS G. BAIRD, *BANKRUPTCY'S UNCONTESTED AXIOMS*, 108 *YALE L.J.* 573, 592 (1998); THOMAS H. JACKSON, *AVOIDING POWERS IN BANKRUPTCY*, 36 *STAN. L. REV.* 725, 726 (1984) [hereinafter Jackson, *Avoiding Powers*]; THOMAS H. JACKSON, *BANKRUPTCY, NON-BANKRUPTCY ENTITLEMENTS, AND THE CREDITORS' BARGAIN*, 91 *YALE L.J.* 857, 866 (1982) [hereinafter Jackson, *Bankruptcy, Non-Bankruptcy*]; THOMAS H. JACKSON, *OF LIQUIDATION, CONTINUATION, AND DELAY: AN ANALYSIS OF BANKRUPTCY POLICY AND NONBANKRUPTCY RULES*, 60 *AM. BANKR. L.J.* 399, 400 (1986) [hereinafter Jackson, *Of Liquidation*]; THOMAS H. JACKSON & ROBERT E. SCOTT, *ON THE NATURE OF BANKRUPTCY: AN ESSAY ON BANKRUPTCY SHARING AND THE CREDITORS' BARGAIN*, 75 *VA. L. REV.* 155, 190 (1989).

18. See THEODORE EISENBERG, *A BANKRUPTCY MACHINE THAT WOULD GO OF ITSELF*, 39 *STAN. L. REV.* 1519 (1987); DONALD R. KOROBKIN, *REHABILITATING VALUES: A JURISPRUDENCE OF BANKRUPTCY*, 91 *COLUM. L. REV.* 717, 721 (1991) [hereinafter Korobkin, *Rehabilitating Values*]; DONALD R. KOROBKIN, *VALUE AND RATIONALITY IN BANKRUPTCY DECISIONMAKING*, 33 *WM. & MARY L. REV.* 333 (1992) [hereinafter Korobkin, *Value and Rationality*]; ROBERT K. RASMUSSEN, *DEBTOR'S CHOICE: A MENU APPROACH TO CORPORATE BANKRUPTCY*, 71 *TEX. L. REV.* 51, 55-59 (1992) (arguing creditors' bargain theory ignores reality); ELIZABETH WARREN, *BANKRUPTCY POLICY*, 54 *U. CHI. L. REV.* 775, 795-97 (1987) [hereinafter Warren, *Bankruptcy Policy*]; ELIZABETH WARREN, *BANKRUPTCY POLICYMAKING IN AN IMPERFECT WORLD*, 92 *MICH. L. REV.* 336 (1993) [hereinafter Warren,

return are circumscribed by the need to minimize the potential for forum shopping and distortion of bargained-for state law entitlements.¹⁹ At times, this suggests a narrow role for the trustee. For instance, she theoretically should not always be able to rearrange rights between creditors.²⁰ Yet applying the same limits to trustee suits against third parties misses the fundamental differences between the aftermath of a Ponzi scheme and a normal Chapter 7 or Chapter 11 bankruptcy case. As I explain in Part Four below, collectivist theory, as applied to a Ponzi scheme, suggests a far broader role.²¹ In a Ponzi bankruptcy, there are few participants, no secured creditors, and little going-concern value.²² Crucially, individual creditors do not have incentives to bring claims against third parties in this situation, even though doing so would maximize the group's total recovery.²³ The costs of racing behavior and uncertainty plague the process. Far from suggesting a narrow role for the trustee, these are precisely the circumstances where the collectivist account dictates that the trustee should have broad powers.

Part Two begins with a brief overview of the mechanics of a Ponzi scheme, with a particular focus on the causes of action typically available against third parties in a bankruptcy case. They range from negligence and deepening insolvency to more traditional theories of aiding and abetting fraud. Part Three considers the courts' construction of Sections 541, 544, and 704 in light of the Supreme Court's holding in *Caplin v. Marine Midland Grace Trust Co. of New York*.²⁴ Part Four presents the collectivist and traditional accounts. The Note then returns to the first principles of bankruptcy law by evaluating whether the courts' approach is in line with either theory of bankruptcy. It finds that the courts' construction of the trustee's power is inconsistent with either the collectivist or traditional bankruptcy theory and misconstrues the trustee's proper role. Finally, Part Five considers some of the remedies that have been suggested, from reinterpreting Section 544 to amending the Code as suggested by the *Caplin* majority. Ultimately, this Note concludes that while amending the Code to provide the trustee with the power to sue third parties on behalf of creditors is a good idea, the

Policymaking]; Elizabeth Warren & Jay Lawrence Westbrook, *Searching for Reorganization Realities*, 72 WASH. U. L.Q. 1257, 1265–86 (1994).

19. See *infra* notes 171–86 & accompanying text.

20. See *infra* notes 209–11 & accompanying text.

21. See *infra* notes 208–28 & accompanying text.

22. See *supra* notes 8–10 & accompanying text.

23. See *infra* notes 218–19 & accompanying text.

24. 406 U.S. 416 (1972).

power has no place in Section 544 and properly requires Congress to either adopt a new Code provision or modify §704.

I. THE MODERN PONZI SCHEME

As the recent financial crisis spiraled outward, the Commodities Futures Trading Commission (“CFTC”) filed twenty-two actions to halt ongoing schemes during the first six months of 2009.²⁵ Though many schemes are smaller, a few stand out for their incredible scope, with losses in the billions. For example, according to the Department of Justice, R. R. Allen Stanford’s scheme totaled seven billion dollars, second only to Madoff’s sixty-eight billion dollar fraud.²⁶ While all of these schemes are fundamentally similar—each using new money to pay out old investors—the larger schemes are more likely to require the involvement of third party facilitators like banks and brokers.²⁷ For instance, while some smaller schemes rely on cash from victims, it would be difficult to operate a securities-based scheme like Madoff’s without the use of banks, and the same holds for Stanford’s scheme, which relied on savings accounts.²⁸ Along with several other characteristics common to all Ponzi schemes, this makes them particularly attractive targets for enterprising trustees. This Part explains the characteristics that make third parties frequent targets for trustees and then examines the causes of action trustees most frequently pursue.

A. *Three Reasons to Target Third Parties*

First, courts consider all Ponzi schemes insolvent from the moment they begin to operate.²⁹ Moreover, every transfer made as a part of a Ponzi scheme has no purpose other than to “hinder, delay

25. See Kevin McCoy, *Recession Forces Unraveling Ponzi Schemes into the Open*, USATODAY.COM (Apr. 17, 2009), http://www.usatoday.com/money/markets/2009-04-17-markets-recession-scheme-ponzi_N.htm.

26. See Zachary Goldfarb and Anita Kumar, *Stanford, 5 Associates Charged with Running \$7 Billion Ponzi Scheme*, WASH. POST. (June 20, 2009), <http://www.washingtonpost.com/wp-dyn/content/article/2009/06/19/AR2009061900078.html>; Leslie Wayne, *Troubled Times Bring Mini-Madoffs to Light*, N.Y. TIMES (Jan. 28, 2009), <http://www.nytimes.com/2009/01/28/business/28ponzi.html>.

27. See *infra* notes 38–45 & accompanying text.

28. See *infra* note 38 & accompanying text.

29. Miriam A. Cherry & Jarrod Wong, *Clawbacks: Prospective Contract Measures in an Era of Excessive Executive Compensation and Ponzi Schemes*, 94 MINN. L. REV. 368, 401 (2009); *Bayou Superfund, LLC v. WAM Long/Short Fund II (In re Bayou Group, LLC)*, 362 B.R. 624 (Bankr. S.D.N.Y. 2007).

or defraud” creditors.³⁰ Because they are insolvent, this makes any transfer made during the course of the scheme technically a fraudulent conveyance.³¹ Under Section 548 of the Bankruptcy Code, these transfers can be voided in excess of the amount transferred by the debtor in good faith.³² Finally, the nature of the scheme means that excepting net-winners and the complicit, on a theoretical level, each investor is defrauded equally on a pro rata basis.³³

The second reason trustees target third parties in Ponzi scheme bankruptcies is the relatively small size of the estate. Although Madoff’s estate eventually grew to over nine billion dollars,³⁴ before the trustee began bringing avoidance actions and clawbacks, the estate was unlikely to offer much solace for the creditors.³⁵ This flows logically from the nature of the scheme: Ponzi schemes fail because their perpetrators run out of money to cover redemptions. Thus, in most cases, there is initially little left over in the estate besides the furniture.³⁶ This leads trustees to look outward at avoidance actions and potential suits against third parties to bolster prospects for recovery to enlarge the estate.³⁷

Lastly, the most complex (and often most harmful) Ponzi schemes are likely to involve deep-pocketed third parties on some level.³⁸ Again, consider the Madoff scheme. Perpetuating such a scheme required recruiting new investors and satisfying old inves-

30. *In re* Manhattan Inv. Fund, 397 B.R. 1, 8 (S.D.N.Y. 2007) (“There is a general rule—known as the ‘Ponzi scheme presumption’—that such a scheme demonstrates ‘actual intent’ as matter of law because ‘transfers made in the course of a Ponzi scheme could have been made for no purpose other than to hinder, delay or defraud creditors.’”) (quoting *In re* Manhattan Fund Ltd., 359 B.R. 510, 517–18 (Bankr. S.D.N.Y. 2007)).

31. Cherry & Wong, *supra* note 29, at 399; Hurt, *supra* note 3, at 971–72; UNIF. FRAUDULENT TRANSFER ACT §§ 4–5 (1984).

32. 11 U.S.C. § 548(a)(1) (2006) (“The trustee may avoid any transfer . . . if the debtor voluntarily or involuntarily . . . incurred such obligation with actual intent to hinder, delay, or defraud any entity to which the debtor was or became, on or after the date that such transfer was made or such obligation was incurred.”).

33. Cherry & Wong, *supra* note 29, at 407.

34. THE MADOFF RECOVERY INITIATIVE, <http://madoff.com/recoveries-04.html> (last visited Feb. 28, 2012) (recording over nine billion dollars in recoveries).

35. *See id.* (noting that the trustee was able to glean less than one billion from the estate before avoidance and clawback actions, less than 1/17th of the losses recognized by the courts).

36. *See supra* note 10.

37. Wasserman, *supra* note 13, at 366–68. *See also* Johnston et al., *supra* note 2, at 45–47.

38. *See* Johnson et al., *supra* note 2, at 34–43 (discussing various doctrines for establishing the liability of financial institutions for their role in Ponzi schemes); Del Quentin Wilber, *Economic Downturn Accelerates Collapse of Ponzi Schemes*, WASH.

tors that their investments were safe and growing. The Madoff trustee estimates that some eight billion dollars of Madoff's investments were recruited through the use of "feeder funds" marketed by major banks.³⁹ Funds like these require administrators and custodians, who are likely to be large and reputable financial services companies.⁴⁰ Moreover, keeping current investors happy requires falsified trade records and accounting records, which implicates additional third parties.⁴¹ For instance, in *Hirsch v. Arthur Andersen*, Ponzi scheme victims sued the auditors who failed to uncover the fraud.⁴² In *Shearson Lehman Hutton, Inc. v. Wagoner*, victims sued the brokers responsible for "churning" the portfolio, a process used to generate the appearance of legitimate trading activity.⁴³ In schemes like Stanford's and Madoff's, both of which involved fictitious savings accounts, third-party banks are often crucial counterparties.⁴⁴ Finally, Ponzi schemes involving hundreds of millions or billions of dollars may rely on banks to warehouse proceeds before distribution.⁴⁵

Taken together, these three factors make lawsuits against third parties, and especially against financial service providers, extremely attractive. The transactions themselves are definitively fraudulent, the paltry estate leaves the trustee with few other options, and third parties—often large, deep-pocketed financial institutions—are necessarily involved. In short, it is unsurprising that so many trustees hear the siren's song of third-party liability.

B. Available Causes of Action

When they do bring suit, trustees can choose from several different state law causes of action. Chief among them are negligence, fraudulent transfer, breach of fiduciary duty, aiding and abetting

POST, Jun. 12, 2009, at B1 (discussing role of banks and new technologies in perpetuating Ponzi schemes).

39. Amended Complaint at 21, *Sec. Investor Prot. Corp. v. Bernard L. Madoff Inv. Sec. LLC*, No. 08-01789 (BRL) (Bankr. S.D.N.Y. Dec. 5, 2010).

40. *Id.* at 137–38.

41. See *Hirsch v. Arthur Andersen & Co.*, 72 F.3d 1085 (2d Cir. 1995) (trustee suit against third party accountants for their complicity in a Ponzi Scheme); *Sharp Int'l Corp. v. KPMG LLP (In re Sharp Int'l Corp.)*, 278 B.R. 28 (Bankr. E.D.N.Y. 2002) (trustee suit against third party accountants for their negligent and reckless failure to uncover and disclose Chapter 11 corporate debtor's fraud).

42. *Hirsch*, 72 F.3d 1085.

43. 944 F.2d 114, 117 (2d Cir. 1991).

44. Zachary A. Goldfarb, *SEC Alleges \$8 Billion Savings Fraud*, WASH. POST, Feb. 18, 2009, at D1.

45. Johnston et al., *supra* note 2, at 43–45 (discussing lawsuits against J.P. Morgan Chase & Co. for its complicity in the Madoff scheme).

fraud, and deepening insolvency.⁴⁶ Though the claims operate under different names, the crux of each allegation is fundamentally the same.⁴⁷ Each is predicated on the claim that professional malfeasance either intentionally or negligently perpetuated the scheme. As I explain in Part Four, the availability of a state law cause of action is essential because minimizing the deviation between underlying state law and bankruptcy is a central goal of the Supreme Court's bankruptcy jurisprudence.⁴⁸ Thus, while I explain each possible cause of action below, it is important to keep in mind that so long as at least one theory is viable, the analysis stands.

For a negligence action to succeed, the third party must (1) owe the creditors a duty of care, (2) breach that duty, and (3) proximately cause damages.⁴⁹ The difficulty with this cause of action is that, except in a few anomalous cases, the third party is unlikely to owe the creditors a duty of care.⁵⁰ For instance, a duty of care may arise when a bank or investment fund directly solicits investments from creditors.⁵¹ Absent such a relationship, however, there is un-

46. *See id.* at 34–43 (discussing theories of liability under which bank liability obtains including breach of fiduciary duty, negligence, negligent misrepresentation, fraudulent transfer, and aiding and abetting fraud).

47. *Cenco Inc. v. Seidman & Seidman*, 686 F.2d 449, 453 (7th Cir. 1982) (Posner, J.) (“[B]reach of contract, negligence, and fraud, when committed by auditors, are a single form of wrongdoing under different names.”).

48. *See Butner v. U.S.*, 440 U.S. 48, 55 (1979)

Property interests are created and defined by state law. Unless some federal interest requires a different result, there is no reason why such interests should be analyzed differently simply because an interested party is involved in a bankruptcy proceeding. Uniform treatment of property interests by both state and federal courts within a State serves to reduce uncertainty, to discourage forum shopping, and to prevent a party from receiving “a windfall merely by reason of the happenstance of bankruptcy.”

(citations omitted).

49. *See* Prosser, *Torts* (4th ed.), § 30; *Integrated Waste Servs., Inc. v. Akzo Nobel Salt, Inc.*, 113 F.3d 296, 299 (2d Cir. 1997). Generally “professional malpractice is a species of negligence.” *Hydro Investors, Inc. v. Trafalgar Power Inc.*, 227 F.3d 8, 15 (2d Cir. 2000) (quoting *Marks Polarized Corp. v. Solinger & Gordon*, 476 N.Y.S.2d 743, 744 (1984)).

50. *See* Prosser, *Torts* (4th ed.), § 30. *See also* *Sharp Int'l Corp v. KPMG LLP (In re Sharp Int'l Corp.)*, 278 B.R. 28, 35 (Bankr. E.D.N.Y. 2002) (finding Trustee had sufficiently pled negligence based on accountant's failure to uncover fraud). *But see* *Sec. Investor Prot. Corp. v. BDO Seidman, LLP*, 222 F.3d 63, 73 (2d Cir. 2000) (establishing a test for negligent misrepresentation when the professional is aware that plaintiffs would rely on their work product in the context of a suit against an accounting firm).

