

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

**VOLUME 66
ISSUE 4**

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 seventieth 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 © 2011 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>

NEW YORK UNIVERSITY ANNUAL SURVEY OF AMERICAN LAW

2010-2011 BOARD OF EDITORS

Editor-in-Chief

DARRYL G. STEIN

*Managing Editors*CLAIRE DECAMPO
JARED ROSCOE*Development Editors*ELISA KAO
MATTHEW M. KARLAN*Note Editors*LINA BENSMAN
JAIME A. MADELL*Executive Article Editors*BRANDON ADONI
MATTHEW D. GIBBS
COLIN GILLESPIE
NICOLLE NONKEN
DANIEL L. PASSESER
JOSH SIMPSON
ROBERT TAYLOR
LERON THUMIM*Senior Articles Editor*

STEPHEN S. MAR

Symposium Editor

ALEXANDRA MITTER

*Article Editors*JONAH C. ABRAMOWITZ
THOMAS ADAMS
ALLA BERENSHTEYN
ALEXANDER S. BIRKHOLOD
MATT BROWN
KATHERINE COMEAU
EVA GARDNER
QAIS GHAFARY
JESSICA GREENBAUM
JASON B. GUMERHAYES HOLDERNESS JR.
ANDREW S. JACOBS
ISABELLE KINSOLVING FARRAR
JOSH LEVY
ED LINTZ
ANDY LUTES
PAUL MILETIC
ALEXANDER F. MINDLIN
MICHAEL MIX
MICHAEL NADLER
PHILIP N. PILMARGIULIA STELLA PREVITI
RYAN G. REIFFERT
SARAH M. SCHOENBACH
ANDREW R. SPERL
JAMES M. THOMPSON
NEIL J. TOOMEY
JASON VINOKUR
TIMES WANG
LENNOX WHITE
MADELYN K. WHITE*Staff Editors*COREY BERMAN
DANIEL BREMMER
ZACHARY BRIERS
ALICE BYOWITZ
EDWARD DECKER
JAMES EBBERSON
BURTON EGGERTSEN
GRETA FAILS
AARON FINCH
MATTHEW GABBARD
STEPHANIE GAMIZ
MATTHEW GIBBONS
ELIJAH GJELTEMA
GEORGE HANG
JEREMY W. HAYSINAYAT ALI HEMANI
IAN HERBERT
ANDERSON HESTON
JESSICA HEYMAN
TIMOTHY-TAYLOR HURLEY
MICHAEL KAHN
FELICITY KOHN
JENNA BASS LEVY
TREVOR MAUCK
MELISSA C. MCGRANE
CHRISTOPHER MATTHEW
MCNERNEY
ROSEMARY MORGAN
EVAN PARZYCH
CASEY PEARLMAN
LAUREN E. PEDLEYVICTORIA PORTNOY
MARTIN RAPHAN
SPENCER RICKS
KATHARINE RODGERS
ARMINE SANAMYAN
MATHEW SCHUTZER
MAXINE SHARAVSKY
ELINA SHEYKH-ZADE
JAKE TRACER
SEAN TYRER
SANDERS WITKOW
BENJAMIN J. WOLFERT
DONNA XU
JAMES YU
DIWEI ZHANG*Faculty Advisor*

HARRY FIRST

SUMMARY OF CONTENTS

ARTICLES

THE COST OF INEQUALITY: SOCIAL DISTANCE, PREDATORY
CONDUCT, AND THE FINANCIAL CRISIS
Raymond H. Brescia 641

DEMOCRACY AT THE CORNER OF FIRST AND FOURTEENTH:
JUDICIAL CAMPAIGN SPENDING AND EQUALITY
James Sample 727

NOTES

PRICE INCLUDES TAX: PROTECTING CONSUMERS FROM TAX-
EXCLUSIVE PRICING
Hayes Holderness 783

REVISING JUDICIAL APPLICATION OF THE SINGLE SUBJECT RULE
TO INITIATIVE PETITIONS
Florin V. Ivan 825

THE POSTER'S PLIGHT: BRINGING THE PUBLIC DISCLOSURE
TORT ONLINE
Jaime A. Madell 895

NEW YORK CITY'S LANDMARKS LAW AND THE RESCISSION
PROCESS
Joachim Beno Steinberg 951

THE COST OF INEQUALITY: SOCIAL DISTANCE, PREDATORY CONDUCT, AND THE FINANCIAL CRISIS

RAYMOND H. BRESCIA*

TABLE OF CONTENTS

I. An Overview of the Mortgage Crisis	647	R
A. How Did We Get Here?	647	R
B. The Present Foreclosure Crisis.....	650	R
II. Inequality and Financial Crises	655	R
A. On Game Theory, Trust and Social Distance.....	655	R
B. Income Inequality, Social Distance and the Foreclosure Crisis.....	669	R
C. Racial Distance and the Foreclosure Crisis	684	R
III. An Alternative Explanation and a Response.....	693	R
A. The CRA: Not Guilty	694	R
B. The Role of Government Sponsored Entities	699	R
IV. Moving Forward: Social Distance and Financial Reform	701	R
A. Social Distance, Dodd-Frank and a New Regulatory Framework	703	R
1. The Bureau of Consumer Financial Protection	705	R
2. Strengthened Fair Lending Enforcement	709	R
3. Reforming Mortgage Laws	710	R

* Assistant Professor of Law, Albany Law School; J.D., Yale Law School (1992); B.A., Fordham University (1989); formerly the Associate Director of the Urban Justice Center in New York City, a Skadden Fellow at The Legal Aid Society of New York, law clerk to the Honorable Constance Baker Motley, and staff attorney at New Haven Legal Assistance Association. I am grateful to the participants in the Albany Law School Scholarship Workshop Series for their helpful comments as I was conceptualizing this piece, especially Timothy D. Lytton and Kirsten Keefe, and those attendees at the mid-year meeting of the Association for American Law Schools in June 2010, where I presented an earlier draft of this piece, who offered helpful comments and critiques, especially Vicki L. Been, Emma Coleman Jordan, and Florence Wagman Roisman. I would also like to thank Thomas Sander of the Saguardo Seminar for comments on an earlier draft. For their invaluable contributions to this project, I would like to thank my research assistants, Natalie Bernardi, Meghan McDonough, and Andrew Martingale; and my indefatigable legal assistant, Fredd Brewer.

4. Weakening Federal Preemption of State Consumer Protection Laws	713	R
B. The Community Reinvestment Modernization Act	716	R
C. Social Capital, Social Distance, Regulatory Reform and Financial Crises	718	R
V. Conclusion	720	R

“An imbalance between rich and poor is the oldest and most fatal ailments of all republics.”

—*Plutarch*¹

“The injunction of Jesus to love others as ourselves is a recognition of self-interest We have to tolerate the inequality as a way to achieving greater prosperity and opportunity for all.”

—*Lord Brian Griffiths, Vice Chairman, Goldman Sachs International*²

The present financial crisis—what is now commonly referred to as “the Great Recession”³—presents, for many, the opportunity to promote a particular economic and socio-political world view about the causes of the crisis. For those on the left, the crisis was the product of a deregulatory philosophy with its roots in Reaganomics, which gained a reckless head of steam during the presidency of George W. Bush.⁴ On the right, many blame government policies as well, but not those that promoted de-regulation. Rather, it was liberal Democrats, like Presidents Carter and Clinton and Representative Barney Frank, who purportedly pushed the banking industry to lend to people of low-income who had no business being homeowners.⁵ In some ways, the financial crisis is like a Rorschach test: what

1. BRUCE JUDSON, *IT COULD HAPPEN HERE: AMERICA ON THE BRINK* 76 (2009) (no citation in original).

2. Simon Clark & Caroline Binham, *Profit ‘Is Not Satanic,’ Barclays CEO Varley Says (Correct)*, BLOOMBERG NEWS (Nov. 3, 2009), http://www.bloomberg.com/apps/news?pid=newsarchive&sid=aGR1F_bjSIZw.

3. See Barry Eichengreen & Kevin H. O’Rourke, *A Tale of Two Depressions*, ADVISOR PERSPECTIVES (Apr. 21, 2009), http://www.advisorperspectives.com/newsletters09/pdfs/A_Tale_of_Two_Depressions.pdf; see also Paul Krugman, *The Great Recession Versus the Great Depression*, N.Y. TIMES, Mar. 20, 2009, <http://krugman.blogs.nytimes.com/2009/03/20/the-great-recession-versus-the-great-depression/>.

4. For the ways that deregulation laid the groundwork for the financial crisis, see SIMON JOHNSON & JAMES KWAK, 13 BANKERS: THE WALL STREET TAKEOVER AND THE NEXT FINANCIAL MELTDOWN 67–72 (2010).

5. For an argument that the government programs were responsible for fueling the financial crisis, see Howard Husock, Op-Ed., *Housing Goals We Can’t Afford*, N.Y. TIMES, Dec. 11, 2008, at A49; see also *Whose [sic] Responsible for Economic Mess?*,

one sees in an ink blot may reveal more about him or her than it does about the picture.

Taking a hard look at some of the economic indicators present in the buildup to the crisis, one such indicator stands out: prior to the crisis, the United States experienced a stunning increase in income inequality.⁶ This increase was reminiscent of a similar increase in income inequality that preceded the great economic crisis of the last century—the Great Depression.⁷ Given this phenomenon, has income inequality itself had an impact on financial markets to such an extent that it exacerbated, or even led to these financial crises?

There are several potential explanations for the connection between rising income inequality and the great strains on the economy that often follow such inequality. Did rising income for certain sectors lead to an ability to use that income to influence policymaking in such a way that favored those sectors? Did such income inequality pressure politicians to promote policies that favored easy access to credit as a way to mollify lower-income constituents who might otherwise grow frustrated with their own stagnating wages in the face of such inequality? These are the types of explanations that some have offered to help explain the link between income inequality and the Great Recession.⁸ This Article offers a third: that both income inequality and racial inequality created greater social distance and that this social distance, in turn, may have led to greater predatory conduct.⁹ This predatory conduct helped to pro-

CITIZEN, May 2009, at 18, *available at* <http://www.scribd.com/doc/15419200/Whose-Responsible-for-Economic-Mess> (arguing that Barney Frank, as well as the Clinton and Carter Administrations, are to blame for the financial crisis).

6. *See infra* Table 3.

7. U.S. CONG. JOINT ECON. COMM., 111TH CONG., INCOME INEQUALITY AND THE GREAT RECESSION 2–3 (2010) (“Income inequality peaked prior to the United States’ two most severe economic crises—the Great Depression and the Great Recession.”).

8. *See, e.g.*, JOHNSON & KWAK, *supra* note 4, at 89–90 (describing increased influence of financial sector on legislative process); *infra* text accompanying notes 83–85 (describing argument concerning use of credit to offset income inequalities).

9. For the most part, this predatory conduct took the form of predatory lending, which has been defined as follows:

[T]he process of engaging in unfair and deceptive lending practices and sales techniques that rely on misrepresentation, threats, unfair pressure, and borrower ignorance. The goal of predatory lending is to coerce or trick homeowners into obtaining loans with interest rates or fees higher than the borrowers’ credit profiles and the market would justify or loans larger than or different from what the borrowers need, want or can afford.

mote the risky lending practices that led to an asset bubble in real estate and left many middle class borrowers saddled with onerous loans tethered to overvalued homes. Once the bubble burst, such borrowers were left with mortgage obligations they could not meet and homes they could no longer afford. Since the financial sector was so heavily tied to the health of the U.S. home mortgage market, the collapse of that market led to an economic “tsunami”¹⁰ that swamped all other economic sectors, leading to double-digit foreclosure and unemployment rates in many states. The information presented here suggests that income and racial inequality may have helped foster an environment where predatory conduct could flourish in the home mortgage market. Such predatory conduct crippled home finances and ultimately spread to all sectors of the economy.

This Article is organized as follows. Part I presents an overview of both the financial crisis and the foreclosure crisis embedded within it. Part II discusses the interplay between social distance and predatory conduct. Part III considers an alternative theory that attempts to explain the connection between income inequality and the present financial crisis. Part IV concludes this Article with an analysis of the financial reform legislation passed by Congress in the summer of 2010 to determine to what extent the reforms passed may address social inequality and social distance. Part IV also provides some thoughts on what social distance theory says about potential policy responses to the present crisis.

The review of the information presented here reveals several compelling findings regarding some of the potential causes of the financial crisis, some of which have not received a great deal of attention to date. First, differences in economic inequality within states correspond to differences in mortgage delinquency rates: i.e., the greater the income inequality in a state, on average, the greater the delinquency rate in that state. Second, the greater the level of generalized trust in a state, the lower that state’s delinquency rate. Third, the higher the social capital in a state, the lower its delinquency rate. Fourth, the higher the median income in a state, the

Kurt Eggert, *Held Up in Due Course: Predatory Lending, Securitization, and the Holder in Due Course Doctrine*, 35 CREIGHTON L. REV. 503, 507 (2002).

10. Alan Greenspan, former Federal Reserve Chairman, testified before a House of Representatives committee that the “current global financial crisis is a ‘once in a century credit tsunami’” *Greenspan Calls Economic Crisis a ‘Credit Tsunami’*, PENNLIVE.COM (Oct. 23, 2008), http://www.pennlive.com/midstate/index.ssf/2008/10/greenspan_calls_economic_crisi.html; see also Habib Siddiqui, NRB Council, USA, *Corporate America Still Doesn’t Get It* (Jan. 17, 2010) (working paper), available at <http://ssrn.com/abstract=1537869>.

higher the delinquency rate in that state. Fifth, an index of a series of indicators—including income inequality within a state, the size of the African American population in a state and the median income of the African American population in that state—reveals a strong correlation between these indicators and delinquency rates. This correlation suggests not that low-income African Americans are to blame for the foreclosure crisis, but rather that middle-class African Americans were targeted for and steered towards loans on unfair terms, precipitating the foreclosures that are now concentrated disproportionately in communities of color.

Some cautionary words regarding the use of data in this Article must precede this analysis. Admittedly, the data and analysis presented here does not constitute a deeply sophisticated and granular evaluation using complex statistical techniques. Rather, the information presented here, and the manner in which it is presented, is meant solely as a “conversation starter,” a way of looking at features of the financial crisis in a light that, hopefully, helps to start a dialogue about not just the forces that may have been at work in the lead up to the crisis, but also those factors that must be taken into account when considering paths out of it. Furthermore, some of the information that I use in this assessment is itself hard to measure: e.g., mortgage delinquency rates and foreclosure rates,¹¹ levels of trust in a community,¹² and levels of social capital in a community.¹³ Additionally, the purpose of using this data is not to prove theorems about cause and effect but to inform a discussion. Others have engaged in more in-depth analysis of some of the issues addressed in this Article, including, for example, whether, in the lead up to the financial crisis, borrowers of color may have been steered into higher cost loans than White borrowers of similar

11. On the challenges to measuring foreclosure and delinquency rates, see *National Delinquency Survey Facts*, MORTGAGE BANKERS ASSOCIATION, 1 (May 2008), <http://www.mbaa.org/files/Research/NDSFactSheet.pdf>. This fact summary of the Association’s survey demonstrates the complexity and challenges in obtaining current and accurate rates, including having to adjust for such things as seasonal trends and intricacies in state foreclosure statutes. *Id.*

12. On the difficulties involved with measuring generalized trust in a community, see Edward L. Glaeser et al., *Measuring Trust*, 115 Q. J. ECON. 811, 811–12 (2000). For a discussion of this topic, see *infra* text accompanying note 99.

13. On measuring social capital, see *Measuring the Dimensions of Social Capital*, WORLD BANK, <http://web.worldbank.org/WBSITE/EXTERNAL/TOPICS/EXTSOCIALDEVELOPMENT/EXTTSOCIALCAPITAL/0,,contentMDK:20305939~menuPK:418220~pagePK:148956~piPK:216618~theSitePK:401015~isCURL:Y,00.html> (last visited July 28, 2010). On the challenges of measuring social capital, see Robert M. Solow, *But Verify—Trust: The Social Virtues and the Creation of Prosperity by Francis Fukuyama*, THE NEW REPUBLIC, Sept. 11, 1995, at 36.

creditworthiness.¹⁴ While this Article attempts to build on that work, its analysis is not intended to establish a causal relationship between any two phenomena. Rather, its purpose is to attempt to draw connections between phenomena and begin a dialogue about potential responses to them.

One of the techniques used most often in analyzing the information presented in this work is to rank U.S. states according to their relation to one another using a range of different characteristics and sources of data: e.g., the levels of generalized trust in different states, median incomes in the states, and crime rates. This Article compares these and other data sets to the mortgage delinquency rates of different states in an attempt to explore the potential connections between the characteristics of the different states and the level of mortgage distress within their borders. Rather than presenting raw data, this Article ranks the states to make the scatter point graphs utilized easier to read visually and to present information that is easier to process and that serves as a better foil for conversation.¹⁵

In these ways, this Article attempts to take an approach similar to that used in other settings. Heather Gerken, in her recent work promoting the use of data analysis to assess the relative success of the machinery of democracy in the United States—i.e., our system of casting votes—takes such an approach.¹⁶ There, she argues for the creation of a “Democracy Index,” one which would help begin a dialogue about election reform and help point the way to accomplishing it. She describes the value of such an approach as follows:

Rather than focusing on necessarily atmospheric judgments about what problems exist, the Index would provide concrete, comparative data on bottom-line results. It would allow us to figure out not just what is happening in a given state or locality, but how its performance compares to similarly situated juris-

14. *See infra* text accompanying notes 140–48.

15. Much of the raw data underlying the rankings is available in the Appendix. Heather Gerken explains the value of rankings as follows: “Rankings get more political traction than the alternatives, precisely because they reduce the data to their simplest form.” HEATHER K. GERKEN, *THE DEMOCRACY INDEX: WHY OUR ELECTION SYSTEM IS FAILING, AND HOW TO FIX IT* 34 (2009) (footnote omitted). Gerken argues that, through ranking, information can be presented in a “highly intuitive, accessible format . . .” *Id.* at 68.

16. *Id.* The recently released “Economic Security Index” is another example of this type of approach. *See* JACOB S. HACKER ET AL., *ECONOMIC SECURITY AT RISK: FINDINGS FROM THE ECONOMIC SECURITY INDEX* (2010), available at <http://www.economicsecurityindex.org/assets/Economic%20Security%20Index%20Full%20Report.pdf> (describing Economic Security Index).

dictions'. It would help us spot, surface, and solve the problems that afflict our system. The Democracy Index would, in short, give us the same diagnostic tool used routinely by corporations and government agencies to figure out what's working and what's not.¹⁷

This Article seeks to apply this approach to the relationship between trust and predatory conduct; the role of social capital in serving as a check on such predatory conduct; and the interplay between race, place and class in the lead up to the financial crisis. The juxtaposition of certain critical pieces of information, as presented in this Article, will hopefully help inform, and perhaps even shape, what should be an ongoing debate about the need for and the contours of financial reform after the passage of the Dodd-Frank bill.

I. AN OVERVIEW OF THE MORTGAGE CRISIS

A. *How Did We Get Here?*

The story of the financial crisis has been told many times, and this Article does not attempt to offer a full recount. Suffice it to say, the most immediate seeds of the crisis were planted in the wake of the collapse of the so-called "dot-com" bubble of the late 1990s. The loss in value in that market sector and the attacks of September 11, 2001 led the Federal Reserve to lower interest rates in an attempt to weaken the recessionary effects of those events.¹⁸ A savings glut in other parts of the world, coupled with lowered interest rates, sent investors looking for higher returns than those they could earn from investing in U.S. Treasury securities.¹⁹ At the same time, innovations in home mortgage finance brought about by the increase in securitization of mortgages created investment vehicles that offered investors what they wanted: high returns from a seemingly safe

17. GERKEN, *supra* note 15, at 59. On the mechanics of and values behind the creation of an effective index, see DAVID ROODMAN, CTR. FOR GLOBAL DEV., BUILDING AND RUNNING AN EFFECTIVE POLICY INDEX: LESSONS FROM THE COMMITMENT TO DEVELOPMENT INDEX (2006), *available at* <http://cgdev.org/content/publications/detail/6661>.

18. Ben S. Bernanke, Chairman, Bd. of Governors of the Fed. Reserve Sys., Monetary Policy and the Housing Bubble, Address at the Annual Meeting of the American Economic Association 2–3 (Jan. 3, 2010), *available at* <http://www.federalreserve.gov/newsevents/speech/bernanke20100103a.htm>.

19. DANIEL GROSS, DUMB MONEY: HOW OUR GREATEST FINANCIAL MINDS BANKRUPTED THE NATION 18–20 (2009).

product.²⁰ Questions about the safety of that product—securitized mortgages—were answered by credit rating agencies that deemed them “AAA” safe, or the safest securities available.²¹

These innovations revolutionized the home mortgage market in the United States by converting the business operations of thousands of home mortgage lenders who transformed their business plan from “lend and hold” to “originate to securitize.”²² Lenders sought out prospective borrowers, not because those lenders could profit from the spread between the interest rate at which they borrowed and the interest rate at which they lent to those customers; rather, their profit came from investment banks that paid those lenders a fee for writing mortgages and selling them to those banks.²³ The model was driven by quantity and not quality.²⁴ As the pool of potential borrowers dried up, new mortgage products were devised that allowed lenders to offer loans through lowered underwriting standards, on terms that few understood and even fewer could afford.²⁵ Exotic adjustable rate mortgages were not to be

20. For an overview of the securitization process and a description of the ways in which mortgage securitization was one of the leading causes of the subprime mortgage crisis, see Kurt Eggert, *The Great Collapse: How Securitization Caused the Subprime Meltdown*, 41 CONN. L. REV. 1257 (2009).

21. See JOHNSON & KWAK, *supra* note 4, at 131; Jason Cox, Judith Faucette & Consuelo Valenzuela Lickstein, *Why Did the Credit Crisis Spread to Global Markets*, in THE E-BOOK ON INTERNATIONAL FINANCE AND DEVELOPMENT Part 5 (Mar. 2010), <http://www.uiowa.edu/ifdebook/ebook2/contents/part5-II.shtml> (“Many mortgage-backed securities received AAA credit ratings, meaning that they were considered among the safest assets on the market.”).

22. As former Securities and Exchange Commissioner Christopher Cox explained before Congress, the “originate to securitize” model was at the heart of the financial crisis because of the risks such an approach encouraged: “When mortgage lending changed from originate-to-hold to originate-to-securitize, an important market discipline was lost. The lenders no longer had to worry about the future losses on the loans, because they had already cashed out.” *The Role of Federal Regulators: Lessons from the Credit Crisis for the Future of Regulation, Hearing Before the H. Comm. on Oversight and Gov’t Reform*, 110th Cong., 26 (2008) (testimony of Christopher Cox, Sec. and Exch. Comm’n).

23. For a description of the new ways that lenders and investment banks could make money through securitization, including the sale of the underlying interest to generate more money as well as the fees associated with that sale, see JOHNSON & KWAK, *supra* note 4, at 76 (“In each case, the revenues available depended on the volume of mortgage-backed securities.”).

24. John Kiff & Paul Mills, *Money for Nothing and Checks for Free: Recent Developments in U.S. Subprime Mortgage Markets* 7 (Int’l Monetary Fund, Working Paper No. 188, 2007), available at <http://www.imf.org/external/pubs/ft/wp/2007/wp07188.pdf>.

25. For an overview of the ways in which lenders lowered underwriting criteria to recruit new customers, see Allan N. Krinsman, *Subprime Mortgage Meltdown*:

feared, however, as lenders promised borrowers the ability to refinance their loans after a short period of time (i.e., before the interest rate adjusted upwards), assuming, of course, that the value of the home securing the mortgage would continue to rise;²⁶ otherwise such an approach would not work. The idea that a single borrower could generate an initial loan and then could refinance that loan, perhaps even multiple times, held out the promise of a steady stream of income for the lender, who would collect a new set of fees with each transaction.²⁷

As home values reached unsustainable heights, fueled by the availability of easy credit,²⁸ lenders began to pull back on the reins, and borrowers could no longer refinance their mortgages.²⁹ As credit dried up, borrowers began to fall behind on their mortgages as their interest rates reset to unaffordable levels through opaque formulas that masked their potential for economic destruction.³⁰ With borrower delinquencies increasing, home prices began to fall and investors grew suspicious of securitized mortgage bonds as investments.³¹ The music stopped, leaving investment banks holding both the securities they maintained in their portfolios and those

How Did It Happen and How Will It End?, J. STRUCTURED FIN., Summer 2007, at 13, 14–15; Souphala Chomsisengphet & Anthony Pennington-Cross, *The Evolution of the Subprime Mortgage Market*, 88 FED. RES. BANK ST. LOUIS REV. 31 (2006); MICHAEL LEWIS, *THE BIG SHORT: INSIDE THE DOOMSDAY MACHINE* 28 (2010).

26. See Jane Birnbaum, *The Affluent, Too, Couldn't Resist Adjustable Rates*, N.Y. TIMES, Mar. 20, 2008, <http://www.nytimes.com/2008/03/20/business/20mortgage.html>. “Today’s ARMs were ‘designed to fail, so you have to refinance It shouldn’t be surprising that values go up and down in this kind of situation. And when you most need to refinance you can’t — the crux of the crunch.” *Id.* (quoting Susan M. Wachter, Professor, Univ. of Pa.).

27. JOHNSON & KWAK, *supra* note 4, at 128; see MARK ZANDI, *FINANCIAL SHOCK: A 360¹/₄ LOOK AT THE SUBPRIME MORTGAGE IMPLOSION, AND HOW TO AVOID THE NEXT FINANCIAL CRISIS* 95 (2009).

28. JOHNSON & KWAK, *supra* note 4, at 130.

29. See Vikas Bajaj & Louise Story, *Mortgage Crisis Spreads Past Subprime Loans*, N.Y. TIMES, Feb. 12, 2008, at A1, *available at* <http://www.nytimes.com/2008/02/12/business/12credit.html?pagewanted=1&r=1> (noting difficulty borrowers have refinancing “[a]s home prices fall and banks tighten lending standards”).

30. Jennifer E. Bethel, Allen Ferrell & Gang Hu, *Law and Economic Issues in Subprime Litigation* 19 (Harvard: John M. Olin Ctr. for Law, Econ., and Bus., Discussion Paper No. 612, 2008), *available at* http://www.law.harvard.edu/programs/olin_center/papers/pdf/Ferrell_et_al_612.pdf.

31. *Wall Street and the Financial Crisis: The Role of High Risk Home Loans, Hearing Before the S. Subcomm. on Investigations* (2010) (statement of Sen. Carl Levin), *available at* <http://levin.senate.gov/newsroom/release.cfm?id=323765> (“[R]elaxed credit guidelines, breakdowns in manual underwriting processes, and inexperienced subprime personnel . . . coupled with a push to increase loan volume and the lack of an automated fraud monitoring tool’ led to deteriorating loans.”).

they had recently constructed for sale to what had become skittish investors.³² With no one to buy these assets, bank balance sheets turned toxic.³³ These toxic assets produced the credit freeze of the fall of 2008,³⁴ which turned the collapse of the subprime mortgage market into a full-blown financial crisis.³⁵ At the heart of the broader financial crisis is a foreclosure crisis, with a national delinquency rate hovering around 10%.³⁶

This well-worn history, however, says little about how economic and social inequality might have helped to precipitate this crisis. What role, if any, did inequality play in the lead up to the foreclosure crisis? For an exploration of this issue, and in an attempt to tease out potential causes of the foreclosure crisis itself, this Article turns next to a detailed overview of the present state of that crisis.

B. *The Present Foreclosure Crisis*

In order to address the role that social inequality may have played in the present financial crisis, an overview of the present state of the foreclosure crisis is in order. Since different states have different foreclosure and delinquency rates, an analysis of these different rates by state may reveal some of the qualities of those states and how those qualities might explain some of the root causes of the foreclosure crisis within those states. One quality analyzed is the extent to which states have different types of social inequality, both

32. See LAWRENCE G. McDONALD & PATRICK ROBINSON, A COLOSSAL FAILURE OF COMMON SENSE: THE INSIDE STORY OF THE COLLAPSE OF LEHMAN BROTHERS 270, 287 (2009).

33. John Parry, *Next Up - Bear Stearns Portfolio Value a Litmus Test for Bonds*, REUTERS (July 2, 2008), <http://www.reuters.com/article/idUSN0148556720080702?pageNumber=1> (“Markets for many of these investments . . . deteriorated further as the outlook for U.S. housing . . . worsened and investors [saw] banks and hedge funds continuing to unload risky assets from balance sheets.”).

34. See Jeannine Aversa, *Banks Borrow Record Amount from Fed*, USA TODAY (Oct. 24, 2008), http://www.usatoday.com/money/economy/2008-10-24-483413484_x.htm (discussing increase in bank borrowing from Fed because of investor flight).

35. See Cox et al., *supra* note 21 (“Since banks did not know what banks held toxic assets . . . it became impossible to tell what banks were going to be able to pay back loans. The credit freeze thus resulted in a drastic reduction in lending Thus commenced the global financial crisis.”). For a detailed description of the height of the financial crisis in September 2008, see generally ANDREW ROSS SORKIN, TOO BIG TO FAIL: THE INSIDE STORY OF HOW WALL STREET AND WASHINGTON FOUGHT TO SAVE THE FINANCIAL SYSTEM — AND THEMSELVES (2009).

36. See Press Release, Mortgage Bankers Ass’n, Delinquencies and Foreclosure Starts Decrease in Latest MBA National Delinquency Survey (Aug. 26, 2010), available at <http://www.mortgagebankers.org/NewsandMedia/PressCenter/73799.htm>.

economic and racial. This analysis may shed light on the extent to which social inequality may have fed into the current foreclosure crisis afflicting community after community throughout the United States.

Any discussion of the foreclosure crisis should begin with a snapshot of the residential real estate market in the U.S. There are approximately seventy-five million owner-occupied residential properties in the United States.³⁷ About 70% of those, roughly fifty-two million properties, have outstanding mortgages on them.³⁸ Of those, roughly one in seven, nearly eight million, are presently in some stage of the foreclosure process, or are at least thirty days delinquent on a mortgage payment.³⁹ Furthermore, one in five mortgaged properties are presently “underwater”—that is, the owner owes more on the mortgage than the property is worth.⁴⁰ The success of efforts to curb the wave of foreclosures has proven elusive. As of the end of November 2010, nearly 550,000 borrowers had entered into permanent modifications through the Obama Administration’s loan modification efforts.⁴¹ The map below shows in graphic form the state of foreclosures across the United States.

37. *Selected Housing Characteristics: 2006-2008*, U.S. CENSUS BUREAU, <http://factfinder.census.gov/> (select “Data Sets”; then “American Community Survey”; then “2006-2008 American Community Survey 3-Year Estimates”; then “Data Profiles”; then “Show Result”) (last visited Dec. 10, 2010).

38. *Id.*

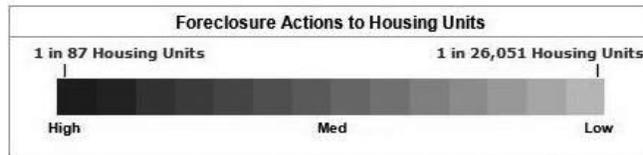
39. E. Scott Reckard, *Foreclosures Will Keep Rising Through 2010, Report Says*, L.A. TIMES, Nov. 20, 2009, at B1, *available at* <http://articles.latimes.com/2009/nov/20/business/la-fi-mortgage-defaults20-2009nov20>.

40. Jonathan Stempel, *One in Five U.S. Mortgage Borrowers are Underwater*, REUTERS, Mar. 4, 2009, <http://www.reuters.com/article/idUSN0349273420090304>. For a discussion of the social norms that prevent borrowers from engaging in “strategic default” when outstanding principal on a mortgage is higher than the value of the underlying property, see Brent T. White, *Underwater and Not Walking Away: Shame, Fear and the Social Management of the Housing Crisis* (Ariz. Legal Studies, Discussion Paper No. 09-35, 2009), *available at* http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1494467.

41. U.S. DEP’T OF THE TREASURY, MAKING HOME AFFORDABLE PROGRAM: SERVICER PERFORMANCE REPORT THROUGH NOVEMBER 2010 (2010), *available at* <http://www.treasury.gov/initiatives/financial-stability/results/MHA-Reports/Documents/Nov%202010%20MHA%20Report.pdf>.

Table 1: U.S. Map by Delinquency Activity⁴²

June 2010 Foreclosure Rate Heat Map



The foreclosure crisis affects not only those saddled with debt they cannot afford. Loss of home value due to foreclosures impacts the prices of homes in a neighborhood considerably, with estimates suggesting that foreclosures reduce the property values of homes in close proximity to the foreclosed property (as far away as one eighth of a mile) from .9% to 1.3% for each foreclosure.⁴³ In 2009 alone, one estimate concluded that U.S. homeowners would lose a

42. *National Real Estate Trends: June 2010 Foreclosure Rate Heat Map*, REALTYTRAC, <http://www.realtytrac.com/trendcenter/default.aspx> (last visited July 28, 2010).

43. A number of studies estimate the costs of foreclosures on neighborhoods and surrounding properties. Immergluck and Smith conducted a study in 2006, examining foreclosures in Chicago from 1997 and 1998, using data originally collected by the Illinois Department of Revenue. See Dan Immergluck & Geoff Smith, *The External Costs of Foreclosure: The Impact of Single-Family Mortgage Foreclosures on Property Values*, 17 HOUSING POL'Y DEBATE 57, 58 (2006). They concluded that “each conventional foreclosure within an eighth of a mile of a single-family home results in a 0.9 percent decline in the value of that home.” *Id.* They also found that, “for the entire city of Chicago, the 3,750 foreclosures that occurred in 1997 and 1998 are estimated to have reduced nearby property values by more than \$598 million, or an average of \$159,000 per foreclosure.” *Id.*

combined \$500 billion in wealth due to the loss of home value as a direct result of neighboring foreclosures in just that year.⁴⁴ Furthermore, it is estimated that before the crisis is over, American homeowners will have lost nearly \$2 trillion in the value of their homes.⁴⁵ Falling home prices reduce the local tax base, starving localities of revenue at a time when they are desperate for funds and when foreclosed and warehoused properties actually require greater municipal services, like police and fire protection.⁴⁶

The foreclosure crisis is adversely impacting certain communities more than others. Certain states, such as Florida, Nevada, Arizona, California, Illinois and Michigan, have been hit the hardest.⁴⁷ Are there ways to explain these differences with reference to macro- or microeconomic trends? Some states, such as California and Florida, experienced a large upswing in property values before property values declined significantly, resulting in an increased number of

A 2008 study by Harding, Rosenblatt, and Yao, using data from thirty-seven Metropolitan Statistical Areas (MSAs) and thirteen states, showed “that foreclosed properties within 300 feet of the subject property create[d] a negative externality effect of approximately 1.3% per distressed property. This contagion discount diminished rapidly with distance and [fell] to .6% at a distance of 500 feet and beyond.” JOHN P. HARDING, ERIC ROSENBLATT & VINCENT W. YAO, THE CONTAGION EFFECT OF FORECLOSED PROPERTIES 14, 20 (2008), *available at* <http://ssrn.com/abstract=1160354>; *see also* John Y. Campbell et al., *Forced Sales and House Prices* (Nat’l Bureau of Econ. Research, Working Paper No. 14866, 2009) (studying foreclosures in Massachusetts and reaching similar findings as Harding et al.), *available at* <http://ssrn.com/abstract=1376188>.

44. CTR. FOR RESPONSIBLE LENDING, SOARING SPILLOVER: ACCELERATING FORECLOSURES TO COST NEIGHBORS \$502 BILLION IN 2009 ALONE; 69.5 MILLION HOMES LOSE \$7,200 ON AVERAGE 1 (2009), *available at* <http://www.responsiblelending.org/mortgage-lending/research-analysis/soaring-spillover-3-09.pdf>. The Center for Responsible Lending (CRL) used the estimate from the 2008 study of Harding et al. that a 0.744% home value decline occurs for each foreclosure within one-eighth of a mile. *Id.* at 2. The national results reported by the CRL were for 56,777 census tracts or similar geographies and only included counties located in MSAs. *Id.* The CRL outlined the foreclosure spillover by state, including data for 2009 foreclosures and anticipated foreclosures from 2009 to 2012. *Id.* at 3.

45. The CRL estimated that “[o]ver the next four years, foreclosures will affect 91.5 million nearby homes, reducing property values \$1.86 trillion in total, or \$20,300 per household.” *Id.* at 2.

46. William C. Apgar & Mark Duda, *Collateral Damage: The Municipal Impact of Today’s Mortgage Foreclosure Boom*, HOMEOWNERSHIP PRESERVATION FOUNDATION, 6–7 (May 11, 2005), http://www.995hope.org/content/pdf/Apgar_Duda_Study_Short_Version.pdf.

47. The delinquency rates for these five states are as follows: Florida (20.43%), Nevada (19.04%), Arizona (13.20%), California (12.49%), Illinois (11.23%), and Michigan (11.13%). MORTGAGE BANKERS ASS’N, NATIONAL DELINQUENCY SURVEY 4 (2010) [hereinafter MBA DATA], *available at* http://media.oregonlive.com/frontporch/other/NDS_Q409.pdf.

underwater properties.⁴⁸ But other high-priced markets that experienced large increases in property values have not had similarly high foreclosure or delinquency rates. For example, in New York City, the first quarter of 2010 saw only 164 foreclosure starts in Manhattan, where home values skyrocketed in the last decade.⁴⁹

Unemployment likely has also played a role in many markets. Unemployment is high in Michigan and Ohio,⁵⁰ where a substantial percentage of both states' residential properties are facing delinquency and entering foreclosure, particularly in cities.⁵¹ But there is also high unemployment in Mississippi and South Carolina, which are not experiencing the same sort of delinquencies and foreclosures as some of the hardest hit states.⁵²

While there is no doubt that price fluctuations and unemployment are driving delinquencies and foreclosures, there are also other forces at work. In the following section, I will address the interplay between income inequality and financial crises. This review reveals that there appears to be a correlation between social inequality and the incidence of foreclosures presently plaguing the states.

48. STANDARD & POOR'S, S&P/CASE-SHILLER HOME PRICE INDICES 2008, A YEAR IN REVIEW 5 (2009), *available at* http://www2.standardandpoors.com/spf/pdf/index/Case-Shiller_Housing_Whitepaper_YearinReview.pdf ("The area traditionally defined as the Sun Belt — Arizona, California, Florida and Nevada — clearly had both the largest run-up in prices since 2000 and has been the hardest hit in the downturn.").

49. Constance Mitchell Ford, *City Foreclosures Climb in First Quarter*, WALL ST. J., May 21, 2010, at A21, *available at* http://online.wsj.com/article/SB10001424052748703559004575256573362821934.html?mod=rss_newyork_main. New York City also experienced a drop in property values. *See* Press Release, Standard and Poor's, Home Prices in the New Year Continue the Trend Set in Late 2009 According to the S&P/Case-Shiller Home Price Indices 3 (Mar. 30, 2010), *available at* http://www.standardandpoors.com/spf/CSHomePrice_Release_033056.pdf. This drop continued into the second quarter. *See* James Comtois, *NY-Area Home Prices Back to 2004 Levels*, CRAIN'S NEW YORK BUSINESS.COM (Aug. 9, 2010, 1:25 PM), http://www.craigslist.com/article/20100809/REAL_ESTATE/100809846.

50. Press Release, U.S. Dep't of Labor: Bureau of Labor Statistics, Regional and State Employment and Unemployment (June 18, 2010) (on file with author). The most recent unemployment numbers are available at <http://www.bls.gov/news.release/laus.nr0.htm>.

51. *See* MBA DATA, *supra* note 47, at 4.

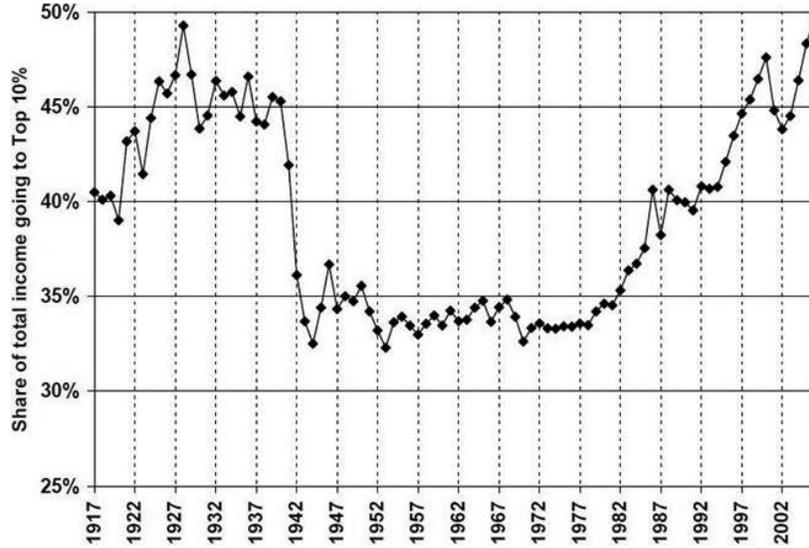
52. *See* U.S. DEP'T OF LABOR: BUREAU OF LABOR STATISTICS, *supra* note 50; MBA DATA, *supra* note 47, at 4.

II. INEQUALITY AND FINANCIAL CRISES

A. *On Game Theory, Trust, and Social Distance*

Table 2 is a timeline of the share of total income enjoyed by the wealthiest 10% of the population in the U.S. over the last ninety years. This share was remarkably high in both the lead up to the Great Depression and the years preceding the Great Recession.

Table 2: Share of Total Income Going to Top 10%⁵³



53. Emmanuel Saez, Econometrics Laboratory Software Archive at the University of California, Berkeley, *Striking it Richer: The Evolution of Top Incomes in the United States (Update with 2007 Estimates)*, 6 fig.1 (Aug. 5, 2009), <http://elsa.berkeley.edu/~saez/saez-UStopincomes-2007.pdf>.

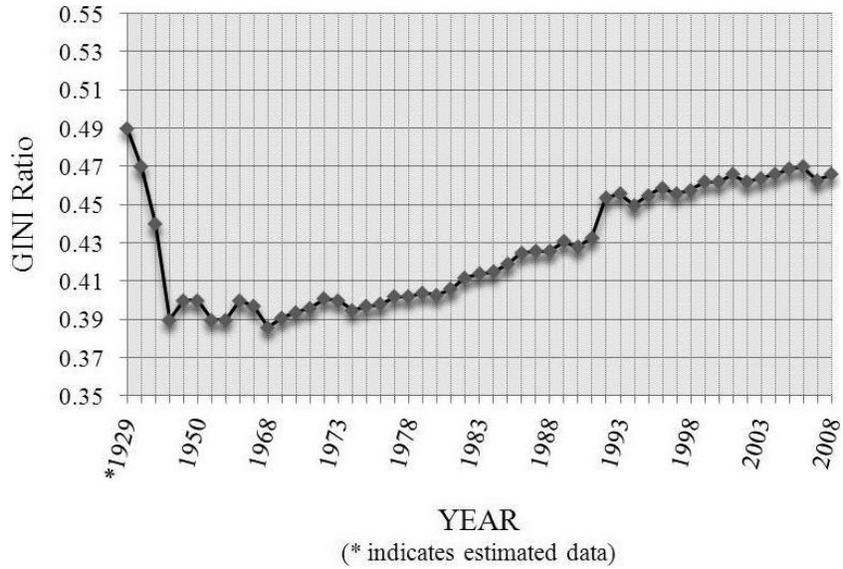
If the past is prologue, one might guess that some event or shock similar to that which occurred near the left hand side of the time line might occur again when this indicator rises again. Similarly, as Table 3 reveals, income inequality, as measured using the GINI Coefficient,⁵⁴ also shows that income inequality has generally increased in the United States since about the early 1970s, reaching levels in the middle of this decade similar to those seen immediately prior to the Great Depression.

54. The U.S. Central Intelligence Agency describes the GINI coefficient as follows:

This index measures the degree of inequality in the distribution of family income in a country. The index is calculated from the Lorenz curve, in which cumulative family income is plotted against the number of families arranged from the poorest to the richest. The index is the ratio of (a) the area between a country's Lorenz curve and the 45 degree helping line to (b) the entire triangular area under the 45 degree line. The more nearly equal a country's income distribution, the closer its Lorenz curve to the 45 degree line and the lower its Gini index, e.g., a Scandinavian country with an index of 25. The more unequal a country's income distribution, the farther its Lorenz curve from the 45 degree line and the higher its Gini index, e.g., a Sub-Saharan country with an index of 50. If income were distributed with perfect equality, the Lorenz curve would coincide with the 45 degree line and the index would be zero; if income were distributed with perfect inequality, the Lorenz curve would coincide with the horizontal axis and the right vertical axis and the index would be 100.

Cent. Intelligence Agency, *The World Factbook: Distribution of Family Income - GINI Index*, THE WORLD FACTBOOK, <https://www.cia.gov/library/publications/the-world-factbook/fields/2172.html> (last visited July 29, 2010). A range of societal ills are linked to having a high GINI coefficient, e.g., poor public health outcomes and civil unrest, to name just two. See U.N. COMM. FOR DEV. POL'Y, IMPLEMENTING THE MILLENNIUM DEVELOPMENT GOALS: HEALTH INEQUALITY AND THE ROLE OF GLOBAL HEALTH PARTNERSHIPS, at 4, 10, U.N. Sales No. E.09.II.A.2 (2009), available at http://www.un.org/esa/policy/devplan/cdppublications/2009cdp_mdghealth.pdf (finding that "when income distribution becomes more unequal, health inequality worsens"); Craig Bradford, *Effects of Globalization on Income Inequality in High Income Countries*, 2 BRYANT ECON. RES. PAPER 7, Spring 2009, at 7, available at <http://web.bryant.edu/~economix/Research%20Papers/Vol%202/Vol%202%20No%202%20Bradford.pdf>. See generally Patrick M. Regan & Daniel Norton, *Greed, Grievance, and Mobilization in Civil Wars*, 49 J. CONFLICT RESOL. 319 (2005), available at <http://jcr.sagepub.com/content/49/3/319.full.pdfhtml> (finding that "[r]esource distribution—as indicated by the extent of political discrimination—is one of the strongest predictors of the onset of violent forms of antistate activity").

Table 3: GINI Timeline⁵⁵



55. For estimated 1929 data, see Eugene Smolensky & Robert Plotnick, *Inequality and Poverty in the United States: 1900 to 1990*, at 6 fig.2, 8–9 (Univ. of Wis.-Madison, Inst. for Res. on Poverty, Discussion Paper No. 998-93, 1993). The authors explain that by utilizing income tax return data from the period, identifying the relationship for 1947 through 1989 (using the concrete GINI data), and then projecting backwards, they were able to obtain the GINI ratios for the first half of the century. *Id.* at 8–9; see also Edward C. Budd, *Introduction to PROBLEMS OF THE MODERN ECONOMY: INEQUALITY AND POVERTY* xiii tbl.1 (Edward C. Budd ed., 1967) (citations omitted). The 1929 data was obtained from a 1958 article which used figures generated by the Brookings Institution. *Id.* at xiii n.2; Selma F. Goldsmith, *The Relation of Census Income Distribution Statistics to Other Income Data*, in 23 NAT'L BUREAU OF ECON. RES., AN APPRAISAL OF THE 1950 CENSUS INCOME DATA 63, 92–94 (1958), available at <http://www.nber.org/chapters/c1050.pdf>. The Institution combined a number of different income statistics for persons and then converted those statistics to a family-unit basis. *Id.* at 94. They also utilized data from federal individual income tax returns. *Id.* Goldsmith adjusted the figures from the Brookings Institution to remove capital gains and losses from her 1929 figure. *Id.* “The adjustments were necessarily rough, but they serve to make the estimates for 1929 more comparable with those for recent years and thereby make it possible to avoid some mistaken conclusions drawn by [those] who compared postwar income distributions directly with the Brookings figures.” *Id.* For 1929 to 1962 data, see Budd, *supra*, at xiii tbl.1. For 1967 to 2008 data, see U.S. CENSUS BUREAU, SELECTED MEASURES OF HOUSEHOLD INCOME DISPERSION: 1967–2008, at 1–2 tbl.A-3, available at <http://www.census.gov/hhes/www/income/data/historical/inequality/IE-1.pdf>; see also Arthur F. Jones Jr. & Daniel H. Weinberg, U.S. CENSUS BUREAU, P60-204, THE CHANGING SHAPE OF THE NATION'S INCOME DISTRIBUTION: 1947–1998 (2000), available at <http://www.census.gov/prod/2000pubs/p60-204.pdf>. The GINI coefficient for 2009 was virtually

Comparing international GINI scores, in 2008 the United States came in relatively high at 0.466 in money income, or 0.451 in equivalence-adjusted income.⁵⁶ That is, the United States has relatively high income inequality, scoring just above Ghana and Turkmenistan, and just below Senegal and Cambodia.⁵⁷

As former Secretary of Labor under President Clinton, Robert Reich, recently explained:

Consider: in 1928 the richest 1 percent of Americans received 23.9 percent of the nation's total income. After that, the share going to the richest 1 percent steadily declined. New Deal reforms, followed by World War II, the GI Bill and the Great Society expanded the circle of prosperity. By the late 1970s the top 1 percent raked in only 8 to 9 percent of America's total annual income. But after that, inequality began to widen again, and income reconcentrated at the top. By 2007 the richest 1 percent were back to where they were in 1928—with 23.5 percent of the total.⁵⁸

Even David Stockman, director of the Office of Management and Budget under President Reagan, in an op-ed piece recently penned for the New York Times, raised concerns about growing income inequality and its impact on markets:

It is not surprising, then, that during the last bubble (from 2002 to 2006) the top 1 percent of Americans — paid mainly

unchanged from 2008. CARMEN DENAVAS-WALT ET AL., U.S. CENSUS BUREAU, P60-238, INCOME, POVERTY AND HEALTH INSURANCE COVERAGE IN THE UNITED STATES: 2009, at 4–5 (2010), available at <http://www.census.gov/prod/2010pubs/p60-238.pdf>.

56. CARMEN DENAVAS-WALT ET AL., U.S. CENSUS BUREAU, P60-236, INCOME, POVERTY, AND HEALTH INSURANCE COVERAGE IN THE UNITED STATES: 2008, at 10 (2009), available at <http://www.census.gov/prod/2009pubs/p60-236.pdf>.

57. See KEVIN WATKINS ET AL., U.N. DEV. PROGRAMME, *Human Development Report 2007/2008*, at 281–83 (2007), available at http://hdr.undp.org/en/media/HDR_20072008_EN_Complete.pdf. The GINI coefficients used by the U.N. reflected GINI data for different years for each country, based on the data available at the time the report was issued. *Id.* The countries were then ranked according to the human development index (a “composite index that measures the average achievements in a country in three basic dimensions of human development: a long and healthy life; access to knowledge; and a decent standard of living”), not by GINI indexes. *Id.* at 225. Using the most recent data available, in 2006 the U.S.'s GINI coefficient actually increased to .464. See BRUCE H. WEBSTER JR. & ALEMAYEHU BISHAW, U.S. CENSUS BUREAU, ACS-08, INCOME, EARNINGS, AND POVERTY DATA FROM THE 2006 AMERICAN COMMUNITY SURVEY 10–11 (2007), available at <http://www.census.gov/prod/2007pubs/acs-08.pdf>. Applying this number to the 2008 report, the U.S. would actually rank 19 spots higher (i.e., worse—with greater income inequality) than presently reflected in the U.N. report.

58. Robert Reich, *Unjust Spoils*, THE NATION, July 19, 2010, available at <http://www.thenation.com/article/36893/unjust-spoils>.

from the Wall Street casino — received two-thirds of the gain in national income, while the bottom 90 percent — mainly dependent on Main Street's shrinking economy — got only 12 percent. This growing wealth gap is not the market's fault. It's the decaying fruit of bad economic policy.⁵⁹

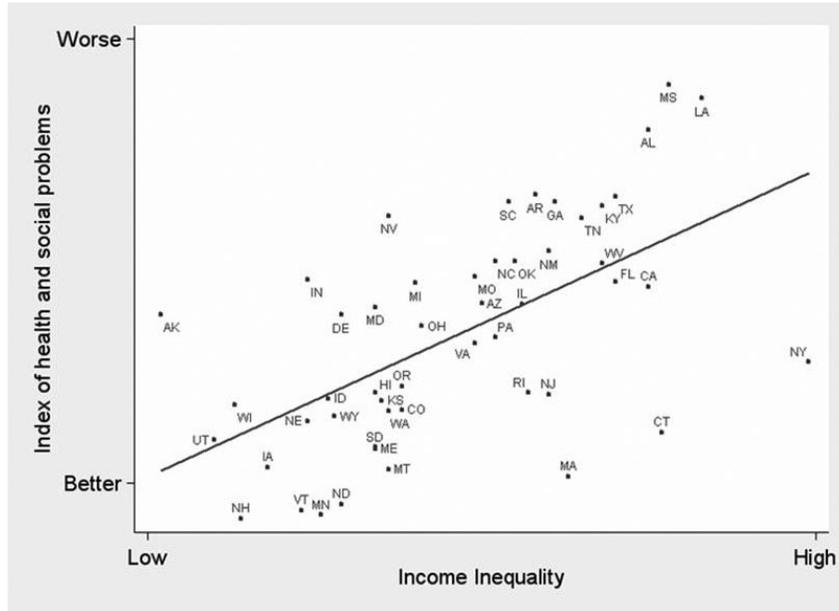
Why concern ourselves with economic inequality, though? Speaking generally, Wilkinson and Pickett, in their recent book *The Spirit Level*, show that, globally, income inequality corresponds strongly with a range of societal ills.⁶⁰ By creating an index of these societal problems, such as crime rates, and measuring them against income inequality, the authors find that nations with higher income inequality have the greatest prevalence of these social ills.⁶¹ State-by-state analysis in the U.S. yields a similar pattern, as Table 4 indicates.

59. David Stockman, *Four Deformations of the Apocalypse*, N.Y. TIMES, July 31, 2010, at WK9, available at <http://www.nytimes.com/2010/08/01/opinion/01stockman.html?pagewanted=1&ref=opinion>.

60. RICHARD WILKINSON & KATE PICKETT, *THE SPIRIT LEVEL: WHY GREATER EQUALITY MAKES SOCIETIES STRONGER* (2009). For a critique of *The Spirit Level*, together with the Authors' response to the critique, see Christopher Snowden, *20 Questions for Richard Wilkinson & Kate Pickett, THE SPIRIT LEVEL DELUSION: FACT-CHECKING THE LEFT'S NEW THEORY OF EVERYTHING* (July 29, 2010), <http://spiritleveldelusion.blogspot.com/2010/04/20-questions-for-richard-wilkinson-kate.html>.

61. This Author makes no claim that the data presented in this graph exposes any causal connection between inequality and the societal problems Wilkinson and Pickett studied and made a part of this index.

Table 4: State-by-State Spirit Level Scores⁶²



A graph with a trend line like this indicates a possible positive correlation between the data represented on the two axes. Here, scoring high on the Wilkinson and Pickett index of social problems correlates with having high income inequality.

But what relationship, if any, does economic inequality have with financial crises? In his seminal work on the Great Depression, Galbraith pointed to the severe income inequality in the United States—what he called the “bad distribution of income”—as one of the five “weaknesses” in the economy that “had an especially intimate bearing on the ensuing disaster.”⁶³

In 1929 the rich were indubitably rich. The figures are not entirely satisfactory, but it seems certain that the 5 per cent of the population with the highest incomes in that year received approximately one third of all personal income. The proportion of personal income received in the form of interest, dividends, and rent — the income, broadly speaking, of the well-to-do — was about twice as great as in the years following the Second World War.⁶⁴

62. WILKINSON & PICKETT, *supra* note 60, at 22 fig.2.4.

63. JOHN KENNETH GALBRAITH, THE GREAT CRASH 1929, at 177 (1997).

64. *Id.* Galbraith explained the impact of this imbalance on consumer spending as follows:

Another view on income inequality comes from Bruce Judson of the Yale School of Management. According to Judson, rising income inequality creates a “governance problem.”⁶⁵ This governance problem begins with the wealthy developing a “sense of entitlement” and insulating themselves from society.⁶⁶ They then:

become less dependent on public services and less connected to the concerns of the rest of society [T]his leads the bulk of those in the top income strata to oppose tax increases that would fund enhanced public amenities [T]hey use their wealth to obtain political influence that solidifies their privileges.⁶⁷

Teddy Roosevelt, discussing the need for an estate tax, described the corrosive effect of income inequality on the political process as follows:

The absence of effective State, and, especially, national, restraint upon unfair money-getting has tended to create a small class of enormously wealthy and economically powerful men, whose chief object is to hold and increase their power. The prime need is to change the conditions which enable these men to accumulate power which is not for the general welfare that they should hold or exercise.⁶⁸

Turning to the present financial crisis, as a description of the forces at work in the lead up to the financial crisis, Judson’s theory—and Roosevelt’s fears—would seem to apply. If one traces the recent rise in income inequality in the U.S. over the last forty years, it corresponds with two phenomena: (1) a greater share of the rising income of the wealthy going to the financial sector, and (2) a

This highly unequal income distribution meant that the economy was dependent on a high level of investment or a high level of luxury consumer spending or both. The rich cannot buy great quantities of bread. If they are to dispose of what they receive it must be on luxuries or by way of investment in new plants and new projects. Both investment and luxury spending are subject, inevitably, to more erratic influences and to wider fluctuations than the bread and rent outlays of the \$25-a-week workman. This high-bracket spending and investment was especially susceptible, one may assume, to the crushing news from the stock market in October of 1929.

Id. at 177–78. For a further discussion of the economic distortions caused by the concentration wealth among top earners, see ROBERT B. REICH, *AFTERSHOCK: THE NEXT ECONOMY AND AMERICA’S FUTURE* 32–38 (2010).

65. BRUCE JUDSON, *IT COULD HAPPEN HERE: AMERICA ON THE BRINK* 77 (2009).

66. *Id.*

67. *Id.*

68. Theodore Roosevelt, *The New Nationalism*, Speech at John Brown Memorial Park (Aug. 31, 1910), available at <http://www.theodore-roosevelt.com/images/research/speeches/trnationalismspeech.pdf>.

push towards deregulation of the financial sector in the service of greater profits. As stated earlier, rising income inequality in the United States commenced in the late 1970s and a disproportionate share of that income was ceded to households earning their income in the financial sector. In 1978, average banker compensation was on par with pay in the private sector overall. By 2007, it was more than twice that of the average private sector employee.⁶⁹ Moreover, since the 1970s, a disproportionate share of profits, in relation to gross domestic product, has gone to the financial sector.⁷⁰

This rising share of income for the financial sector corresponds with the era of financial deregulation that began in the 1980s. That era brought about laws and policies that accomplished a range of “innovations”: lifting interest rate caps on mortgages, which ushered in the creation of subprime mortgage products;⁷¹ permitting investment banks to take on more debt, lowering margin requirements;⁷² preempting state anti-predatory lending laws that allowed a wide range of banking entities to operate with light federal regulation despite state efforts to rein in risky lending;⁷³ lowering the wall between investment banks and commercial banks through the repeal of the Glass-Steagall Act;⁷⁴ and prohibiting the regulation of derivatives through passage of the Commodity Fu-

69. JOHNSON & KWAK, *supra* note 4, at 60–61 (citation omitted).

70. John Bellamy Foster & Fred Magdoff, *Financial Implosion and Stagnation: Back to the Real Economy*, MONTHLY REV., Dec. 2008, at 12 chart 2 (citation omitted), available at <http://www.monthlyreview.org/081201foster-magdoff.php>.

71. Depository Institutions Deregulation and Monetary Control Act of 1980, 12 U.S.C. § 226 (2006).

72. For a description of the SEC’s efforts permitting certain investment banks to increase their leverage, see Stephen Labaton, *Agency’s ‘04 Rule Let Banks Pile Up New Debt*, N.Y. TIMES, Oct. 3, 2008, at A1, available at <http://www.nytimes.com/2008/10/03/business/03sec.html>; Joseph E. Stieglitz, *Capitalist Fools*, VANITY FAIR, Jan. 2009, at 48. For a description of the risks of increased leverage on the financial system, see Eric S. Rosengren, President & CEO Fed. Reserve Bank of Bos., *Could a Systemic Regulator Have Seen the Current Crisis?*, Before the 10th Seoul International Financial Forum (Apr. 5, 2009), available at <http://www.bos.frb.org/news/speeches/rosengren/2009/041509.pdf>.