51. This may arise out of a vehicle like a feeder fund, which refers potential investors to the scheme in exchange for fees. *See, e.g., S.E.C. v. Cohmad Sec. Corp.*, No. 09-CV-5680, 2010 WL 363844, at *3 (S.D.N.Y. Feb. 2, 2010) (finding a duty for

likely to be such a claim. Moreover, if an action does exist, it would only accrue to those individual creditors who were actually induced to invest by the third party. As I explain in Part Three, trustees attempting to bring these individual-based actions have not had success in court and the theoretical basis for a collective bankruptcy process does not support making such actions available to the trustee.⁵²

Fraudulent transfer claims rely on the Uniform Fraudulent Transfers Act,⁵³ as embodied in Section 548 of the Code.⁵⁴ As noted above, Ponzi schemes are technically insolvent from their inception. But for a transfer to be fraudulent, it must actually be a transfer, as opposed to a bailment.⁵⁵ Taken together, these two requirements doom most fraudulent transfer actions against third-party service providers. The payments for services, which are more likely to be considered transfers, are unlikely to be of a magnitude to make a dent in losses running into the billions. Although Section 548 may play a role in recovery, it is primarily clawback actions that redistribute the costs of the fraud between net winners and losers.⁵⁶

a Madoff feeder fund, however dismissing under Rule 9(b) for failing to plead facts “with particularity . . . constituting fraud”).

52. See *infra* notes 92–168 & accompanying text.

53. UNIF. FRAUDULENT TRANSFER ACT § 4(a)(1) (1984) (“A transfer made . . . by a debtor is fraudulent as to a creditor . . . if the debtor made the transfer or incurred the obligation: [w]ith actual intent to hinder, delay, or defraud any creditor of the debtor.”).

54. “The trustee may avoid any transfer . . . if the debtor voluntarily or involuntarily . . . made such transfer or incurred such obligation with actual intent to hinder, delay, or defraud any entity to which the debtor was or became, on or after the date that such transfer was made or such obligation was incurred, indebted.” 11 U.S.C. § 548(a)(1) (2006).

55. Section 548 only applies to transfers of interests in property and thus cannot be applied to a bailment. *Id.*; *Golden v. The Guardian (In re Lenox Healthcare, Inc.)*, 343 B.R. 96, 100 (Bankr. D.Del. 2006) (noting that an exercise of avoidance powers requires “a transfer of property of the estate”). See also *Maxwell v. Penn Media (In re marchFirst)*, No. 01-B-24742, 2010 WL 4027723, at *8 (Bankr. N.D. Ill. Oct. 14, 2010) (“[T]here was no such transfer in this case because marchFirst made the payments to Penn with NCS funds it was holding as agent or bailee.”).

56. See *Bayou Superfund, LLC v. WAM Long/Short Fund II (In re Bayou Grp., LLC)*, 396 B.R. 810, 827 (Bankr. S.D.N.Y. Oct. 16, 2008) (discussing the “equitable” role of Section 548 in effecting “Congress’ determination that under limited circumstances creditors must share equally in the insolvency. . . .”); Karen E. Nelson, *Turning Winners into Losers: Ponzi Scheme Avoidance Law and the Inequity of Clawbacks*, 95 MINN. L. REV. 1456, 1475 (2010) (discussing use of Section 548 in Ponzi schemes to clawback fictitious profits). In a clawback action the trustee relies on Section 548(a)(1), which empowers the trustee to void transfers if he can show that (1) less than reasonably equivalent value was received, and (2) the transferee

This distribution of losses between the creditors is the subject of constant attention and debate among both judges and academics⁵⁷ and is beyond the scope of this Note.

Several courts have embraced deepening insolvency as a cause of action, starting with the Seventh Circuit in *Schact v. Brown*.⁵⁸ Though exact definition is elusive, the essence of this cause of action is that a third party's conduct allowed the corporation to prolong its existence, incurring additional debt and wasting additional resources.⁵⁹ In *Schact*, Judge Posner wrote that courts that reject deepening insolvency as a cause of action:

[R]est upon a seriously flawed assumption, i.e., that the fraudulent prolongation of a corporation's life beyond insolvency is automatically to be considered a benefit to the corporation's interests. This premise collides with common sense, for *the corporate body is ineluctably damaged by the deepening of its insolvency, through increased exposure to creditor liability*.⁶⁰

Despite Judge Posner's forceful words, not every court accepts the theory and it remains a question of state law. Thus, although the Third Circuit accepted deepening insolvency as a stand-alone cause of action in *Official Committee of Unsecured Creditors v. R.F. Lafferty*,⁶¹ a result that has been endorsed by the Ninth⁶² and Second

was insolvent at the time of the transfer. 11 U.S.C. § 548(a)(1) (2006); *Bayou Superfund, LLC v. WAM Long/Short Fund II (In re Bayou Grp., LLC)*, 396 B.R. at 827 (applying section 548(a)(1) to trustee's Ponzi scheme clawback action).

57. See generally Adler, *supra* note 17 (critiquing the use of loss distribution amongst creditors as a risk-sharing mechanism); Barry E. Adler, *An Equity-Agency Solution to the Bankruptcy-Priority Puzzle*, 22 J. LEGAL STUD. 73 (1993); Douglas G. Baird & Donald S. Bernstein, *Absolute Priority, Valuation Uncertainty, and the Reorganization Bargain*, 115 YALE L.J. 1930 (2006) (discussing deviations from absolute priority and distributions of equity in Chapter 11 cases); James W. Bowers, *Whither What Hits the Fan?: Murphy's Law, Bankruptcy Theory, and the Elementary Economics of Loss Distribution*, 26 GA. L. REV. 27, 44–68 (1991); Lynn M. LoPucki, *The Unsecured Creditor's Bargain*, 80 VA. L. REV. 1887 (1994).

58. 711 F.2d 1343, 1350 (7th Cir. 1983), *cert. denied*, 464 U.S. 1002 (1983) (“[T]he corporate body is ineluctably damaged by the deepening of its insolvency, through increased exposure to creditor liability.”).

59. See *id.*; *Smith v. Arthur Andersen, LLP*, 421 F.3d 989, 1003–04 (9th Cir. 2005) (finding a “cognizable harm” when an insolvent corporation's existence was prolonged); *Official Comm. of Unsecured Creditors v. R.F. Lafferty & Co.*, 267 F.3d 340, 350 (3d Cir. 2001) (recognizing deepening insolvency based on continued corporate existence); *Allard v. Arthur Andersen & Co.*, 924 F. Supp. 488, 494 (S.D.N.Y. 1996) (same); *Feltman v. Prudential Bache Sec.*, 122 B.R. 466, 473 (S.D. Fla. 1990) (same).

60. *Schacht*, 711 F.2d at 1350 (emphasis added) (citations omitted).

61. 267 F.3d 340 (3d Cir. 2001).

62. *Cf. Smith*, 421 F.3d at 1003–04.

Circuits,⁶³ acceptance is by no means universal.⁶⁴ In the Delaware Chancery Courts, then-Vice Chancellor Strine refused to recognize it.⁶⁵ Moreover, commentators have questioned the logic of accepting deepening insolvency as a stand-alone cause of action, as opposed to merely as a theory of damages.⁶⁶ Even if accepted, the application of deepening insolvency to a Ponzi schemes remains unclear. For instance, *Schacht* distinguishes between *Cenco*-like corporations, which are established purely as vehicles for fraud, and legitimate corporations that can be harmed, as in *Schacht* itself.⁶⁷ As a theoretical matter this would seem to conflict with the application of deepening insolvency actions to Ponzi schemes, as was the case in *R.F. Lafferty*.⁶⁸ Where the corporation's sole interest is in perpetuating a fraud, it is difficult to see how deepening insolvency does the corporation any harm.

Finally, even if one accepts deepening insolvency as a cause of action there still exists the subsidiary question of a standard of conduct. The Ninth Circuit has recognized liability predicated on a negligence standard.⁶⁹ However, in *In re CitX*, the Third Circuit took a harder line by requiring proof of something more, but left the ultimate standard unclear.⁷⁰

Regardless of these uncertainties, deepening insolvency remains attractive from the trustee's point of view, as theoretically al-

63. Allard, 924 F. Supp. at 494. *But see* Official Comm. of Unsecured Creditors v. Blomen (*In re* Hydrogen, LLC), 431 B.R. 337, 357 (Bankr. S.D.N.Y. 2010) (construing New York law to reject deepening insolvency as a stand-alone cause of action).

64. *See, e.g.*, Bondi v. Citigroup, Inc., No. BER-L-10902-04, 2005 WL 975856, at *16 (N.J. Super. Ct. Law Div. 2005), *appeal granted* by 878 A.2d 850 (N.J. 2005); *In re* Hydrogen, LLC, 431 B.R. at 357 (construing New York law to reject deepening); Limor v. Buerger (*In re* Del-Met Corp.), 322 B.R. 781, 807 (Bankr. M.D. Tenn. 2005) (discussing the lack of a clear definition of deepening insolvency); *In re* VarTec Telecom, Inc., 335 B.R. 631 (Bankr. N.D. Tex. 2005) (construing Texas law); Schnellling v. Crawford (*In re* James River Coal Co.), 360 B.R. 139, 178–180 (Bankr. E.D. Va. 2007) (construing Virginia law).

65. *Trenwick Am. Litig. Trust v. Ernst & Young, LLP*, 906 A.2d 168, 205 (Del. Ch. 2006).

66. *Cf.* William Bates III, *Deepening Insolvency: Into the Void*, 24 AM. BANKR. INST. J. 1, 1 (2005) (criticizing the lack of coherence in the theory of deepening insolvency).

67. *Schacht v. Brown*, 711 F.2d 1343, 1347–48 (7th Cir. 1983), *cert. denied* 464 U.S. 1002 (1983).

68. *See* Official Comm. of Unsecured Creditors v. R.F. Lafferty & Co., 267 F.3d 340, 343–44 (3d Cir. 2001) (applying deepening insolvency to a Ponzi scheme).

69. *Smith v. Arthur Andersen, LLP*, 421 F.3d 989 (9th Cir. 2005).

70. *In re CitX Corp., Inc.*, 448 F.3d 672, 681 (3d Cir. 2006) (requiring more than negligence and some level of fraudulent intent).

most all of the costs incurred by the firm during its prolonged existence can be recovered.⁷¹

The trustee has two other potential causes of action: (1) derivative suits based on breach of the fiduciary duty the third party owes the debtor⁷² and (2) actions against the third party for aiding and abetting the debtor's fraud.⁷³ These two represent the trustee's best hopes for recovery and are the subject of Part Two. Often times they go together. For instance, in *In re CitX*, the trustee brought suit against an individual accountant and his firm for their failure to detect the former CEO's securities fraud and misappropriation of company resources.⁷⁴

However, it is important to distinguish between actions brought by creditors derivatively, which are brought *on behalf* of the corporation, and actions brought by creditors in their independent capacity because state law may bar derivative actions. As the Eighth Circuit noted in *In re Senior Cottages of America*, "If the corporation owned a cause of action against the principal who breached a duty, it follows that it also owns the cause of action for aiding and abetting the principal's breach."⁷⁵ The aiding and abetting breach of fiduciary duty action is brought by a party on behalf of the corporation.⁷⁶ If malfeasance would be imputed to the corporation, it

71. Bates, *supra* note 66 (noting that you can recover for a great deal of payments, basically every cost of the firm).

72. See *Whitney v. Citibank, N.A.*, 782 F.2d 1106, 1115 (2d Cir. 1986) (noting that to prove a claim of aiding and abetting breach of fiduciary duty requires: 1) a breach of fiduciary duty, 2) defendant knowingly inducing or participating in the breach, and 3) damages resulting from the breach).

73. See *Mazzaro de Abreu v. Bank of America Corp.*, 525 F. Supp. 2d 381, 387 (S.D.N.Y. 2007) ("[T]o state a claim for aiding and abetting fraud . . . a plaintiff must allege: (1) the existence of an underlying fraud; (2) actual knowledge of the fraud by the aider and abettor; and (3) substantial assistance by the aider and abettor in the achievement of the underlying fraud.").

74. 448 F.3d 672, 676 (3d Cir. 2006).

75. *Moratzka v. Morris (In re Senior Cottages of Am., LLC)*, 482 F.3d 997, 1002 (8th Cir. 2007).

76. See *Delgado Oil Co., Inc. v. Torres*, 785 F.2d 857, 861 (10th Cir. 1986) (holding that claims common to all creditors belong to the trustee and can only be brought on behalf of the corporation); *Ford Motor Credit Co. v. Minges*, 473 F.2d 918, 920-21 (4th Cir. 1973) (holding that the right of action for directors' negligent mismanagement may only be maintained in the name of corporation or its receiver if it is insolvent); *N. Am. Catholic Educ. Programming Found., Inc. v. Gheewalla*, 930 A.2d 92, 101-02 (Del. 2007) (holding that creditors may maintain a derivative action on behalf of the corporation for breach of fiduciary duty); *Dana Molded Products, Inc. v. Brodner*, 58 B.R. 576, 580 (N.D.Ill. 1986) (explaining that creditor may not bring action against corporation's fiduciaries where the creditor is harmed indirectly).

would therefore also be imputed to the party bringing the suit derivatively, including a trustee.⁷⁷ In these cases, the doctrine of *in pari delicto*, which translates as “in equal fault,” blocks the suit.⁷⁸ Likewise, when the aiding and abetting cause of action is brought derivatively, the same rule can apply.⁷⁹ I explain the courts’ application of *in pari delicto* in these situations in Part Three.

Although *in pari delicto* may bar derivative suits, it cannot bar an aiding and abetting fraud action brought directly by or on behalf of the creditors. This distinction makes aiding and abetting fraud the most viable state law cause of action. The trustee’s suit against HSBC in the Madoff litigation is a textbook example of such a claim.⁸⁰ The aiding behavior varies, but in the securities fraud context, one of the best examples is churning, which involves third-party brokers helping to simulate the appearance of trading activity through economically meaningless trades.⁸¹ The downside to an aiding and abetting claim is that, unlike a breach of fiduciary duty action, it requires meeting the pleading standard for fraud claims.⁸² For example, the court in *In re Alphastars* required that the plaintiff raise a strong inference of fraud by proving both motive and oppor-

77. See *Cenco Inc. v. Seidman & Seidman*, 686 F.2d 449 (7th Cir. 1982) (describing how agency principles impute a manager’s fraud to the corporation); RESTATEMENT (THIRD) OF AGENCY § 5.03 (2006) (“[N]otice of a fact that an agent knows or has reason to know is imputed to the principal if knowledge of the fact is material to the agent’s duties to the principal . . .”). See also Wasserman, *supra* note 13, at 373–74. In bankruptcy, the trustee succeeds to the estate. 11 U.S.C. § 541(a) (2006). Thus fraud that would be imputed to the debtor corporation is also imputed to the trustee. See *E.F. Hutton & Co., Inc. v. Hadley*, 901 F.2d 979, 987 (11th Cir. 1990) (imputing fraud of the corporation to the trustee).

78. See *infra* notes 98–101 & accompanying text.

79. See *infra* note 101 & accompanying text.

80. *Picard v. HSBC Bank PLC*, 454 B.R. 25, 28 (S.D.N.Y. 2011) (“[T]he trustee . . . seeks to recover under various common law theories such as unjust enrichment, aiding and abetting fraud and aiding and abetting breach of fiduciary duty.”).

81. This is known as “churning.” See *Hays & Co. v. Merrill Lynch, Pierce, Fenner & Smith, Inc.*, 885 F.2d 1149, 1150 (3d Cir. 1989) (trustee in bankruptcy for corporation brought churning claim against securities broker); *DeRance, Inc. v. PaineWebber Inc.*, 872 F.2d 1312, 1314–19 (7th Cir. 1989) (not-for-profit corporation brought breach of contract, fraud, and breach of fiduciary duty claims based in part on alleged churning of its account); *M & B Contracting Corp. v. Dale*, 795 F.2d 531, 532–33 (6th Cir. 1986) (corporation brought churning claim against securities broker).

82. See *Bondi v. Bank of Am. (In re Parmalat)*, 383 F. Supp. 2d 587, 593 (S.D.N.Y. 2005) (applying Rule 9(b) pleading standard to trustee’s claim against third party for aiding and abetting debtor’s fraud).

tunity.⁸³ Finally, the *prima facie* case itself is daunting as a plaintiff must prove that a fraud occurred, that the third party had sufficient knowledge of the violation, and that the third party provided substantial assistance.⁸⁴ Despite these challenges, the claim is available as a matter of state law.

As noted at the beginning of this section and as the court in *Cenco* recognized, all of these claims are conceptually similar.⁸⁵ At their core, these suits allege that third-party providers either aided fraud or breached their fiduciary duties to the debtor. Although the facts of each individual suit may be similar, state law raises barriers to many of these claims. Not every state court recognizes deepening insolvency as a stand-alone theory; negligence and breach of fiduciary duty require a relationship (requirements do vary by state and circuit); and the doctrine of *in pari delicto* is likely to bar derivative actions, as explained in Part Three. To succeed, the trustee must rely on a cause of action that runs directly to the creditors. Aiding and abetting liability meets this requirement and should be available to the trustee. Indeed trustees have succeeded in bringing such claims outside of bankruptcy. This raises the question addressed in Part Three: why have bankruptcy courts refused to allow the trustee to bring such claims during Chapter 7 proceedings?