73. For a description of federal bank regulators’ efforts to preempt state anti-predatory lending laws, see Julia Patterson Forrester, *Still Mortgaging the American Dream: Predatory Lending, Preemption, and Federally Supported Lenders*, 74 U. CIN. L. REV. 1303, 1339–49 (2006); Christopher L. Peterson, *Federalism and Predatory Lending: Unmasking the Deregulatory Agenda*, 78 TEMP. L. REV. 1, 68–84 (2005).

74. On the connection between the repeal of the Glass-Steagall Act and the financial crisis, see Joseph Karl Grant, *What the Financial Services Industry Puts Together Let No Person Put Asunder: How the Gramm-Leach-Bliley Act Contributed to the 2008-2009 American Capital Markets Crisis*, 73 ALB. L. REV. 371 (2010).

tures Modernization Act.⁷⁵ In the words of Simon Johnson and James Kwak: “The Wall Street banks are the new American oligarchy—a group that gains political power because of its economic power, and then uses that political power for its own benefit.”⁷⁶ These commentators describe the power of the financial sector due to its growing share of income as follows:

The unprecedented amounts of money flowing through the financial sector, increasingly concentrated in a handful of megabanks, were the foundation of the new financial oligarchy Wall Street used an arsenal of other, completely legal weapons in its rise to power. The first was traditional capital: money, which yielded its influence directly via campaign contributions and lobbying expenses. The second was human capital: the Wall Street veterans who came to Washington to shape government policy and shape a new generation of civil servants. The third, and perhaps most important, was cultural capital: the spread and ultimate victory of the idea that a large, sophisticated financial sector is good for America.⁷⁷

And the banks did not stop throwing their financial weight around once the financial crisis hit; to the contrary, their lobbying reached a fevered pitch. According to one study, the six largest banks and their trade associations spent nearly \$600 million lobbying Congress on financial reform.⁷⁸ Moreover, some companies, like AIG, continued to devote resources to lobbying even after receiving federal bailout funds,⁷⁹ and some, like Bank of America,

75. Commodity Futures Modernization Act of 2000, 7 U.S.C. §§ 1-27f (2006). For a description of the passage of the CFMA, see JUDSON & KWAK, *supra* note 4, at 9. For a description of the role of credit default swaps, a form of derivative, in promoting risky lending practices during the build up to the financial crisis, see RICHARD A. POSNER, *A FAILURE OF CAPITALISM: THE CRISIS OF '08 AND THE DESCENT INTO DEPRESSION* 56–60 (2009).

76. JOHNSON & KWAK, *supra* note 4, at 6.

77. *Id.* at 89–90. For a further description of how this “governance problem” may have played out in the transformation of the U.S. economy over the last forty years, see JACOB S. HACKER & PAUL PIERSON, *WINNER-TAKE-ALL POLITICS: HOW WASHINGTON MADE THE RICH RICHER—AND TURNED ITS BACK ON THE MIDDLE CLASS* 66–72 (2010).

78. KEVIN CONNOR, INST. FOR AMERICAN'S FUTURE, *BIG BANK TAKEOVER: HOW TOO-BIG-TO-FAIL'S ARMY OF LOBBYISTS HAS CAPTURED WASHINGTON* 5 (2010), *available at* <http://www.ourfuture.org/files/documents/big-bank-takeover-final.pdf>.

79. Elizabeth Williamson, *AIG Still Lobbies to Relax Oversight Rules: After Receiving Federal Aid, Insurer Focuses on Laws Aimed at Keeping Tabs on Mortgage Originators*, WALL ST. J., Oct. 16, 2008, at A6, *available at* http://online.wsj.com/article_email/SB122409634988737247-1MyQjAxMDI4MjE0NjAxOTY2Wj.html.

actually increased such expenditures after receiving bailout assistance from the federal government.⁸⁰

It is hard to argue that a “governance problem” does not lie at the heart of the connection between financial sector influence, deregulation, and the financial crisis. Lifting many of the constraints on the financial sector created a shadow banking system that operated without a net (at least one that was not taxpayer financed), taking on greater and greater risk, and operating with little oversight.⁸¹ The practices that deregulation authorized—excessive leverage, an unregulated market in derivatives, exotic mortgage products—were at the center of an interconnected and overextended financial sector, a sector that ultimately collapsed under its own weight.⁸² While the governance problem seems to explain a great deal, it may not be the only potential connection between rising income inequality and this financial crisis.

An alternate theory that may explain this connection comes from former International Monetary Fund economist, now at the University of Chicago, Raghuram Rajan, who posits that growing economic inequality in the United States led to the present financial crisis because politicians used easy access to credit to mollify people of low income, in an effort offset the possibility of discontent that might arise from growing inequality.⁸³ He argues that this

80. Paul Kiel, *Bailed-Out Companies Spend Millions to Lobby Congress*, PROPUBLICA (July 22, 2009), <http://www.propublica.org/article/bailed-out-companies-spend-millions-to-lobby-congress>.

81. For a discussion of the interplay between de-regulation and the financial crisis, see Todd J. Zywicki & Joseph D. Adamson, *The Law and Economics of Subprime Lending*, 80 U. COLO. L. REV. 1, 5–7 (2009); Sally Pittman, Comment, *ARMs, But No Legs to Stand on: “Subprime” Solutions Plague the Subprime Mortgage Crisis*, 40 TEX. TECH L. REV. 1089, 1093–94 (2008); Patricia A. McCoy & Elizabeth Renuart, *The Legal Infrastructure of Subprime and Nontraditional Home Mortgages*, in *BORROWING TO LIVE: CONSUMER AND MORTGAGE CREDIT REVISITED* 110 (Nicolas P. Retsinas & Eric S. Belsky eds., 2008).

82. See Luis A. Aguilar, Commissioner, U.S. Sec. & Exch. Comm’n, Speech at Compliance Week 2010: Market Upheaval and Investor Harm Should Not be the New Normal (May 24, 2010), *available at* <http://www.sec.gov/news/speech/2010/spch052410laa-1.htm> (stating that “[t]he financial crisis and its enormous costs to society were the direct result of years of deregulation This market breakdown and the difficulty in determining how and why it occurred are yet further stark reminders of the dangers of weak oversight of our tightly interconnected financial markets”).

83. RAGHURAM RAJAN, *FAULT LINES: HOW HIDDEN FRACTURES STILL THREATEN THE WORLD ECONOMY* (2010). Rajan explains this phenomenon as follows: because it is hard to impact income disparities:

[P]oliticians have looked, or been steered into looking, for other, quicker ways to mollify their constituents. We have long understood that it is not in-

push from elected officials is reflected in such laws as the Community Reinvestment Act (CRA)⁸⁴ and the behavior of the Government-Sponsored Entities (GSEs), Fannie Mae and Freddie Mac. According to Rajan, the CRA and the affordable housing goals of the GSEs—promoted by Congress, and the Clinton and Bush Administrations—forced banks to make risky loans to unworthy borrowers.⁸⁵ Similarly, the prevalence of risky lending in communities of color, as discussed below, has led some to blame those same communities for the financial crisis.⁸⁶ This Article will return to these and other arguments in Part III, *infra*. In that section, a review of the hard facts on these issues—for example, that 94% of the riskiest lending carried out during the inflation of the housing bubble took place beyond the reach of the CRA—renders such arguments unconvincing.

A third potential explanation of the connection between financial crises and social inequality is that social inequality increases social distance. And increased social distance reduces trust within a community, which leads to non-cooperative and predatory conduct.⁸⁷ Adam Smith,⁸⁸ John Stuart Mill,⁸⁹ Kenneth Arrow,⁹⁰ Amartya Sen,⁹¹ and Niall Ferguson⁹² extol the importance of trust

come that matters but consumption. Stripped to its essentials, the argument is that if somehow the consumption of middle-class householders keeps up, if they can afford a new car every few years and the occasional exotic holiday, perhaps they will pay less attention to their stagnant monthly paychecks.

Id. at 8–9.

84. 12 U.S.C. §§ 2901–08 (2006).

85. RAJAN, *supra* note 83, at 32–45. For a description of the policies of the Clinton and Bush Administrations in promoting so-called “affordable housing goals” through the GSEs, see ALYSSA KATZ, *OUR LOT: HOW REAL ESTATE CAME TO OWN US* (2009).

86. See discussion *infra* Part II.C.

87. See *infra* Part II.B.

88. ADAM SMITH, *AN INQUIRY INTO THE NATURE AND CAUSES OF THE WEALTH OF NATIONS* 292 (R.H. Campbell & A.S. Skinner eds., Oxford Univ. Press 1976) (describing importance of trust in extensions of credit and bank transactions).

89. JOHN STUART MILL, *PRINCIPLES OF POLITICAL ECONOMY* 111 (W.J. Ashley ed., new ed. 1909) (stressing importance of trust in “conjoint” human action).

90. KENNETH J. ARROW, *Gifts and Exchanges*, 1 *PHIL. & PUB. AFFS.* 343, 357 (1972).

91. Amartya Sen, *Adam Smith’s Market Never Stood Alone*, *FIN. TIMES*, Mar. 11, 2009, available at <http://www.ft.com/cms/s/0/f8e0b1be-0ddc-11de-8ea3-0000779fd2ac.html> (defending Adam Smith as offering a more nuanced view of capitalism than those that assert he stood solely for unfettered free market, and affirming that “an economy needs other values and commitments such as mutual trust and confidence to work efficiently”).

92. NIALL FERGUSON, *THE ASCENT OF MONEY: A FINANCIAL HISTORY OF THE WORLD* 29–30 (2008) (describing money as “trust inscribed”).

in economic exchanges. It is difficult to engage in any economic activity without trusting, at least to some degree, a counterparty's willingness to honor their contracts; the currency system—that the money changing hands is valid and redeemable; the financial system—that the method of payment can convey value; and the courts and the legal system—that they will police and punish breaches of trust. As Arrow posits, “Virtually every commercial transaction” has an element of trust in it.⁹³

Since trust is at the heart of cooperative economic exchange, does a lack of trust diminish the capacity for trusting, and raise the prospect of untrustworthy behavior? Furthermore, what are some of the forces that might diminish trust, and lead to a lack of trustworthiness? Research resulting from the prisoner's dilemma and similar games, as well as field studies, reveals that where there is greater social distance between game participants, participants are less trusting and each is more willing to take advantage of other participants.⁹⁴ For example, when studying rubber traders in Singapore and Malaysia in the 1960s, Janet Landa unearthed a complex hierarchy of social relations among these traders.⁹⁵ The extent to which a trader trusted another trader depended on the level of social distance between the two of them; those closest in kinship were trusted more, and as social distance increased, trust decreased.⁹⁶

93. Arrow, *supra* note 90, at 357.

94. See, e.g., Elizabeth Hoffman et al., *Social Distance and Other-Regarding Behavior in Dictator Games*, 86 AM. ECON. REV. 653, 658 (1996) (finding increase in non-cooperative behavior as social distance widens). Summarizing a body of research on group dynamics, Elizabeth Chamblee Burch describes conduct within groups as follows:

Group members exhibit other-regarding preferences—trust, reciprocity, and altruism—toward other members. Their fairness considerations change based on whether the situation involves another group member (inclusionary concerns) or individuals outside the group (exclusionary concerns). Cohesive group members are more likely to cooperate with one another and care about the collective outcome, and less likely to exit the group when doing so benefits the individual rather than the group. These theory-based insights suggest that group membership plays a pivotal role in attitude changes, particularly when group identity is salient and relevant to the attitudinal issue. Shared histories, implicit feedback, and trust, for example, offer insights about whether individuals will cooperate or defect.

Elizabeth Chamblee Burch, *Litigating Groups*, 61 ALA. L. REV. 1, 17 (2009) (citations omitted).

95. Janet T. Landa, *A Theory of the Ethnically Homogeneous Middleman Group: An Institutional Alternative to Contract Law*, 10 J. LEGAL STUD. 349, 350 (1981).

96. *Id.* at 352.

Like it or not, individuals trust those who they perceive to be similar, whom they perceive as being members of the same group.⁹⁷ At the same time, we are more prone to take advantage of individuals who are less like us, whom we perceive to not share certain characteristics as ourselves.⁹⁸

Since trust is critical to economic exchange, and a lack of trust can make such exchange more difficult, is there a way to measure the relative trust within a society or community? One commonly accepted way to measure the presence of trust in a community is to review responses to the General Values Survey, which asks, among many other questions, whether “most people can be trusted.”⁹⁹ Looking at responses to this question, in the U.S., generalized trust has declined in recent years as income inequality has increased, as Table 5 indicates.

97. See studies cited *infra* note 98. The findings of the studies support the proposition “that levels of cooperation tend to be higher the stronger the ties between actors in the exchange.” Nancy R. Buchan et al., *Swift Neighbors and Persistent Strangers: A Cross-Cultural Investigation of Trust and Reciprocity in Social Exchange*, 108 AM. J. SOC. 168, 200 (2002) (citation omitted).

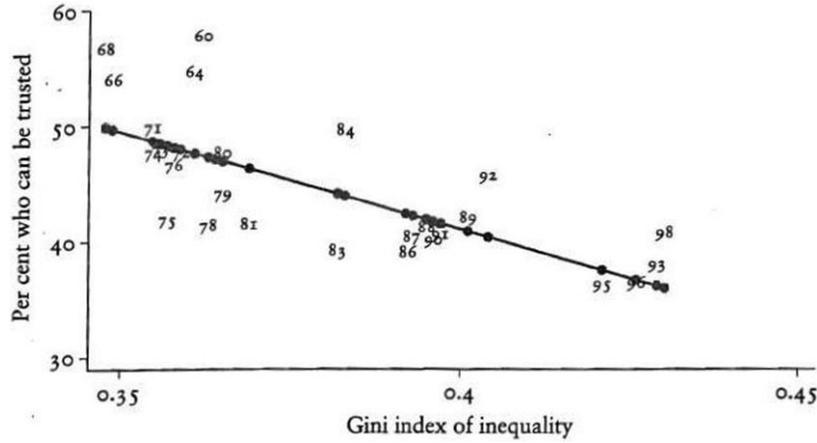
98. For a discussion of the relationship between social distance and cooperation, see Buchan et al., *supra* note 97; Iris Bohnet & Bruno S. Frey, *The Sound of Silence in Prisoner's Dilemma and Dictator Games*, 38 J. ECON. BEHAV. & ORG. 43 (1999); Michael W. Macy & John Skvoretz, *The Evolution of Trust and Cooperation Between Strangers: A Computational Model*, 63 AM. SOC. REV. 638 (1998). As one commentator remarks:

Because we actively, and sometimes unconsciously, participate in the preservation of our perceptions and preferences, situations of great conflict and social distance are especially troublesome. We perceive our enemies to be evil, distant, strange, unapproachable, unfamiliar, distasteful, and unknowable. Moreover, we actively resist any evidence to the contrary.

David Sally, *Game Theory Behaves*, 87 MARQ. L. REV. 783, 791 (2004) (citation omitted).

99. See C. MÓNICA CAPRA ET AL., ATTITUDINAL AND BEHAVIORAL MEASURES OF TRUST: A NEW COMPARISON I (2008), available at <https://www.gate.cnrs.fr/IMG/pdf/Capra.pdf>. The finding of their study “supports the continued use of the GSS . . . questions to measure trust for policy purposes . . .” *Id.* at 2.

Table 5: On Income Inequality and Trust¹⁰⁰



The numbers on this chart indicate the level of trust in the U.S. in a given year in the late twentieth century, together with income inequality in those same years. Here we see a negative correlation: as income inequality has increased in the United States over the last few decades, trust has declined.

If we believe inequality and social distance reduce cooperative behavior and encourage rent-seeking, to what extent might the financial crisis bear such a theory out? Turning first to economic inequality, the theory suggests there would be less cooperative behavior and higher incidents of predatory conduct in communities in which there is higher income inequality.¹⁰¹ The following section addresses these questions.

100. ERIC M. USLANER, THE MORAL FOUNDATIONS OF TRUST 187 fig.6-6 (2002). The methodology used in creating this graph was fairly straightforward, comparing the level of trust in the United States in a given year to the level of inequality in that year, revealing an apparent correlation between the two. *Id.* Without more, no claim that one causes the other is made by Uslander or this Author.

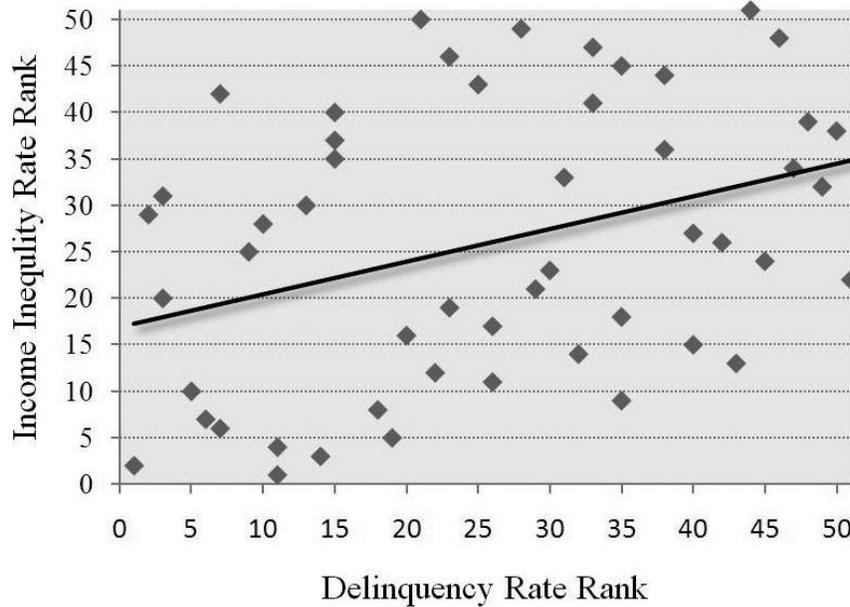
101. There are obviously many bases for social distance. This paper focuses on the ways that two such bases appear to have played out in the lead up to the present financial crisis: social distance based on income and that based on race. For a discussion of gender and subprime lending, see NAT'L COUNCIL OF NEGRO WOMEN, INCOME IS NO SHIELD, PART III, ASSESSING THE DOUBLE BURDEN: EXAMINING RACIAL AND GENDER DISPARITIES IN MORTGAGE LENDING (2009), available at <http://www.nrc.org/images/stories/pdf/research/nrc%20nosheild%20june%2009.pdf>.

B. Income Inequality, Social Distance and the Foreclosure Crisis

If delinquency rates are one of the key symptoms of the financial crisis and income inequality may lie at the heart of the financial crisis, does a community’s level of economic inequality correspond to its delinquency rate?

As Table 6 reveals, there appears to be some correlation between income inequality by state at the height of the subprime mortgage frenzy—2006, the last year in which such information is available from the U.S. Census Bureau—and present delinquency rates by state.

Table 6: Income Inequality and Delinquency Rates by State¹⁰²



102. Tables 6 through 9, 12, 13, and 15 were generated by the author. For the income inequality data used, see WEBSTER JR. & BISHAW, *supra* note 57, at 11 tbl.5. For the delinquency rate data, see MBA DATA, *supra* note 47, at 4. The measurement of the statistical significance of this and the other author-generated tables, the “P-Value” of each such table, is set forth in *infra* Appendix, Data Set 3. It is generally accepted that a P-Value that is below .05, which means there is no more than a 5% chance that a particular outcome could happen by chance, reveals statistical significance. With this data set, which compares each state’s delinquency rate rank to the level of inequality within each state, reveals a P-Value of .011. *Id.* One of the reasons for requiring such a low P-Value is to make up for sampling error, which is not an issue with this data set, because no sampling was done: the universe of state-by-state data was used, which might suggest that one could be

This graph shows the rank of states in terms of their ranking on the GINI index, lowest income inequality to highest, together with their rank in terms of their delinquency rates (combining properties in foreclosure with properties where the borrower is more than ninety days past due on his or her mortgage) as of December 2009. This graph shows that there is a statistically significant correlation between income inequality and delinquency rates.

Using 2006 data on income inequality and 2009 foreclosure data actually makes some sense. Many of the problems in the mortgage market stem from activities in 2005 and 2006, when underwriting standards were loosened to maintain mortgage volume to feed the securitization market.¹⁰³ Furthermore, subprime adjustable rate mortgages written in 2006 often had two- or three-year teaser rates so that the monthly payments on mortgages written in those years likely did not increase dramatically until 2008 or 2009, when the foreclosure crisis started to hit the hardest.¹⁰⁴

Admittedly, many other forces might also be driving this correlation between income inequality and delinquency rates. It would be easy to posit that the higher the number of people living in poverty, the higher the delinquency rate is likely to be. But the facts do not support this theory.

Table 7 suggests that the poverty rate in a given state may have little influence on delinquency rates. A higher poverty rate appears to have little correlation with delinquency rates.¹⁰⁵

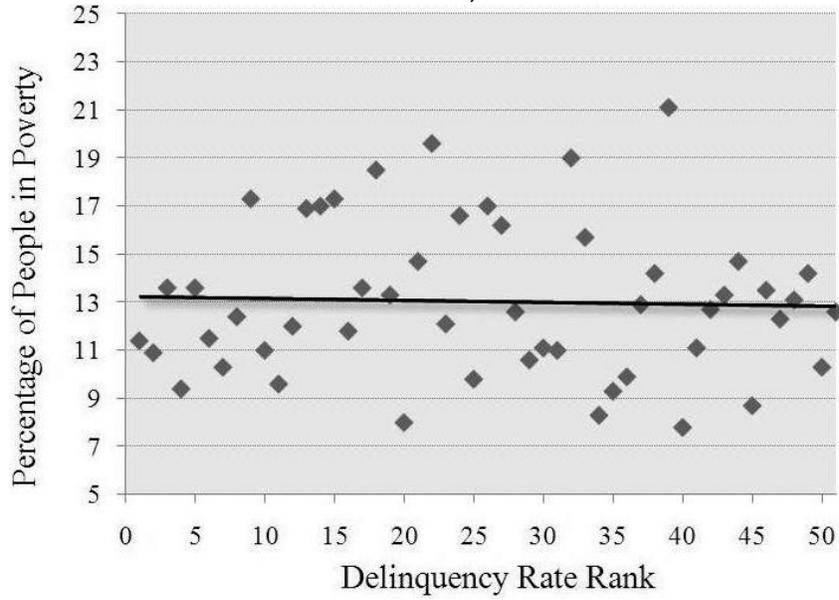
more comfortable with the statistical significance of a data set with a higher P-Value in this setting. On the importance of P-Value and its value in settings where sampling is used, see Michael D. Green, *Expert Witnesses and Sufficiency of Evidence in Toxic Substances Litigation: The Legacy of Agent Orange and Bendectin Litigation*, 86 Nw. U. L. REV. 643, 682–86 (1992).

103. See OFFICE OF THE COMPTROLLER OF THE CURRENCY, SURVEY OF CREDIT UNDERWRITING PRACTICES 4 (2006), available at <http://www.occ.gov/publications/publications-by-type/survey-credit-underwriting/pub-survey-cred-under-2006.pdf>.

104. See *A Second Mortgage Disaster on the Horizon?*, CBS NEWS (Dec. 14, 2008), <http://www.cbsnews.com/stories/2008/12/12/60minutes/main4666112.shtml>. At the end of 2008, the “loans made back in the heyday [started] to reset, causing the mortgage payments to go up and homeowners to default.” *Id.*

105. Indeed, the P-Value for this graph is quite high (.8), revealing no meaningful statistical significance. See *infra* Appendix, Data Set 3. Since the poverty rate data and the median income data revealed in this and the next table tend to create a “spread” between the highest and lowest states that can be graphically displayed in a way that is similar to a ranking, no ranking was used in this table or the following one, which displays median income within each state.

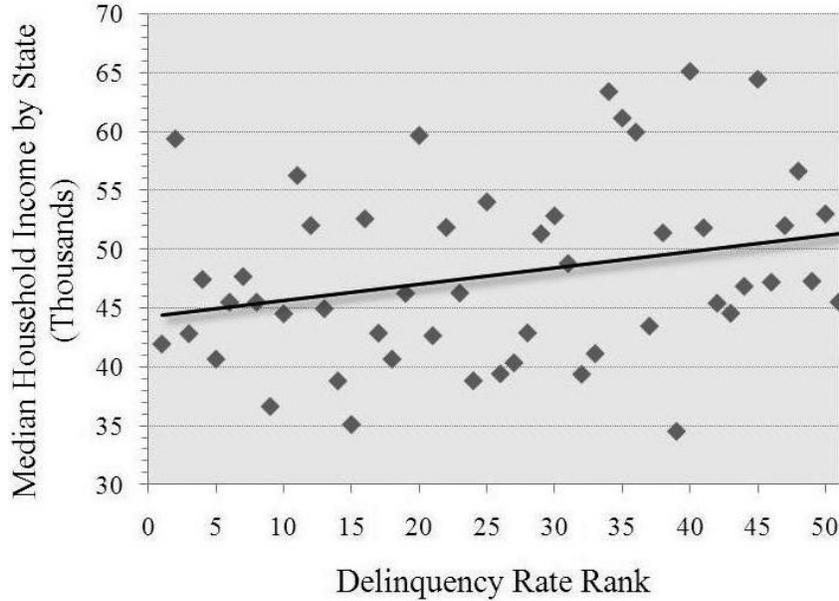
Table 7: Poverty Rate¹⁰⁶



Digging deeper into the delinquency data, however, reveals an interesting phenomenon about the impact of the foreclosure crisis; the higher the median income in a state, the higher the delinquency rate in that state, as reflected in Table 8.

106. For the data on poverty rates by state, see WEBSTER JR. & BISHAW, *supra* note 57, at 21 tbl.9. For the delinquency data used, see MBA DATA, *supra* note 47, at 4.

Table 8: Median Income by State¹⁰⁷



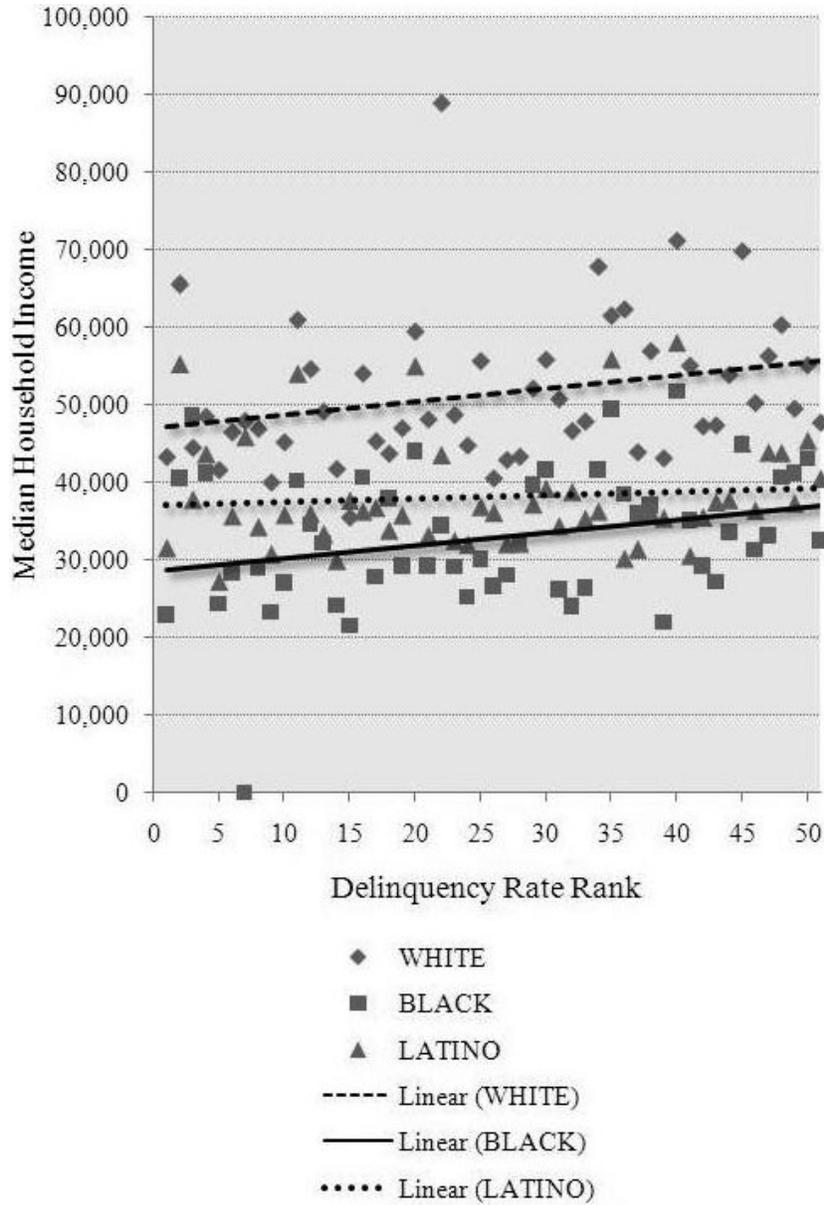
It would be easy to say that it is income inequality that is driving this correlation, but median income in a state does not dictate income inequality. There are states with high income inequality but low median income, like Louisiana, and states where the opposite is true, like New Hampshire, where residents enjoy a high median income while the state as a whole ranks low on the income inequality scale.¹⁰⁸

Interestingly, when we look at delinquency rates compared to median income by race, we actually see that for Whites, African Americans, and to a lesser extent, Latinos, the higher the median income by state for each racial classification, the higher the delinquency rate.

107. For the data on the median household income by state, see WEBSTER JR. & BISHAW, *supra* note 57, at 4 tbl.2 & 5 fig.1. The data on delinquency rate ranks by state can be found at MBA DATA, *supra* note 47, at 4. Admittedly, this data generates a P-Value of .056, placing it just above the level generally recognized as indicating statistical significance. As Table Nine indicates, when this is data is separated by race, we see new P-Values generated for each racial group measured. These values reveal a correlation between African American median income within a state and the delinquency rate in that state.

108. Louisiana's median income is \$23,986, while its inequality rate is 0.475, ranking it forty-eighth (third highest) in the U.S. In contrast, New Hampshire's median income is \$43,933, while its inequality rate is 0.417, ranking it third lowest in the United States.

Table 9: Median Income by Race¹⁰⁹

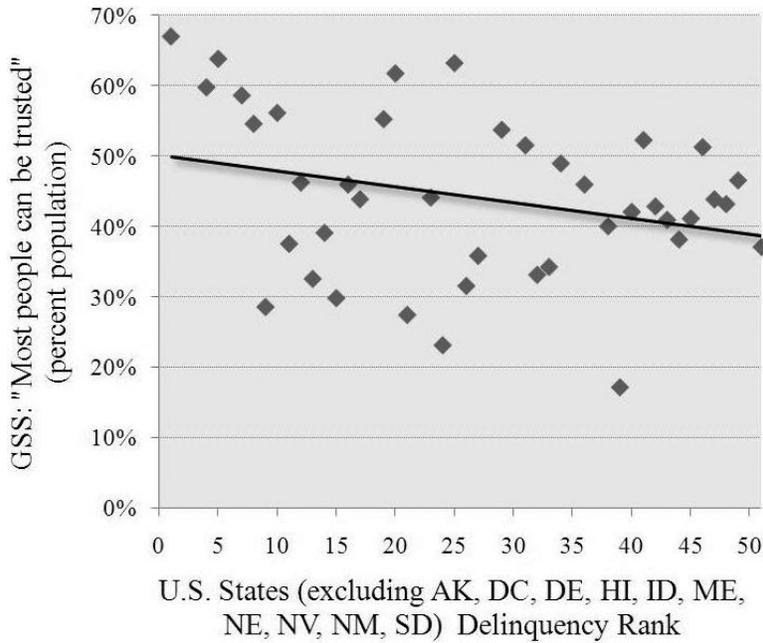


109. For the data on the median household income by race, see *American Factfinder*, U.S. CENSUS BUREAU (2007), <http://factfinder.census.gov/> (select “Data Sets”; then “American Community Survey”; then “2006 American Community Survey”; then “Selected Population Profiles”; then apply the geographic filters).

These graphs raise the question: how might we explain this difference in delinquency rates in relation to levels of inequality by state? As stated earlier, a feature of societies with greater economic inequality is that they have lower levels of trust: that is, members of the community are less willing to trust other members of that community.¹¹⁰ Research also suggests that lower levels of trust correspond with lower levels of trustworthiness, as discussed further below.¹¹¹

Looking at how states with high trust fare in terms of delinquencies, states with higher levels of trust have *lower* delinquency rates. Table 10 reveals the possible negative correlation between trust and delinquencies.

Table 10: Trust and Delinquencies¹¹²



For the data on delinquency rate ranks, see MBA DATA, *supra* note 47, at 4. As stated above, see *supra* note 107, once median income is separated by race, the relationship between median income of the African American population in a state and the delinquency rate in that state seems to indicate a positive, and statistically significant, correlation. See *infra* Appendix, Data Set 3.

110. See USLANER, *supra* note 100, at 187 fig.6-6.

111. See *infra* text accompanying notes 113–22.

112. For the trust data, see data from ROBERT D. PUTNAM, BOWLING ALONE: THE COLLAPSE AND REVIVAL OF AMERICAN COMMUNITY (2000) [hereinafter BOWLING

What, then, does generalized trust itself have to do with predatory conduct? Arguably, this measure of trust also identifies the level of *trustworthiness* in a community. That is, when an individual responds to a question that asks whether he or she trusts others, what the respondent may actually be saying is whether that individual considers him or herself trustworthy. Numerous studies lend credence to this supposition.

In one such study, researchers assessed the cooperative tendencies of undergraduates who were paired together as “senders” and “receivers.”¹¹³ Senders were given a sum of money and invited to share this sum with a receiver; whatever sum the sender sent was matched by the researchers, and receivers were encouraged to share what they received with the senders by returning a portion of what they received.¹¹⁴ The optimal outcome for both participants involved the sender sharing his or her entire grant with the receiver, and the receiver splitting what he or she received evenly and sending one of the divided portions back to the sender.

Prior to conducting the study, the researchers measured the level of trust among the study’s participants. The answers revealed that the level of trust among senders bore little relation to the amount they shared with their partners.¹¹⁵ Instead, the extent and level of cooperation by the *receiver* was directly related to his or her level of trust, leading the researchers to conclude that “the standard trust questions may be picking up trustworthiness rather than trust.”¹¹⁶ They also implied from these findings that, in order to find out if someone is trustworthy, one should ask the person whether he or she trusts others.¹¹⁷

Lest we think these findings are relevant in ivory tower settings only, a similar study of residents of rural Bangladesh reached simi-

ALONE DATA], available at <http://www.bowlingalone.com/data.htm>. For the delinquency data used, see MBA DATA, *supra* note 47, at 4. States excluded above are states for which the GSS Data is not available. Admittedly, the P-Value for this graph shows a roughly one in fourteen chance that this data could appear randomly (i.e., a P-Value of .07), which is considered in the range of P-Values that are “marginally statistically significant,” that is, having a P-Value between .1 and .05. On marginal statistical significance, see Kimberly West-Faulcon, *The River Runs Dry: When Title VI Trumps State Anti-Affirmative Action Laws*, 157 U. PA. L. REV. 1075, 1136 n.211 (2009).

113. See Glaeser et al., *supra* note 12, at 812. For an earlier version of this type of study, see Joyce Berg et al., *Trust, Reciprocity, and Social History*, 10 GAMES & ECON. BEHAV. 122 (1995).

114. Glaeser et al., *supra* note 12, at 812.

115. *Id.* at 826.

116. *Id.* at 833.

117. *Id.*

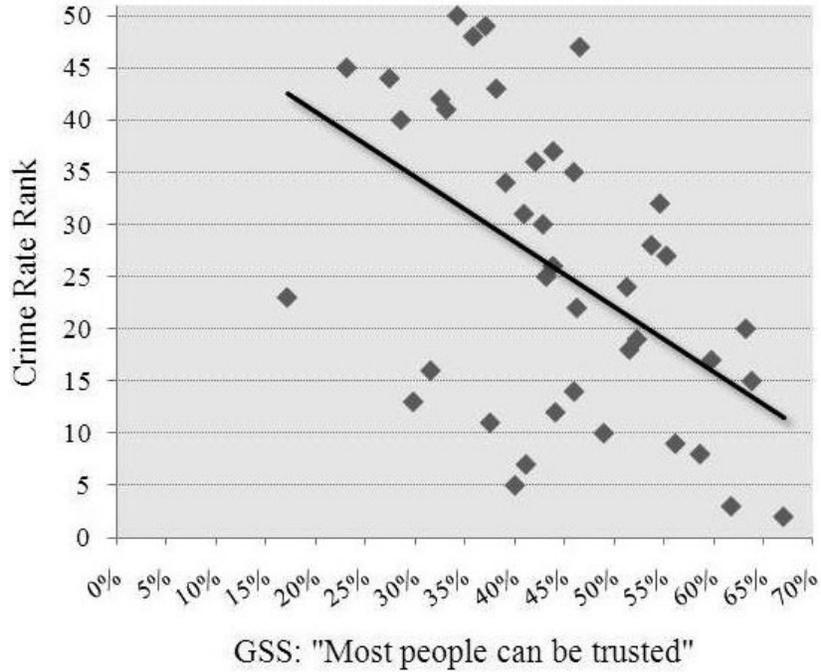
lar outcomes, leading the researchers there to conclude that “stated trust is a better predictor of the amount sent back by the receivers, than of the amount sent by the senders in the trust game.”¹¹⁸ Similarly, analyzing cross-country “lost wallet experiments,” research concluded that residents of countries with high levels of generalized trust were more likely to return lost wallets than residents of nations with lower levels of generalized trust.¹¹⁹ These findings show that people are more trustworthy—i.e., they return lost wallets with greater frequency—in countries where trust is high. Thus, trustworthy conduct can be found in nations with high levels of trust. In other words, when one finds high levels of trust somewhere, one is also going to find high levels of trustworthiness.¹²⁰ Going further, one indicator of trustworthiness—the level of crime in a community—corresponds to levels of trust in a particular community, as Table 11 implies. Lower levels of trust correspond to a higher crime rate.

118. Olof Johansson-Stenman et al., *Trust, Trust Games and Stated Trust: Evidence from Rural Bangladesh* 22 (Goteborg Univ. Dep’t of Econ., Working Paper No. 166, 2005), available at <http://gupea.ub.gu.se/bitstream/2077/2758/1/gunwpe0166.pdf>.

119. Stephen Knack, *Trust, Associational Life and Economic Performance* 18–22 (Mar. 19, 2000) (prepared for the HRDC-OECD International Symposium on the Contribution of Investment in Human and Social Capital to Sustained Economic Growth and Well-Being), available at <http://www.oecd.org/dataoecd/6/31/1825662.pdf> (analyzing Readers Digest study involving lost wallets in several nations and U.S. cities, and showing that rate of return of such wallets corresponded to levels of generalized trust in those areas).

120. There is an obvious feedback loop between trust and trustworthiness; one is more likely to be trusting the more one observes trustworthy behavior. Conversely, the more one is the victim of predatory conduct, the less likely he or she is to trust others. This phenomenon poses particular challenges in the wake of the financial crisis in terms of rebuilding trust in financial institutions. Those burned by such institutions are less likely to be trusting of those institutions, and others like them, in the future. In a fascinating study of the effects of the Bernie Madoff scandal on trust in financial institutions, researchers found a drop in trust in such institutions in communities hardest hit by Madoff’s Ponzi scheme. Luigi Guiso, *A Trust-Driven Financial Crisis: Implications for the Future of Financial Markets* 8–10 (Eur. Univ. Inst., Working Paper No. 2010/07, 2009), available at http://Cadmus.eui.eu/bitstream/handle/1814/13657/ECO_2010_07.pdf?sequence=3.

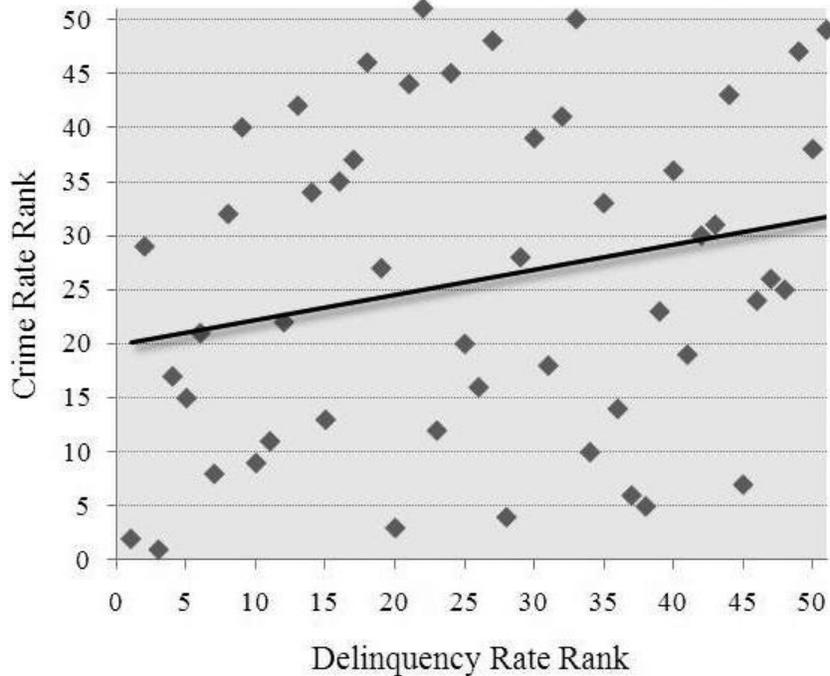
Table 11: Trust and Crime Rate Rank¹²¹



If indicators of trust are also (or perhaps actually) indicators of trustworthiness, then it makes sense that states with lower levels of trust have a greater prevalence of delinquencies—that is, states where people are less trustworthy are also states in which uncooperative and predatory conduct is more likely to take place. As Table 12 reveals, taking general crime rates, set forth in the previous table, and comparing them to delinquency rates, there is a similar correlation between the two data sets. This makes sense since, as shown before, lower levels of trust correspond to higher delinquency rates.

121. The P-Value for this graph is particularly strong. See *infra* Appendix, Data Set 3. For the GSS data on trustworthiness, see BOWLING ALONE DATA, *supra* note 112. For crime rate data, see U.S. DEP'T OF JUSTICE, 2008 CRIME IN THE UNITED STATES, tbl.5 (2008), available at http://www2.fbi.gov/ucr/cius2008/data/table_05.html.

Table 12: Delinquency Rate Rank and Crime Rate Rank¹²²



This potential correlation between crime rates, levels of trust and trustworthiness, and delinquencies suggests that the presence of low trust/trustworthiness, which brings with it higher crime rates, may be correlated with high delinquencies and a greater likelihood of predatory conduct. This correlation implies the possibility that such predatory conduct may have been one of the driving forces behind such delinquencies. States with higher crime rates, by definition, are states in which predatory conduct is more prevalent. It is not a stretch to suggest then that such states were also states where predatory *lending* was likely to flourish. Since there is a possible correlation between trust, trustworthiness, crime rates, and delinquencies, it is also not a stretch to presume that such predatory conduct is likely to have helped to fuel some of the behavior behind the delinquencies that are at the root cause of the foreclosure crisis.

122. For the P-Value of this Table, see *infra* Appendix, Data Set 3. As with Table Ten, *supra*, the P-Value for this table is within the range of “marginal statistical significance.” For delinquency rate ranks, see MBA DATA, *supra* note 47, at 4. For crime rate data, see U.S. DEP’T OF JUSTICE, *supra* note 121.

Another way that social scientists determine the level of trust within a given community is to look at the relative level of “social capital” within that community.¹²³ As Robert Putnam, the author of the landmark work “Bowling Alone,” describes it, social capital is manifest in the “social networks and the . . . norms of reciprocity and trustworthiness” associated with such networks.¹²⁴

These networks and norms facilitate cooperative behavior by generating feelings of mutual obligation towards other members of a network.¹²⁵ Communities that have high levels of social capital are better off economically and have lower crime rates; the residents of those communities report higher levels of life satisfaction.¹²⁶ Levels of social capital also tend to correspond with levels of trustworthiness because of the symbiotic relationship between trust, trustworthiness and social capital.¹²⁷

Returning to the foreclosure crisis, according to commentators from different points on the political spectrum—from Nobel Prize-winning economist Joseph Stiglitz¹²⁸ and Nobel Peace Prize recipient Muhammad Yunus,¹²⁹ to Federal Deposit Insurance Corpora-

123. The level of social capital within a community is sometimes considered a “proxy” for the level of trust in that community, and vice versa. See, e.g., Francis Fukuyama, *Social Capital and Civil Society* 4–5 (Int’l Monetary Fund, Working Paper No. 00/74, 2000), available at <http://www.imf.org/external/pubs/ft/wp/2000/wp0074.pdf> (describing interplay between trust and social capital); PIPPA NORRIS, *DEMOCRATIC PHOENIX* 166 (2002) (describing the same); Luigi Guiso et al., *The Role of Social Capital in Financial Development*, 94 *AM. ECON. REV.* 526, 528 (2004) (characterizing “high-social-capital areas as those with high levels of generalized trust”).

124. Robert D. Putnam, *E Pluribus Unum: Diversity and Community in the Twenty-first Century*, 30 *SCANDINAVIAN POL. STUD.* 137, 137 (2007).

125. On the ways in which social capital develops trust among the members of a network, see James S. Coleman, *Social Capital in the Creation of Human Capital*, 94 *AM. J. SOC.* S95, S119 (Supp. 1988).

126. For an overview of the benefits of social capital, see Michael Woolcock, *The Place of Social Capital in Understanding Social and Economic Outcomes*, 2 *ISUMA: CAN. J. POL’Y RES.* 1 (2001).

127. See Robert D. Putnam, *Bowling Alone: America’s Declining Social Capital*, 6 *J. DEMOCRACY* 65 (1995), reprinted in *CULTURAL METAPHORS: READINGS, RESEARCH TRANSLATIONS, AND COMMENTARY* 109, 111 (Martin J. Gannon ed., 2001).

128. JOSEPH E. STIGLITZ, *FREEFALL: AMERICA, FREE MARKETS, AND THE SINKING OF THE WORLD ECONOMY* 14 (2010).

129. When asked why Grameen Bank is expanding and its microlending portfolio has a 99% success rate when, in the wake of the financial crisis, other financial institutions are pulling back on credit, Yunus attributed this to the relationship between the bank and its customers:

One thing I would say [sic] that very close relationship between the lender and the borrower. That relationship disappeared in a conventional bank. You lend the money to borrower, you know her or you came to know here [sic].

tion (FDIC) Chair Sheila Bair¹³⁰—one of the root causes of the financial crisis was the breakdown of the traditional borrower-lender relationship. In the traditional relationship, the lender was concerned with the long-term viability of the borrower because bank income and profitability hinged on the dependability of the borrower over the long run.¹³¹ These long-term bonds and commitments were likely infused with aspects of social capital.¹³² With securitization came the breakdown of the traditional borrower-lender relationship, as lenders were looking to “originate to securitize,” keeping loans on their books for days, or even hours, just until they could sell them to investment banks to be packaged and sold off as securities.¹³³ Stiglitz describes this transformation as follows:

Securitization, the hottest financial-products field in the years leading up to the collapse, provided a textbook example of the

And you give the loan. And then you sell the document to somebody else. The person who is now owning it doesn't know this person anymore. And this has been sold, and sold, many times, with nothing to do with the person who still owes the money.

So, the relationship has disappeared completely. So we have to go back to the original concept of banking. They stole that relationship. I think that is a very important element that we have to com [sic] back to.

Quest Means Business (CNN television broadcast May 19, 2010) (transcript available at <http://transcripts.cnn.com/TRANSCRIPTS/1005/19/qmb.01.html>) (interview by Maggie Lake with Muhammad Yunus, Founder, Grameen Bank, in New York, NY).

130. Sheila C. Bair, Chairman, Fed. Deposit Ins. Corp., Possible Responses to Rising Mortgage Foreclosures Before the Committee on Financial Services, U.S. House of Representatives (Apr. 17, 2007) [hereinafter Blair Testimony], *available at* <http://www.fdic.gov/news/news/speeches/archives/2007/chairman/spapr1707.html>.

131. Frank A. Hirsch, Jr., *The Evolution of a Suitability Standard in the Mortgage Lending Industry: The Subprime Meltdown Fuels the Fires of Change*, BNET (Mar. 2008), http://findarticles.com/p/articles/mi_6779/is_12/ai_n28511999/ (asserting that “most lenders, prior to the recent increase in subprime mortgages, were unwilling to make a loan in which they doubted the borrower’s ability to repay”).

132. Of course, these bonds could easily have their down side, and bankers could play favorites, give in to stereotypes, and discriminate with impunity. At the same time, such character assessment could come with close and frequent contact with one’s banker. In Frank Capra’s classic film, the community banker’s bedtime story, “It’s a Wonderful Life,” Lionel Barrymore’s Potter, the model of a corrupt, profit-driven banker, criticizes what he sees as the Bailey Building & Loan’s favoritism in its underwriting approach: “You see, if you shoot pool with some employee here, you can come and borrow money.” *IT’S A WONDERFUL LIFE* (Liberty Films 1946).

133. On the “originate to securitize” or “originate to distribute” model, see Kiff & Mills, *supra* note 24, at 11–16.

risks generated by the new innovations, for it meant that the relationship between lender and borrower was broken In the Frankenstein laboratories of Wall Street, banks created new risk products (collateralized debt instruments, collateralized debt instruments squared, and credit default swaps . . .) without mechanisms to manage the monster they had created. They had gone into the moving business—taking mortgages from the mortgage originator, repackaging them, and moving them onto the books of pension funds and others—because that was where the fees were the highest, as opposed to the “storage business,” which had been the traditional business model for banks (originating mortgages and then holding on to them).¹³⁴

While the initial assessment of creditworthiness was likely colored by the depth of the relationship between the prospective borrower and his or her bank, once the deal was consummated, a temporary economic setback would be forgiven more easily when the borrower could communicate with his or her banker and seek forbearance directly, as opposed to dealing with a servicer call center in another state or country.¹³⁵

134. STIGLITZ, *supra* note 128, at 14.

135. Caroline Baum, *Paulson Goes to Washington, Loses Way*, BLOOMBERG (Dec. 4, 2007), <http://www.bloomberg.com/apps/news?pid=newsarchive&sid=aFLLy4LlxMc>. As Sheila Bair has explained: “[T]he increased complexity of the structure and the different interests of the various securitization parties can make credit workout strategies more complicated than in a direct borrower/lender relationship.” Bair, *supra* note 130. Stiglitz describes the transformation of mortgage renegotiation dynamics as follows:

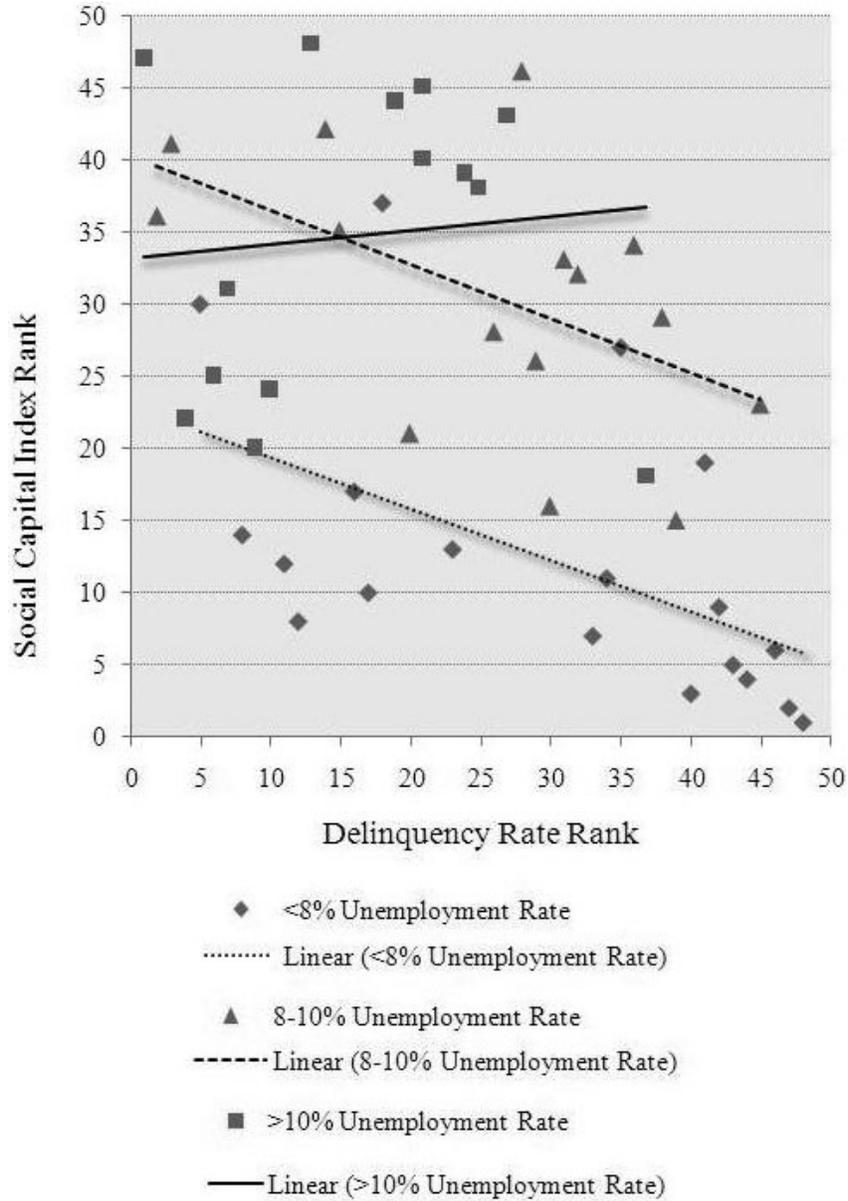
Banks with long-standing relations with the community had an incentive to treat borrowers who got into trouble well; if there was a good chance that borrowers would catch up on their payments if they were given some time, then the bank would give them the time they needed. But the distant holders of the mortgages had no interest in the community and no concern about having a reputation as a good lender.

STIGLITZ, *supra* note 128, at 95.

If there is a connection between social capital and delinquency rates, the level of social capital in a community will also impact delinquency rates. Comparing a state's relative social capital rank, using Putnam's social capital index,¹³⁶ with delinquency rate rank by state, reveals a correlation between high levels of social capital and low delinquency rates, at least with respect to states that are not experiencing above-average unemployment rates, as Table 13 indicates.

136. Putnam's social capital index includes fourteen indicators, including those that measure the level of participation in civic organizations, the level of trust in a community, and the number of non-profit organizations per 1000 residents in a community. See BOWLING ALONE DATA, *supra* note 112. Admittedly, this social capital index was generated when *Bowling Alone* was published, and has not been updated. Using this data is helpful, though, to show that social capital was high or low in a particular state prior to the changes in the economy, most notably, the rise in home values, which marked the middle part of the last decade.

Table 13: Social Capital and Unemployment Data¹³⁷



137. For the social capital data, see BOWLING ALONE DATA, *supra* note 112. For the delinquency data used, see MBA DATA, *supra* note 47, at 4. For unemployment data, see U.S. Dep’t of Labor: Bureau of Labor Statistics, *Unemployment Rates for States*, LOCAL AREA UNEMPLOYMENT STATISTICS, <http://www.bls.gov/lau/lastrk09>.

For the most part, the higher the social capital, the lower the delinquency rate. Here, the data is separated for states with similar unemployment rates. In states with average or below-average unemployment rates, we see a strong negative correlation between social capital and foreclosure rates. In states with above-average unemployment rates, there is a slight positive correlation. This may either be a statistical anomaly, or social capital may have no positive effects on foreclosure rates in the face of extreme economic distress, like having double-digit unemployment within a state.

The findings with respect to a generally negative correlation between foreclosure rates and social capital correspond to similar findings reached by the Corporation for National and Community Service (CNCS) in its study of foreclosure rates and volunteering (another reflection of social capital) in large U.S. cities in 2008 and 2009. "Looking at the relationship between foreclosure and volunteer rates, [CNCS] found that cities with higher foreclosure rates tended to have lower volunteer rates. [The] findings show that a one percent decrease in foreclosure rates would be associated with a 1.2 percent increase in volunteer rate."¹³⁸

If these correlations are present, then it is worth discussing whether the absence of social capital may have been one of the driving forces behind the foreclosure crisis.

C. Racial Distance and the Foreclosure Crisis

Another type of social distance can come about as a result of distinctions based on race. Again, according to the theory, racial differences can lead to the same sort of predatory conduct that occurs where there is greater income inequality.¹³⁹ For this to hold true in the financial crisis, we would have to see an increase in the occurrence of predatory conduct in communities of color.

Looking at lending during the heart of the subprime mortgage frenzy in 2006, the Federal Reserve's analysis of Home Mortgage Disclosure Act (HMDA) data reveals that nearly 54% of the home

htm (last modified Mar. 8, 2010). For the P-Values for this chart, see *infra* Appendix, Data Set 3. As stated above, there is no statistically significant impact of social capital on delinquencies in states with high unemployment.

138. CORP. FOR NAT'L & CMTY. SERV., ISSUE BRIEF, VOLUNTEERING IN AMERICA 2010: NATIONAL, STATE AND CITY INFORMATION 8–9 (2010), available at <http://www.volunteeringinamerica.gov/assets/resources/IssueBriefFINALJune15.pdf>.

139. See *supra* text accompanying notes 94–98. For an overview of research on social distance and race relations, see Ralph Richard Banks & Richard Thompson Ford, (*How*) *Does Unconscious Bias Matter?: Law, Politics, and Racial Inequality*, 58 EMORY L.J. 1053, 1069–72 (2009), and sources cited therein.

purchase loans made to African Americans in that year had subprime features, compared to just under 18% for Whites, a nearly three-to-one ratio.¹⁴⁰

Controlling for borrower characteristics,¹⁴¹ the Fed found that in 2006, over 30% of African American borrowers and 24% of Latino borrowers, compared to just under 18% of White borrowers of comparable creditworthiness, received subprime loan products.¹⁴² Thus, an African American borrower of similar creditworthiness to a White borrowers was 75% more likely to receive a subprime loan product than the White borrower, and a Latino borrower was 36% more likely to take out a subprime loan than a White borrower with a similar economic profile.

A study done by the New York Times on lending in the New York City metropolitan region found the prevalence of lending patterns even more striking, particularly as it related to working-class

140. See Robert B. Avery et al., *The 2006 HMDA Data*, 93 FED. RES. BULL. A73, A95 (2007) [hereinafter 2006 HMDA Data], available at <http://www.federalreserve.gov/pubs/bulletin/2007/pdf/hmda06final.pdf>.

141. The authors controlled for borrower characteristics by utilizing a “matching procedure” in assessing the 2004 and 2005 HMDA data. *Id.* This procedure allowed the Fed to measure “differences in denial rates by comparing applications for a specific loan product filed by applicants who differ by race, ethnicity, or sex but who are matched on the basis of the limited set of items in the HMDA data.” Robert B. Avery et al., *New Information Reported under HMDA and Its Application in Fair Lending Enforcement*, 91 FED. RES. BULL. 344, 387 (2005) (emphasis omitted), available at <http://www.federalreserve.gov/pubs/bulletin/2005/3-05hmda.pdf>. The HMDA data can be manipulated, allowing

individuals to be matched by loan type and purpose, type of property securing the loan, lien status, owner-occupancy status, property location (for example, same MSA or even same census tract), income relied on for underwriting, loan amount, and time of year when the loan was made as well as by whether the loan involved a co-applicant.

Id. at 372. In 2006, a control for lender was also added. 2006 HMDA Data, *supra* note 140, at A95.

142. *Id.* at A96 tbl.11. A subprime loan is a loan made to a borrower who possesses certain features:

Subprime borrowers typically have weakened credit histories that include payment delinquencies, and possibly more severe problems such as charge-offs, judgments, and bankruptcies. They may also display reduced repayment capacity as measured by credit scores, debt-to-income ratios, or other criteria that may encompass borrowers with incomplete credit histories. Subprime loans are loans to borrowers displaying one or more of these characteristics at the time of origination or purchase.

OFFICE OF THE COMPTROLLER OF THE CURRENCY ET AL., EXPANDED GUIDANCE FOR SUBPRIME LENDING PROGRAMS 2 (2001), available at <http://www.federalreserve.gov/boarddocs/press/boardacts/2001/20010131/attachment.pdf>.