II.

CASE LAW AND FORMALISTIC SHORTCOMINGS

If legislators and judges worked from a blank slate determining whether the trustee should have the power to sue third parties,

83. O'Connell v. Arthur Andersen LLP (*In re* Alphastar Insurance Group Ltd.), 383 B.R. 231, 257 (Bankr. S.D.N.Y. 2008).

84. See *Lerner v. Fleet Bank, N.A.*, 459 F.3d 273, 286 (2d Cir. 2006) (setting out *prima facie* case for aiding and abetting fraud). Proving substantial assistance can also be a challenge as it may require either affirmative assistance or failure to act in the face of a duty. See *In re Sharp Int'l Corp.*, 403 F.3d 43, 52 (2d Cir. 2005) (finding that mere knowledge of fraud is insufficient without duty to disclose); *Cromer Fin. Ltd. v. Berger*, 137 F. Supp. 2d 452, 470 (S.D.N.Y. 2001) (including knowledge of the fraud in aiding and abetting liability when there is a duty to disclose). The plaintiff must also prove knowledge negligence does not suffice for liability. RESTATEMENT (SECOND) OF TORTS § 876(b) (1977) (stating that a person must "know" another's conduct constitutes a breach of duty to be subject to liability for acting in concert). Some courts consider proof of scienter and the level of assistance in tandem, requiring less assistance when scienter is clear. See *In re Enron Corp. Sec., Derivative & "ERISA" Litigation*, 511 F. Supp. 2d 742, 803 (S.D.Tex. 2005) (quoting *Witzman v. Lehrman & Flom*, 601 N.W.2d 179, 188 (Minn. 1999)).

85. See *Cenco Inc. v. Seidman & Seidman*, 686 F.2d 449, 453 (7th Cir. 1982) (Posner, J.).

then finding a solution would be a relatively straightforward task. Considering the availability of an aiding and abetting fraud or deepening insolvency action as explained in Part One, policy-makers would then simply ask whether the principles of bankruptcy discussed in Part Three warrant that power. Indeed, this is what I do in this Part Three. Practitioners, however, do not work from a blank slate. While trustees continue to claim that the current Code already authorizes them to bring these claims, this Part explores their arguments, which focus on Sections 541 and 544 of the Code. I begin with Section 541 and explain how *in pari delicto* and standing analysis act to block the trustee from relying on this Section to bring a claim against third parties. I then consider the trustee's prospects for an aiding and abetting action using Section 544. Although it is a closer case, I conclude that courts have correctly construed the Code and Supreme Court precedent to block these actions as well.

A. *The Failure of Section 541*

On its face, Section 541(a) is simple. If a debtor has an interest in property pre-petition, then that interest becomes property of the estate post-petition.⁸⁶ In short, the trustee stands in the shoes of the debtor. This undoubtedly includes causes of action that the debtor could have brought.⁸⁷ Thus Section 704 entrusts the trustee to “collect and reduce to money the property of the estate for which [the] trustee serves”⁸⁸ This power has been interpreted to implicitly extend to bringing causes of action that belong to the debtor.⁸⁹

86. “The commencement of a case. . . creates an estate. Such estate is comprised of all the following property . . . all legal or equitable interests of the debtor in property as of the commencement of the case.” 11 U.S.C. § 541(a) (2006).

87. See *Sierra Switchboard Co. v. Westinghouse Electric Corp.*, 789 F.2d 705, 709 (9th Cir. 1986) (construing § 541(a) broadly to include all causes of action and discussing legislative history of the Bankruptcy Act supporting this understanding); *Howe v. Richardson (In re Howe)*, 232 B.R. 534, 537 (B.A.P. 1st Cir. 1999) (“The inclusion of choses in action within the sweep of § 541 represented a significant departure from the more limited composition of the bankruptcy estate under the Bankruptcy Act.”). See also 5 COLLIER ON BANKRUPTCY ¶ 541.03, at 541–15 (Alan N. Resnick & Henry J. Somme eds., LexisNexis, 16th ed. 2008).

88. 11 U.S.C. § 704(a) (2006). See also *Koch Refining v. Farmers Union Cent. Exch., Inc.*, 831 F.2d 1339, 1348 (7th Cir. 1987) (“It is axiomatic that the trustee has the right to bring any action in which the debtor has an interest, including actions against the debtor’s officers and directors for breach of duty or misconduct.”) (citing *Pepper v. Litton*, 308 U.S. 295, 307 (1939)).

89. This logically follows from the fact that the trustee, as the debtor’s successor, is the real party in interest. See *Wieburg v. GTE Sw. Inc.*, 272 F.3d 302, 306 (5th Cir. 2001) (“Because the claims are property of the bankruptcy estate, the Trustee

Specifically, courts recognize that the trustee's authority to bring suit on behalf of the debtor extends to breaches of fiduciary duty by insiders and third parties:

For example, these sections give the trustee authority to bring an action for damages on behalf of a debtor corporation against corporate principals for alleged misconduct, mismanagement, or breach of fiduciary duty, because these claims could have been asserted by the debtor corporation, or by its stockholders in a derivative action.⁹⁰

This makes sense in light of underlying state law, which, as discussed in Part One, would as an initial matter allow a debtor to bring these actions. In fact, this is how it works when the trustee attempts to bring an action against corporate principals for breach of fiduciary duty.⁹¹ However, it is important to reiterate the distinction between claims that belong to the creditors and claims that belong to the debtor. The trustee stands in the shoes of the debtor under Section 541. Therefore, the claim must belong to the debtor and not the creditors individually. This is a question of state law, and certain claims—like veil-piercing or fraudulent conveyance—may belong to the creditors, meaning they could not be the subject of a Section 541 suit by the trustee.⁹² However, at a minimum, a claim for breach of fiduciary duty that a third party owed the corporation should be available to the debtor.⁹³ But *in pari delicto* and standing analysis have led courts to reject Section 541 actions in these situations.

is the real party in interest with exclusive standing to assert them.”) (citations omitted); *Feist v. Consol. Freightways Corp.*, 100 F.Supp.2d 273, 274–75 (E.D.Pa. 1999), *aff'd*, 216 F.3d 1075 (3d Cir. 2000), *cert. denied*, 532 U.S. 920 (2001).

90. *Mixon v. Anderson (In re Ozark Rest. Equip. Co., Inc.)*, 816 F.2d 1222, 1225 (8th Cir. 1987) (citing *Pepper v. Litton*, 308 U.S. 295, 307 (1939)).

91. *See, e.g., Mediators, Inc. v. Manney (In re Mediators, Inc.)*, 105 F.3d 822, 826–27 (2d Cir. 1997) (“We agree that a bankruptcy trustee, suing on behalf of the debtor under New York law, may pursue an action for breach of fiduciary duty against the debtor’s fiduciaries.”).

92. *See In re Ozark*, 816 F.2d at 1225 (“Where . . . ‘the applicable state law makes such obligations or liabilities run to the corporate creditors personally, rather than to the corporation, such rights of action are not assets of the estate under Section 541(a) that are enforceable by the trustee [under Section 704(1)].”). However, there are states where this is not the case and the corporation can essentially pierce its own veil. *See, e.g., Koch Ref.*, 831 F.2d at 1346 (finding that both Indiana and Illinois state law allow trustee to pierce the veil under § 541).

93. *See In re Mediators*, 105 F.3d at 826–27.

Courts have done this for several reasons. In the first instance, the court may decide to impute the debtor's fraud to the trustee.⁹⁴ In general, a corporation is presumed to have all of the knowledge of its employees who act as its agents.⁹⁵ The corollary to this is that the employees' actions are also considered those of the corporation.⁹⁶ As such, when the agent commits fraud, the corporation does as well.⁹⁷ This bit of agency law is critical to a Section 541 claim because it means that the trustee who stands in the shoes of the debtor and sues a third party for aiding and abetting fraud is technically suing his own accomplice. The doctrine of *in pari delicto*, which means "equally at fault,"⁹⁸ bars such suits as a matter of public policy, preventing one wrongdoer from collecting from another.⁹⁹ As Judge Posner wrote in *Cenco Inc. v. Seidman & Seidman*, "[a] Participant in a fraud cannot also be a victim entitled to recover damages, for he can't have relied on the truth of the fraudulent representations."¹⁰⁰ The upshot is that the trustee cannot use Section 541 to sue third-party service providers for breach of fiduci-

94. See *E.F. Hutton & Co., Inc. v. Hadley*, 901 F.2d 979 (11th Cir. 1990) (imputing fraud of the corporation to the trustee); *Cenco Inc. v. Seidman & Seidman*, 686 F.2d 449 (7th Cir. 1982) (describing how agency principles impute a manager's fraud to the corporation and then to the trustee); *In re Mediators*, 105 F.3d 822 (same).

95. See *In re Sharp Int'l Corp.*, 278 B.R. 28, 36 (Bankr. E.D.N.Y. May 20, 2002) ("The general rule is that knowledge acquired by an agent acting within the scope of his agency is imputed to his principal and the latter is bound by such knowledge although the information is never actually communicated to it . . . Underlying the rule is the presumption that an agent has discharged his duty to disclose to his principal 'all the material facts coming to his knowledge with reference to the subject of his agency.'" (quoting *Center v. Hampton Affiliates, Inc.*, 66 N.Y.2d 782, 784 (1985)). See also RESTATEMENT (THIRD) OF AGENCY § 5.03 (2006) ("[N]otice of a fact that an agent knows or has reason to know is imputed to the principal if knowledge of the fact is material to the agent's duties to the principal.").

96. See RESTATEMENT (THIRD) OF AGENCY § 5.03 cmt. c. (2006) ("Imputation recognizes that an organization constitutes one legal person and that its link to the external world is through its agents . . .").

97. This is corporate criminal liability. See *N.Y. Cent. & Hudson River R.R. Co. v. U.S.*, 212 U.S. 481, 494–95 (1909) ("[W]e see no good reason why corporations may not be held responsible for and charged with the knowledge and purposes of their agents, acting within the authority conferred upon them.").

98. BLACK'S LAW DICTIONARY 806 (8th ed. 2004).

99. See *Bateman Eichler, Hill Richards, Inc. v. Berner*, 472 U.S. 299, 306 (1985) ("The defense is grounded on two premises: first, that courts should not lend their good offices to mediating disputes among wrongdoers; and second, that denying judicial relief to an admitted wrongdoer is an effective means of deterring illegality.") (citations omitted).

100. *Cenco Inc. v. Seidman & Seidman*, 686 F.2d 449 (7th Cir. 1982).

ary duty or aiding and abetting the fraud. This result has largely been ratified by the circuits.¹⁰¹

The textbook example of such an instance is *R.F. Lafferty*.¹⁰² In *R.F. Lafferty*, the Third Circuit was confronted with a Ponzi scheme operating through fraudulent debt certificates.¹⁰³ When the scheme collapsed, the corporations entered bankruptcy and the trustee sought to bring deepening insolvency actions against third parties that helped issue the securities.¹⁰⁴ The Third Circuit held that because Section 541 places the trustee in the debtor's shoes at the moment the petition is filed, the fraud was necessarily imputed to the trustee and any defense that could have been raised against the debtor was available against the trustee regardless of whether the trustee was an innocent successor.¹⁰⁵ Thus the third party could assert *in pari delicto* against the trustee. While the court recognized the seemingly inequitable result, it also joined the vast majority of circuits that recognize that the text of Section 541 requires evaluation *at the moment of the petition*.¹⁰⁶ The Pennsylvania Supreme Court

101. See, e.g., Official Comm. of Unsecured Creditors of PSA, Inc. v. Edwards, 437 F.3d 1145, 1151 (11th Cir. 2006); Grassmuck v. Am. Shorthorn Ass'n, 402 F.3d 833, 836 (8th Cir. 2005); Official Comm. of Unsecured Creditors v. R.F. Lafferty & Co., 267 F.3d 340, 356–57 (3d Cir. 2001); *In re Dublin Sec., Inc.*, 133 F.3d 377, 381 (6th Cir. 1997); *In re Hedged-Invs. Assocs. Inc.*, 84 F.3d 1281, 1285 (10th Cir. 1996); *Hirsch v. Arthur Andersen & Co.*, 72 F.3d 1085, 1094 (2d Cir. 1995).

102. 267 F.3d 340 (3d Cir. 2001).

103. *Id.* at 344.

104. *Id.*

105. See *id.* at 356–57.

106. See *In re Hedged Investments*, 84 F.3d at 1285

We emphasize § 541(a)(1) limits estate property to the debtor's interests 'as of the commencement of the case.' This phrase places both temporal and qualitative limitations on the reach of the bankruptcy estate. In a temporal sense, it establishes a clear-cut date after which property acquired by the debtor will normally not become property of the bankruptcy estate.

See also *Shearson Lehman Hutton, Inc. v. Wagoner*, 944 F.2d 114, 118 (2d Cir. 1991) (“[W]e must determine what claims HMK possessed against Shearson *before* HMK went bankrupt.”); *R.F. Lafferty*, 267 F.3d at 356 (“[T]he explicit language of section 541 directs courts to evaluate defenses as they existed at the commencement of the bankruptcy.”); *Caplin v. Marine Midland Grace Trust Co.*, 406 U.S. 416, 429 n.19 (1972); 5 COLLIER ON BANKRUPTCY, ¶ 541.04, at 541–12 (Alan N. Resnick & Henry J. Somme eds., LexisNexis, 16th ed. 2008) at 541–42 (15th ed. 1989) (“Under section 541(a) . . . the bankruptcy estate consists of all of the debtor's legal and equitable property interests that existed as of the commencement of the case, that is, as of the time that the bankruptcy petition. . . is filed.”); This result has been muddied by the decisions that consider the power of a *receiver* to bring a very similar action. See *Scholes v. Lehman*, 56 F.3d 750, 754 (7th Cir. 1995) (“That reason falls out now that Douglas has been ousted from control of and beneficial interest in the corporations. The appointment of the receiver re-

later narrowed *R.F. Lafferty*, but nonetheless recognized that *in pari delicto* may apply.¹⁰⁷ Thus, while it seems unfair to ignore the differences between the trustee and the debtor, the Code's "at the commencement of the case"¹⁰⁸ language leaves no other choice.

There is, of course, an exception to this rule and an exception to that exception. When the fraud operates in a manner that is adverse to the corporation's interests, the so-called "adverse interest" exception may apply and block the application of *in pari delicto*.¹⁰⁹ Like *in pari delicto*, adverse interest is a logical outgrowth of agency law.¹¹⁰ When the agent is not acting for the principal's benefit, the agent's actions are not imputed to the corporation. Likewise, when a third party aids an agent in defrauding a corporation, the corporation is not "in equal fault." A district court reached this result in

moved the wrongdoer from the scene. The corporations were no more Douglas's evil zombies."). In that context, without the bar imposed by Section 541, the courts may distinguish the receiver from the guilty debtor and refuse to apply *in pari delicto*. See *FDIC v. O'Melveny & Myers*, 61 F.3d 17, 19 (9th Cir. 1995) ("[D]efenses based on a party's unclean hands or inequitable conduct do not generally apply against that party's receiver.") (citations omitted).

107. See Official Comm. of Unsecured Creditors of Allegheny Health Educ. & Research Found. v. PriceWaterhouseCoopers, LLP, 989 A.2d 313, 332 (Pa. 2010) ("We agree with the Seventh Circuit to the extent it has held that *in pari delicto* may be available as a defense in some cases arising in the corporate auditing context, across the broader part of the spectrum of the various common-law causes of action which may be asserted.").

108. 11 U.S.C. § 541(a)(1) (2006) ("[The following comprises the estate:] [e]xcept as provided in subsections (b) and (c)(2) of this section, all legal or equitable interests of the debtor in property as of the commencement of the case.").

109. See *In re Sharp Int'l Corp.*, 278 B.R. 28, 36 (E.D.N.Y. May 20, 2002) ("Even if the *Wagoner* rule applies, a trustee may have standing to assert a claim against the corporation's third party professionals under the adverse interest exception to that rule. The adverse interest exception applies where 'the officer acted entirely in his own interests and adversely to the interests of the corporation.' 'The theory is that 'where an agent, though ostensibly acting in the business of the principal, is really committing a fraud for his own benefit, he is acting outside of the scope of his agency, and it would be most unjust to charge the principal with knowledge of it.'") (quoting *Wight v. BankAmerica Corp.*, 219 F.3d 79, 87 (2d Cir. 2000)) (citations omitted). See also RESTATEMENT (SECOND) OF AGENCY § 282(1) (1958) ("A principal is not affected by the knowledge of an agent in a transaction in which the agent secretly is acting adversely to the principal and entirely for his own or another's purposes . . .").

110. See *supra* notes 95–97 & accompanying text. Acts that are intended to defraud the corporation or are otherwise adverse to the corporation's interest cannot be said to fall within the scope of employment. See *In re Mediators, Inc.*, 105 F.3d 822, 827 (2d Cir. 1997) ("Under New York law, the adverse interest exception rebuts the usual presumption that the acts and knowledge of an agent acting within the scope of employment are imputed to the principal.").

In re Sharp.¹¹¹ However, the exception is narrow and the adversity of interests must be complete.¹¹²

It would seem that the adverse interest exception would solve the trustee's problem in most Ponzi bankruptcy cases. After all, a Ponzi scheme clearly benefits the individual fraudster and not the corporation, which makes no real profit off the scheme.¹¹³ However, as explained in Part One, Ponzi schemes are typically operated by individual actors. Unlike corporate frauds like Enron, which operate on a massive scale and involve many agents and many parts of the corporation, a Ponzi scheme like Madoff's can inflict billions of dollars in damages with just a few employees. This distinction has given rise to an exception to the adverse interest exception called the "sole actor exception."¹¹⁴ To understand how this works, take for example an instance where the veil could be pierced under an instrumentality rationale.¹¹⁵ In such a situation there can be no adverse interest. Everything done by the agent is done by the corporation and everything that benefits the agent benefits the corporation. As explained above, given the nature of the scheme,

111. 278 B.R. at 41.

112. *In re Mediators*, 105 F.3d at 827.

113. *See supra* notes 1, 29–31, 35 & accompanying text.

114. *See In re Mediators*, 105 F.3d at 826–27

[T]he adverse interest exception does not apply to cases in which the principal is a corporation and the agent is its sole shareholder . . . where the principal and agent are one and the same, the adverse interest exception is itself subject to an exception styled the 'sole actor' rule . . . Where, as here, a sole shareholder is alleged to have stripped the corporation of assets, the adverse interest exception to the presumption of knowledge cannot apply.

See also *Grasmueck v. Am. Shorthorn Ass'n*, 402 F.3d 833, 838 (8th Cir. 2005) ("The sole actor doctrine is an established principle of agency law.").

115. The instrumentality test is one of several alternative theories used to pierce the veil. *See Dzikowski v. Friedlander*, 411 B.R. 434, 446 (Bankr. S.D. Fla. 2009) (quoting *Zaist v. Olson*, 227 A.2d 552, 576 (Conn. 1967) ("The instrumentality rule requires proof of three elements: (1) Control, not merely majority or complete stock control, but complete domination, not only of finances but of policy and business practice *in respect to the transaction attacked* so that the corporate entity as to this transaction had at the time no separate mind, will or existence of its own; (2) that such control must have been used by the [principal] to commit fraud or wrong, to perpetrate the violation of a statutory or other positive legal duty, or a dishonest or unjust act in contravention of [defendant's] legal rights; *and* (3) that the aforesaid control and breach of duty must proximately cause the injury or unjust loss complained of."). *See also* 1 WILLIAM D. FLETCHER, FLETCHER CYCLOPEDIA ON THE LAW OF PRIVATE CORPORATIONS § 41 (1983) ("[Alter ego] fastens liability on the individual who uses a corporation merely as an instrumentality to conduct his own personal business, and such liability arises from fraud or injustice perpetrated not on the corporation but on third persons dealing with the corporation.").

the sole actor exception often acts to reinstate *in pari delicto* and block trustees from using Section 541 to sue third parties.¹¹⁶

The Second Circuit has complicated things even further by sometimes applying *in pari delicto* as a part of standing analysis. In *Shearson Lehman Hutton, Inc. v. Wagoner*, the trustee sued a broker for aiding and abetting breach of fiduciary duty and fraud by churning accounts and disregarding standard verification and compliance rules.¹¹⁷ The court wrote that, unlike other circuits:

In our analysis of the question presented, the “case or controversy” requirement coincides with the scope of the powers the Bankruptcy Code gives a trustee, that is, if a trustee has no power to assert a claim because it is not one belonging to the bankrupt estate, then he also fails to meet the prudential limitation that the legal rights asserted must be his own.¹¹⁸

In *Shearson*, the estate did not have a claim because *in pari delicto* applied. Most of the other circuits that have confronted this line of reasoning have rejected the Second Circuit’s rationale, often explicitly.¹¹⁹ However, although courts in the Second Circuit have taken a few tenuous steps towards distinguishing standing and *in pari delicto*, *Shearson* remains the law of the circuit.¹²⁰ The distinction also remains important because under *Shearson* the trustee truly cannot bring the claim at all, while under *R.F. Lafferty*, the obstacle is an affirmative defense that the defendant must assert.¹²¹

116. See *supra* notes 109–115 & accompanying text.

117. 944 F.2d 114, 119 (2d Cir. 1991).

118. *Id.* at 118.

119. See, e.g., *In re Senior Cottages of Am., LLC*, 482 F.3d 997, 1003–04 (8th Cir. 2007); Official Comm. of Unsecured Creditors of PSA, Inc. v. Edwards, 437 F.3d 1145, 1149–50 (11th Cir. 2006); *Baena v. KPMG, LLP*, 453 F.3d 1, 6–10 (1st Cir. 2006); Official Comm. of Unsecured Creditors v. R.F. Lafferty & Co., 267 F.3d 340, 346 (3d Cir. 2001); *In re Dublin Sec., Inc.*, 133 F.3d 377, 379–80 (6th Cir. 1997).

120. See *In re Parmalat*, 383 F. Supp. 2d 587, 595–99 (S.D.N.Y. 2005) (narrowing *Wagoner* to applications of New York law and applying *in pari delicto* separately). See also *Senior Cottages*, 482 F.3d at 1003 (8th Cir. 2007) (discussing recent trends away from the *Wagoner* rule in the Second Circuit).

121. Compare *Wagoner*, 944 F.2d at 117 (“Because standing is jurisdictional under Article III of the United States Constitution . . . it is a threshold issue in all cases since putative plaintiffs lacking standing are not entitled to have their claims litigated in federal court.”) (citations omitted), with *R.F. Lafferty*, 267 F.3d 340, 361 (3d Cir. 2001) (Cowen, J., dissenting) (“[A]s a general matter, the ultimate merits of an affirmative defense do not raise questions about a plaintiff’s standing, or else the moment the court was poised to rule in favor of the defendant on the affirmative defense, the court would lose jurisdiction . . .”).

B. Section 544(a)(1) Actions

When *in pari delicto* bars a trustee from suing from the shoes of a debtor, the trustee may attempt to stand in the shoes of a creditor using Section 544(a)(1) of the Code. Under Section 544(a)(1), the trustee is empowered to bring all of the actions a hypothetical lien creditor could bring:

As the Tenth Circuit has explained, “Congress has fashioned a legal fiction [that] permits the trustee . . . to assume the guise of a creditor with a judgment against the debtor. Under that guise, the trustee may invoke whatever remedies provided by state law to judgment lien creditors to satisfy judgments against the debtor.”¹²²

This “strong arm power” is another potential avenue through which trustees have tried to assert claims for aiding and abetting fraud or breach of fiduciary duty.¹²³ In addition, Section 544(b) allows the trustee to stand in the shoes of a particular unsecured creditor.¹²⁴ A claim against a third party is naturally at home within Section 544, which typically is used to avoid transfers to third parties that delay, hinder, or defraud creditors.¹²⁵ It is also a more natural fit because an action brought on behalf of the creditors does not rely on the relationship between the third party and the corporation. Thus the context of the third party’s actions is less important. *In pari delicto* cannot bar the trustee when he stands in the shoes of the victims. Unfortunately, under the Supreme Court’s de-

122. *Picard v. JPMorgan Chase & Co.*, 460 B.R. at 93 (S.D.N.Y. 2011) (quoting *Zilkha Energy Co. v. Leighton*, 920 F.2d 1520, 1523 (10th Cir. 1990)).

123. *See, e.g., Mixon v. Anderson (In re Ozark Rest. Equip. Co.)*, 816 F.2d 1222, 1225 (8th Cir. 1987) (rejecting Trustee’s attempt to use strong arm power to bring breach of fiduciary duty and alter ego causes of action), *cert. denied*, 484 U.S. 848 (1987); *Picard*, 460 B.R. at 93 (rejecting Trustee’s § 544 claim against broker dealers); *Hill v. Gibson Dunn & Crutcher, LLP (In re MS55, Inc.)*, No. 06-CV-01233, 2007 WL 2669150 (D.Colo. Sept. 6, 2007) (allowing Trustee’s § 544 claim for aiding and abetting breach).

124. “Except as provided in paragraph (2), the trustee may avoid any transfer of an interest of the debtor in property or any obligation incurred by the debtor that is voidable under applicable law by a creditor holding an unsecured claim that is allowable under section 502 of this title or that is not allowable only under section 502(e) of this title.” 11 U.S.C. § 544(b)(1) (2006).

125. *See In re Ozark*, 816 F.2d at 1229 (“In this vein, we note that Sections 544(a) and (b) are flavored with the notion of the trustee having the power to avoid ‘transfers’ of the debtor, as were its predecessors, sections 70c and e of the Act.”); *Picard*, 460 B.R. at 93 (“The purpose of this section is to allow the trustee in bankruptcy to cut off any secret or unperfected liens on debtor property that would bind the debtor itself, but not the debtor’s judgment creditor, who typically enjoys top priority under state creditor/debtor laws.”).

cision in *Caplin v. Marine Midland Grace Trust Co.*¹²⁶ and the lower courts' subsequent interpretations of that case, the Section 544 avenue has been closed off as well.

In *Caplin* the Supreme Court addressed the equivalent of Section 544 under the old Bankruptcy Act ("Act").¹²⁷ The trustee, standing in the shoes of bondholders, sued the indenture trustee of a bankrupt real estate developer.¹²⁸ The trustee alleged that the indenture trustee willfully or negligently signed off on "grossly overvalued appraisals"¹²⁹ and failed to ensure compliance with the terms of the indenture and allowed the debtor to engage in transactions that resulted in "great financial losses."¹³⁰ As the Court noted, a Section 541 claim would not work as the indenture trustee and the debtor were likely *in pari delicto*, meaning the trustee had no choice but to make his claim from the shoes of the creditors via Section 544.¹³¹ Relying on Second Circuit precedent interpreting Section 567 of the Act, the trustee argued that Congress intended to remedy collective action problems and reduce litigation by empowering the trustee to bring suits on behalf of the creditors.¹³²

In a 5–4 decision, the Court rejected the trustee's arguments.¹³³ The Court raised three objections. First, Section 567 gave the trustee responsibility to collect money and investigate fraud, but it did not explicitly authorize the trustee to bring suit. The Court construed this as evidence of congressional intent to give the trustee a limited role.¹³⁴ Second, the Court believed subrogation would be required in the event of any recovery.¹³⁵ If this were the case, the insurer would have a claim for contribution from the estate for at least a portion of the amount paid out to the bondholders. Thus there would be little point to bringing the actions because any amount received from the insurer would lead to an equal increase in the claims against the estate, leaving the creditors exactly where they started. Finally, the Court was concerned with potentially divergent interests among the creditors.¹³⁶ The Court believed disagree-

126. 406 U.S. 416 (1972).

127. *Id.* at 423–28 (discussing sections 567 and 587 of Chapter X of the Bankruptcy Act).

128. *Id.* at 419–20.

129. *Id.*

130. *Id.*

131. *Id.* at 430.

132. *Caplin v. Marine Midland Grace Trust Co.*, 406 U.S. 416, 427 (1972).

133. *Id.* at 428–32.

134. *Id.* at 428.

135. *Id.* at 430.

136. *Id.* at 432.

ment over which claims to pursue and how much to demand in damages made individual actions likely, regardless of whether the trustee brought the claim, in the absence of any procedure to bind the creditors to the outcome of the trustee's suit.¹³⁷ In short, while the discharge makes bankruptcy a mandatory collective procedure, the Court was unsure the same would be the case when the trustee sues third parties. The Court has not revisited trustee claims against third parties on behalf of the creditors since *Caplin*. This complicates matters in view of the overhaul of the Code in 1978, six years after *Caplin*. However, the circuit courts have heard similar cases since and addressed some of the issues arising from *Caplin*.

The Eighth Circuit was one of the first to do so. In *In re Ozark Restaurant Equipment Co.*, a three-judge panel reviewed a district court decision holding that, despite the broader powers granted to the trustee by Section 544, *Caplin* still did not permit her to bring an action on behalf of the creditors against a third party.¹³⁸ In *Ozark*, the trustee sought to pierce the veil, a cause of action that belongs to the creditors individually under the applicable state law.¹³⁹ This meant that Section 541 was unavailable.¹⁴⁰ In applying *Caplin* to the new Section 544, the Eighth Circuit considered in turn whether any of the three *Caplin* rationales applied. It found that although the trustee's powers under the new Section 544 were broad, *Caplin*'s concern that an action by the trustee could not bind the creditors applied. This meant that the concern that empowering the trustee would lead to duplicative litigation remained:

If the trustee in the instant case was allowed to pursue and recover on the alter ego cause of action on behalf of Ozark's creditors, there obviously would be questions as to which creditors were bound by the settlement. This is because the trustee is not the real party in interest, and thus does not have the power to bind the creditors to any judgment reached in the litigation.¹⁴¹

Ozark is also notable because of its treatment of the Code's legislative history. In *Caplin*, the Court invited Congress to consider

137. *Id.*

138. *Mixon v. Anderson (In re Ozark Rest. Equip. Co., Inc.)*, 816 F.2d 1222, 1229 (8th Cir. 1987).

139. *Id.* at 1225 ("Thus, the obligations and liabilities of an action to pierce the corporate veil in Arkansas do not run to the corporation, but to third parties, e.g., creditors of the corporation.").

140. *Id.*

141. *Id.* at 1230 (citations omitted).

expanding the reach of Section 567.¹⁴² As the court in *Ozark* noted, Congress in fact proposed a Section 544(c),¹⁴³ which would have given the trustee precisely this power.¹⁴⁴ The *Ozark* court found the eventual deletion of this provision “extremely noteworthy”¹⁴⁵ and treated it as positive congressional approval of the *Caplin* Court’s construction of the trustee’s role. The court therefore continued to narrowly construe the trustee’s power under Section 544 and the Section 704 power to collect monies. In doing so, it reaffirmed the three rationales listed in *Caplin* and added its own: Congress’ rejection of Section 544(c).

Like the Eighth Circuit, the other circuits to apply *Caplin* under the Code have continued to inquire whether any of the three rationales—subrogation, preemption, and construction of the trustee’s role—require the court to dismiss the trustee’s claim. Like the *Ozark* court, they have also placed great emphasis on Congress’s exclusion of the proposed Section 544(c).¹⁴⁶ Because these courts

142. *Caplin v. Marine Midland Grace Trust Co.*, 406 U.S. 416, 434–35 (1972) (“Congress could determine that the trustee in a reorganization was so well situated for bringing suits against indenture trustees that he should be permitted to do so. In this event, Congress might also determine that the trustee’s action was exclusive, or that it should be brought as a class action on behalf of all debenture holders, or perhaps even that the debenture holders should have the option of suing on their own or having the trustee sue on their behalf Whatever the decision, it is one that only Congress can make.”).

143. The proposed § 544(c) would have read:

(c)(1) The trustee may enforce any cause of action that a creditor, a class of creditors, an equity security holder, or a class of equity security holders has against any person, if-

(2) If the trustee brings an action on such cause of action-

(3) A judgment in any such action brought by the trustee binds all creditors or equity security holders that could have brought an action on such cause of action. Any recovery by the trustee, less any expense incurred by the trustee in effecting such recovery, shall be for the benefit only of such creditors or equity security holders.

In re Ozark, 816 F.2d at 1228 n.9. The legislative history gives no reason for the exclusion of the proposed section. *See id.* at 1228, n.10.

144. *Mixon v. Anderson (In re Ozark Rest. Equip. Co.)* 816 F.2d 1222, 1227–28 (8th Cir. 1987).

145. *Id.* at 1228.