African Americans.¹⁴³ Indeed, after studying lending patterns in the city and comparing borrowers of similar incomes, the researchers concluded as follows: “[T]he hardest blows rain down on the backbone of minority neighborhoods: the black middle class. In New York City, for example, black households making more than \$68,000 a year are almost five times as likely to hold high-interest subprime mortgages as are whites of similar — or even lower — incomes.”¹⁴⁴

Deeper analysis of HMDA data reveals that 169 mortgage lenders failed in 2007, following the collapse of the housing market.¹⁴⁵ With these lenders, who were most likely engaged in the riskiest lending (hence their closing their doors in 2007), a disproportionate share of their subprime lending in 2006 was directed towards African American borrowers when compared to lending in the industry as a whole. Indeed, 74% of the loans made by these lenders to African Americans in 2006 were subprime loans, and 63% of the loans to Latinos were subprime, compared to an industry average of 54% and 47% respectively, as Table 14 indicates.¹⁴⁶ Thus, it appears that the riskiest subprime lending in 2006 was carried out by lenders that focused such lending on borrowers of color.

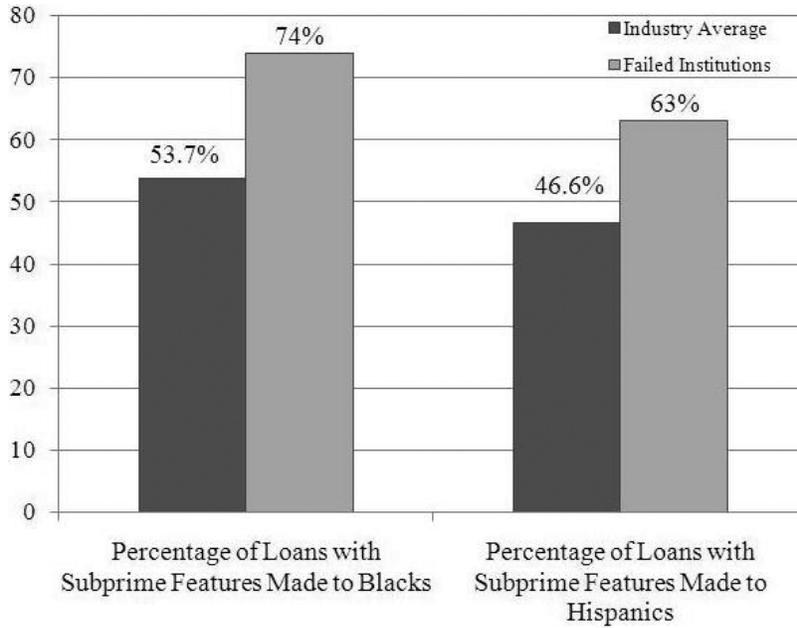
143. Michael Powell & Janet Roberts, *Minorities Hit Hardest as New York Foreclosures Rise*, N.Y. TIMES, May 16, 2009, at A1.

144. *Id.*

145. Robert B. Avery et al., *The 2007 HMDA Data*, 94 FED. RES. BULL. A107, A109 (2008), available at <http://www.federalreserve.gov/pubs/bulletin/2008/pdf/hmda07final.pdf>.

146. For the failed institutions data, see *id.* at A126 tbl.12. For the industry average data see 2006 HMDA Data, *supra* note 140, at A95–96.

Table 14: Data on 2007 Failed Institutions¹⁴⁷



What does all of this information say about economic and racial inequality and the financial crisis? States with greater inequality are also states in which predatory conduct is likely to be more prevalent. Is it possible then, that in states with greater inequality, their higher delinquency rates may be a function of the fact that predatory conduct, in the form of predatory lending, was more prevalent? Similarly, mortgage lenders concentrated a disproportionate share of their subprime lending in African American and Latino communities. A fair number of these borrowers saddled with subprime loans would have qualified for prime loans, which they would have stood a better chance of paying back, sparing themselves the ordeal of foreclosure.¹⁴⁸ The lenders that pedaled toxic products in

147. The data on failed institutions by race was obtained from Avery et al., *supra* note 145, at A126 tbl.12. The data on the 2006 industry average was obtained from the 2006 HMDA Data, *supra* note 140, at A95–96.

148. Estimates place the number of subprime borrowers who would have qualified for prime loans between 35% and 61%. See *The Community Reinvestment Act: Thirty Years of Accomplishments but Challenges Remain*, Hearing Before the H. Comm. on Fin. Servs., 110th Cong. 4 (2008) (prepared testimony of Michael S. Barr, Professor, Univ. of Mich. Law Sch.), available at <http://financialservices.house.gov/hearing110/barr021308.pdf> (arguing that 35% of subprime borrowers would have qualified for prime loans); Rick Brooks & Ruth Simon, *Subprime Debacle Traps Even Very Credit-Worthy: As Housing Boomed, Industry Pushed Loans to a Broader Market*,

those communities are the ones who should be held largely responsible for what ensued, not their victims.

Indeed, as a growing body of evidence strongly suggests, the problem in the mortgage market was not risky borrowers but predatory loans. One study by the Center for Community Capital at the University of North Carolina compared the performance of risky subprime loans with loans where strong underwriting guidelines were used.¹⁴⁹ The Center identified borrowers of similar profiles that entered into either the risky loans or the more stable loans.¹⁵⁰ That study found that for loans made in 2004, the subprime loans were four times more likely to enter into default than the more stable loan products; in 2006, they were 3.3 times as likely.¹⁵¹ Imagine the amount of suffering that could be alleviated if foreclosure rates were reduced by 67% or even 75%.

A critical aspect of the theory about social distance is that greater social distance leads to non-cooperative, predatory conduct. Is it possible, then, that the problem was borrower fraud and not just predatory lending? Were borrowers behaving in an untrustworthy fashion? Certainly some were, but according to FBI estimates roughly 80% of losses due to mortgage fraud involved lender and/or broker misconduct.¹⁵² Indeed, an *internal study* from 2005 conducted by Washington Mutual of just one of its offices, in Montebello, California, revealed that 83% of the mortgages generated by that office involved some form of fraud on the part of bank officials.¹⁵³

WALL ST. J., Dec. 3, 2007, at A1 (reporting that 61% of subprime loans originated in 2006 “went to people with credit scores high enough to often qualify for conventional[, i.e., prime,] loans with far better terms”). In his recent book, Gary Rivlin puts it succinctly: “The problem wasn’t the people [receiving the loans] but the product they were being sold.” GARY RIVLIN, *BROKE USA: FROM PAWNSHOPS TO POVERTY INC.—HOW THE WORKING POOR BECAME BIG BUSINESS* 133 (2010).

149. Lei Ding et al., *Risky Borrowers or Risky Mortgages: Disaggregating Effects Using Propensity Score Models* 1 (Univ. of N.C. Ctr. for Cmty. Capital, 2010), available at <http://www.ccc.unc.edu/documents/Risky.Disaggreg.5.17.10.pdf>.

150. *Id.* at 15–16.

151. *Id.* at 28.

152. U.S. DEP’T OF JUSTICE: FED. BUREAU OF INVESTIGATION, *FINANCIAL CRIMES REPORT TO THE PUBLIC D1–D12* (2005), available at http://www.fbi.gov/stats-services/publications/fcs_report2005/financial-crimes-report-to-the-public-2005-pdf.

153. On April 13, 2010, the Permanent Subcommittee on Investigations, headed by Senator Carl Levin, began holding hearings aimed at examining “some of the causes and consequences of the recent financial crisis.” Memorandum from Senator Carl Levin, Subcomm. Chairman, to Members of the Permanent Subcomm. on Investigations 1 (Apr. 13, 2010) [hereinafter Levin Memo], available at

In one recently filed lawsuit under the Fair Housing Act, former employees of the defendant Wells Fargo, an aggressive subprime lender,¹⁵⁴ alleged the following:

Wells Fargo's Memphis branches targeted African Americans for subprime loans because employees held negative views of African Americans. [Former bank employee] Taylor explains that "[t]he prevailing attitude was that African American customers weren't savvy enough to know they were getting a bad loan, so we would have a better chance of convincing them to apply for a high-cost, subprime loan."

<http://levin.senate.gov/newsroom/supporting/2010/PSI.LevinCoburnmemo.041310.pdf>.

In particular, the subcommittee focused on a case study involving Washington Mutual Bank (WaMu). The hearings revealed that, in 2005, WaMu itself had conducted an internal investigation that turned up a number of red flags regarding its lending practices. David Heath, *WaMu Executives Knew of Rampant Mortgage Fraud and Failed to Act*, HUFFINGTON POST (Apr. 12, 2010), http://www.huffingtonpost.com/2010/04/12/wamu-executives-knew-of-r_n_534800.html. The investigation uncovered that:

[L]oans originated from two top loan offices in southern California contained an extensive level of fraud Despite fraud rates in excess of 58% and 83% at those two offices, no steps were taken to address the problems, and no investors who purchased loans originated by those offices were notified . . . of the fraud problem.

Levin Memo, *supra*, at 4. James Vanasek was one of the former WaMu executives who testified before the subcommittee. Despite assertions of other executives that they had not been aware of the impending housing crisis early on, Vanasek "said he realized in 2004 that 'the industry was in some degree of difficulty . . .'" Kirsten Grind, *WaMu Hearing Begins*, PUGET SOUND BUS. J. (Apr. 13, 2010), http://seattle.bizjournals.com/seattle/blog/2010/04/wamu_hearing_begins.html.

Another 2007 internal investigation of one of the southern California WaMu offices uncovered a fraud rate of 62%. Levin Memo, *supra*, at 4. WaMu, along with its affiliate, Long Beach Mortgage Company, compounded this problem by "creating misplaced incentives that encouraged high volumes of risky loans but little or no incentives to ensure high quality loans that complied with the bank's credit requirements." *Id.* at 2, 4-5. In addition to the 2005 and 2007 internal investigations, a number of other documents were uncovered which included discussions among the Board of Directors as well as various emails, all of which clearly illustrated the higher risk lending strategies that were rampant throughout the company from 2003 to 2008. *Wall Street and the Financial Crisis: The Role of High Risk Home Loans, Hearing Before the S. Permanent Subcomm. on Investigations*, 111th Cong. at 2-9 (Apr. 13, 2010) (exhibit list), available at http://hsgac.senate.gov/public/_files/Financial_Crisis/041310Exhibits.pdf. One internal report did conclude that "[t]hroughout the process, red flags were over-looked, process requirements were waived, and exceptions to policy were granted." Heath, *supra* (internal quotations omitted).

154. Michael Powell, *Suit Accuses Wells Fargo of Steering Blacks to Subprime Mortgages in Baltimore*, N.Y. TIMES, June 7, 2009, at A16.

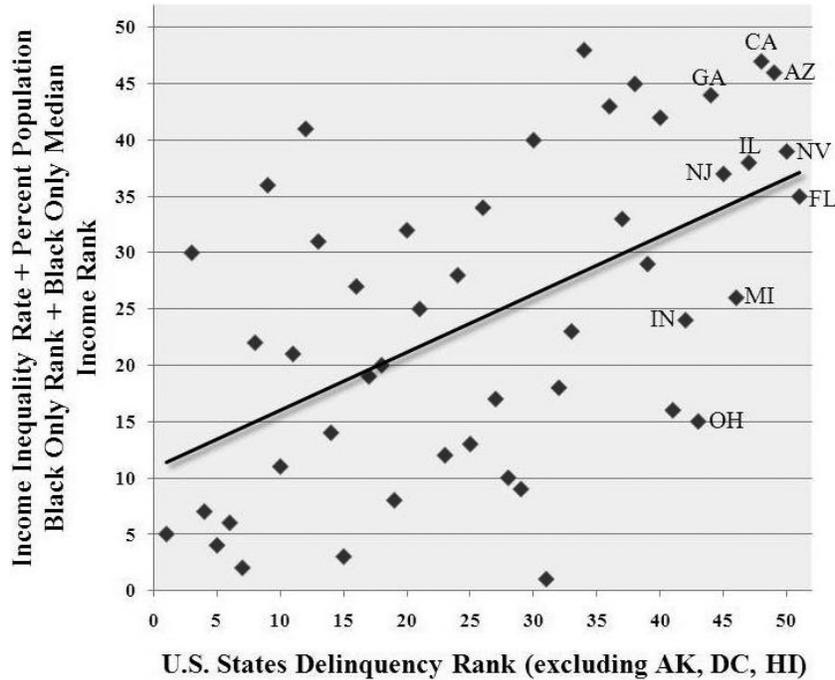
Likewise, Thomas[, another former bank employee,] explains that “[i]t was generally assumed that African American customers were less sophisticated and intelligent and could be manipulated more easily into a subprime loan with expensive terms than white customers.”¹⁵⁵

Assuming these allegations are true, it appears likely that greater social distance, both in terms of economic and racial inequality, created the conditions in which predatory conduct flourished.

Finally, putting together a number of gauges of social distance caused by economic and racial inequality, Table 16 combines a number of factors outlined above that likely played out in harmful ways in the financial crisis: greater economic inequality, median income of the African American community in a state, and percentage of the African American community in a particular state. Putting these indicators together generates a graph that supports the theory that the African American middle class was targeted for subprime loans.

155. First Amended Complaint for Declaratory and Injunctive Relief and Damages at 32–33, *City of Memphis v. Wells Fargo Bank* (W.D. Tenn. Apr. 7, 2010) (No. 2:09-cv-02857x), 2010 WL 1506670. Similarly, a CitiFinancial branch manager described that lender’s approach as follows: “[t]he more gullible the consumer appeared . . . the more [additional costs] . . . I would try to include in the loan.’ By ‘gullible,’ she explained, she meant the very young or the very old, minorities and those who ‘appeared uneducated, inarticulate.’” Rivlin, *supra* note 148, at 152.

Table 15: Targeting the Black Middle Class¹⁵⁶



This data analysis is consistent with the findings of a joint report by the U.S. Department of Housing and Urban Development (HUD) and U.S. Department of the Treasury issued in 2000 on predatory lending¹⁵⁷ and later studies conducted during the subprime mortgage market’s heyday.¹⁵⁸ In the HUD–Treasury report, these agencies found that predatory lending was more prevalent in the subprime market than the prime market and that such sub-

156. This Table has the strongest P-Value of the tables generated for this study. See *infra* Appendix, Data Set 3. For the data on black middle class income, see US CENSUS BUREAU, *supra* note 109. The data on delinquency rate ranks by state can be found at MBA DATA, *supra* note 47, at 4. The P-Value for this graph as a whole, as well as for each of the separate data sets used and their relation to the delinquency rates with respect to each data set is set forth in the Appendix, Data Set 3, *infra*. In creating this index, these indicators were given equal weight. On the selection and weighting of criteria in the creation of an index, see Roodman, *supra* note 17, at 5–7.

157. U.S. DEP’T OF HOUS. & URBAN DEV. & U.S. DEP’T OF TREASURY, CURBING PREDATORY HOME MORTGAGE LENDING 47–48 (2000) [hereinafter HUD-TREASURY REPORT], available at <http://www.huduser.org/Publications/pdf/treasrpt.pdf>.

158. See *infra* note 163 and accompanying text.

prime lending was prevalent in communities of color.¹⁵⁹ Such communities, HUD and Treasury found, tended to be underserved by what the agencies called “traditional prime lenders,” which they identified as banks, thrifts, and credit unions, all of which are subject to more federal oversight than mortgage finance companies.¹⁶⁰ In a later report, the National Community Reinvestment Coalition described the effects of a legacy of discrimination in communities of color as paving the way for subprime lenders to flourish in those communities:

[M]inority borrowers and communities that received disproportionately high numbers of subprime loans had historically lower homeownership rates, as they were systematically excluded from home mortgages due to bank-redlining and discrimination until the early 1980s. Then, when financial institutions rapidly expanded their lending to minority households it was associated with the use of high-cost, or otherwise unfair and abusive products. The resulting high density of subprime loans “increases the risk of foreclosures and negative spillover effects like declines in property values and increasing crime rates.” Hence, a larger proportion of homeowners are facing foreclosures in predominantly minority neighborhoods.¹⁶¹

A legacy of discrimination in certain communities,¹⁶² that translated into a lower homeownership rate in those communities, meant such areas were fertile ground for the expansion of subprime lending that occurred in the last decade. With few prime banking alternatives, middle and lower-middle class borrowers of color were targeted for subprime loans and steered into loans with less favorable terms than they might have otherwise accessed had they had better credit alternatives available to them.¹⁶³ The thin

159. HUD-TREASURY REPORT, *supra* note 157, at 16, 47–48.

160. *Id.* at 18.

161. Tamara Jayasundera et al., *Foreclosure in the Nation's Capital: How Unfair and Reckless Lending Undermines Homeownership* 19 (Nat'l Cmty. Reinvestment Coal., 2010) (citations omitted), available at http://www.ncrc.org/images/stories/pdf/research/ncrc_foreclosure_paper_final.pdf.

162. On the history of housing discrimination in the United States, see DAN IMMERGLUCK, CREDIT TO THE COMMUNITY: COMMUNITY REINVESTMENT AND FAIR LENDING POLICY IN THE UNITED STATES 87–108 (2004); Adam Gordon, *The Creation of Homeownership: How New Deal Changes in Banking Regulation Simultaneously Made Homeownership Accessible to Whites and Out of Reach for Blacks*, 115 YALE L.J. 186 (2005).

163. For an analysis of lending patterns in seven metropolitan areas, which showed 40% of the activities of subprime lenders concentrated in communities of

market for mainstream financial institutions in communities of color—the legacy of credit redlining—is precisely what narrowed the options available to the residents of such communities. The correlation between borrower median income and the size of the African American population with foreclosure rates in a state would suggest that the larger the African American middle class in a community, and the wealthier it was, the more likely that subprime lending was prevalent in that community and the more likely that the African American middle class was targeted for loans on less than favorable terms.¹⁶⁴ These less than favorable terms led to the foreclosure crisis now plaguing those very same communities, with devastating effects.

III. AN ALTERNATIVE EXPLANATION AND A RESPONSE

As described above, Rajan and others have argued that government policies—most notably the demands of the CRA and the affordable housing goals of the GSEs—were responsible for the housing bubble and the financial collapse that followed the bursting of that bubble.¹⁶⁵ In order for these arguments to hold true, their proponents would have to show that it was CRA oversight that forced subprime lenders to engage in risky lending, that the GSE

color, while only 10% of their lending taking place in white communities, see CAL. REINVESTMENT COAL. ET AL., *PAYING MORE FOR THE AMERICAN DREAM: THE SUBPRIME SHAKEOUT AND ITS IMPACT ON LOWER-INCOME AND MINORITY COMMUNITIES* 5 (2008), available at <http://www.woodstockinst.org/publications/download/paying-more-for-the-american-dream-%11-the-subprime-shakeout-and-its-impact-on-lower-%11income-and-minority-communities>.

164. Another potential explanation for the prevalence of foreclosures in communities with larger African American populations is the theory that communities with greater population heterogeneity have lower social capital. The logical conclusion to be drawn from such a theory is that this lowered social capital also tends to correspond with higher foreclosure rates, as we have already seen. If the correlation between social capital, trust, and cooperative behavior is well-founded, then the potential correlation between population heterogeneity and predatory conduct is certainly cause for alarm. This controversial issue is addressed in Robert Putnam's relatively recent work on the subject, Putnam, *E Pluribus Unum*, *supra* note 124. For a response to Putnam, see Casey J. Dawkins, *Reflections on Diversity and Social Capital: A Critique of Robert D. Putnam's "E Pluribus Unum: Diversity and Community in the Twenty-First Century the 2006 Johan Skytte Prize Lecture"*, 19 HOUS. POL'Y DEBATE 207 (2008). For a description of the ways in which mortgage lenders actually used social capital networks to gain entrée into communities of color, see Creola Johnson, *The Magic of Group Identity: How Predatory Lenders Use Minorities to Target Communities of Color*, 17 GEO. J. POVERTY L. & POL'Y 165 (2010).

165. See RAJAN, *supra* note 83, at 34–37, 43–45.

policies were driving the subprime mortgage market, and that GSE underwriting standards were weaker than those of lenders acting on their own, selling bundled mortgages on the strictly private market. As the following discussion shows, not one of these factual suppositions enjoys much support.

A. *The CRA: Not Guilty*

Given our previous discussion of income inequality and its potential connection to financial crises, it is particularly salient to discuss the Community Reinvestment Act (CRA) because it was originally designed to protect the interests of low- and moderate-income communities.¹⁶⁶ As a result, it was originally crafted to combat income inequality, or at least the distortions that income inequality creates on credit markets.¹⁶⁷ If the previous information about income inequality and the foreclosure crisis suggests that the public should be concerned with the impact of such inequality on housing finance policy, then there is no better place to start than to look at the CRA to see its impact on causing or failing to prevent the financial crisis. That some have attempted to lay the blame for the crisis on the CRA only strengthens the need to analyze its purpose, scope, successes, and failures.

Under the CRA, federal bank regulators are to use their authority “to *encourage* [financial] institutions to help meet the credit needs of the local communities in which they are chartered,” so long as this goal can be carried out “consistent with the safe and sound operation of such institutions.”¹⁶⁸ The CRA was enacted to combat two practices that Congress found to undermine the economic health of low- and moderate-income communities: bank redlining (excluding certain communities from access to capital) and capital exportation (taking deposits from a community while refusing to lend within that community).¹⁶⁹ The policies of the CRA were also seen as a quid pro quo for the substantial benefits that

166. The original language of the CRA was amended during debates over its passage to include consideration of low- and moderate-income communities, as opposed to focusing on the needs of banks’ “primary service area”. H.R. REP. NO. 95-634, at 75-76 (1977) (Conf. Rep.).

167. On the distortions on credit markets affecting low- and moderate-income communities, see Michael S. Barr, *Credit Where It Counts: The Community Reinvestment Act and Its Critics*, 80 N.Y.U. L. REV. 513, 533-44 (2005).

168. 12 U.S.C. § 2901(b) (2006) (emphasis added).

169. 123 CONG. REC. 17,630 (1977) (statement of Sen. William Proxmire).

banks received: exclusive charter powers, deposit insurance, and discount loans.¹⁷⁰

So, what is it about this law that makes it a favorite target of those who would lay the blame for the crisis on it and those minority communities which the CRA supposedly directs banks to serve? Under the CRA,¹⁷¹ bank regulators grade the financial institutions the CRA covers on those institutions' success in meeting the credit needs of their communities, including low- and moderate-income neighborhoods, consistent with safe and sound banking practices.¹⁷² Bank regulators are then supposed to take into account that grade when a bank makes an application to its regulators to take certain actions, like merging with another bank or opening a bank branch.¹⁷³ If the CRA were some kind of sword of Damocles hanging over the heads of banks, forcing them to make unwise loans, we would likely see two things: first, the regulators giving banks poor CRA ratings; and second, the regulators rejecting bank

170. See 123 CONG. REC. 1,958 (1977) (statement of Sen. William Proxmire). For a further discussion of the connection between government support for banks and the CRA, see Barr, *supra* note 167, at 616–24 (2005); Allen J. Fishbein, *The Community Reinvestment Act After Fifteen Years: It Works, but Strengthened Federal Enforcement Is Needed*, 20 FORDHAM URB. L.J. 293, 293 (1993) (citation omitted).

171. Under the CRA, the following federal banking agencies supervise different sectors of the banking industry: the Office of the Comptroller of the Currency oversees national banks; the Board of Governors of the Federal Reserve System oversees state chartered banks which are members of the Federal Reserve System and bank holding companies; the Federal Deposit Insurance Corporation regulates state chartered banks, savings banks that are not members of the Federal Reserve System and the deposits insured by the FDIC; and the Office of Thrift Supervision (OTS) with respect to savings associations, the deposits of which are insured by the FDIC and savings and loan holding companies. 12 U.S.C. § 2902(1) (2006). With the dissolution of the OTS through the Dodd-Frank financial reform legislation, it is likely that the OCC will assume CRA responsibilities over many of the former OTS-regulated entities. Dodd-Frank Wall Street Reform and Consumer Protection Act, Pub. L. No. 111-203, § 312, 124 Stat. 1376, 1521–23 (2010).

172. 12 U.S.C. § 2906(a)(1) (2006). Regulators issue one of four “grades” to the institutions they oversee based on those institutions’ success in “meeting community credit needs” consistent with the CRA: “outstanding,” “satisfactory,” “needs to improve,” or “substantial noncompliance.” *Id.* § 2906(b)(2).

173. Under the CRA, a federal bank must take into account a covered bank’s CRA record whenever the bank applies to its regulator for a “deposit facility.” 12 U.S.C. § 2903(a)(2) (2006). The statute defines such an application to include a request seeking approval for any of the following: “a charter for a national bank or Federal savings and loan association; . . . deposit insurance in connection with a newly chartered . . . bank”; the establishment of a branch or other facility that will accept deposits; the relocation of a home or branch office; or the merger, consolidation, or acquisition of another regulated financial institution in certain circumstances. *Id.* § 2902(3).

applications on CRA grounds with great frequency. In reality, in the lead up to the financial crisis, neither of these occurred.

First, in terms of regulators issuing bank CRA grades, in 2005, a year in which many financial institutions engaged in risky lending practices, 99% of all banks covered by the CRA received a grade of “satisfactory” or “outstanding” in terms of meeting their CRA obligations.¹⁷⁴ Second, in terms of bank applications, during the 15 year span from 1985 to 1999, only eight bank applications were denied by bank regulators *on any grounds*, a figure that might seem significant until one learns that there were over 92,000 applications subject to CRA review that were filed during that period.¹⁷⁵ Analysis of the Federal Reserve’s record shows that, from 1988 through 2007, the Federal Reserve denied only eight bank applications of the more than 13,000 filed with it for unsatisfactory efforts to meet community needs or to provide consumer protection.¹⁷⁶

It is hard to imagine that the enforcement mechanisms, as applied by the regulators, were a meaningful check on bank behavior. Putting aside whether the CRA, as structured and enforced, was a serious check on the behavior of the banks to which the CRA applied, the truth of the matter is that the overwhelming majority of the riskiest bank behavior during the subprime mortgage market’s heyday took place beyond the CRA’s reach. As stated previously, the legislative history of the CRA reveals that it was enacted as a response to two forces in the market: redlining and credit exportation. As a result, the CRA covers only depository institutions and concerns itself with the geographic focus of bank activities as they relate to those places where a covered financial institution does the bulk of its business.¹⁷⁷ The CRA was thus conceived in a very different time, before internet and global banking, when the neighborhood bank was where one deposited one’s money and sought access to credit. Given this history, the CRA was ill-equipped to handle the subprime tsunami that hit in the early part of this decade. A review of its scope reveals that the CRA exempted much of the riski-

174. *The Community Reinvestment Act: Thirty Years of Accomplishments, but Challenges Remain: Hearing Before the H. Comm. on Fin. Servs.*, 110th Cong. 194 (2008) (statement of John Taylor, President and CEO, Nat’l Cmty. Reinvestment Coal.).

175. See BART, *supra* note 167, at 586 (citing TREASURY DEP’T, APPLICATIONS SUBJECT TO CRA THAT WERE SUBJECT TO CRA THAT WERE PROTESTED ON CRA GROUNDS (2000)).

176. *Foreclosures at the Front Step of the Federal Reserve Bank of Cleveland: Hearing Before the Subcomm. on Domestic Policy of the H. Comm. on Oversight and Gov’t Reform*, 110th Cong. 63–64 (2007) (statement of Sandra Braunstein, Director, Div. of Consumer and Cmty. Affairs, Fed. Reserve Sys.).

177. See *infra* notes 178–84 and accompanying text.

est lending that took place in this time from its coverage in a number of ways. As a result of this narrow CRA scope, it is hard to argue that the CRA forced financial institutions that it did not even cover to do anything.

According to the statute, the CRA only applies to “regulated financial institutions,”¹⁷⁸ which the statute describes as “insured depository institution[s].”¹⁷⁹ This is the largest exemption, which, according to Fed Chairman Bernanke, excluded during the recent expansion of the market as much as two-thirds of mortgage lending from the CRA’s reach.¹⁸⁰ That is the share of mortgage lending during that time that was carried out by standalone mortgage companies, companies that did not take deposits and thus were not covered by the CRA.¹⁸¹ A second exemption permits depository institutions, at their discretion, to exempt the lending of their non-bank affiliates from CRA coverage, even if they are themselves covered by the law.¹⁸² Thus, a bank like Bank of America can exempt its subsidiary, Countrywide, from CRA coverage if it chooses.

Third, the requirements of the CRA only apply to bank activities within each bank’s “assessment areas” and regulators assess a bank’s lending in low- and moderate-income communities only within those assessment areas.¹⁸³ A bank’s assessment area includes where it has offices and branches, and where a “substantial” portion of a bank’s lending occurs.¹⁸⁴ Regulators do not review lending for

178. 12 U.S.C. § 2901(a)(1) (2006).

179. *Id.* § 2902(2). Under the CRA, the definition of “insured depository institution” is that set forth in 12 U.S.C. § 1813, which provides that such an institution is “any bank or savings association the deposits of which are insured by the [Federal Deposit Insurance] Corporation . . .” *Id.* § 1813(c)(2).

180. Ben S. Bernanke, *The Community Reinvestment Act: Its Evolution and New Challenges*, Speech at the Community Affairs Research Conference 5 (Mar. 30, 2007), available at <http://www.federalreserve.gov/newsevents/speech/bernanke20070330a.htm>; see also ROBERT E. LITAN ET AL., *THE COMMUNITY REINVESTMENT ACT AFTER FINANCIAL MODERNIZATION: A BASELINE REPORT 70–72* (2000), available at http://www.novoco.com/low_income_housing/resource_files/research_center/crareport.pdf (analyzing lending data from 1992 and 1998, and finding that two-thirds of the increase in lending by institutions not covered by the CRA to low- and moderate-income communities was attributable to lenders specializing in subprime, while only 15% of these loans were originated by CRA-covered institutions).

181. Bernanke, *supra* note 180 (noting that companies owned by banks or bank holding companies can be nondepository institutions).

182. See Richard D. Marsico, *Subprime Lending, Predatory Lending, and the Community Reinvestment Act Obligations of Banks*, 46 N.Y.L. SCH. L. REV. 735, 738–39 (2003) (citing 12 C.F.R. § 25.22(a)(1), (c)(1) (2010)).

183. See 12 C.F.R. §§ 25.22, 25.41 (2010).

184. See *id.* § 25.41 (2010).

its consistency with CRA goals when lending activity is conducted by covered financial institutions and occurs outside their assessment areas. Thus, even a covered institution can engage in predatory practices beyond the reach of the CRA simply by doing a less than substantial amount of lending in communities where they have no branches. Given the age of internet banking, this is fairly easy to do. Another way of looking at this is to say that the CRA was irrelevant to lending that took place outside of covered financial institutions' designated CRA assessment areas; as a result, risky lending that occurred outside of these areas could not have been driven by the CRA simply because the CRA did not cover such lending.

These exclusions from the CRA—i.e., that it covers only financial institutions that take deposits, bank subsidiaries that do not themselves take deposits at those parent banks' discretion, and lending in a covered bank's assessment area—leave gaping holes in CRA coverage. These gaps in coverage were particularly relevant to the expansion of subprime lending in the lead up to the financial crisis. In fact, a study by the Federal Reserve showed that at least 94% of all subprime lending during the height of the mortgage mania was outside the scope of the CRA.¹⁸⁵ Moreover, numerous studies have shown that CRA lending was actually as profitable and viable as other stable loans on banks' ledgers, and potentially less likely to enter into foreclosure than subprime loans.¹⁸⁶

185. Memorandum from Glenn Canner & Neil Bhutta, Div. of Research and Statistics, Bd. of Governors of the Fed. Res. Sys., to Sandra Braunstein, Dir., Consumer & Cmty. Affairs Div., Bd. of Governors of the Fed. Reserve Sys. 3 (Nov. 21, 2008), *available at* http://www.federalreserve.gov/newsevents/speech/20081203_analysis.pdf. It is also important to note that the 94% figure might actually *underestimate* the number of subprime loans that the CRA covered during the buildup of the subprime mortgage market: i.e., that figure might be lower than 6%. The Canner & Bhutta report did not differentiate between loans made by CRA-covered institutions and loans made by their affiliates. Thus, this percentage includes the lending carried out by affiliates of CRA-covered entities, which, as mentioned above, would only have been covered by the CRA at the discretion of the parent institution.

186. *See* Canner & Bhutta, *supra* note 185, at 10 tbl.7. A study by the Federal Reserve Bank of San Francisco of the loans issued in California from 2004 through 2006, the height of the boom in that state, revealed that mortgage loans made to borrowers of similar creditworthiness by financial institutions covered by the CRA, in those institutions' assessment areas, were half as likely to enter foreclosure as loans made by independent mortgage lenders not covered by the CRA. Elizabeth Laderman & Carolina Reid, *CRA Lending During the Subprime Meltdown*, in FED. RESERVE BANKS OF BOS. AND S.F., *REVISITING THE CRA: PERSPECTIVES ON THE FUTURE OF THE COMMUNITY REINVESTMENT ACT* 115, 118, 122 (2009), *available at* http://www.frbsf.org/publications/community/cra/revisiting_cra.pdf. On the profitability of lending covered by the CRA prior to the buildup of the subprime mortgage

In the end, the indictment against the CRA is not that it was too strong, but that it was too weak. Built to fight the civil rights battles of the twentieth century, in which communities were red-lined and excluded from traditional bank services and access to capital, it was ill-equipped to withstand the cascade of subprime lenders that flourished in certain communities. Such lenders succeeded precisely because of a legacy of discrimination in those communities that meant fewer legitimate bank options were available there.¹⁸⁷ Remember those 169 failed lending institutions discussed above:¹⁸⁸ One hundred sixty-seven of them, roughly 99%, were beyond the reach of the CRA because they were non-depository institutions and thus not covered by the CRA.¹⁸⁹ It is hard to lay the blame on the CRA when 94% of subprime lending took place outside its scope, and 99% of the riskiest lenders were not even covered by it.

B. *The Role of Government Sponsored Entities*

Turning to the GSEs, the theory goes that politicians drove those entities to increase their share of the subprime securitization pool, which led to an unsustainable expansion of the mortgage market, reaching borrowers that were poor credit risks.¹⁹⁰ But the facts concerning the GSEs' share of the market do not bear this out. Indeed, as the securitization market expanded in 2004 and 2005, both the size and the share of the GSEs' stake in that market actually declined. As the market started to dry up in 2006, their share got even smaller, as Table 17 reveals.

market, see Bd. OF GOVERNORS OF THE FED. RESERVE SYS., *THE PERFORMANCE AND PROFITABILITY OF CRA-RELATED LENDING* 45–51 (2000), available at <http://www.federalreserve.gov/boarddocs/surveys/craloansurvey/cratext.pdf>.

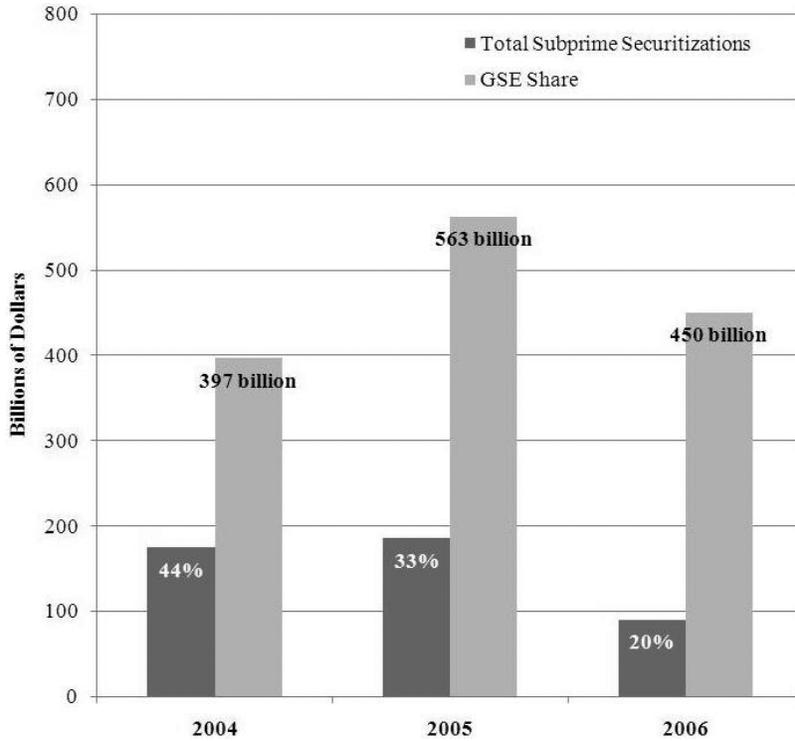
187. See, e.g., Kathleen C. Engel & Patricia A. McCoy, *The CRA Implications of Predatory Lending*, 29 *FORDHAM URB. L.J.* 1571, 1583–84 (2001) (noting that failure of traditional lenders to serve communities of color allows predatory lenders to charge supracompetitive rates); see also Barr, *supra* note 167, at 534–40 (2005) (providing economic reasons for failure of mortgage market to serve certain communities).

188. See *supra* text accompanying notes 145–46.

189. See Avery et al., *supra* note 145, at A123.

190. See RAJAN, *supra* note 83, at 32–45.

Table 17: GSE Share of All Subprime Securities, 2004–2006.¹⁹¹



Furthermore, although the GSEs *did* lower their underwriting standards in the early part of this decade, recent research by the National Community Reinvestment Coalition reveals that securitized subprime loans on the books of private investors, as compared to those on the GSEs’ ledgers, were more than twice as likely to go into default.¹⁹² Thus, if risky lending and poor underwriting standards are to blame for the present foreclosure crisis, then had the underwriting standards of the securitizations that the GSEs invested in been utilized across the board, foreclosure rates may have been significantly reduced.¹⁹³

191. See Carol D. Leonnig, *How HUD Mortgage Policy Fed the Crisis: Subprime Loans Labeled ‘Affordable’*, WASH. POST (June 10, 2008), <http://www.washingtonpost.com/wp-dyn/content/article/2008/06/09/AR2008060902626.html>.

192. See Jayasundera et al., *supra* note 161, at 3, 18 tbl.6.

193. See *id.* at 20. A related issue recently reached the Supreme Court, where the Court upheld the authority of state attorneys general to enforce their respective states’ fair lending laws. *Cuomo v. Clearing House Ass’n*, 129 S. Ct. 2710, 2717, 2722 (2009) (holding that National Bank Act only preempts sovereign’s “vistorial powers” and not its ability to enforce its own fair-lending laws).

It seems likely that non-depository institutions and private investors were driving the riskiest lending, not a political push generated by Washington through either the CRA or the GSEs.

IV. MOVING FORWARD: SOCIAL DISTANCE AND FINANCIAL REFORM

The preceding analysis raises important questions about the impact of racial and economic inequality on the health of national economies and the sustainability of certain economic practices. Given these impacts, efforts to rein in risky practices should take into account the way social inequalities, and the social distance they engender, can increase predatory conduct and weaken markets. There are several straightforward, macroeconomic approaches that could help to reduce social distance. Recalibrating marginal tax rates so as to increase taxes on the wealthy as a way to flatten differences between the haves and have-nots is one obvious way to reduce the social distance that comes from greater economic and social inequality. If economic inequality itself is not subject to direct assault, given political realities, the gravity of Judson's "governance problem" can be reduced by reducing the influence of money on the political process, though the present makeup of the U.S. Supreme Court holds a firm line against campaign finance reform.¹⁹⁴ Short of a constitutional amendment, the outsized influence of money on politics seems firmly in place. Similarly, governmental affirmative action strategies aimed at reducing economic inequality between Whites and people of color have raised fatal constitutional questions.¹⁹⁵ Yet with respect to racial inequality and the way it has

194. André Douglas Pond Cummings, *Procuring "Justice"?: Citizens United, Caperton, and Partisan Judicial Elections*, 95 IOWA L. REV. BULL. 89 (2010), available at http://www.uiowa.edu/~ilr/bulletin/ILRB_95_cummings.pdf. The author notes that especially since the holding in the *Citizens United* case, the Court has "[made] permissible a corporations direct expenditure of general treasury funds to explicitly endorse or malign a candidate for public office immediately prior to an election." *Id.* at 97 (citing *Citizens United v. Fed. Election Comm'n*, 130 S. Ct. 876, 964 (2010) (Stevens, J., concurring in part and dissenting in part)). This unwillingness to tighten campaign financing practices has been an identifiable trend in the U.S. Supreme Court since Chief Justice Roberts took his position. See Richard Briffault, WRTL and Randall: *The Roberts Court and the Unsettling of Campaign Finance Law*, 68 OHIO ST. L.J. 807, 807-08 (2007) (stating that two decisions made in Roberts' first year as Chief Justice "may constitute a pivotal moment in the Court's evolving campaign finance jurisprudence").

195. For a discussion of the constitutional barriers to public affirmative action strategies, see Jed Rubenfeld, *Affirmative Action*, 107 YALE L.J. 427, 428-30 (1997). The author asserts that "[t]he ultimate constitutional question presented by race-

played out in the mortgage arena, the legal infrastructure that might combat reverse redlining exists, but application of the laws against discrimination in lending during the last decade seems slow to catch up to invidious practices. This is due not only to the inadequacy of enforcement mechanisms and resources, but also to the fact that securitization permits downstream investors to claim protection from charges of discrimination in the origination of the underlying mortgages by invoking the “holder in due course” doctrine.¹⁹⁶

Since economic and racial inequality seem to bear some relation to financial crises, short of direct efforts to reduce such inequality—efforts that might seem politically infeasible or constitutionally suspect—to what extent does the recent financial reform legislation that Congress passed in the summer of 2010 address these forms of social inequality, if at all? The remainder of this section is devoted to this question, as well as to the question of whether there are additional legislative or market-oriented fixes, not covered in this legislation, that might rein in the effects of social inequality.

In July 2010, by a close—yet bi-partisan—vote, Congress passed what became known as the Dodd-Frank bill, named for Connecticut Senator Chris Dodd and Massachusetts Representative Barney Frank, both Democrats. Formally known as the “Dodd-Frank Wall Street Reform and Consumer Protection Act,” the legislation runs over 2000 pages and covers a wide range of topics, including increasing transparency in derivatives trading, improving regulatory oversight over financially important institutions, and strengthening regulatory authority to wind down institutions that pose systemic risks to the financial system.¹⁹⁷ At the same time, few in the general

based affirmative action—by no means an easy question—is whether whites’ equal protection rights are violated when the government purposefully acts to assist blacks and other minorities by granting them special opportunities.” *Id.* at 429.

196. Put in its most basic form, the holder in due course doctrine protects a purchaser of a financial instrument who buys that instrument for value without notice of any defects in prior transactions related to that instrument. For a description of the holder in due course doctrine, see Vern Countryman, *The Holder in Due Course and Other Anachronisms in Consumer Credit*, 52 TEX. L. REV. 1, 2–3 (1973). For criticisms of the doctrine, specifically as it relates to the securitization of subprime mortgages, see Kurt Eggert, *Held Up in Due Course: Predatory Lending, Securitization, and the Holder in Due Course Doctrine*, 35 CREIGHTON L. REV. 503 (2002); Siddhartha Venkatesan, Note, *Abrogating the Holder in Due Course Doctrine in Subprime Mortgage Transactions to More Effectively Police Predatory Lending*, 7 N.Y.U. J. LEGIS. & PUB. POL’Y 177 (2003).

197. Dodd-Frank Wall Street Reform and Consumer Protection Act, Pub. L. No. 111-203, 124 Stat. 1376 (2010).

public appear to have faith in the legislation to prevent future financial crises from occurring.¹⁹⁸ Moreover, the legislation does little to address the problem of economic and racial inequality and the social distance they create. In this section, those aspects of the Dodd-Frank legislation that appear to relate in some way to the inequalities discussed in previous sections are analyzed.

What follows is an exploration of some of the issues implicated by the analysis set forth above, and an attempt to identify potential areas of focus for regulators as they begin the difficult process of developing the regulations necessary to implement Dodd-Frank, as well as continue the investigation of predatory conduct that occurred during the lead up to the financial crisis. This analysis also identifies potential areas of focus for the new Consumer Financial Protection Bureau. Each of these topics will be discussed, in turn, below.

In addition, as this Article goes to print, Congress is expected to consider what could amount to wide-ranging amendments to the CRA, changes that would expand its coverage and close the loopholes that left much of the riskiest lending outside its coverage. This Article discusses these proposed amendments below. Finally, this section concludes with a discussion of the role that social capital can play in improving financial industry practices, with particular emphasis on its ability to serve as a consumer-driven lever for reform of the financial sector, regardless of the relative effectiveness of Dodd-Frank to rein in the riskiest, crisis-inducing practices of the shadow banking system.

A. *Social Distance, Dodd-Frank, and a New Regulatory Framework*

Not surprisingly, perhaps, Dodd-Frank explicitly says next to nothing about social inequality. Only in section 1204, entitled “Expanded Access to Mainstream Financial Institutions,” is any reference made to low- and moderate-income individuals. This section authorizes the Secretary of the Treasury to undertake the following:

198. A Bloomberg poll conducted on the eve of the passage of the Dodd-Frank legislation found the following:

Almost four out of five Americans surveyed in a Bloomberg National Poll this month say they have just a little or no confidence that the measure being championed by congressional Democrats will prevent or significantly soften a future crisis. More than three-quarters say they don't have much or any confidence the proposal will make their savings and financial assets more secure. Rich Miller, *Wall Street Fix Seen Ineffectual by Four of Five in U.S.*, BLOOMBERG, July 13, 2010, <http://www.bloomberg.com/news/2010-07-13/wall-street-fix-from-congress-seen-ineffectual-by-four-out-of-five-in-u-s.html>.

to establish a multiyear program of grants, cooperative agreements, financial agency agreements, and similar contracts or undertakings to promote initiatives designed—

(1) to enable low- and moderate-income individuals to establish one or more accounts in a federally insured depository institution that are appropriate to meet the financial needs of such individuals; and

(2) to improve access to the provision of accounts, on reasonable terms, for low- and moderate-income individuals.¹⁹⁹

Apart from this lone explicit mention, there are also other provisions in the legislation that will assist low- and moderate-income communities. For example, the legislation authorizes the Treasury Secretary to assess the risk to consumers from consumer financial products and services²⁰⁰ and to promote programs that will offer alternatives to high-cost, small dollar loans;²⁰¹ adopts mortgage protections geared towards ensuring that mortgage originators offer loan products that are appropriate for the borrowers they serve;²⁰² and places limits on mortgage broker compensation to ensure that brokers are not given incentives to steer borrowers into higher cost and less favorable loans.²⁰³ Each of these provisions will undoubtedly assist lower-income Americans. Moreover, the Bureau of Consumer Financial Protection will likely be inclined towards ensuring that lower-income Americans are not preyed upon by financial institutions using opaque and predatory products.²⁰⁴ The creation of this entity is probably the most significant aspect of the Dodd-Frank legislation that will impact social inequality, and the following section provides an overview of the bureau's powers and functions.

At the same time, Dodd-Frank does little to rein in executive compensation at regulated financial institutions. It is hard to think of a practice that did more to spur economic inequality, and the

199. Dodd-Frank Wall Street Reform and Consumer Protection Act § 1204(a)(2).

200. *Id.* § 1024(b)(1)(C).

201. *Id.* § 1205(a).

202. *Id.* § 1402–33.

203. Such incentives, typically called “Yield Spread Premiums,” often encouraged brokers to try to convince borrowers to accept more expensive loan terms because the broker was paid a higher commission for doing so. On the operation of these premiums, see Howell E. Jackson & Laurie Burlingame, *Kickbacks or Compensation: The Case of Yield Spread Premiums*, 12 STAN. J.L. BUS. & FIN. 289 (2007).

204. A discussion of the Consumer Financial Protection Bureau follows *infra* Part IV.A.1.

Judsonian governance problem mentioned earlier,²⁰⁵ than the growth of financial compensation for financial industry executives.²⁰⁶ What Dodd-Frank does do in this area is direct the development of regulations that will prohibit “excessive compensation” that could lead to excessive risk-taking at a large percentage of financial sector institutions.²⁰⁷ These regulations will cover a range of institutions, including bank holding companies, registered broker dealers, and credit unions, among others.²⁰⁸ In many ways, these new standards adopt similar controls to provisions already in existence related to excessive compensation controls on insured depository institutions.²⁰⁹

Other provisions of the legislation related to compensation include non-binding “say on pay” votes for shareholders,²¹⁰ disclosure requirements on public companies,²¹¹ and compensation clawback provisions for companies that must restate earnings.²¹² It is too early to determine whether these regulations will have teeth and whether regulators will enforce them vigorously.

The following discussion provides an overview of those other aspects of the Dodd-Frank legislation, beyond the modest compensation controls described above, that relate most directly to social inequality. In this discussion, this Article attempts to provide an overview of both the legislation’s specific provisions as well as the critical issues that will face regulators, the financial industry, and consumers as the legislation moves from passage to implementation.

1. The Bureau of Consumer Financial Protection

Dodd-Frank calls for the creation of an independent sub-agency within the Federal Reserve System that will have as its primary function maintaining oversight of federal consumer protection laws and monitoring the financial system for abusive and unfair practices. To be known as the “Bureau of Consumer Finan-

205. *See supra* text accompanying notes 65–67.

206. For a discussion of the potential impacts of executive compensation structures on the financial crisis, see Mark Hulbert, *Did Bankers’ Pay Add to This Mess?*, N.Y. TIMES, Sept. 27, 2009, at BU6, *available at* <http://www.nytimes.com/2009/09/27/business/27stra.html>, and the sources cited therein.

207. Dodd-Frank Wall Street Reform and Consumer Protection Act § 956(b).

208. *Id.* § 956(e).

209. *See* 12 U.S.C. § 1831p-1(c) (2006).

210. Dodd-Frank Wall Street Reform and Consumer Protection Act § 951.

211. *Id.* § 953.

212. *Id.* § 954.

cial Protection”²¹³ (or, as it is more commonly referred to, the Consumer Financial Protection Bureau, or CFPB), the Bureau will draw its funding from the Federal Reserve System,²¹⁴ and its director will be nominated by the President and confirmed by the U.S. Senate.²¹⁵

Section 1021 of Dodd-Frank lays out the CFPB’s purpose, objectives, and functions. According to the legislation, the bureau shall “seek to implement and, where applicable, enforce Federal consumer financial law consistently for the purpose of ensuring that all consumers have access to markets for consumer financial products and services and that markets for consumer financial products and services are fair, transparent, and competitive.”²¹⁶ The Bureau is tasked with enforcing a range of federal financial protection laws, including the Truth in Lending Act and the Equal Credit Opportunity Act.²¹⁷

213. *Id.* § 1011(a).

214. *Id.* § 1017(a).

215. *Id.* § 1011(b)(3).

216. Dodd-Frank Wall Street Reform and Consumer Protection Act § 1021(a).

217. According to Dodd-Frank, the consumer protection laws that the bureau has responsibility over include those laws designated as “enumerated consumer laws,” which are the following:

- (A) the Alternative Mortgage Transaction Parity Act of 1982 (12 U.S.C. 3801 et seq.);
- (B) the Consumer Leasing Act of 1976 (15 U.S.C. 1667 et seq.);
- (C) the Electronic Fund Transfer Act (15 U.S.C. 1693 et seq.), except with respect to section 920 of that Act;
- (D) the Equal Credit Opportunity Act (15 U.S.C. 1691 et seq.);
- (E) the Fair Credit Billing Act (15 U.S.C. 1666 et seq.);
- (F) the Fair Credit Reporting Act (15 U.S.C. 1681 et seq.), except with respect to sections 615(e) and 628 of that Act (15 U.S.C. 1681m(e), 1681w);
- (G) the Home Owners Protection Act of 1998 (12 U.S.C. 4901 et seq.);
- (H) the Fair Debt Collection Practices Act (15 U.S.C. 1692 et seq.);
- (I) subsections (b) through (f) of section 43 of the Federal Deposit Insurance Act (12 U.S.C. 1831t(c)–(f));
- (J) sections 502 through 509 of the Gramm-Leach-Bliley Act (15 U.S.C. 6802–6809) except for section 505 as it applies to section 501(b);
- (K) the Home Mortgage Disclosure Act of 1975 (12 U.S.C. 2801 et seq.);
- (L) the Home Ownership and Equity Protection Act of 1994 (15 U.S.C. 1601 note);
- (M) the Real Estate Settlement Procedures Act of 1974 (12 U.S.C. 2601 et seq.);
- (N) the S.A.F.E. Mortgage Licensing Act of 2008 (12 U.S.C. 5101 et seq.);
- (O) the Truth in Lending Act (15 U.S.C. 1601 et seq.);
- (P) the Truth in Savings Act (12 U.S.C. 4301 et seq.);
- (Q) section 626 of the Omnibus Appropriations Act, 2009 (Public Law 111–8); and

“The Bureau is authorized to exercise its authorities under Federal consumer financial law” to ensure that consumers are given effective information about consumer products and services,²¹⁸ and “consumers are protected from unfair, deceptive, or abusive acts and practices and from discrimination.”²¹⁹ The Bureau is also empowered to address “unduly burdensome regulations . . . in order to reduce unwarranted regulatory burdens.”²²⁰ It should also ensure fair competition among depository and non-depository institutions and that “consumer financial products and services operate transparently and efficiently to facilitate access and innovation.”²²¹ The “primary functions” of the Bureau will include engaging in consumer education; “collecting, investigating, and responding to consumer complaints;” monitoring markets to “identify risks to consumers and the proper functioning of such markets;” supervising compliance with and enforcing consumer protection laws; and issuing rules, orders, and guidance on consumer protection laws.²²²

With respect to rulemaking, the CFPB will have the primary enforcement authority over the federal consumer protection laws over which it has jurisdiction.²²³ At the same time, the Financial Stability Oversight Council,²²⁴ also created by Dodd-Frank,²²⁵ has the ability to stay or set aside any rules promulgated by the Bureau

(R) the Interstate Land Sales Full Disclosure Act (15 U.S.C. 1701).

Id. §§ 1002(12)(A)–(R), 1002(14), 1011(a).

218. *Id.* § 1021(b).

219. *Id.* § 1021(b)(2).

220. *Id.* § 1021(b)(3).

221. *Id.* § 1021(b)(4), (5).

222. *Id.* § 1021(c)(1)–(5).

223. *Id.* § 1025(c).

224. The Council will include the following voting members:

(A) the Secretary of the Treasury, who shall serve as Chairperson of the Council;

(B) the Chairman of the Board of Governors;

(C) the Comptroller of the Currency;

(D) the Director of the Bureau;

(E) the Chairman of the Commission;

(F) the Chairperson of the Corporation;

(G) the Chairperson of the Commodity Futures Trading Commission;

(H) the Director of the Federal Housing Finance Agency;

(I) the Chairman of the National Credit Union Administration Board; and

(J) an independent member appointed by the President, by and with the advice and consent of the Senate, having insurance expertise.

Id. § 111(b)(1).

225. *Id.* § 111(a).

on an affirmative vote of two thirds of the members of the Council.²²⁶

In addition, the Bureau can hold hearings and adjudicative proceedings concerning alleged violations of consumer protection laws,²²⁷ can file litigation in its own name,²²⁸ and is to create a fund to collect penalties and fines for violations of consumer protection laws to help to compensate victims of such violations.²²⁹

It is apparent that the focus of the CFPB will be to minimize information asymmetries, monitor the functioning of markets and the use of dangerous financial products, enforce federal consumer protection laws, and prosecute violations of these laws. These powers will likely inure to the benefit of individuals of low and moderate income the most, or those most susceptible to deceptive practices and those most in need of assistance guarding against such practices. In this way, the Bureau has the potential to be a strong counterweight to the significant financial and informational advantages that well-heeled institutions and intermediaries have over less sophisticated consumers. With so much primary authority over federal consumer protection laws, for such things as rulemaking and enforcement, it is critical that the CFPB is staffed with individuals who will take their role seriously, and will be aggressive and conscientious in carrying out the mandate and mission of the entity. Resting such broad, primary and comprehensive consumer protection authority in a single entity runs the risk that such power will not be wielded well or, worse, will give consumers a false sense of security, leading the way for deceptive practices that are more effective in taking advantage of consumers. “Cognitive regulatory capture” of the CFPB by individuals inclined to favor industry interests could, in fact, do more harm than good.²³⁰ Thus, ensuring that

226. *Id.* § 1023(c)(3)(A) (setting forth the two-thirds requirement). With respect to the specific authority to overturn a bureau’s rule, Section 1023(a) provides as follows:

Review of Bureau Regulations.—On the petition of a member agency of the Council, the Council may set aside a final regulation prescribed by the Bureau, or any provision thereof, if the Council decides, in accordance with subsection (c), that the regulation or provision would put the safety and soundness of the United States banking system or the stability of the financial system of the United States at risk.

Id. § 1023(a).

227. *Id.* § 1053(a).

228. *Id.* § 1054(b).

229. *Id.* § 1017(d).

230. For a discussion of cognitive regulatory capture, in which regulators adopt the mindset and interests of the entities they regulate, see WILLEM H.

the agency is staffed by individuals with independence and who possess the ability to navigate and coordinate the efforts of dozens of federal agencies will be an essential first step in guaranteeing the success of the CFPB in carrying out its purposes and helping to mitigate the effects of social distance. One aspect of this mission—the CFPB’s fair lending enforcement powers—will obviously be essential in addressing the effects of social distance. This authority will now be considered.

2. Strengthened Fair Lending Enforcement

Dodd-Frank mandates that the Bureau create an Office of Fair Lending and Equal Opportunity.²³¹ The powers of this Office include any that the Bureau Director may delegate to it, including the following:

- (A) providing oversight and enforcement of Federal laws intended to ensure the fair, equitable, and nondiscriminatory access to credit for both individuals and communities that are enforced by the Bureau, including the Equal Credit Opportunity Act and the Home Mortgage Disclosure Act;
- (B) coordinating fair lending efforts of the Bureau with other Federal agencies and State regulators, as appropriate, to promote consistent, efficient, and effective enforcement of Federal fair lending laws;
- (C) working with private industry, fair lending, civil rights, consumer and community advocates on the promotion of fair lending compliance and education; and
- (D) providing annual reports to Congress on the efforts of the Bureau to fulfill its fair lending mandate.²³²

Along a similar vein, earlier this year the Department of Justice (DOJ) announced the creation of a new Fair Lending Unit within the Department’s Civil Rights Division that will focus on discriminatory lending occurring now and in the lead up to the financial crisis.²³³

Given the fact that unfair lending was prevalent during the mortgage mania of the last decade, there is an obvious need for this

BUITER, LESSONS FROM THE NORTH ATLANTIC FINANCIAL CRISIS 37 (2008), *available at* <http://www.newyorkfed.org/research/conference/2008/rmm/buiter.pdf>.

231. Dodd-Frank Wall Street Reform and Consumer Protection Act § 1013(c)(1).

232. *Id.* § 1013(c)(2)(A)–(D).

233. Thomas E. Perez, Assistant Att’y Gen., Testimony Before the H. Subcomm. on the Constitution, Civil Rights, and Civil Liberties (Apr. 29, 2010), *available at* <http://www.justice.gov/crt/speeches/2010/crt-speech-100429.html>.

type of focus for the CFPB, as well as for the DOJ. These two entities should join forces with others both inside and outside the federal government to conduct a comprehensive assessment of lending practices during the lead up to the financial crisis. With the outstanding portfolios of Fannie Mae and Freddie Mac, those of the banks the FDIC has taken over in the last two years, and the records the Department of the Treasury has amassed through its mortgage modification program, there is a wealth of data available to the CFPB and the DOJ to identify discriminatory lending patterns. If many of the lenders responsible for the worst abuses have since dissolved, at least such investigations will help to uncover the extent to which discrimination may have played a part in the buildup of the mortgage market. Such information is critical to help develop a deeper appreciation for the causes of the crisis; to date, little has been done to root out such information, short of the efforts of a few plaintiff-side attorneys and intrepid state attorneys general.

Placing primary enforcement authority over such laws as the Equal Credit Opportunity Act and the Home Mortgage Disclosure Act with the CFPB means there is now a federal financial regulator with a primary focus on fair lending. This will help address one of the critical causes of social distance that played itself out with ferocity in the lead up to the financial crisis: economic inequality that falls along racial lines. In order to minimize the harshest consequences of this social distance, robust enforcement of the fair lending laws is necessary. Once again, in order for the CFPB to wield this authority in an effective way, it must be staffed by committed, competent and professional bureaucrats who will enforce both the letter and the spirit of the fair lending laws.