146. *See, e.g., Grede v. Bank of New York Mellon*, 598 F.3d 899, 901–02 (7th Cir. 2010) (“*Caplin* gave three reasons for its conclusion that a bankruptcy trustee may not pursue third-party claims None of these reasons applies to suit by a liquidation trustee on assigned claims.”); *E.F. Hutton & Co. Inc., v. Hadley*, 901 F.2d 979, 985–86 (11th Cir. 1990) (“On the facts of this case, however, we approve the reasoning of . . . the Eighth Circuit In rewriting the bankruptcy laws in 1978, Congress considered and rejected a provision that expressly would have overruled *Caplin*.”); *St. Paul Fire & Marine Ins. Co. v. Pepsi Co., Inc.*, 884 F.2d 688,

have not applied *Caplin* mechanically and because the underlying state cause of action continues to vary, the results have been mixed. However, in cases analogous to those in a Ponzi scheme where the trustee brings *tort* claims, *Caplin* and *Ozark* have usually been followed.¹⁴⁷

Confusion occurs because the line between a debtor's claim and a creditor's claim is not always so clear. In particular, when trustees have brought alter-ego actions, a subspecies of veil-piercing actions generally, the circuits have been split. In *Koch Refining v. Farmers Union Central Exchange, Inc.*, the trustee brought suit on behalf of trade creditors who had sold petroleum to the debtor arguing that he could bring an action to pierce the debtor's veil under Section 544 and Illinois state law.¹⁴⁸ The Court found that Illinois allowed either the creditors or the debtor corporation to bring a veil-piercing action.¹⁴⁹ According to the court, an action that belongs to the corporation but can be brought by the creditors derivatively is a *general* claim as opposed to a *personal* claim, which properly belongs to each creditor individually.¹⁵⁰ The *Koch* court's distinction between general and personal claims, between those that accrue to the debtor's estate and also to the creditors *on behalf* of the corporation and those that arise solely out of direct injury to the individual creditors, explains why this line of Section 544 cases

700–701 (2d Cir. 1989) (assessing the applicability of each of *Caplin's* rationales to the trustee's claim against third parties); *Williams v. California 1st Bank*, 859 F.2d 664, 666 (9th Cir. 1988) (“The Bank contends that *Caplin* and its progeny control this case. We agree . . . Evaluating the Trustee's claim in light of the three concerns that informed the Court's holding in *Caplin* reveals that substantially the same problems exist.”).

147. See *Picard v. JPMorgan Chase & Co.*, 460 B.R. 84, 95 (S.D.N.Y. 2011) (“For these reasons, the Trustee's theory has been rejected in numerous persuasive cases, including cases from at least three Circuit Courts of Appeals.”). But see *Hill v. Gibson Dunn & Crutcher, LLP (In re MS55, Inc.)*, No. 06-CV-01233, 2007 WL 2669150 (D.Colo. Sept. 6, 2007) (disagreeing with *Ozark's* application of *Caplin* to § 544).

148. *Koch Ref., v. Farmers Union Cent. Exch., Inc.*, 831 F.2d 1339, 1340 (7th Cir. 1987).

149. *Id.* at 1345–46.

150. *Id.* at 1348–49.

However, the trustee has no standing to bring *personal* claims of creditors. A cause of action is “personal” if the claimant himself is harmed and no other claimant or creditor has an interest in the cause. But allegations that could be asserted by any creditor could be brought by the trustee as a representative of all creditors. If the liability is to all creditors of the corporation without regard to the personal dealings between such officers and such creditors, it is a general claim.

Id.

is of little help to the trustee in a Ponzi scheme.¹⁵¹ As explained below, it is also broadly consistent with *Caplin's* preemption rationale.

Distinguishing between personal and general claims can be difficult, as there seems to be a rather arbitrary line separating the two. However, the circuits that have confronted the distinction have embraced it.¹⁵² The *Koch* court itself did not provide much help: "to determine whether an action accrues individually to a claimant or generally to the corporation, a court must look to the injury for which relief is sought and consider whether it is peculiar and personal . . . or general and common."¹⁵³ In *Delgado*, the Tenth Circuit asked whether the injury was common to all creditors and treated all of them equally.¹⁵⁴ This broad definition would seem to include some Ponzi scheme claims. Likewise, in *St. Paul Fire and Marine Insurance Co. v. Pepsico, Inc.*, the Second Circuit accepted the distinction but added to the confusion by referencing the *Caplin* rationale: When the claim is general and the injury indirect the automatic stay operates to bar anyone but the trustee from asserting it.¹⁵⁵ The corollary is that all the creditors are bound by the result, making *Caplin's* concern for preemption irrelevant.¹⁵⁶

The general claims rationale would seem to apply to torts that do equal harm to all the creditors. However, most courts continue to exclude such actions when they arise out of Ponzi schemes. In *E.F. Hutton & Co., Inc. v. Hadley*, the Eleventh Circuit drew this distinction based on whether *in pari delicto* would apply to the debtor.¹⁵⁷ In *Hirsch v. Arthur Andersen & Co.*,¹⁵⁸ the Second Circuit reached the same conclusion but relied instead on the distinction between fraud committed against the corporation itself, where the

151. See Picard, 460 B.R. at 96–97 (determining that the trustee was asserting claims *on behalf* of the creditors, while general claims are asserted *on behalf* of the corporation).

152. See, e.g., *E.F. Hutton & Co., Inc. v. Hadley*, 901 F.2d 979, 986–87 (11th Cir. 1990) (finding trustee was bringing a specific claim and discussing *Koch Refining*); *St. Paul Fire & Marine Ins. Co. v. Pepsi Co., Inc.*, 884 F.2d 688, 700–01 (2d Cir. 1989) (distinguishing between specific and general claims); *Delgado Oil Co., Inc. v. Torres*, 785 F.2d 857, 861 (10th Cir. 1986) (finding the automatic stay applies to general claims); *Cissell v. Am. Home Assurance Co.*, 521 F.2d 790, 793 (6th Cir. 1975) (finding the trustee may only pursue general claims); *Picard*, 460 B.R. 84 (collecting cases).

153. *Koch Ref.*, 831 F.2d at 1349.

154. *Delgado Oil*, 785 F.2d at 861.

155. *St. Paul*, 884 F.2d at 700–01.

156. *Id.*

157. 901 F.2d 979, 987 (11th Cir. 1990).

158. 72 F.3d 1085 (2d Cir. 1995).

debtor and the creditors on its behalf might have an action, and fraud committed against the individuals, like in a Ponzi scheme.¹⁵⁹ While a trustee can bring a claim for the former, the latter is barred by *Caplin*.¹⁶⁰ Most recently, the Madoff trustee, Irving Picard, made the same argument. In *Picard v. JPMorgan Chase & Co.*,¹⁶¹ a bankruptcy court rejected his claims and elaborated on the differences between a tort claim arising out of a Ponzi scheme and a truly general claim.¹⁶² According to the court in *JPMorgan Chase & Co.*, a general claim requires not only that all the creditors benefit but also that they all benefit *equally*.¹⁶³

A few enterprising trustees have sought to circumvent the limits of Section 544(a) by relying on assignment. This method has had some success but its requirements make it somewhat impracticable. In *Grede v. Bank of New York Mellon*, the Seventh Circuit recognized that if the creditors assigned their claims to a trustee, the trustee would have standing to bring those claims after a Chapter 7 liquidation was confirmed.¹⁶⁴ When the claims are assigned, *Caplin's* preemption concern is not relevant provided that the court concludes the trustee is the real party in interest and that there is no potential for subrogation after confirmation. Nonetheless, for this to succeed, the assignments would have to be unconditional,¹⁶⁵ making recovery purely *pro rata*, and the action would likely need to wait until after the bankruptcy plan is confirmed.¹⁶⁶ Yet even then, in *Williams v. California 1st Bank*, the Ninth Circuit refused to allow the

159. *Id.* at 1093–95.

160. *Id.*

161. 460 B.R. 84 (S.D.N.Y. 2011).

162. See *id.* at 96 (“A trustee may maintain only a general claim. His right to bring a claim ‘depends on whether the action vests in the trustee as an assignee for the benefit of creditors or, on the other hand, accrues to specific creditors.’ Requiring an equal right and interest in all creditors is the only justification for vesting exclusive standing in the trustee. Where, as here, the right to relief and the benefits of relief are peculiar to individual or groups of creditors, the right is not a generalized one that belongs to the debtor’s estate.”) (citation omitted).

163. *Id.*

164. 598 F.3d 899, 901 (7th Cir. 2010), *cert. denied*, 131 S. Ct. 418 (2010).

165. See *In re Bogdan*, 414 F.3d 507, 511–13 (4th Cir. 2005) (distinguishing *Williams* based on unconditional assignment and finding that *Caplin's* concerns do not apply to unconditional assignments); *Mukamal v. Bakes*, 383 B.R. 798, 811–14 (S.D. Fl. 2007) (finding partial assignment was insufficient to ameliorate *Caplin*, and leaving unresolved whether *Bogdan's* absolute assignment avoidance of *Caplin* would be followed). Cf. *In re Bennett Funding Group, Inc.*, 336 F.3d 94, 99–100 (2d Cir. 2003) (finding assignment did not give the trustee standing).

166. See *Grede v. Bank of New York Mellon*, 598 F.3d 899, 902 (7th Cir. 2010) (explaining that because the Bankruptcy case was over the Code’s limits on trustee standing were no longer applicable).

trustee to proceed, finding that the creditors were still the “real parties in interest.”¹⁶⁷ In addition to these hurdles, the practical challenges of convincing every creditor to voluntarily and unconditionally assign their claims to the trustee while knowing that a claim might not brought for some time make it unlikely that a trustee would be able to rely on *Grede*.

Since *Caplin*, Congress has overhauled and amended the Code. However, the 5–4 opinion continues to shape the scope of the trustee’s powers under Section 544(a)(1). Under *Caplin*, *Ozark*, and, most recently, *JPMorgan Chase & Co.*, the trustee’s role is narrow. A literal reading of Section 704, confining the trustee to collecting monies for the estate, and a focus on the avoidance power in the second sentence of Section 544(a)(1) circumscribes the trustee’s ability to stand in the shoes of creditors. Even where *Caplin*’s three concerns do not apply, the courts have construed the general-versus-specific claims distinction narrowly. The result is what the Court in *Ozark* termed the “flavor” of Section 544(a)(1): a limited addition to the trustee’s ability to expand the estate under Section 541 through avoiding transactions.¹⁶⁸

C. Conclusion

Courts’ analyses of Sections 541 and 544 leave little room for a trustee to recover from third parties in Ponzi schemes. The text of Section 541 requires the courts to consider whether the fraud would have been imputed at the moment of the petition. In almost every case, the answer is yes as a result of the application of *in pari delicto*. Defeating *in pari delicto* requires relying on Section 544 and arguing that the creditors are the real parties in interest. However, *Caplin* and *Ozark* teach that the courts are wary of expanding the trustee’s power through the backdoor, so to speak. Reading Section 544 in the context of the other avoiding powers and the history of the trustee’s role, courts are only willing to grant the trustee standing when she is actually acting “on behalf of the debtor in possession.”¹⁶⁹ Setting aside the circular nature of this dilemma, one has to ask whether this makes any sense. The trustee’s powers should carry her no further than they can be justified by the background principles of bankruptcy. The purpose of Part Three is to explore

167. See *Williams v. California 1st Bank*, 859 F.2d 664, 666–67 (9th Cir. 1988) (finding that *Caplin*’s concerns applied to block the trustee even after the close of the bankruptcy case).

168. See *Mixon v. Anderson (In re Ozark Rest. Equip. Co., Inc.)*, 816 F.2d 1222, 1229 (8th Cir. 1987).

169. See *supra* notes 124–26 & accompanying text.

whether the result of Sections 541 and 544 is consistent with the theoretically appropriate scope of the trustee's role given the first principles and purposes of bankruptcy law.

III. FUNDAMENTALS OF BANKRUPTCY AND APPLICATION TO CREDITOR ACTIONS AGAINST THIRD PARTIES

Part Three addresses the theoretical justifications for having a bankruptcy forum at all and analyzes whether these justifications support giving the trustee standing to sue third parties in the context of Ponzi schemes. There are two rationales for the bankruptcy forum: the “creditors’ bargain” and the “traditionalist” accounts.¹⁷⁰ I begin with an explanation of each theory. I then argue that either model supports granting trustees standing to sue third parties given the peculiar circumstances of a Ponzi scheme.

A. *Theoretical Justifications*

1. The Theory of the Creditors’ Bargain

Leaving aside distributional questions, the “creditors’ bargain” theory begins from the assumption that the purpose of bankruptcy should be to maximize the creditors’ collective benefits.¹⁷¹ In a world free of transaction costs, the creditors would agree *ex ante* to the loss-distribution regime that maximizes welfare.¹⁷² Thomas Jackson proposes three core reasons why the creditors in such a hypothetical world would choose a collective proceeding: (1) reduction of strategic costs; (2) increased aggregate asset pools; and (3) administrative efficiencies.¹⁷³

In the absence of a functioning collective action regime, each unsecured creditor has incentives to race to the courthouse and

170. See *supra* notes 17–20 & accompanying text. The classification of Warren and others’ theory as “traditionalist” is borrowed from *Bankruptcy’s Uncontested Axioms*. Baird, *supra* note 17, at 575. However, while Professor Baird refers to his school of thought as “proceduralist,” I borrow Thomas H. Jackson’s moniker for the heuristic that forms the basis for the proceduralist theory—the creditors’ bargain. See generally THOMAS H. JACKSON, *THE LOGIC AND LIMITS OF BANKRUPTCY LAW* (1986); Adler, *supra* note 17, at 444–45 (applying the creditors’ bargain and explaining the heuristic).

171. See *supra* note 17 & accompanying text.

172. See Jackson, *Bankruptcy, Non-Bankruptcy*, *supra* note 17, at 860 (characterizing the creditors’ bargain as an *ex ante* agreement amongst the creditors).

173. *Id.* at 862–68; THOMAS H. JACKSON, *THE LOGIC AND LIMITS OF BANKRUPTCY LAW*, 7–19 (1986); Jackson, *Avoiding Powers*, *supra* note 17, at 728 n.10.

secure a lien as soon as it believes a debtor is likely to default.¹⁷⁴ The “race to the courthouse” may not involve an actual courthouse. Instead, creditors may demand collateral, repayment, or increased interest payments.¹⁷⁵ In many cases, this is a costly endeavor because, assuming the debtor does not willingly give up collateral, which may be a voidable preference, the creditor will have to pay court fees and attorney’s fees. Moreover, if the creditors all rush to demand their money or secure a judgment, the debtor’s business is likely to suffer in the end, leaving a smaller estate for the creditors. As Jackson notes, this race is likely to be worse in a homogenous creditor pool where claims are differentiated only by the order in which they are filed.¹⁷⁶ The unpredictable results of this strategic behavior also increase risk *ex ante* by increasing the variance of a given creditor’s expected return and ultimately result in higher *ex ante* borrowing costs for everyone. The creditors’ bargain theory postulates that creditors may in part prefer a collective collection procedure to prevent the costs of this strategic behavior.

As noted above, the race to the courthouse also may reduce the size of the estate itself. When creditors demand collateral or repayment when they believe the debtor is likely to fail, the result is a self-fulfilling prophecy: faced with an onslaught from its creditors, the debtor fails before it is economically unviable.¹⁷⁷ The size of the estate is likely to suffer further if no collective proceeding ever occurs. In the absence of a bankruptcy proceeding, each creditor will eventually reduce their claim to a judgment, obtain a writ of execution, and establish a lien on whatever property is left. The problem is that the estate is frequently greater than the sum of its parts. A piecemeal individual collection regime may destroy any going-concern value still in the business.¹⁷⁸ Thus, when the debtor is economically viable, the hypothetical creditors would agree *ex ante* to a

174. See Jackson & Scott, *supra* note 17, at 169–70 (discussing the incentives for creditors to coerce payment from debtors when bankruptcy becomes foreseeable); Jackson, *Avoiding Powers*, *supra* note 17, at 758–59 (discussing the role of avoidance powers in counteracting these incentives).

175. See Jackson, *Avoiding Powers*, *supra* note 17, at 772 (discussing use of collateral as a means to improve priority during a race).

176. See Jackson, *Bankruptcy, Non-Bankruptcy*, *supra* note 17, at 863–64.

177. *Id.* at 860–65.