3. Reforming Mortgage Laws

Dodd-Frank imposes sweeping new restrictions on all home mortgage loans. Section 1411 of the legislation imposes the following requirement on residential mortgage lending:

[N]o creditor may make a residential mortgage loan unless the creditor makes a reasonable and good faith determination based on verified and documented information that, at the time the loan is consummated, the consumer has a reasonable ability to repay the loan, according to its terms, and all applica-

ble taxes, insurance (including mortgage guarantee insurance), and assessments.²³⁴

This requirement is similar to the suitability requirement applied in the securities context.²³⁵ A “reasonableness” requirement could go a long way towards offsetting the asymmetries of information that plagued the mortgage market during the last decade. Such asymmetries left borrowers at the mercy of mortgage brokers and originators who could mask the true import of mortgage terms, including critical interest rate adjustment formulas.

The mortgage context is one in which the design of the laws does not match the needs of many consumers. In this context, to date, we have relied heavily on a disclosure regime, one which places a great deal of somewhat murky information at the disposal of the borrower.²³⁶ Such a regime clearly failed; it was and is ill-suited to the task of reining in the modern, at times exotic, mortgage market.

If one traces the history of the development of other areas of property law, the law has been transformed to take into account the realities of the day.²³⁷ The warranty of habitability is perhaps the best example. In the landlord–tenant context, history shows that

234. Dodd-Frank Wall Street Reform and Consumer Protection Act § 1411(a)(2) (amending Chapter 2 of the Truth in Lending Act, 15 U.S.C. § 1631 et seq. (2006)).

235. For an overview of the suitability requirement with respect to the securities industry, see *Suitability*, SEC. & EXCH. COMM’N, <http://www.sec.gov/answers/suitability.htm> (last modified Oct. 16, 2009).

236. For an analysis of the failure of the mortgage lending disclosure regime to protect against abuses in the subprime mortgage market, see Jeff Sovern, *Preventing Future Economic Crises Through Consumer Protection Law or How the Truth in Lending Act Failed the Subprime Borrowers* (St. John’s Univ. Sch. of Law, Legal Studies Research Paper No. 10-0189, 2010) available at <http://ssrn.com/abstract=1531781>.

237. That our property laws are constantly evolving is beyond question, as Justice Stevens, in his dissenting opinion in *Lucas v. S.C. Coastal Council*, 505 U.S. 1003 (1992), made clear:

The human condition is one of constant learning and evolution—both moral and practical. Legislatures implement that new learning; in doing so they must often revise the definition of property and the rights of property owners. Thus, when the Nation came to understand that slavery was morally wrong and mandated the emancipation of all slaves, it, in effect, redefined “property.” On a lesser scale, our ongoing self-education produces similar changes in the rights of property owners: New appreciation of the significance of endangered species, see, e. g., *Andrus v. Allard*, 444 U.S. 51 (1979); the importance of wetlands, see, e. g., 16 U. S. C. § 3801 et seq.; and the vulnerability of coastal lands, see, e. g., 16 U. S. C. § 1451 et seq., shapes our evolving understandings of property rights.

Id., at 1069–1070 (Stevens, J., dissenting).

the landlord-tenant relationship changed from one in which a jack-of-all-trades farmer entered into a relationship with a land owner, primarily for the use of farmland, to one in which an unskilled tenant rented property in an urban setting.²³⁸ The law had to adapt to take into account the inability of the tenant to tend to the quality of his or her own housing and the inequity of bargaining power between the landowner and the tenant.²³⁹

In the mortgage context, the disclosure regime, one in which prospective borrowers received incomprehensible information with references to terms like “LIBOR” and the use of other methods to calculate the interest rate on an adjustable rate mortgage, clearly failed and many borrowers accepted loans that they could not understand with terms that were unsustainable.²⁴⁰

238. See *Javins v. First Nat'l Realty*, 428 F.2d 1071, 1078-79 (D.C. Cir. 1970), *cert. den.*, 400 U.S. 925 (1970).

239. For a history of the development of the warranty of habitability, see Mary Ann Glendon, *The Transformation of American Landlord-Tenant Law*, 23 B.C. L. REV. 503, 523-24 (1982). The sociological and economic forces at work that justified the warranty of habitability's transformation of the landlord-tenant relationship are perhaps best summed up by Judge J. Skelly Wright's opinion in the landmark *Javins v. First Nat'l Realty* case. 428 F.2d 1071 (D.C. Cir. 1970). In his opinion, Judge Wright explained those forces as follows:

Today's urban tenants, the vast majority of whom live in multiple dwelling houses, are interested, not in the land, but solely in “a house suitable for occupation.” Furthermore, today's city dweller usually has a single, specialized skill unrelated to maintenance work; he is unable to make repairs like the “jack-of-all-trades” farmer who was the common law's model of the lessee. Further, unlike his agrarian predecessor who often remained on one piece of land for his entire life, urban tenants today are more mobile than ever before. A tenant's tenure in a specific apartment will often not be sufficient to justify efforts at repairs. In addition, the increasing complexity of today's dwellings renders them much more difficult to repair than the structures of earlier times. In a multiple dwelling repair may require access to equipment and areas in the control of the landlord. Low and middle income tenants, even if they were interested in making repairs, would be unable to obtain any financing for major repairs since they have no long-term interest in the property.

Id. at 1078-79 (citations omitted).

240. Mortgage disclosure laws were ill-suited to prepare borrowers with little knowledge of sophisticated financial instruments to understand the terms of their mortgages. See Aaron Smith, Note, *A Suitability Standard for Mortgage Brokers: Developing a Common Law Theory*, 17 GEO. J. ON POVERTY L. & POL'Y 377, 389 (2010) (“The vast majority of consumers would likely be considered unsophisticated in the details of complex modern mortgages, with the possible exception of people who still get standard fixed rate mortgages. Modern circumstances demand that mortgage brokers not take advantage of this understandable reliance.”). For example, some adjustable mortgage rates were pegged to the London Interbank Offered Rate (LIBOR), the floating rate at which banks borrow money from other banks on the London wholesale money market. See Alex M. Johnson, Jr., *Preventing a Return En-*

Moving forward, it is critical to recognize that a more robust body of laws is necessary to protect borrowers in the real world. The “reasonable ability to repay” standard embodied in Dodd-Frank requires that mortgage originators and brokers must ensure that a particular loan is appropriate for a particular borrower.²⁴¹ This is a logical and necessary step to ensure that the mortgage feeding frenzy fed by imbalances in the level of sophistication and knowledge between borrowers and lenders does not occur again. Admittedly, this would not protect any of the millions of borrowers presently underwater, but it might help prevent the type of predatory conduct that appears to have been so prevalent before the crisis.

4. Weakening Federal Preemption of State Consumer Protection Laws

Another important reform accomplished through Dodd-Frank is the partial rollback of efforts by several federal banking regulators to preempt state consumer protection laws. In the 1990s, the Office of Thrift Supervision (OTS) was authorized to promulgate regulations pursuant to the Home Owners’ Loan Act (HOLA)²⁴² that could preempt any state laws affecting “the operations of federal savings associations.”²⁴³ The regulations could preempt any state law that imposed licensing requirements on OTS-regulated entities and attempted to regulate the interest rates such entities could charge.²⁴⁴ In 2003, the OTS issued letters declaring that federal laws and regulations preempted state predatory lending laws as they affected OTS-regulated entities in a range of states.²⁴⁵

agement: Eliminating the Mortgage Purchasers’ Status as a Holder-In-Due-Course: Properly Aligning Incentives Among the Parties, 37 PEPP. L. REV. 529, 572 nn.152–54 (2010) (providing relevant language present in Freddie Mac Single-Family Uniform Instrument for Multistate Initial Interest Adjustable Rate Notes). For a discussion of the connection between LIBOR and adjustable rate mortgages, see Bruce Yandle, *Lost Trust: The Real Cause of the Financial Meltdown*, 14 INDEP. REV. 341, 348 (2010); see also Olaf Clemens, *Accounting Discretion, Securitization, and the Subprime Crisis: An Accounting-based Analysis of the Subprime Market* 22 (Jan. 15, 2010), available at <http://www.fma.org/NY/Papers/AccountingDiscretionSecuritizationandSubprimeCrisis.pdf>.

241. Dodd-Frank Wall Street Reform and Consumer Protection Act, Pub. L. No. 111-203, §§ 1402–03, 124 Stat. 1376, 2038–41 (2010).

242. 12 U.S.C. §§ 1461–70 (2006).

243. 12 C.F.R. § 560.2(a) (2010).

244. *Id.* § 560.2(b).

245. Those states were: Georgia (*see* OFFICE OF THRIFT SUPERVISION, DEP’T OF THE TREASURY, P-2003-1, PREEMPTION OF GEORGIA FAIR LENDING ACT (2003), available at <http://files.ots.iteas.gov/56301.pdf>); New York (*see* OFFICE OF THRIFT SUPERVISION, DEP’T OF THE TREASURY, P-2003-2, PREEMPTION OF NEW YORK PREDATORY

In January 2004, the Office of the Comptroller of the Currency (OCC) followed the lead of the OTS and issued its own regulations that preempted state laws that “obstruct, impair, or condition a national bank’s ability to fully exercise its Federally authorized real estate lending powers”²⁴⁶ This regulation preempted any state laws that related to such matters as the terms of mortgages, escrow account requirements for loans, loan-to-value ratios, and a host of other mortgage-related issues.²⁴⁷ The OCC also strengthened its examination powers, declaring that it had the exclusive powers to visit, examine, and inspect the banks under its supervision.²⁴⁸

The question of regulatory preemption ultimately reached the Supreme Court in 2006, with the Court siding with the federal regulators, finding that “state regulators [could not] interfere with the ‘business of banking’ by subjecting [federally regulated] national banks or their OCC-licensed operating subsidiaries to [state] audits and surveillance under rival oversight regimes.”²⁴⁹ This decision of the Court, as one state attorney general pronounced, “took 50 sheriffs off the beat at a time when lending was becoming the Wild West.”²⁵⁰

Even with a strong and independent consumer protection bureau, states should be authorized to regulate financial sector conduct within their borders. If the delinquency and other data discussed here shows anything, it is that predatory conduct occurred at different rates and with different virulence in different states. As such, states must have the tools at their disposal to regulate lending and other financial services when such services threaten their citizens’ financial security.

Under Dodd-Frank, federal law as it relates to national banks will only preempt state law under certain pre-conditions. First, a state consumer law will be preempted if its application will discrimi-

LENDING LAW (2003), available at <http://files.ots.treas.gov/56302.pdf>); New Jersey (see OFFICE OF THRIFT SUPERVISION, DEP’T OF THE TREASURY, P-2003-5, PREEMPTION OF NEW JERSEY PREDATORY LENDING ACT (2003), available at <http://files.ots.treas.gov/56305.pdf>); and New Mexico (see OFFICE OF THRIFT SUPERVISION, DEP’T OF THE TREASURY, P-2003-6, PREEMPTION OF NEW MEXICO HOME LOAN PROTECTION ACT (2003), available at <http://files.ots.treas.gov/56306.pdf>).

246. 12 C.F.R. § 34.4(a) (2010).

247. *Id.*

248. *Id.* § 7.4000(a).

249. *Watters v. Wachovia Bank, N.A.*, 550 U.S. 1, 21 (2007).

250. Jo Becker et al., *White House Philosophy Stoked Mortgage Bonfire*, N.Y. TIMES, Dec. 21, 2008, at A1 (quoting Roy Cooper, Att’y Gen. of N.C.).

nate against national banks in favor of state chartered banks.²⁵¹ Second, preemption will apply if the state law “prevents or significantly interferes with” the national bank’s exercise of its charter powers.²⁵² In this way, the legislation returns federal preemption to its state prior to OCC and OTS preemption, consistent with the Supreme Court’s decision in *Barnett Bank of Marion County, N.A. v. Nelson*.²⁵³ Or, third, state law will be preempted as it applies to national banks if other federal laws expressly preempt the application of state laws to such national banks.²⁵⁴ Thus, if a state law does not meet any of these criteria, state enforcement authorities can apply it to national banks. The law also exempts subsidiaries of national banks (unless those subsidiaries themselves are chartered as national banks) from the preemptive effect of Dodd-Frank.²⁵⁵

Curtailing the ability of federal regulators to preempt state laws will go a long way towards permitting state legislatures, and the state executive branch officials who enforce state laws, to adapt their laws and to provide oversight over their state’s needs. Different markets call for different responses, as well as experimentation with different types of consumer protection regimes. States should be allowed to carry out consumer protection activities without facing the shielding effects of federal preemption unless important issues of federalism—like protecting federally chartered institutions from unfair competition from state-chartered institutions—are implicated by a particular state’s practices towards national banks acting within its borders.

251. Dodd-Frank Wall Street Reform and Consumer Protection Act, Pub. L. No. 111-203, § 1044(a), 124 Stat. 1376, 2015 (2010) (creating 12 U.S.C. § 5136C).

252. *Id.*

253. *See* 517 U.S. 25, 33 (1996).

254. Dodd-Frank Wall Street Reform and Consumer Protection Act § 1044(a). For an analysis of the preemption pre-conditions, see Linda Singer et al., *Breaking Down Financial Reform*, 14 J. CONSUMER & COM. L. 2, 10–11 (2010); *see also Barnett*, 517 U.S. at 31–33.

255. Dodd-Frank describes the application of preemption to subsidiaries of national banks as follows:

Notwithstanding any provision of this title or section 24 of Federal Reserve Act (12 U.S.C. 371), a State consumer financial law shall apply to a subsidiary or affiliate of a national bank (other than a subsidiary or affiliate that is chartered as a national bank) to the same extent that the State consumer financial law applies to any person, corporation, or other entity subject to such State law.

Dodd-Frank Wall Street Reform and Consumer Protection Act § 1044(a).

B. The Community Reinvestment Modernization Act

After passage of the Dodd-Frank legislation, Congress considered addressing CRA reform. In March of 2009, Rep. Eddie Johnson introduced in the U.S. House of Representatives the Community Reinvestment Modernization Act of 2009 (CMRA).²⁵⁶ The goals of the CMRA are to close the CRA's loopholes and to extend its coverage to a range of financial industry actors not currently covered by the CRA.²⁵⁷ The changes proposed to the CRA are sweeping, and Congress is unlikely to adopt them without a serious clash over the purposes and goals of the CRA, as well as an airing of the accusations that it was to blame for the crisis. Indeed, Rep. Jeb Hensarling has introduced competing legislation that would repeal the CRA entirely.²⁵⁸ While it is unlikely that CRA repeal could get through Congress or past a presidential veto, some form of CRA reform is possible, and the following is a description of the key elements of the CMRA.

The CMRA attempts to accomplish seven major changes to the CRA. First, it calls for the repeal of regulations enacted in the last decade that imposed a lighter CRA burden on mid-sized banks.²⁵⁹ Second, it would overhaul the ratings process by adding to the range of ratings that can be assigned to institutions covered by the CRA (adding grades of "low satisfactory" and "high satisfactory"),²⁶⁰ requiring the generation of a separate CRA grade for each financial institution's assessment area,²⁶¹ expanding the definition of CRA assessment area²⁶² and requiring regulators to downgrade banks found to have engaged in predatory lending and other unfair prac-

256. H.R. 1479, 111th Cong. (2009).

257. The CMRA has the following "purposes":

- (1) To enhance the availability of financial services to citizens of all economic circumstances and in all geographic areas.
- (2) To enhance the ability of financial institutions to meet the capital and credit needs, and needs for other banking and financial services of all citizens and communities, including and especially minority and low- and moderate-income communities and populations.
- (3) To ensure that community reinvestment keeps pace with developments in the financial industry and with the affiliation of banks, securities firms, and other financial service providers, as provided by the Gramm-Leach-Bliley Act.

Id. § 3.

258. Fair Access to Credit and Job Creation Act of 2010, H.R. 5038, 111th Cong. (2010).

259. H.R. 1479 § 101.

260. *Id.* § 103(a).

261. *Id.*

262. *Id.*

tices.²⁶³ Third, it would require any financial institution receiving a failing CRA grade (i.e., “Low Satisfactory,” “Needs to Improve,” or “Substantial Noncompliance”) in any assessment area to submit a corrective action plan to its regulator.²⁶⁴ Fourth, it would expand the types of financial institutions covered by the CRA to include non-depository mortgage lenders,²⁶⁵ insurance companies,²⁶⁶ securities companies²⁶⁷ and credit unions.²⁶⁸ It would also make CRA coverage of affiliates of covered institutions mandatory.²⁶⁹ Fifth, it would impose new reporting and disclosure requirements on covered institutions.²⁷⁰ Sixth, it would make concern for communities of color an explicit purpose of the CRA, whereas now the CRA only addresses bank services to “low- and moderate-income communities” with no explicit mention of the race of those communities.²⁷¹ Seventh, it would enhance the regulators’ ability to conduct public hearings on financial institution applications for which CRA review is appropriate.²⁷²

These changes are not only welcome, but are necessary if the CRA—perhaps the strongest federal legislation designed to lower barriers to credit for low- and moderate-income communities—is to serve its core purposes. Moreover, short of a progressive tax structure and funding for public education, the CRA serves as one of the few federal laws specifically designed to lower social distance, to the extent that such distance is created and strengthened by an entrenched financial system that remains resistant to overcoming a legacy of racial discrimination. As discussed above, the CRA’s many loopholes, and its anachronistic focus on narrowly drawn assess-

263. *Id.* § 104(a).

264. *Id.* § 103(a).

265. *Id.* § 108(a), (b)(3).

266. *Id.* § 109(a).

267. *Id.* § 107(a).

268. *Id.* § 111.

269. *Id.* § 102, 110.

270. *See, e.g.*, H.R. 1479 § 105 (expanding Equal Credit Opportunity Act to include reporting requirements for small business lending); *see also id.* § 203 (requiring development of reporting methods for insurance companies).

271. *Id.* § 2(3).

272. *Id.* § 303. More recently, a bill was introduced in the House of Representatives, the American Community Investment Act of 2010, H.R. 6334, 111th Cong. (2010), which attempts to accomplish many of the same goals of the CRMA in a more streamlined fashion. *See* Press Release, Rep. Gutierrez et al., Rep. Gutierrez Initiates Effort to Modernize Community Reinvestment Act: In Introducing American Community Investment Reform Act of 2010 (Sept. 29, 2010), *available at* http://www.gutierrez.house.gov/index.php?option=com_content&task=view&id=617&Itemid=55.

ment areas and depository institutions, meant that it was ill-suited to prevent the rush of subprime lending in communities that the CRA was ostensibly designed to protect. In order to serve the CRA's purposes, and make it relevant to banking in the twenty-first century, the changes proposed through the CMRA seem wise and necessary.²⁷³

*C. Social Capital, Social Distance, Regulatory Reform, and
Financial Crises*

Financial institutions that did not engage in risky lending or enter the market in exotic derivatives during the buildup to the crisis have weathered it much better than the large commercial and investment banks.²⁷⁴ Without a massive infusion of Troubled Asset Relief Program (TARP) funds, many more banks would have certainly collapsed.²⁷⁵ The faceless, nameless, opaque transactions that brought down some of the most venerable financial institutions were devoid of social capital—ties of trust and reciprocity that reduce the risk of non-cooperative or predatory conduct.

At the same time, financial transactions infused with social capital—the type of lending that occurred at community banks and other financial institutions that were “lower to the ground”—performed better than the innovators of high finance. One recent study of microlending revealed that increased contact between participants in meetings of borrowers decreased both social distance and default rates by participants in the program.²⁷⁶ Is there a way to

273. For a further discussion of the CMRA and additional ways to strengthen the CRA, see Raymond H. Brescia, *A CRA for the 21st Century: Congress Considers the Community Reinvestment Modernization Act of 2009*, 28 BANKING & FIN. SERVICES POL'Y REP., Oct. 2009, at 1.

274. See Raymond H. Brescia, *Trust in the Shadows: Law, Behavior, and Financial Re-Regulation*, 57 BUFF. L. REV. 1361, 1419–22 (2009) (and citations contained therein).

275. See Donny Wise, *Rapid Reaction via TARP Diverts Collapse at the Cost of Accountability*, EXAMINER.COM (Apr. 11, 2010, 4:30 AM), <http://www.examiner.com/consumer-and-banking-in-washington-dc/rapid-reaction-via-tarp-diverts-collapse-at-the-cost-of-accountability> (“As the economic crisis spread rapidly in September 2008, the government intervention in the financial system via TARP was a necessity, and the quick action saved the economic and societal collapse of the United States.”).

276. Benjamin Feigenberg et al., *Building Social Capital through Microfinance* 3–4 (Nat'l Bureau of Econ. Research, Working Paper No. 16018, 2010), available at <http://www.nber.org/papers/w16018>. Similarly, another study of microlending found that a high level of trust between borrower and lender improved repayment rates. Asif Dowla, *In Credit We Trust: Building Social Capital by Grameen Bank in Bangladesh*, 35 J. SOCIO-ECON. 102, 108 (2006). A study of domestic credit unions

take such lessons to scale and capture the benefits of social capital in financial transactions?

A grassroots movement is afoot that attempts to do just that. The “Move Your Money” campaign was hatched by Arianna Huffington and others in late 2009 as a way to bring community banking values back to retail banking.²⁷⁷ This effort is designed to convince retail banking consumers to utilize the banking services of financial institutions with strong ties to local communities.²⁷⁸ Using a risk assessment tool developed by Institutional Risk Analytics,²⁷⁹ the Move Your Money Campaign claims that over 2 million participants, moving \$5 billion in assets, have ended their relationship with their large banking institutions.²⁸⁰ Institutional and governmental investors have also adopted similar tactics. Municipalities, universities, unions, and public pension funds have either pledged to invest and bank with community-based institutions or have attempted to exact compliance with demands for behavioral changes from the larger institutions, like improved performance in terms of modifying mortgages.²⁸¹

found that longer-term credit union members had better repayment rates on micro-loans than new members. See Marva Williams, *Cooperative Credit: How Community Development Credit Unions are Meeting the Need for Affordable, Short-Term Credit*, WOODSTOCK INST., 17 (May 2007), http://www.economicintegrity.org/pdf/cooperativecredit_may2007_williams-1.pdf.

277. Arianna Huffington & Rob Johnson, *Move Your Money: A New Year's Resolution*, HUFFINGTON POST, Dec. 29, 2009, http://www.huffingtonpost.com/arianna-huffington/move-your-money-a-new-yea_b_406022.html. For a critique of the Move Your Money campaign, see Martha C. White, *Does Boycotting Big Banks Make Sense? No. (Sorry, Arianna)*, SLATE: THE BIG MONEY (Jan. 4, 2010), available at <http://www.thebigmoney.com/articles/hey-wait-minute/2010/01/04/does-boycotting-big-banks-make-sense?page=full>.

278. Huffington & Johnson, *supra* note 277 (The movement advocates that “[i]t’s time for Americans to move their money out of these reckless behemoths [B]ig banks are the core of the problem. We need to return to the stable, reliable, people-oriented approach of America’s community banks.”).

279. See *Move Your Money, Special Report: Big Four Banks Deposits Change Survey*, INSTITUTIONAL RISK ANALYTICS, http://us1.irabankratings.com/MoveYourMoney/IRADeposits_MYM_TBTF.asp (last visited Dec. 29, 2010).

280. Dennis Santiago, *For Biggest Banks, Deposits Remain Sticky*, HUFFINGTON POST, May 10, 2010, http://www.huffingtonpost.com/dennis-santiago/for-biggest-banks-deposit_b_569655.html (dividing net deposit outflows for major banks by assumed average account value of \$2,500).

281. See, e.g., Steven Greenhouse, *New York Presses Banks on Foreclosures*, N.Y. TIMES, July 13, 2010, at B3, available at <http://www.nytimes.com/2010/07/14/business/14foreclose.html> (detailing efforts of elected official and union leaders to pressure banks to improve loan modification efforts); Steve Flynn, *A Proposal for New Thinking on University Endowments and Community Investment*, RESPONSIBLE ENDOWMENTS COALITION BLOG (July 20, 2010, 5:52 PM), <http://www.endowment>

Efforts to break up the large banks by putting asset caps on such banks—by, for example, pegging their size to no larger than a percentage of U.S. gross national product—failed during the creation of financial reform legislation.²⁸² Instead, a consumer movement—both individual and institutional—may stand a better chance than regulation of re-infusing financial transactions with elements of social capital, by driving the reform of institutions, both large and small, to take less risk, maintain better customer relations, reinvest in local communities and meet local priorities. Such efforts are well under way, yet it would be premature to attempt to gauge their success in meeting such goals.

V. CONCLUSION

If we are to believe that a lack of trustworthiness and an excess of predatory conduct helped feed the subprime mortgage bonfire, it is essential to restore checks and balances to the financial system that would restrain such conduct and incentivize trustworthy behavior in mortgage lending and other financial practices. We also need to reduce social distance, both economic and racial. To do so would require both a greater enforcement of the laws on the books—the creation of the CFPB and the DOJ's recent announcement of the creation of a new fair lending unit are also helpful—as well as a strengthening of laws like the Community Reinvestment Act by closing loopholes that permitted predatory lenders to thrive.

We also have to recognize the role that law can play in fostering trustworthy conduct. Behavioral economists like Dan Ariely tell us that without rules, people cheat.²⁸³ Members of the law and economics school, like Richard Posner, tell us that the microeconomic incentives in place in the lead up to the financial crisis helped to

ethics.org/news-media/archives/305 (urging investing university endowments with community banks).

282. See *Implications of the 'Volcker Rules' for Financial Stability: Hearings Before the S. Banking Comm.*, 111th Cong. 5 (2010) (testimony of Simon Johnson, Professor, MIT Sloan Sch. of Mgmt.), available at http://banking.senate.gov/public/index.cfm?FuseAction=Files.View&FileStore_id=98845464-f06e-4f20-b1bd-cbce9a87803c (opposing asset caps on banks pegged to GDP). The Brown-Kaufman amendment would have adopted such asset caps as part of Dodd-Frank, but the amendment failed. Tim Fernholz, *On the Death of Brown-Kaufman*, TAPPED: THE GROUP BLOG OF THE AM. PROSPECT (May 7, 2010, 4:16 PM), http://www.prospect.org/csnc/blogs/tapped_archive?base_name=on_the_death_of_brownkaufman&month=05&year=2010.

283. See DAN ARIELY, *PREDICTABLY IRRATIONAL: THE HIDDEN FORCES THAT SHAPE OUR DECISIONS* 213 (revised ed. 2009).

2011] SOCIAL DISTANCE AND THE FINANCIAL CRISIS 721

encourage risky behavior and brought about the crisis.²⁸⁴ The analysis presented in this Article suggests that an exploration of the role that greater social distance plays in decreasing trust and increasing predatory conduct deserves its place at the table in conversations about how to reform mortgage markets, and markets in general.