178. *Id.* See generally Douglas G. Baird, *The Uneasy Case for Corporate Reorganizations*, 15 J. LEGAL. STUD. 127, 133–40 (1986) [hereinafter Baird, *Uneasy Case*] (discussing the case for preservation of going concern value in the choice between chapter 7 and chapter 11 filings). Preservation of going-concern value is also essential to the traditionalist account. See, e.g., Warren, *Bankruptcy Policy*, *supra* note 18, at 798 (discussing preservation of going concern value as a part of the rehabilitative goals of bankruptcy law).

collective regime that preserves any going-concern value of the debtor in the event of a financial failure.¹⁷⁹

Finally, the creditors should prefer a collective regime because of the administrative efficiencies it is likely to generate.¹⁸⁰ Each creditor faces the same costs of a court proceeding. A collective proceeding reduces duplicative expenses reserving as much of the estate as possible to satisfy the creditors' claims.¹⁸¹

According to the creditors' bargain theory, there are no other reasons to have a collective bankruptcy procedure. The creditors do not achieve this desirable outcome without bankruptcy because it is impossible to contract with every potential future creditor, making government, in the form of the Code, necessary to effectuate the creditors' bargain.¹⁸² The corollary to this idea is that bankruptcy should only disturb state-based entitlements when it is necessary to achieve the goals of the system in the first place.¹⁸³ To otherwise

179. See Jackson, *Bankruptcy, Non-Bankruptcy*, *supra* note 17, at 864–65. However, creditors' bargain theorists do suggest that this will rarely be the case. See THOMAS H. JACKSON, *THE LOGIC AND LIMITS OF BANKRUPTCY LAW*, 11–17 (1986). Some commentators have also suggested that the business may be preserved to avoid the transactional costs of a sale or piecemeal liquidation. See *generally* Baird, *Uneasy Case*, *supra* note 186, at 136.

180. See Jackson, *Bankruptcy, Non-Bankruptcy*, *supra* note 17, at 866 (discussing the reduction in duplicative litigation in a collective proceeding). Thus, a central concern in the literature discussing the trustee's role is that providing him with expansive powers will increase administrative costs. See *generally* Bryan D. Hull, *A Void in Avoidance Powers? The Bankruptcy Trustee's Inability to Assert Damages Claims on Behalf of Creditors Against Third Parties*, 46 U. MIAMI L. REV. 263, 301 (1991) (discussing the increase in administrative fees likely if Congress had passed Section 544(c) or if the trustee was otherwise empowered).

181. See Jackson, *Bankruptcy, Non-Bankruptcy*, *supra* note 17, at 866.

182. See Jackson & Scott, *supra* note 17, at 203. See also *id.* at 156 n.2 (“One of the touchstones of this model is that . . . collective regime such as bankruptcy is sometimes necessary as a device to induce cooperative behavior”); Jackson, *Bankruptcy, Non-Bankruptcy*, *supra* note 17, at 866 (“As a result, the creditors themselves cannot be expected to negotiate this agreement, even though it would be in their joint interest to do so. A federal bankruptcy rule solves this problem by making available a mandatory collective system after insolvency has occurred.”).

183. See Jackson & Scott, *supra* note 17, at 160–62 (describing the importance of preserving state law entitlements to entice secured creditors to join the hypothetical bargain and to reduce strategic behavior). This also implicitly follows from the rejection of rehabilitation rationale for the collective proceeding which would require additional deviations from state law. See *id.* at 164 (“A central premise of the simple creditors' bargain is that redistribution in bankruptcy is inconsistent with the maximizing objectives of the collective. Insolvency is seen as a foreseeable risk that is borne individually by the various claimants of any business enterprise. Thus the model assumes, *inter alia*, that none of the risks of business failure will be shared among claimants of different classes, except as otherwise explicitly agreed.”).

disturb entitlements is to reshape contractually agreed-to rights that do not face the same collective action problems that lead to the creditors' bargain.¹⁸⁴ Disturbing entitlements can lead to forum shopping and may undermine the certainty that is an important justification for the collective system in its own right.¹⁸⁵ In the creditors' bargain, a particular policy or Code provision can be judged on whether the creditors would hypothetically agree *ex ante* to include the particular policy or provision in their bargain.¹⁸⁶

2. The Traditional Account

Instead of asking how to maximize creditor welfare, the traditionalist account considers loss distribution on a grander scale.¹⁸⁷ Traditionalists approach bankruptcy with a broader perspective by considering the rights of the creditors, the debtor, and other affected constituencies, such as employees.¹⁸⁸ Like the creditors' bargain theory, bankruptcy at its core remains a collection procedure to traditionalists but while the basis of that procedure may be pre-bankruptcy state entitlements, the traditionalist account accepts that deviation may be necessary to accomplish distributional goals.¹⁸⁹

Traditionalists do not ignore the value of ending the race to the courthouse. These theorists recognize that a collective procedure is necessary to preserve value by reducing strategic behavior, eliminating redundant transaction costs, and preserving any going-

184. *Id.* at 160–61.

185. *Id.* at 162–63. *See also* Douglas G. Baird, *Loss Distribution, Forum Shopping, and Bankruptcy: A Reply to Warren*, 54 U. CHI. L. REV. 815, 825–28 (1987) [hereinafter Baird, *Loss Distribution*] (explaining how creating an additional forum for enforcement with different rules can lead to costly strategic behavior).

186. *See, e.g.*, Baird, *Loss Distribution*, *supra* note 185, at 827 (explaining bankruptcy's recognition of state priority through the creditors' bargain); Jackson & Scott, *supra* note 17, at 164–66 (explaining risk-sharing in bankruptcy through the creditors' bargain); Adler, *supra* note 17 (evaluating risk-sharing theory through the creditors' bargain).

187. *See* Warren, *Policymaking*, *supra* note 18, at 346 (“Federal law creates a multifaceted, integrated system to cope with the competing concerns of a wider range of interested parties in more complicated relationships and more distressed circumstances.”).

188. *See* Warren, *Bankruptcy Policy*, *supra* note 18, 789–90 (explaining bankruptcy's concern for distribution and employees).

189. *See* Warren, *Policymaking*, *supra* note 18, at 353 (“Every distribution that benefits a particular creditor at the expense of the collective estate represents a considered judgment to depart from the norm in a particular instance. Equality—and deliberate deviations from equality—stand at the center of bankruptcy policy.”).

concern value.¹⁹⁰ However, because they believe that the Code accomplishes this through collectivizing the process under the aegis of a bankruptcy court, the judge or trustee may have enhanced power under the traditionalist account.¹⁹¹ Indeed, a central difference between the two theories is the traditionalist recognition that bankruptcy judges can, and probably should, have a great deal of discretion.¹⁹² The traditionalists believe that this discretion is built into the Code by providing for abbreviated trials, emergency orders, and restricted notifications.¹⁹³

Traditionalists also recognize the role of distributional concerns in the Code:

Congress was acutely aware of the wider effect of a business failure on the surrounding community, and it adopted the 1978 Bankruptcy Code specifically to ameliorate those harmful effects—that is, to redistribute the benefits that would stem from some creditors’ collection rights to other parties who did not enjoy those rights.¹⁹⁴

Elizabeth Warren proposes that this solicitude towards distributional effects is at least in part premised on the idea that the Code should protect those unable to do so. Thus Warren argues that the Code should favor employees.¹⁹⁵ Moreover, the idea that “equity is equality” motivates the decision to treat like creditors the same regardless of where they finish in the race to the courthouse.¹⁹⁶ Even attempts to preserve going-concern value can be motivated by distributional concerns: “The economy of an entire town can be disrupted when a large factory closes. . . . Some believe that preventing such consequences is worth the costs of trying to keep the firm running and justifies placing burdens on a firm’s secured creditors.”¹⁹⁷

190. Compare Warren, *Bankruptcy Policy*, *supra* note 18, at 792–93 (discussing estate value maximization), with Baird, *Loss Distribution*, *supra* note 185, at 815–17 (discussing traditional theory and the commonalities between the two camps).

191. See Warren, *Bankruptcy Policy*, *supra* note 18, at 805–08 (discussing the expansive authority of judges in the traditional account); see also Baird, *supra* note 17, at 590–95 (discussing the broader role for judges in the traditional account).

192. See Warren, *Policymaking*, *supra* note 18, at 351–52 (discussing the “enormous discretion” of bankruptcy judges).

193. *Id.* at 348.

194. *Id.* at 355.

195. *Id.* at 355–57 (explaining bankruptcy’s protections for employees through reorganization and rehabilitation).

196. *Id.* at 357.

197. Warren, *Bankruptcy Policy*, *supra* note 18, at 798.

3. Internalizing Costs and the Moralism Account

In addition to the theories outlined above, two other policies also motivate bankruptcy: (1) internalization of costs, and (2) private monitoring.¹⁹⁸ The Code internalizes costs by operating as a self-supporting system.¹⁹⁹ It also limits the spillover of costs by prioritizing tax payments to the government ahead of distributions to private creditors.²⁰⁰ Finally, by making the decision to opt into the system a private choice, the Code allows debtors, and, in limited situations, creditors, to decide between the two forums as they see fit.²⁰¹ This saves costs and ensures that the decision is made by the party with the best information. This is so even though incentives dictate that at certain times, creditors must be able to force the debtor's hands.²⁰²

The creditors' bargain and, to a lesser extent, the traditionalist account, have been criticized by moralism commentators.²⁰³ These theorists argue that an economic analysis misses the social context and moral values that drive important sections of the Code.²⁰⁴ For moralists, bankruptcy is a means to distribute losses, calm chaos, and rehabilitate the debtor.²⁰⁵ However, once we move beyond the realm of the individual debtor, moralism concerns become less relevant.

Similarly, beyond the context of an individual debtor, the differences between the traditionalist and creditors' bargain accounts are less noticeable than they first appear. Both theories conclude

198. See Warren, *Policymaking*, *supra* note 18, at 343–44.

199. *Id.* at 361.

200. *Id.* at 382 n.124 (discussing IRS priority).

201. See 11 U.S.C. § 301 (2006) (voluntary petitions).

202. See 11 U.S.C. § 303 (2006) (conditions for involuntary petitions). There is some debate on whether it is normatively desirable for bankruptcy to be a contractual default rule or a mandatory rule. Compare Rasmussen, *supra* note 18, at 55–63 (critiquing the creditors' bargain conception of chapter 11 as mandatory and arguing that it is better analyzed as a default contractual term), with Baird, *Uneasy Case*, *supra* note 186, at 135 (explaining that bankruptcy must be mandatory to effectuate the creditors' bargain).

203. See generally Korobkin, *Rehabilitating Values*, *supra* note 18, at 721; Korobkin, *Value and Rationality*, *supra* note 18; Donald R. Korobkin, Contractarianism and the Normative Foundations of Bankruptcy Law, 71 TEX. L. REV. 541 (1993).

204. See Korobkin, *Rehabilitating Values*, *supra* note 18, at 721 (offering a “values-based” account for bankruptcy). Other scholars have also criticized the creditors' bargain's economic analysis as unrealistic. See, e.g., Eisenberg, *supra* note 18, at 1522, 1529 (arguing that the creditors' bargain oversimplifies things); Rasmussen, *supra* note 18, at 55–59 (same).

205. See Korobkin, *Rehabilitating Values*, *supra* note 18, at 765–66.

that bankruptcy should work from the baseline of pre-bankruptcy state entitlements, that collective action and strategic behavior are important rationales for bankruptcy, and that the bankruptcy process should remain a procedural collection system.²⁰⁶ Indeed, when faced with a decision that does not involve distribution between creditors but rather between creditors and a third party, the two theories are very likely to reach the same result.

B. When the Two Theories Agree: Ponzi Schemes

The open question after *Ozark* and *JPMorgan Chase & Co.* is not whether the trustee currently has the power to bring tort claims against third parties on behalf of the creditors: she does not. But that does not mean that the trustee should not be able to do so. The concerns in *Caplin* are relevant, as concerns about preemption and subrogation do affect whether it is in the interests of the debtor and creditors for the trustee to take on this role.²⁰⁷ And under either conception of bankruptcy, there are conditions where we would not want the trustee to have this authority. For instance, a Chapter 11 reorganization complicates the analysis by introducing the relationship of the debtor to the creditors. However, policy-makers and judges should analyze the problem in the context of a modern Ponzi scheme: few participants, a liquidation proceeding, and no secured creditors. Under these conditions, either theory of bankruptcy suggests the trustee should have the power to bring tort claims against third parties for aiding and abetting the schemer's fraud.

1. Trustee Suits and the Creditors' Bargain

According to the creditors' bargain theory, the creditors as a group hypothetically agree to a collective process because it ultimately leads to the largest estate possible by preventing strategic behavior and curing collective action problems.²⁰⁸ The trustee and her avoidance powers play an important role in this model because it is through the mechanism of the trustee that the rights of creditors vis-a-vis each other are sorted out.²⁰⁹ Meanwhile, the trustee

206. See Baird, *Loss Distribution*, *supra* note 185, at 815–17 (discussing traditional theory and the commonalities between the two camps).

207. See *supra* notes 146–47 and accompanying text. See also Hull, *supra* note 188, at 280–81 (recognizing the linkage between the theoretical underpinnings for bankruptcy and the *Caplin* concerns).

208. See *supra* notes 17–18 & accompanying text.

209. See Jackson, *Avoiding Powers*, *supra* note 17, at 726 (distinguishing avoidance powers based on whether they adjust rights between creditors or between

saves the individual creditors the costs and uncertainty of a race and, through the role of a “hypothetical lien creditor,” is able to avoid transfers no individual creditor could necessarily reach.²¹⁰ However, whereas scholars, notably Thomas Jackson, have explored whether the trustee should be able to bring claims that rearrange rights between the creditor and the debtor,²¹¹ little attention has been devoted to whether the trustee rearranging rights between creditors and third parties is consistent with the creditors’ bargain theory.

Unlike many of the trustee’s avoidance powers, the power to sue third-party creditors does not alter pre-bankruptcy state entitlements. Part One explained that the tort creditors in a Ponzi scheme have several available causes of action. Negligence and breach of fiduciary duty are difficult to make out because they require a duty between the creditors themselves and the third parties.²¹² However, secondary liability predicated on the third party’s relationship with the debtor has been recognized as a viable cause of action²¹³ even by the courts that have found that the trustee does not have standing.²¹⁴ Likewise, while deepening insolvency is controversial, it remains a stand-alone cause of action in many states.²¹⁵ Thus allowing the trustee to bring the claim leaves state law entitlements as they were outside of bankruptcy. Relatedly, the absence of any secured creditors means that state property rights are also unaffected by the additional powers.²¹⁶ Empowering the trustee modifies the procedures used. It does not modify the underlying substantive state-based entitlements. This makes empowering the trustee consistent

creditors and the debtor); *see also* David Gray Carlson, *Bankruptcy’s Organizing Principle*, 26 FLA. ST. U. L. REV. 549, 551–53 (1999) (presenting avoidance, instead of successorship, as the centerpiece of the author’s theory of bankruptcy).

210. *See* 11 U.S.C. § 544(a) (2006); Jackson & Scott, *supra* note 17, at 179–82 (discussing the prototypical use of § 544). However, because trustee expenses are administrative expenses entitled to priority there is a fine line between savings created by a collective process and potentially duplicative and costly litigation. *See* Steve H. Nickles & Edward S. Adams, *Tracing Proceeds to Attorneys’ Pockets (and the Dilemma of Paying for Bankruptcy)*, 78 MINN. L. REV. 1079, 1166–77 (1994). Thus a central distinction Jackson makes is between actions that creditors could already bring outside of bankruptcy, like fraudulent conveyances, and what he considers to be “true” avoidance powers like § 544(a) and § 547(a) preferences. *See* Jackson, *Avoiding Powers*, *supra* note 17, at 726.

211. *See* Jackson, *Avoiding Powers*, *supra* note 17.

212. *See supra* notes 49–52 & accompanying text.

213. *See supra* notes 86–90 & accompanying text.

214. *See supra* notes 91–101 & accompanying text.

215. *See supra* notes 58–71 & accompanying text.

216. *See supra* notes 182–83 & accompanying text.

with the creditors' bargain theory's goal of limiting the deviation from state law except where necessary to effectuate a collective proceeding.²¹⁷

It is relatively easy to see how empowering the trustee would ultimately increase the creditors' collective recovery by ameliorating collective action problems and reducing strategic behavior. With the exception of the Tenth Circuit, the automatic stay does not apply to these actions.²¹⁸ Thus each creditor can choose whether and when to bring suit. This choice is governed by the costs, in terms of legal fees, and by the expected payoff. Both of these factors are dependent in part on the bankruptcy process and the actions of other creditors. If other creditors bring suit first, the hypothetical claimant may benefit from their success through offensive collateral estoppel or by their loss as an indication of their own probability of success. At the same time, if their claim is completely satisfied by the bankruptcy process, their expected payoff may be minimal. Moreover, if the trustee chooses to claw back net winners, the payoff may be altered even further.²¹⁹ The result is that no individual creditor has incentives to bring suit quickly, if at all.

Importantly, the decision not to bring suit is only indirectly based on the probability of success. When the trustee is responsible for the decision whether to bring the claim, these concerns are less important. When the trustee is empowered, the automatic stay applies; thus she cannot benefit from a wait-and-see approach.²²⁰ Moreover, clawbacks and net-winners are no longer relevant. Any recovery by the trustee will go into the general pool to be paid out to creditors as determined by the other parts of the bankruptcy proceeding.²²¹ Every creditor benefits and each benefits according to his priority and pro rata share.²²² Thus a collective remedy solves

217. See *supra* notes 183–86 & accompanying text.

218. See *supra* notes 152–63 & accompanying text.

219. See Nelson, *supra* note 56, at 1458–63 (discussing the application § 544 and § 548 in Ponzi schemes to clawback “net winners”). See also *Pickard v. Estate of Chais (In re Bernard L. Madoff Inv. Sec. LLC)*, 445 B.R. 206 (Bankr. S.D.N.Y. 2011) (finding that trustee could claw back fictitious profits as fraudulent transfers).

220. See *Hill v. Gibson Dunn & Crutcher, LLP (In re MS55, Inc.)*, No. 06-CV-01233, 2008 WL 2358699 (D.Colo. June 6, 2008) (explaining that the stay would apply if the trustee has the power to bring claims under § 544(a)).

221. See 11 U.S.C. § 550 (2006).

222. See *Picard v. JPMorgan Chase & Co.*, 460 B.R. 84, 96 (S.D.N.Y. 2011) (finding that recovery would follow the order of priority established by other parts of the code, but requiring that recovery be equal before the trustee could be empowered). The Court's reasoning here is not immediately clear as there is no reason why an exactly equal right to payment would make the claim any less “peculiar” to the individual or on a theoretical level implicate any reason why the

the collective action problem and reduces incentives for strategic behavior.

According to the creditors' bargain theory, the collective process also serves to reduce transaction costs.²²³ In the Ponzi scheme context, that interest is clearly served. In the absence of a trustee, each individual creditor would need to independently evaluate its prospects for recovery and hire counsel. These efforts are likely to be duplicative. Just as bankruptcy saves the costs of levying on a debtor's property individually,²²⁴ empowering the trustee saves the costs of bringing suit. The unique position of the trustee is also likely to increase the savings. With access to substantially more information than an individual creditor, the trustee can take advantage of the work already done and reduce the costs of discovery.²²⁵ Moreover, he will not need to spend additional resources identifying claimants as that process has already occurred.²²⁶ A collective suit improves efficiency by reducing the number of proceedings, reaping economies of scale, and saving money.

In some situations, the creditors might also prefer the trustee to evaluate the merits of the claim. In a liquidation proceeding, the trustee has control over the estate and has access to all of its records.²²⁷ She also acts as a clearinghouse for all of the claims.²²⁸ This allows her to evaluate the expected payoff of a claim in terms of its probability of success and its expected damages more effi-

creditors would prefer to go it alone. After all, even those at the back of the priority line benefit from expanding the estate, regardless of whether that benefit is equal. *See id.* ("It is true that allowing the Trustee to pursue claims that belong properly to individual creditors would accrue to the benefit of all creditors by augmenting the bankruptcy estate.").

223. *See* Jackson, *Bankruptcy, Non-Bankruptcy*, *supra* note 17, at 866.

224. *See supra* notes 174–81 & accompanying text.

225. *See* Warren, *Policymaking*, *supra* note 18, at 347 ("Savings are also realized by imposing stiff requirements on debtors to cooperate with creditors' efforts to monitor a troubled business. A few examples illustrate the point. Following a bankruptcy filing, the debtor must reveal detailed information about the past operation of the business and its projected business activities."); *see also* *Mixon v. Anderson (In re Ozark Rest. Equip. Co., Inc.)*, 816 F.2d 1222, 1225 (8th Cir. 1987) (acknowledging that empowering the trustee would save time and money).

226. *See* 11 U.S.C. § 704(a)(5) (2006) (empowering trustee to investigate claims).

227. *See* 11 U.S.C. § 704(a)(1) (2006) (empowering trustee in liquidations); 6 COLLIER ON BANKRUPTCY ¶ P. 704.05[2] (Matthew Bender & Co., 16th ed. 2012) (explaining the trustee's duty to account for the debtor's estate and keep records). *But see* 11 U.S.C. § 363(b)(1) (2006) (limiting trustee's ability to dispose of property outside the normal course of business).

228. 6 COLLIER ON BANKRUPTCY ¶ 704.02[3] (Matthew Bender & Co., 16th ed. 2012) (discussing trustee's duties to expeditiously close the estate pursuant to § 704); *id.* at ¶ 704.08 (discussing trustee's duties to object to improper claims).

ciently than an individual litigant. She has access to better information and may already be in a position to subpoena documents from third parties.

Caplin's three concerns can also be evaluated through the lens of the creditors' bargain theory as reasons why the creditors might not want the trustee to have the power to bring these sorts of claims. While the claim is not general in the sense used by the *Koch* court in that the injury is not indirect, it is functionally general in the sense of the term as used by the *Delgado* court, which makes *Caplin's* preemption rationale longer relevant.²²⁹ Every creditor could make the same argument and each has the same probability of success. The only difference between their claims is the amount of damages, but this is fundamentally no different than the *pro rata* distribution the bankruptcy process uses to satisfy unsecured claims. Thus, under the Tenth and Second Circuits' approach, the stay should operate to block any of the creditors from bringing the claim, making the trustee the sole party able to bring it. If the stay operates, then logically the individual creditors are also bound by the result of the trustee's action.²³⁰ Thus the creditors would be preempted from bringing suit individually but, as explained above, they are likely to prefer this state of affairs. Additionally, the Securities Investor Protection Corporation ("SIPC") requires assignment. To the extent any of the creditors wish to recover through insurance, they therefore already have to accept preemption.²³¹ Assignment also helps mitigate the argument that the trustee is not the real party in interest.²³²

Subrogation is not an issue because unlike *Caplin*, a Ponzi scheme is unwound as a SIPA liquidation, not a reorganization.²³³

229. See *St. Paul Fire & Marine Ins. Co. v. Pepsi Co., Inc.*, 884 F.2d 688, 700 (2d Cir. 1989) (finding preemption wasn't possible where the trustee is the only party with standing as bankruptcy law would bind the creditors); *Delgado Oil Co., Inc. v. Torres*, 785 F.2d 857, 861 (10th Cir. 1986) (finding the automatic stay applies to general claims); *Hill v. Gibson Dunn & Crutcher, LLP (In re MS55, Inc.)*, No. 06-CV-01233, 2007 WL 2669150, at *13 (D.Colo. Sept. 6, 2007) (finding *Caplin's* preemption rationale unpersuasive where the automatic stay applies).

230. See *supra* notes 152–56 & accompanying text.

231. See 15 U.S.C. §§ 78fff-2(c)(1)-(3) (2006) (requiring recovered customer property to be used to repay SIPC prepayments and treating recovered transfers as customer property).

232. See *supra* notes 165–67 & accompanying text.

233. In *Caplin*, the Court found it plausible that the business would be reorganized, giving the debenture trustee (the third party being sued) a claim for subrogation against the corporation. See *Caplin v. Marine Midland Grace Trust Co.*, 406 U.S. 416, 430–32 (1972). However, given the realities of liquidating a Ponzi scheme, where claims far outstrip assets, this seems unlikely to come to pass. See

Moreover, to the extent *Caplin* relied on the potential harm subrogation posed to equityholders as residual claimants, that concern is inapt under the creditors' bargain theory, where equity is irrelevant.²³⁴

Finally, *Caplin*'s construction of the trustee's role is inconsistent with the creditors' bargain theory. As noted above, the Court in *Caplin* gave the trustee a narrow role, based merely on collecting information and money for the estate.²³⁵ In the creditors' bargain theory, the trustee is charged with implementing the hypothetical agreement through the collective process.²³⁶ So long as she is not reordering pre-bankruptcy entitlements between the creditors in the process, there is no reason to construe her role so narrowly. Additionally, while the court in *Ozark* rejected the trustee's claim, it found that Congress had intended to substantially enlarge the trustee's powers under the Code as compared to the trustee's powers under the old Bankruptcy Act.²³⁷ There is no clear basis for treating the Court's decision in *Caplin*, which is arguably inconsistent with congressional intent and entailed interpreting a statute that has since been replaced, as binding.

The result of this analysis is that hypothetical creditors would likely agree to empower the trustee to bring these claims as a way to reduce collective action problems, generate efficiencies, and expand the estate. Meanwhile, the *Butner* principle,²³⁸ which states

Logan v. JVK Real Estate Servs. (*In re Bogdan*), 414 F.3d 507, 514–15 (4th Cir. 2005) (remarking that subrogation was very unlikely in the context of a liquidation, as it would require recovery of more than the amount of claims against the debtor and even then would likely be barred by *in pari delicto*). Moreover, under SIPA, SIPC would have first crack at any of these potential assets ahead of any third parties asserting subrogation rights. See 15 U.S.C. § 78fff-2(c) (2006). Even if this was not the case, as the dissent in *Caplin* pointed out, a subrogee typically needs clean hands, a condition no party found liable for aiding and abetting fraud could satisfy. See *Caplin*, 406 U.S. at 440 (Douglas, J., dissenting). Finally, the subrogation rationale has less force when the trustee stands in the creditors', rather than the debtor's, shoes. See *Williams v. California 1st Bank*, 859 F.2d 664, 667 (9th Cir. 1988) (acknowledging that the third party's subrogation defense applies to actions by the debtor, but not necessarily to actions on behalf of the creditors).

234. See *supra* notes 172–81 & accompanying text; *Caplin*, 406 U.S. at 438 (“Whether conditions have changed so as to leave some equity for the old stockholders, we do not know . . . In some cases the elimination of one entire class of creditors or a pro rata reduction in their claims would give stockholders a chance to participate in the plan.”).

235. See *Caplin*, 406 U.S. at 428–29.

236. See *supra* notes 180–81 & accompanying text.

237. See *Mixon v. Anderson (In re Ozark Rest. Equip. Co., Inc.)*, 816 F.2d 1222, 1227 (8th Cir. 1987).

238. See *supra* note 48 & accompanying text.

that state law entitlements should be respected in a bankruptcy proceeding, remains unaffected. In addition, analyzing *Caplin's* tripartite rationale as reasons why the hypothetical creditors would not want the suit brought shows that none of these three factors suggests that the trustee should not have standing. Given that empowering the trustee serves the purposes of the creditors' bargain theory and does not alter substantive state law, the theory confirms that the trustee should be empowered.

2. Traditionalist Conceptions

Recognizing the distinctions between a corporate liquidation and an individual bankruptcy is central to understanding why the traditional explanation for bankruptcy also supports empowering the trustee. The traditionalist account stresses the fresh start policy as an important rationale for deviating from a free-for-all state collection system.²³⁹ However, in a corporate liquidation, the "fresh start" policy is irrelevant.²⁴⁰ By definition, a Ponzi scheme has no going-concern value: in short, there is nothing to preserve and thus no independent interest of the debtor. Yet at least one rationale for limiting the trustee's power to sue on behalf of the creditors stems from the conflict of interests between creditors and a trustee acting as the debtor-in-possession.²⁴¹ In such a situation, the role of the trustee, and by extension the Code, is to further the interests of the creditors and potentially other affected constituencies but not the debtor.²⁴²

Distributional concerns are also prominent in the traditionalist account.²⁴³ In a typical corporate bankruptcy, there are secured creditors, unsecured creditors, employees, future tort claimants, customers, and past employees. All parties have some stake in the business even if that stake may not be recognized as a formal claim

239. See Warren, *Policymaking*, *supra* note 18, at 341–42.

240. See *id.* (drawing a distinction between business bankruptcy and consumer bankruptcy). See also Thomas H. Jackson, *The Fresh-Start Policy in Bankruptcy Law*, 98 HARV. L. REV. 1393, 1396 (1985) ("The fresh-start policy is thus substantively unrelated to the creditor-oriented distributional rules that give bankruptcy law its general shape and complexity.").

241. See *Caplin v. Marine Midland Grace Trust Co.*, 406 U.S. 416, 433 nn. 21–22 (1972) (expressing concern that the trustee may stand on both sides of the suit).

242. See Warren, *Policymaking*, *supra* note 18, at 344–52 (prioritizing enhancement of value in business bankruptcies).

243. See *supra* notes 187–97 & accompanying text.

under the Code.²⁴⁴ The Code accommodates these stakeholders by accelerating contingent and unmatured claims to force future claimants into²⁴⁵ and through the Chapter 11 process,²⁴⁶ which may preserve jobs by preserving going-concern value.

In a Ponzi scheme bankruptcy, none of these concerns is relevant. When the “sole actor” exception applies, there are few other constituencies to consider. Ponzi schemes have few employees and even fewer who are innocent of wrong-doing.²⁴⁷ More importantly, the mechanism for considering their jobs—preserving the business—is untenable. There are also few, if any, secured creditors.²⁴⁸ With the exception of a situation in which a secured creditor is also a target for the trustee, secured claims should be unaffected by these suits. The collateral that served as a security interest is untouched and any deficiency claim bifurcated by Section 506 benefits from the expansion of the estate.²⁴⁹ Finally, bankruptcy serves as a mechanism to distribute losses.²⁵⁰ If the trustee can make out an aiding and abetting claim against a third party, then their complicity and concomitant benefit from the scheme suggest they should share some of the costs.

244. See Warren, *Policymaking*, *supra* note 18, at 355–57 (discussing the Code’s protections for parties with no legal rights in terms of distributing losses in bankruptcy); Warren, *Bankruptcy Policy*, *supra* note 18, at 808–11 (discussing the primary importance of apportioning losses in bankruptcy).

245. See 11 U.S.C. § 502 (2006) (including contingent and unmatured liabilities as claims).

246. See Christopher W. Frost, *Bankruptcy Redistributive Policies and the Limits of the Judicial Process*, 74 N.C. L. REV. 75, 78 (“While employees, communities, and other business dependents may not have traditional claims, the failure of businesses nevertheless implicates their interests. Thus, an analysis of the reorganization process that excludes the interests of these constituencies is too narrow. On this view reorganization can only be understood as a broadly inclusive mechanism of social policy intended to distribute the social costs of business failure.”).

247. See *supra* notes 27–28 & accompanying text.

248. See *supra* notes 8–10 & accompanying text.

249. See 11 U.S.C. § 506(a)(1) (2006) (“An allowed claim of a creditor secured by a lien on property in which the estate has an interest, or that is subject to setoff under section 553 of this title, is a secured claim to the extent of the value of such creditor’s interest in the estate’s interest in such property, or to the extent of the amount subject to setoff, as the case may be, and is an unsecured claim to the extent that the value of such creditor’s interest or the amount so subject to setoff is less than the amount of such allowed claim.”).

250. See Baird, *Loss Distribution*, *supra* note 185, at 816–18 (discussing the traditional account and the importance of loss distribution therein); Warren, *Bankruptcy Policy*, *supra* note 18, at 788 (“By giving the debtor business an opportunity to reorganize, the bankruptcy scheme acknowledges the losses of those who have depended on the business and redistributes some of the risk of loss from the default.”).

The traditionalist account also recognizes that a collective proceeding can serve to calm chaos.²⁵¹ Preventing a mad dash to levy the assets of a debtor does not just reduce strategic behavior but it also calms contagion and prevents fear from rippling outward into the markets.²⁵² SIPC insurance is also designed to reduce the impact of collapse on *ex ante* investment.²⁵³ However, when a Ponzi scheme collapses and there are few initial assets in the estate, the process is unlikely to engender confidence among investors and victims. Empowering the trustee furthers the goals of the broader statutory scheme by consolidating the proceeding and providing certainty and fixed liabilities faster than would otherwise be the case.

As with an analysis of empowering the trustee under the creditors' bargain theory, the traditional theory also supports allowing suit against third parties. This conclusion stems from an appreciation of the differences between a Ponzi scheme bankruptcy and a personal bankruptcy or Chapter 11 reorganization. Constituencies to consider are narrower, the debtor itself is out of the picture entirely, and distribution amongst the creditors is unaffected. Indeed, to the extent any of the concerns specific to the traditionalist account are relevant, they suggest that empowering the trustee to vindicate the rights of third parties is also consistent with this theory of bankruptcy.

3. Forum Shopping and Incentives

Although expanding the trustee's power does not alter the substantive state-based entitlement, it may still have worrisome implications. Regardless of the underlying theoretical model, changes to the bankruptcy process may be undesirable to the extent that they incentivize forum shopping and increase moral hazard.²⁵⁴ Indeed,

251. See *supra* notes 190–93 & accompanying text.

252. See Korobkin, *Rehabilitating Values*, *supra* note 18, at 764–66 (discussing the potential for contagion without an orderly bankruptcy process). See generally Warren, *Policymaking*, *supra* note 18 (describing the importance of bankruptcy as a self-contained system that limits spillover and provides counterparties with time to uncouple themselves from failing businesses).

253. See Hurt, *supra* note 3, at 969–71 (discussing the role of SIPC compensation in Ponzi schemes); Pozza, Jr., et al., *supra* note 7, at 122–23 (explaining Congress' goal of "restoring investor confidence" in enacting SIPA).

254. See Baird, *Loss Distribution*, *supra* note 185, at 826 (evaluating forum shopping as a cost of creating an additional avenue of enforcement); Jackson, *Avoiding Powers*, *supra* note 17, at 730–31 n.17 (discussing the potential for forum-shopping); Jackson, *Of Liquidation*, *supra* note 17, at 400 (discussing the role of the *Butner* principle in limiting forum-shopping).

critics of allowing trustees to sue third parties have suggested that forum shopping could result. However, given the peculiar facts of Ponzi schemes, neither of these concerns are valid.

Moral hazard occurs when an *ex post* remedy or result provides a disincentive to seeking contractual protection or to monitor.²⁵⁵ To use a timely example, when banks are “too big to fail” they may have incentives to make riskier bets, knowing that a share of the risk at the left tail of the probability distribution is cut off.²⁵⁶ In the case of a Ponzi scheme, the concern is that empowering the trustee will cause investors to decrease their due diligence and monitoring.

This concern is misplaced. Any decrease in investors’ incentives to monitor is likely to be offset by the increased incentives for third parties to monitor. This may in fact be the most efficient model as large banks and other counterparties that are repeat players have inside information and access to more sophisticated monitoring mechanisms.²⁵⁷ Additionally, moral hazard may presuppose that investors have an ability to contract for protection. In the case of creditor-debtor relations, this is true. For instance, a creditor can seek seniority from other creditors or security from a debtor. What they cannot do is seek protection from a third party with which they have no direct relationship. The Ponzi schemer has a fiduciary duty to his investors but, as explained above, the claim against third parties sounds in tort.²⁵⁸ Thus *ex ante* contracting for protection is not an option. In fact, given the challenges to individual claims explained above, a collective remedy may actually ameliorate a third party’s perverse incentives not to monitor, thereby serving the incentive purpose of tort liability in the first place.²⁵⁹

255. See Adler, *supra* note 17, at 473–74 (discussing the potential for risk sharing to create perverse investment incentives similar to moral hazard); Rasmussen, *supra* note 18, at 96 (discussing the potential for moral hazard in Chapter Eleven cases).

256. See Governor Daniel K. Tarullo, Fed. Reserve Bd., *Confronting Too Big to Fail*, Speech at the Exchequer Club (Oct. 21, 2009) (transcript available at <http://www.federalreserve.gov/newsevents/speech/tarullo20091021a.htm>) (describing too-big-to-fail as “classic” moral hazard).

257. This could be considered similar to a form of gatekeeper liability. See generally Jennifer Arlen & Reinier Kraakman, *Controlling Corporate Misconduct: An Analysis of Corporate Liability Regimes*, 72 N.Y.U. L. REV. 687 (1997) (describing the effects of different liability regimes on corporate incentives to take care). Gatekeeper liability can also make up for insufficient enforcement mechanisms. See Reinier H. Kraakman, *Corporate Liability Strategies and the Costs of Legal Controls*, 93 YALE L.J. 857, 889 (1984).

258. See *supra* notes 46–84 & accompanying text.

259. On the incentive goals of tort liability, see generally ROBERT COOTER & THOMAS ULEN, *LAW AND ECONOMICS* 3–4 (2d ed. 1987) (discussing incentive effects

An underlying premise of the *Butner* principle is that a deviation from state law may lead to undesirable forum shopping.²⁶⁰ When a debtor can get a better deal in bankruptcy, it disrupts the contractual relationship the parties initially agreed to and may lead to strategic behavior.²⁶¹ However, forum shopping is also not an issue in the context of a Ponzi scheme. A Ponzi bankruptcy is almost certainly a liquidation.²⁶² The debtor has no choice in the matter and already has incentives to avoid bankruptcy and the attendant collapse of the scheme for as long as possible.²⁶³ For forum shopping to be a concern, there have to be several options available, but a class action in state court is barred by the Securities Litigation Uniform Standards Act (SLUSA)²⁶⁴ and SIPC already

of tort liability); WILLIAM M. LANDES & RICHARD A. POSNER, *THE ECONOMIC STRUCTURE OF TORT LAW* 312–16 (1987); STEVEN SHAVELL, *ECONOMIC ANALYSIS OF ACCIDENT LAW* 297–98 (1987); Guido Calabresi, *Some Thoughts on Risk Distribution and the Law of Torts*, 70 *YALE L.J.* 499, 517–19 (1961); R. H. Coase, *The Problem of Social Cost*, 3 *J.L. & ECON.* 1, 44 (1960); John C.P. Goldberg, *Twentieth-Century Tort Theory*, 91 *GEO. L.J.* 513, 544–55 (2003) (discussing the modern economic deterrence theory of tort law, which places an emphasis on *ex ante* incentives); ROBERT COOTER & THOMAS ULEN, *LAW AND ECONOMICS* 3–4 (2d ed. 1987) (discussing incentive effects of tort liability); WILLIAM M. LANDES & RICHARD A. POSNER, *THE ECONOMIC STRUCTURE OF TORT LAW* 312–16 (1987).

260. *See supra* notes 48, 254 & accompanying text.

261. *See supra* note 254 & accompanying text.

262. The scheme only ends when the vehicle runs out of cash and the authorities put an end to the scheme. *See supra* notes 8–10 & accompanying text. Forum shopping is generally more of a concern in the context of a reorganization, which managers may use to avoid unfavorable contracts. *See Baird, Uneasy Case, supra* note 186, at 133–40; Jackson, *Of Liquidation, supra* note 17, at 400 (assessing reorganization and noting the potential for forum shopping); Lynn M. LoPucki & William C. Whitford, *Venue Choice and Forum Shopping in the Bankruptcy Reorganization of Large, Publicly Held Companies*, 1991 *Wis. L. Rev.* 11 (1991).

263. *See supra* notes 9–16 & accompanying text.

264. *See* 15 U.S.C. § 78bb(f)(1) (2006):

No covered class action based upon the statutory or common law of any State or subdivision thereof may be maintained in any State or Federal court by any private party alleging—

(A) a misrepresentation or omission of a material fact in connection with the purchase or sale of a covered security; or

(B) that the defendant used or employed any manipulative or deceptive device or contrivance in connection with the purchase or sale of a covered security.

SLUSA has also been consistently applied to Ponzi schemes, including the Madoff case. *See Croscill, Inc. v. Gabriel Capital, LP (In re Merkin)*, 817 F. Supp. 2d 346, 359 (S.D.N.Y. 2011) (finding SLUSA applied to claims alleging debtor aided and abetted Madoff's fraud and collecting cases). Given that SLUSA applies, the Trustee must meet the pleading requirements of the PSLRA. *See Picard v. Kohn, No.*

requires assignment.²⁶⁵ To the extent plaintiffs are able to bring suit outside of bankruptcy, the potential for undesirable forum shopping may actually be greater. Outside of bankruptcy, plaintiffs would have the option of filing suit in any number of different venues. For instance, while the locus of the Madoff fraud was in New York, plaintiffs have avoided that state's Martin Act, which limits securities fraud claims,²⁶⁶ by filing in Connecticut.²⁶⁷

The only other option, a class action under Federal Rule of Civil Procedure 23, would entail the same pleading standards under the Private Securities Litigation Reform Act (PSLRA).²⁶⁸ In addition, the substantive and procedural hurdles claimants would face in a hypothetical class action make it an unattractive option.

The recent experiences of Madoff victims who have filed class actions are illustrative. In *In re Merkin*, victims sued a feeder fund that had invested 100% of their proceeds in Madoff vehicles.²⁶⁹ They argued that the fund's managers had failed to conduct due diligence and had been consciously reckless to the fraud.²⁷⁰ However, the district court rejected the securities fraud claims and

11-CV-1181, 2012 WL 566298, at *6 (S.D.N.Y. Feb. 22, 2012) (finding that the PSLRA scienter requirements applied to certain of trustee's claims).

265. See *supra* note 231 & accompanying text.

266. See N.Y. C.L.S. GEN. BUS. LAW, ART. 23-A, § 352-c(1)(a)-(c) (Consol. 1955):

1. It shall be illegal and prohibited for any person, partnership, corporation, company, trust or association, or any agent or employee thereof, to use or employ any of the following acts or practices:

(a) Any fraud, deception, concealment, suppression, false pretense or fictitious or pretended purchase or sale;

(b) Any promise or representation as to the future which is beyond reasonable expectation or unwarranted by existing circumstances;

(c) Any representation or statement which is false, where the person who made such representation or statement: (i) knew the truth; or (ii) with reasonable effort could have known the truth; or (iii) made no reasonable effort to ascertain the truth; or (iv) did not have knowledge concerning the representation or statement made

267. See, e.g., *Levinson v. PSCC Servs., Inc.*, No. 3:09-CV-00269, 2010 WL 5477250 (D. Conn. Dec. 29, 2010).

268. See *Tellabs, Inc. v. Makor Issues & Rights, Ltd.*, 551 U.S. 308, 324 (2007) (explaining that class actions covered by the PSLRA must raise an inference of scienter at least as compelling as any opposing inference of non-fraudulent intent); *Merrill Lynch, Pierce, Fenner & Smith Inc. v. Dabit*, 547 U.S. 71 (2006) (finding SLUSA and the PSLRA apply broadly to securities fraud class actions); see also *In re Tremont Sec. Law, State Law & Ins. Litig.*, 703 F. Supp. 2d 362, 370 (S.D.N.Y. 2010) (applying PSLRA standard to a class action complaint filed against Madoff's auditors).

269. 817 F. Supp. 2d 346, 361 (S.D.N.Y. 2011).

270. *Id.* at 350-52.

found that the common law breach of fiduciary duty and aiding and abetting fraud claims were barred by SLUSA.²⁷¹ In addition, most courts to consider the matter have found that the Martin Act bars Ponzi scheme-related securities fraud claims in the state of New York.²⁷² Even when SLUSA and the Martin Act can be avoided through artful pleading, the result may be a class that is too narrow and fact-specific to provide a substantial recovery for victims as a whole. For instance, in *Levinson v. PSCC Services, Inc.*, two Madoff victims were able to circumvent SLUSA by arguing that two banks had violated their custodial agreements and the Connecticut Unfair Trade Practices Act.²⁷³ This is precisely the sort of piece-meal, inefficient litigation that necessitates a collective solution. Empowering the trustee is more efficient but it does not affect incentives to choose bankruptcy over other forums.

4. Conclusion

The result of this analysis is that, in the circumstances of a Ponzi scheme, both the collective and traditional account for bankruptcy support empowering the trustee. In a corporate liquidation, where no individual has incentives to bring the claim and all benefit equally from the bankruptcy proceeding, it is both more efficient and consistent with policy for the trustee to bring such claims. Moreover, the nature of the scheme renders the fresh start and distributional concerns irrelevant. Finally, concerns about forum shopping and perverse incentives are misplaced; the proceeding is involuntary and the impetus to monitor is undisturbed.

CONCLUSION

There are several ways to empower the trustee. The courts to do so have focused on an expansive interpretation of Section 544(a)(1).²⁷⁴ However, as explained in Part Two, this is inconsistent with the “flavor” of that section.²⁷⁵ Located within the part of the

271. *Id.* at 359–61.

272. *Id.* at 361 (“The vast majority of courts in this district have held that the Martin Act preempts New York state law claims brought by investors seeking to recover losses related to the Madoff scandal.”).

273. *Levinson v. PSCC Servs., Inc.*, No. 3:09-CV-00269, 2010 WL 5477250, at *5 (D. Conn. Dec. 29, 2010).

274. *See Hill v. Gibson Dunn & Crutcher, LLP (In re MS55, Inc.)*, No. 06-CV-01233, 2007 WL 2669150 (D.Colo. Sept. 6, 2007) (allowing Trustee’s § 544 claim for aiding and abetting breach).

275. *See Mixon v. Anderson (In re Ozark Rest. Equip. Co., Inc.)*, 816 F.2d 1222, 1229 (8th Cir. 1987).

Code defining the estate and the avoidance powers, Section 544 is an imperfect fit. There is no transfer to avoid, and, as explained in the analysis of Section 541, only a tenuous link to the estate itself.²⁷⁶ Finally, Congress' rejection of Section 544(c) suggests that Congress did not intend for the trustee to have this power as a part of the avoidance powers.²⁷⁷

Other trustees have sought to locate this power within Section 105(a),²⁷⁸ which broadly empowers the court.²⁷⁹ However, such a construction of Section 105(a) is inconsistent with the usual approach. The courts have construed Section 105(a) to empower the trustee only where it furthers a purpose in a collateral section of the Code.²⁸⁰ In this case, that would necessarily be either Sections 541 or 544, neither of which supports the trustee's argument. Thus Section 105(a) is likewise an undesirable solution.

The last and most desirable option would be to amend Section 704 itself. As explained above, *Caplin* and *Ozark* rejected the trustee's claims in part through an interpretation of Section 544 and in part based on the courts' understanding of the trustee's role.²⁸¹ That role is defined by Section 704.²⁸² As currently written, there is

276. See *supra* notes 86–93 & accompanying text.

277. See *supra* notes 143–45 & accompanying text.

278. See *In re Ozark*, 816 F.2d at 1230.

279. See 11 U.S.C. § 105(a) (2006):

The court may issue any order, process, or judgment that is necessary or appropriate to carry out the provisions of this title. No provision of this title providing for the raising of an issue by a party in interest shall be construed to preclude the court from, sua sponte, taking any action or making any determination necessary or appropriate to enforce or implement court orders or rules, or to prevent an abuse of process.

See also *U.S. v. Energy Res. Co.*, 495 U.S. 545, 549 (1990) (quoting 11 U.S.C. 105(a) (2010)) (“The Code . . . states that bankruptcy courts may ‘issue any order, process, or judgment that is necessary or appropriate to carry out the provisions’ of the Code. Th[is] statutory directive [is] consistent with the traditional understanding that bankruptcy courts, as courts of equity, have broad authority to modify creditor-debtor relationships.”).

280. See *Jamo v. Katahdin Fed. Credit Union*, 283 F.3d 392, 403 (1st Cir. 2002) (“But section 105(a) does not provide bankruptcy courts with a roving writ, much less a free hand. The authority bestowed thereunder may be invoked only if, and to the extent that, the equitable remedy dispensed by the court is necessary to preserve an identifiable right conferred elsewhere in the Bankruptcy Code.”); *United States v. Sutton*, 786 F.2d 1305, 1308 (5th Cir. 1986) (“[Section 105 does not] authorize the bankruptcy courts to create substantive rights that are otherwise unavailable under applicable law, or constitute a roving commission to do equity.”).

281. See *supra* notes 127–47 & accompanying text.

282. See *supra* notes 88–90 & accompanying text.

no doubt that Section 704 does not empower the trustee. Her job is to “collect and reduce to money the property of the estate” and to “investigate the financial affairs of the debtor.”²⁸³ As noted above, the cause of action does not belong to the debtor and *Caplin* construes the “investigate” function to be a reporting requirement.²⁸⁴ However, by amending Section 704 to include the same language as Section 544(c),²⁸⁵ Congress can harmonize the Code and the theoretical role of the trustee.

Faced with the rubble Ponzi schemes leave behind, the case law and the Code deny the trustee the power to sue third parties on behalf of the creditors even when such suits are in the best interests of all creditors and even when such suits are the most efficient way to maximize the estate. The courts and Congress should take stock of whether this outcome makes sense in light of the unique characteristics of a Ponzi scheme liquidation and the principles of why we have a bankruptcy forum in the first place. While a legislative solution would be both the cleanest and clearest solution, it may be unrealistic to hope for such reform in the near future. Left to choose between two evils, manipulation of the Code and an inequitable and theoretically unjustified result, courts and trustees should not shy away from confronting the theoretical basis for a broader role for the trustee.

283. 11 U.S.C. § 704(a) (2006).

284. *See Caplin v. Marine Midland Grace Trust Co.*, 406 U.S. 416, 428 (1972).

285. The proposed § 544(c) would have read:

(c) (1) The trustee may enforce any cause of action that a creditor, a class of creditors, an equity security holder, or a class of equity security holders has against any person, if-

(2) If the trustee brings an action on such cause of action-

(3) A judgment in any such action brought by the trustee binds all creditors or equity security holders that could have brought an action on such cause of action. Any recovery by the trustee, less any expense incurred by the trustee in effecting such recovery, shall be for the benefit only of such creditors or equity security holders.

See Mixon v. Anderson (In re Ozark Rest. Equip. Co., Inc.), 816 F.2d 1222, 1228 n.9 (8th Cir. 1987).