284. See POSNER, *supra* note 75, at 75–112.

APPENDIX A

Data Set 1:

STATE	MEDIAN HOUSEHOLD INCOME	MEDIAN HOUSEHOLD INCOME BY RACE					
	MHI by State (Thousands)	White	White Rank	Black	Black Rank	Latino	Latino Rank
Alabama	38,783	71,195	2	51,726	1	58,093	1
Alaska	59,393	61,616	7	49,486	2	55,909	2
Arizona	47,265	65,641	5	40,499	13	55,322	3
Arkansas	36,599	59,560	10	43,993	5	55,048	4
California	56,645	61,067	8	40,267	14	54,087	5
Colorado	52,015	47,970	28	0	51	45,899	6
Connecticut	63,422	55,225	15	43,027	6	45,392	7
Delaware	52,833	69,836	3	44,866	4	45,049	8
District of Columbia	51,847	56,379	12	33,151	24	43,877	9
Florida	45,495	60,415	9	40,709	11	43,805	10
Georgia	46,832	48,528	26	41,170	10	43,660	11
Hawaii	61,160	88,969	1	34,484	22	43,547	12
Idaho	42,865	47,749	30	32,554	25	40,510	13
Illinois	52,006	55,907	13	41,640	8	39,199	14
Indiana	45,394	46,691	35	23,986	46	38,776	15
Iowa	44,491	44,494	40	48,623	3	37,786	16
Kansas	45,478	53,936	19	33,563	23	37,692	17
Kentucky	39,372	35,454	51	21,554	50	37,631	18
Louisiana	39,337	47,412	31	27,140	38	37,439	19
Maine	43,439	49,566	23	41,198	9	37,288	20
Maryland	65,144	52,179	20	39,744	15	37,157	21
Massachusetts	59,963	55,724	14	30,120	29	36,850	22
Michigan	47,182	45,331	37	27,808	37	36,738	23
Minnesota	54,023	50,309	22	31,276	28	36,372	24
Mississippi	34,473	54,118	18	40,659	12	36,222	25
Missouri	42,841	67,852	4	41,648	7	36,217	26
Montana	40,627	57,021	11	37,107	18	36,147	27
Nebraska	45,474	40,521	49	26,595	40	36,098	28
Nevada	52,998	54,690	17	34,563	21	35,941	29
New Hampshire	59,683	45,222	38	27,017	39	35,789	30
New Jersey	64,470	47,036	33	29,293	31	35,744	31
New Mexico	40,629	46,525	36	28,423	35	35,639	32
New York	51,384	47,291	32	29,309	30	35,484	33
North Carolina	42,625	43,139	45	21,969	49	35,378	34
North Dakota	41,919	47,902	29	26,473	41	35,313	35
Ohio	44,532	50,794	21	26,161	42	34,332	36

2011] SOCIAL DISTANCE AND THE FINANCIAL CRISIS 723

Oklahoma	38,770	46,946	34	29,076	34	34,220	37
Oregon	46,230	43,731	42	38,006	17	33,789	38
Pennsylvania	46,259	49,207	24	32,159	26	33,354	39
Rhode Island	51,814	48,179	27	29,243	32	33,187	40
South Carolina	41,100	48,768	25	29,111	33	32,454	41
South Dakota	42,791	43,347	43	32,074	27	32,085	42
Tennessee	40,315	42,919	46	28,067	36	32,049	43
Texas	44,922	44,788	39	25,203	43	31,930	44
Utah	51,309	43,317	44	22,931	48	31,516	45
Vermont	47,665	43,915	41	35,941	19	31,343	46
Virginia	56,277	40,009	50	23,265	47	30,704	47
Washington	52,583	55,133	16	35,183	20	30,499	48
West Virginia	35,059	62,443	6	38,496	16	30,140	49
Wisconsin	48,772	41,741	47	24,119	45	29,838	50
Wyoming	47,423	41,604	48	24,365	44	27,165	51

724 NYU ANNUAL SURVEY OF AMERICAN LAW [Vol. 66:641

Data Set 2:

STATE	INCOME INEQUALITY		PERCENT OF STATE POPULATION: BLACK		TRUST	POVERTY
	Inequality Rate	Inequality Rank	% Black Pop.	% Black Pop. Rank	GSS: "Most people can be trusted"	Percentage of people in poverty
Alabama	0.475	46	7.2	24	23%	16.6
Alaska	0.402	3	.	.	.	10.9
Arizona	0.45	29	37.3	48	47%	14.2
Arkansas	0.458	35	29.6	46	29%	17.3
California	0.475	42	26.2	43	43%	13.1
Colorado	0.438	28	31.5	47	46%	12
Connecticut	0.477	49	16.7	39	49%	8.3
Delaware	0.429	15	28.6	44	.	11.1
District of Columbia	0.549	51	.	.	.	19.6
Florida	0.47	43	3.2	13	37%	12.6
Georgia	0.461	38	21.2	42	38%	14.7
Hawaii	0.434	18	.	.	.	9.3
Idaho	0.427	6	7.3	25	.	12.6
Illinois	0.456	40	11.4	31	44%	12.3
Indiana	0.424	13	15.5	38	43%	12.7
Iowa	0.418	7	14.9	37	56%	11
Kansas	0.435	20	13.2	33	55%	12.4
Kentucky	0.468	35	14.8	36	32%	17
Louisiana	0.483	48	1.8	11	33%	19
Maine	0.434	10	19.4	40	.	12.9
Maryland	0.434	14	28.7	45	42%	7.8
Massachusetts	0.463	38	14.6	35	46%	9.9
Michigan	0.44	23	10.2	29	51%	13.5
Minnesota	0.426	11	6.0	23	63%	9.8
Mississippi	0.478	45	11.8	32	17%	21.1
Missouri	0.449	26	7.3	25	44%	13.6
Montana	0.436	9	8.7	27	64%	13.6
Nebraska	0.424	11	4.6	19	.	11.5
Nevada	0.436	15	20.4	41	.	10.3
New Hampshire	0.414	3	14.0	34	62%	8
New Jersey	0.46	33	3.2	13	41%	8.7
New Mexico	0.46	32	0.4	1	.	18.5
New York	0.499	50	11.3	30	40%	14.2
North Carolina	0.452	33	5.7	21	27%	14.7

2011] SOCIAL DISTANCE AND THE FINANCIAL CRISIS 725

North Dakota	0.429	15	9.2	28	67%	11.4
Ohio	0.441	26	5.5	20	41%	13.3
Oklahoma	0.455	35	3.6	16	39%	17
Oregon	0.438	23	0.8	5	55%	13.3
Pennsylvania	0.452	30	1.0	8	44%	12.1
Rhode Island	0.457	21	1.6	10	52%	11.1
South Carolina	0.454	40	5.8	22	34%	15.7
South Dakota	0.434	19	3.3	15	.	13.6
Tennessee	0.465	44	0.6	3	36%	16.2
Texas	0.47	47	1.0	8	33%	16.9
Utah	0.41	1	2.2	12	54%	10.6
Vermont	0.423	5	4.0	17	59%	10.3
Virginia	0.449	31	0.5	2	38%	9.6
Washington	0.436	22	4.4	18	46%	11.8
West Virginia	0.468	25	0.8	5	30%	17.3
Wisconsin	0.413	7	0.7	4	52%	11
Wyoming	0.428	2	0.9	7	60%	9.4

Data Set 3:

TABLE	P-Value
Table 6: Income Inequality and Delinquency Rates by State	0.011224
Table 7: Poverty Rate	0.800896
Table 8: Median Income by State	0.056502
Table 9: Median Income by Race	White: 0.061754
	Black: 0.048303
	Latino: 0.501179
Table 10: Trust and Delinquencies	0.070695
Table 11: Trust and Crime Rate Rank	0.000678
Table 12: Delinquency Rate Rank and Crime Rate Rank	0.097418
Table 13: Social Capital and Unemployment Data	<8 Unemployment Rate: 0.017062
	8-10 Unemployment Rate: 0.047184
	>10 Unemployment Rate: 0.752530
Table 15: Targeting the Black Middle Class	Overall: 0.0000583 or 5.83×10^{-5} State Level Inequality and Delinquencies: 0.011224 Percentage African American Population and Delinquencies: 0.039541 African American Median Income and Delinquencies: 0.048303

DEMOCRACY AT THE CORNER OF FIRST AND FOURTEENTH: JUDICIAL CAMPAIGN SPENDING AND EQUALITY

JAMES SAMPLE*

Save for the notorious crack/powder distinction, it's hard to think of a line that has been subjected to more withering criticism over the years than *Buckley's* expenditure/contribution distinction.¹

—Pamela S. Karlan

In invalidating some of the existing checks on campaign spending, . . . the majority in *Citizens United* has signaled that the problem of campaign contributions in judicial elections might get considerably worse and quite soon. . . . [I]f both [unions and corporations] unleash their campaign spending monies without restrictions, then I think mutually-assured destruction is the most likely outcome.²

—Justice Sandra Day O'Connor

ABSTRACT

The Supreme Court recently decided in Caperton v. A.T. Massey Coal Co. that substantial independent expenditures in support of a judicial candidate present threats to judicial impartiality similar to those posed by direct contributions. This Article posits that the Caperton holding, guaranteeing due process of law in state courts, presents a compelling state interest justifying the regulation of spending in judicial elections.

* Associate Professor of Law, Hofstra University School of Law; J.D., Columbia Law School, 2003; B.A., Boston College, 1995. Sincere thanks to William Andersen, Mary Clark, Amanda Frost, Rick Hasen, Michael Gerhardt, Charlie Geyh, Justin Levitt, Darryl Stein, and Jay Tidmarsh for helpful comments. By way of disclosure, I served as counsel on certiorari and merits stage amicus briefs in support of the petitioners in *Caperton v. A.T. Massey Coal Co.*, 129 S. Ct. 2252 (2009).

1. Pamela S. Karlan, *New Beginnings and Dead Ends in the Law of Democracy*, 68 OHIO ST. L.J. 743, 747 (2007).

2. Adam Liptak, *Former Justice O'Connor Sees Ill in Election Finance Ruling*, N.Y. TIMES, Jan. 27, 2010, at A16, available at <http://www.nytimes.com/2010/01/27/us/politics/27judge.html>.

The Supreme Court’s landmark decision in Buckley v. Valeo is understood to hold that only an “anti-corruption” rationale can justify campaign finance regulations. Buckley drew a rigid distinction between political campaign “expenditures” and “contributions,” holding that the anti-corruption interest justifies regulating only the latter. This Article asserts that the contribution–expenditure distinction is particularly counter-productive in the judicial election context, precisely because due process of law is fundamental to the courts to a degree unmatched by the risk of corruption in the constituent branches.

The Article starts by documenting the exponential increases in campaign cash and the newly central roles played by massive, and often highly secretive, independent expenditure campaigns in high-profile judicial elections over the past decade. It then asserts that Caperton’s approach is therefore a refreshing rejection of formalism in these compelling and new circumstances. If embraced more widely, the norms that inform the Caperton approach can mitigate the emerging constitutional crisis in our state courts that Justice O’Connor, among others, so aptly describes.

TABLE OF CONTENTS

Introduction	729	R
I. The Contributions vs. Expenditures Conundrum	733	R
A. <i>Buckley</i> and the Birth of a Distinction	733	R
B. The Unintended Consequences of <i>Buckley</i>	735	R
C. The Rising Tide of Independent Expenditures . . .	738	R
II. Campaign Cash and the New Politics of Judicial Elections	744	R
A. Everything’s Bigger in Texas	745	R
B. 2000–2009: Cash in the Courtroom Goes National	746	R
C. There is no truth. There is only perception	749	R
D. Super Spenders	750	R
E. Case Study: <i>Avery v. State Farm</i>	752	R
III. “ <i>Caperton</i> Contributions” and Due Process	756	R
A. Facts That Push Formalism Beyond Its Limits	757	R
B. <i>Caperton</i> and the <i>Buckley</i> Conundrum	760	R
1. Framing the Issue Before the Court	760	R
2. Sub Silentio Genius	765	R
IV. Judicial Campaign Expenditure Limits: An Exception Born of Exceptionalism	774	R
A. The Due Process Interest in Limiting Judicial Campaign Finance	774	R
B. Expenditure Limits and the Equality Rationale . .	778	R
V. Conclusion	780	R

INTRODUCTION

On June 8, 2009, Professor Rick Hasen of Loyola Law School, indisputably among the nation's leading election law scholars, did something rather remarkable on his widely read Election Law blog: he suggested that a Supreme Court Justice had made an inadvertent error in an opinion, radically changing its meaning.³ His post called attention to *Caperton v. A.T. Massey Coal Co.*,⁴ which was decided that day and involved the constitutional standards dictating when a judge must recuse himself based on campaign contributions. Hasen found it very curious, "given the key distinction in campaign finance law between contributions and expenditures,"⁵ that Justice Kennedy framed the issue as follows: "The basis for the motion was that the justice had received *campaign contributions* in an extraordinary amount from, and through the efforts of, the board chairman and principal officer of the corporation found liable for the damages."⁶ Professor Hasen suggested that this was an unwitting error: "Perhaps this inadvertent equating of contributions and expenditures will disappear when this opinion is finalized for the U.S. Reports."⁷

Not so. Rather than imagining away a key aspect of the Court's opinion in *Caperton* as a scrivener's error, the Court should be taken at its word: at least with respect to judicial elections, substantial independent expenditures can present a significant risk of corruption and a significant challenge to due process, both of which justify greater state regulation of campaign expenditures. Indeed, as Justice Stevens wrote one year later, "In *Caperton* . . . we accepted the premise that, at least in some circumstances, independent expenditures on candidate elections will raise an intolerable specter of *quid pro quo* corruption."⁸ Justice Stevens makes explicit what this Article asserts was implicit in *Caperton*, when he notes that "[t]he reason the Court so thoroughly conflated expenditures and contributions, one assumes, is that it realized that some expenditures may be functionally equivalent to contributions."⁹

3. Rick Hasen, *Initial Thoughts on Caperton v. Massey: First Meaningful Constitutional Limits on Excess of Judicial Elections*, ELECTION LAW BLOG (June 8, 2009, 7:58 AM), <http://electionlawblog.org/archives/013784.html>.

4. 129 S. Ct. 2252 (2009).

5. Hasen, *supra* note 3.

6. *Id.* (quoting *Caperton*, 129 S. Ct. at 2256–57) (emphasis added by Hasen).

7. *Id.*

8. *Citizens United v. FEC.*, 130 S. Ct. 876, 967 (2010) (Stevens, J., concurring in part and dissenting in part).

9. *Id.* at 968.

Love them or loathe them, judicial elections are here to stay. Accordingly, lest due process be reduced to a mere parchment promise, scholars, lawyers, and judges can no longer afford the luxury of further indulging in the perpetual, exclusively binary love/loathe debate over whether judicial appointment or election is preferable.¹⁰ Instead, it is increasingly clear that the realities reflected in the two quotes above are now on a collision course. This Article argues that the decision in *Caperton* suggests a way forward. *Caperton* recognizes the possibility for influence and the perception thereof in the judicial electoral context; and, by implication, supports the notion that the cognate threat to due process is so far heightened as to create an independent state interest in regulating not only campaign contributions, but also expenditures. Independent expenditures in the context of judicial elections are thus regulable in this view.

In the past decade, “in high court contests across America, cash has become king. Would-be justices must raise millions from individuals and groups with business before the courts. Millions more are spent by political parties and special-interest groups, much of it undisclosed.”¹¹ This Author has previously asserted that *Caperton* “is correct in result; correct in its narrowness; and correct in calling on courts to be more rigorous in recusal than due process requires.”¹² This Article goes further by arguing that one of the most controversial aspects of *Caperton*—the manner in which Justice Kennedy’s opinion for the majority conflates direct campaign contributions and independent expenditures for the purposes of due process analysis—is specifically among the most compelling aspects of the decision and provides a model for prospective judicial campaign reform.

Part I of this Article summarizes the consequences of two enduring aspects of the Court’s landmark decision in *Buckley v. Valeo*:¹³ the contribution-expenditure distinction noted above and the rejection of state interests in regulating campaign finance other than that in limiting corruption or the appearance of corruption. It

10. See Charles Gardner Geyh, *The Endless Judicial Selection Debate and Why it Matters for Judicial Independence*, 21 GEO. J. LEGAL ETHICS 1259, 1279 (2008) (noting that changing election systems “can be a worthy goal and one well worth pursuing . . . but not at the expense of ignoring shorter-term remedies that can make a bad system better in the interim”).

11. JAMES SAMPLE ET AL., *THE NEW POLITICS OF JUDICIAL ELECTIONS, 2000–2009: DECADE OF CHANGE 5* (Charles Hall, ed. 2010).

12. James Sample, *Caperton: Correct Today, Compelling Tomorrow*, 60 SYRACUSE L. REV. 293, 293 (2010).

13. 424 U.S. 1 (1976).

also describes the effects of these two aspects of *Buckley*, including the rise of independent expenditures. As Professor Neuborne puts it, “[I]n simple economic terms, the *Buckley* Court limited supply (contributions), while leaving demand (expenditures) free to grow without limit. The predictable effect has been to increase the pressures on candidates to satisfy the ever-increasing demand for campaign cash.”¹⁴

Against a background of that supply-demand asymmetry—arguably counterproductive enough as it is in the legislative and executive contexts—Part II of this Article examines the exponential increases in campaign cash and the newly central roles played by independent expenditures in high-profile judicial elections over the past decade. The dynamics of increasing expenditures and contributions in judicial elections¹⁵ were not at issue in *Buckley*, a case that involved federal legislation applicable to Congressional and Presidential contests. Moreover, this Article notes that those dynamics were utterly inconceivable when *Buckley* was decided in 1974, a time when judicial elections were generally uncontested and were at most low-salience, low-cost, and sleepy affairs.

Part III, picking up on Justice Stevens’s assertions, first coins the term “*Caperton* contributions” to describe the equation of independent expenditures with contributions in the judicial elections context. Then, in light of the new normal in judicial elections, Part III asserts that Justice Kennedy’s *sub silentio* treatment of independent expenditures as contributions is a refreshing rejection of formalism in a circumstance where, if the rigid distinction had been controlling, the focus of analysis would have been \$1000 in campaign contributions as opposed to \$3 million in expenditure support. It is worth noting that, without Justice Kennedy’s vote, that precise head-in-the-sand result would have obtained in the Supreme Court. Although the majority opinion in *Caperton* does not elaborate on the reasoning behind the interchangeable use of the terms, there can be little doubt that the Justices in the majority, all of whom are exceedingly well versed in the Court’s historical and recent campaign finance jurisprudence, fully intended the conflation. The Court’s approach to the distinction was not only correct on the facts, it was also a constitutionally and structurally appropriate mode of analysis that accommodated both the First Amendment and the Due Process Clause in the judicial campaign context.

14. BURT NEUBORNE, CAMPAIGN FINANCE REFORM & THE CONSTITUTION: A CRITICAL LOOK AT *Buckley v. Valeo* 18 (Brennan Center for Justice 1998), available at http://brennan.3cdn.net/f124fc7ebf928fb019_hqm6bn3w0.pdf.

15. See SAMPLE ET AL., *supra* note 11.

Caperton reflects a realpolitik grounding that ought to extend to making the constitutional case for a whole range of legitimate judicial campaign finance regulation, including, for instance, so-called “trigger matching funds”¹⁶ in response to independent expenditures supportive of a candidate’s adversary in the growing number of states adopting public financing systems for their elected judiciaries.

Extending the theoretical premise, Part IV argues that not only is the formal dichotomy of unlimited expenditures and limited contributions a particularly bad fit for modern-day judicial elections; the long-disfavored equality or “leveling” rationale is an especially *good* fit for these examples (for better or worse) of American exceptionalism.¹⁷ Furthermore, this Article asserts that Cass Sunstein’s argument that “[i]n rejecting the claim that controls on financial expenditures could be justified as a means of promoting political equality, . . . *Buckley* might well be seen as the modern-day analogue of the infamous and discredited case of *Lochner v. New York*”¹⁸ applies *a fortiori* to the rejection of expenditure limits in campaigns for judicial offices. Unlike their legislative and executive branch counterparts, judicial offices constitutionally entail both structural and individual rights commitments to impartiality and equality under law. Accordingly, the compelling state interest in the reality and appearance of impartial courts, and, in extreme instances, individual litigants’ due process rights, should be construed to meet the exacting strict scrutiny analysis required for expenditure limitations to pass constitutional muster under the *Buckley* framework.

16. Also termed “rescue funds” or “fair-fight funds,” these additional public grants are “triggered” when a privately funded opponent’s or independent speaker’s expenditures exceed a designated monetary threshold. See Memorandum from the Brennan Center of Justice, H.R. 1826: A Response (Dec. 17, 2009), available at http://www.brennancenter.org/content/resource/h.r._1826_a_response/. See also M. Colleen Connor, *Raising Arizona: Strengthening Express Advocacy Regulation Through the Citizens Clean Elections Act*, 34 ARIZ. ST. L.J. 507, 524 (2002) (noting that trigger matching funds depend on expenditures being fully disclosed and disclosure being strictly enforced).

17. See Herbert M. Kritzer, *Law is the Mere Continuation of Politics by Different Means: American Judicial Selection in the Twenty-First Century*, 56 DEPAUL L. REV. 423, 431 (2007) (indicating that “[t]he United States is almost unique in its use of elections in the judicial selection and retention process”).

18. Cass R. Sunstein, *Political Equality and Unintended Consequences*, 94 COLUM. L. REV. 1390, 1397 (1994).

I.
THE CONTRIBUTIONS VS. EXPENDITURES
CONUNDRUM

A. *Buckley and the Birth of a Distinction*

A generation ago, in *Buckley v. Valeo*,¹⁹ the Supreme Court, in analyzing the Federal Election Campaign Act,²⁰ issued a fractured per curiam ruling that remains the foundational, if much-criticized, landmark of campaign finance law to this day.²¹ As Professor Hasen notes, the Court “decided a number of important legal issues . . . including upholding the challenged disclosure provisions, upholding the constitutionality of the voluntary public financing system for presidential elections, and striking down the process for choosing members of the [Federal Election Commission].”²² However, the most enduring, or at least enduringly controversial, aspect of the decision is the Court’s distinction between campaign contributions and campaign expenditures.

In *Buckley*, the Supreme Court famously—or infamously—upheld campaign contribution limits as a means of preventing corruption and the appearance of corruption²³ and simultaneously struck down limits on expenditures by individuals and groups as violating the First Amendment.²⁴ While the Court’s ruling has been criticized to the point where it has been made clear that a majority of the Supreme Court would overrule it on one ground or another,²⁵ *Buckley* remains the starting point of any campaign finance analysis.

19. 424 U.S. 1 (1976).

20. See Federal Election Campaign Act Amendments of 1974, Pub. L. No. 93-443, §§ 101, 202, 310, 403, 88 Stat. 1263, 1263–68, 1275–76, 1280, 1291 (1974) (codified as amended at 2 U.S.C. §§ 431–56 (2006)), *invalidated in part by* *Buckley v. Valeo*, 424 U.S. 1 (1976).

21. *Buckley*, 424 U.S. at 1.

22. Richard L. Hasen, *The Nine Lives of Buckley v. Valeo*, in *FIRST AMENDMENT STORIES* (Richard W. Garnett & Andrew Koppelman eds., forthcoming 2011) (manuscript at 19), available at <http://ssrn.com/abstract=1593253>.

23. *Buckley*, 424 U.S. at 29, 84.

24. *Id.* at 16–17, 19.

25. See, e.g., Richard Briffault, *Nixon v. Shrink Missouri Government PAC: The Beginning of the End of the Buckley Era?*, 85 MINN. L. REV. 1729, 1757 (2001) (noting “the rejection of critical elements of *Buckley* by a majority of the justices” at the time, and that “Justices Kennedy, Scalia, Thomas, and Stevens . . . expressly called either for overruling of *Buckley* entirely or for overruling key elements of the decision” while Justices Breyer and Ginsburg “expressed the hope that *Buckley* could be salvaged through significant reinterpretation, including the modification of . . . the contribution/expenditure distinction”). Indeed, writing as of 2001, Professor Briffault noted that “[o]nly Justices Souter, O’Connor and Chief Justice Rehnquist raised no questions about the continuing states of *Buckley*,” an observation that is

In addition to establishing the contribution–expenditure dichotomy, the Court’s distinction was driven by the Court’s exclusive focus on the anti-corruption interest.

The Court’s emphasis on the anti-corruption rationale resulted from the specific concern that “large contributions are given to secure a political quid pro quo from current and potential officeholders.”²⁶ That focus on quid pro quos inevitably led to the conclusion that independent expenditures in support of a candidate’s campaign failed to pose the same risk of corruption. Such expenditures by definition excluded coordination, much less a quid pro quo. Arguably, however, the Court’s mistake was not the focus on corruption. Instead, the Court’s error was its singular focus on quid pro quo exchanges as the only basis for that corruption. With respect to contributions, the Court specifically found that the interest “in alleviating the corrupting influence of large contributions” justified limiting the size of contributions and upholding FECA’s disclosure provisions.²⁷

Effectively, the Court found that “a limit on the amount of campaign contributions only marginally restricted a contributor’s ability to send a message of support for a candidate”²⁸ whereas “campaign expenditures were core political speech” thus necessitating “exacting scrutiny.”²⁹ Thirty-five years after the Court’s decision, the contribution–expenditure distinction is well-worn black-letter shorthand. Yet, rather than merely taking the bar on expenditure limitations at face value, it is worth noting that the Court, of course, was not considering the concepts in the abstract. Rather, it was considering the expenditure limitations in FECA itself. In that regard, even in 1970s dollars, FECA’s specific limits were extraordinarily stringent.³⁰

particularly telling given that (a) that doesn’t necessarily mean they didn’t have such questions and (b) that none of those three remains on the Court today. *Id.*

26. *Buckley*, 424 U.S. at 26.

27. *Id.* at 55.

28. Hasen, *supra* note 22, at 19–20.

29. *Id.*

30. FECA carefully regulated political expenditures with a series of caps, all of which the Court ultimately struck down. Campaigns were subject to stringent expenditure limits. Presidential campaigns were capped at \$10 million for the primaries, and \$20 million for the general election. Senate campaigns were limited to 8 cents a voter for the primaries, and 12 cents a voter for the general election. House campaigns were limited to \$70,000 for the primaries and \$70,000 for the general elections. These spending limits were indexed annually for inflation. Finally, the independent spending of individuals was limited to \$1,000 in support of a federal candidate. For example, Voter Jones could take out a newspaper ad supporting Candidate Smith if Jones’ costs were \$1000 or less. Candidates were permitted to

In recent years, even the most ardent proponents of campaign finance regulation have acknowledged that “FECA’s spending limits were set at an unreasonably low level” and that, accordingly, the “Court was correct to note that the \$1,000 ceiling on independent expenditures was a de facto ban on political participation, and FEC’s \$70,000 limit for House races was also unreasonably low.”³¹ Whether substantially higher limits would have generated a different result is now a matter of pure speculation.³² Moreover, whatever the result in *Buckley* might have been with higher limits, the Court’s decision in *Citizens United v. Federal Elections Commission*³³ is just the latest indication that the current Supreme Court would not be solicitous of spending limits—at least not in legislative and executive races.

B. *The Unintended Consequences of Buckley*

Buckley’s singular focus on the anti-corruption rationale manifests in several aspects: (1) framing the campaign spending debate; (2) restraining both campaign regulation proponents and opponents; and, most importantly, (3) shaping political campaigning itself. Scholars who are critical of *Buckley* argue that “expenditures on behalf of a candidate can create some of the dangers of contributions. Candidates often know who spends money on their behalf and for this reason, an expenditure may in some contexts give rise to the same reality and appearance of corruption.”³⁴ Assuming *ar-*

spend up to \$50,000 of their own money on a presidential campaign, \$35,000 on a Senate campaign, and \$25,000 on a House campaign. See NEUBORNE, *supra* note 14, at 9.

31. *Id.* at 16.

32. In *Buckley* itself, Justice White would have sustained the spending limits as “essential to prevent transparent and widespread evasion of the contribution limits” while Justice Marshall would have sustained the limits on personal expenditures in service of the “reality and appearance of equal access to the political arena.” *Buckley*, 424 U.S. at 262 (White, J., concurring in part and dissenting in part); *id.* at 287 (Marshall, J., concurring in part and dissenting in part). For a detailed and fascinating discussion of the individual justices’ positions in *Buckley*, both in the opinion itself and from the pertinent historical papers, see Hasen, *supra* note 22.

33. *Citizens United v. FEC*, 130 S. Ct. 876 (2010). With respect to predictions, an insightful student note indicates that “[o]ver the past twenty years, the Court’s fractured and anomalous jurisprudence has led commentators to repeatedly conclude that *Buckley*’s fall was imminent” but “[t]he present Court seems certain not to overrule *Buckley* in favor of a deferential approach to regulation.” J. Robert Abraham, Note, *Saving Buckley: Creating a Stable Campaign Finance Framework*, 110 COLUM. L. REV. 1078, 1092 (2010) (internal footnotes omitted).

34. Sunstein, *supra* note 18, at 1395 (citing examples in DAN CLAWSON ET AL., MONEY TALKS: CORPORATE PACS AND POLITICAL INFLUENCE 75–79 (1992)).

guendo that the Court's failure to recognize this dynamic was in error, it occurred against a political landscape bearing little relation to the expensive, orchestrated independent expenditure campaigns of the Swift Boat and MoveOn era,³⁵ much less the And For the Sake of the Kids judicial elections era.³⁶

As Professor Neuborne notes, this "is where the *Buckley* Court suffers most from having been without a factual record. Enormous independent expenditures were not part of the fictional record the Court considered, mostly because they were not yet part of America's political process."³⁷ Because such expenditures had become a part of America's presidential and congressional political process since a generation ago, their continued absence from judicial elections until only very recently was even more pronounced.³⁸ Conversely, in the era of costly campaigns, conducted mostly via television advertisements, little doubt exists that elected "officials can be influenced by who spends money on their behalf, just as they can be influenced by who directly contributes money to them. The perception of corruption might be generated by large expenditures for a candidate, just as it can be caused by large contributions."³⁹

From a *realpolitik* perspective, the contributions–expenditures distinction is a lesson in unintended consequences. On the bright side, those unintended consequences have become a rich source of colorful metaphor. Professors Karlan and Issacharoff note that the "result is an unceasing preoccupation with fundraising. The effect is much like giving a starving man unlimited trips to the buffet table but only a thimble-sized spoon with which to eat: chances are great that the constricted means to satisfy his appetite will create a singular obsession with consumption."⁴⁰ Similarly, Josh Rosenkranz writes that the unlimited demand for money resulting from a lack of limits on expenditures, combined with the restricted supply resulting from contribution limits, turns "decent, honest politicians" into "junkies . . . caught in the political equivalent of an arms race in which neither side feels safe to disarm unilaterally because each

35. *See infra* notes 67, 68, 75.

36. *See infra* note 159.

37. NEUBORNE, *supra* note 14, at 16.

38. *See infra* Part II.B.

39. Erwin Chemerinsky, *Preserving an Independent Judiciary: The Need for Contribution and Expenditure Limits in Judicial Elections*, 74 CHI-KENT L. REV. 133, 143–44 (1998).

40. Samuel Issacharoff & Pamela S. Karlan, *The Hydraulics of Campaign Finance Reform*, 77 TEX. L. REV. 1705, 1711 (1999).

candidate lives in mortal fear of being buried by the other's spending."⁴¹

Buckley's construction of the contribution–expenditure distinction thus manifests a paradoxical conundrum: the FECA system that was designed to reduce the influence of money in politics, warped by *Buckley*, has instead exponentially increased the import of chasing every last dollar.⁴² Furthermore, since an increasing number of contribution sources were inevitably required, the marginal sources of the contributions were inevitably further and further afield from a candidate's core, philosophically aligned base. Consequently, at least in some instances, such as when a candidate seeking contributions from groups and individuals with a particular interest about which the candidate himself or herself would otherwise be neutral or even apathetic, the dynamic almost assuredly results in the very type of express or implied quid pro quo exchanges that *Buckley* deemed most pernicious.

The exclusivity of the anti-corruption interest defines the rhetorical push and pull of the campaign finance regulatory debate as well. Professors Issacharoff and Karlan skeptically note that “calls for reform all stem from the assertion that money corrupts the electoral process.”⁴³ They note that this characterization dominates discussions of reform despite the indeterminate nature of what is really meant by “corruption.”⁴⁴ Professors Issacharoff and Karlan acknowledge that while this may be “primarily as a response to *Buckley* . . . it may well be that even in the absence of the *Buckley* imperative, arguments for campaign finance reform would use images of corruption because of their rhetorical power.”⁴⁵

Whether sourced in its “rhetorical power” or not, the focus on anti-corruption leads to two concrete analytical problems. The first problem, as Professors Persily and Lammie have noted, is the “trap that has snared the current Court and those who have analyzed the campaign finance decisions: namely, the alternating tendencies of the word ‘corruption’ to mean everything and nothing.”⁴⁶ The sec-

41. E. Joshua Rosenkranz, *Faulty Assumptions in “Faulty Assumptions”: A Response to Professor Smith’s Critiques of Campaign Finance Reform*, 30 CONN. L. REV. 867, 889 (1998).

42. See *infra* Part II.B.

43. Issacharoff & Karlan, *supra* note 40, at 1707–08.

44. *Id.* at 1708.

45. *Id.* at 1719.

46. Nathaniel Persily & Kelli Lammie, *Perceptions of Corruption and Campaign Finance: When Public Opinion Determines Constitutional Law* 119, 122 (U. Penn. Inst. for Law & Econ. Research Paper 04-22, U. Penn. Law School, Public Law Working Paper 53, 2004), available at <http://ssrn.com/abstract=595721>.

ond problem is that “the unique position of ‘*appearance* of corruption’ . . . has more to do with the difficulties of proving actual corruption . . . than the importance of the state interest in combating such negative perceptions.”⁴⁷

Were the Supreme Court to exorcise the “appearance of corruption” state interest from the campaign finance jurisprudence, few campaign finance regulations would pass constitutional scrutiny. Antireformists might greet this development with cheer, but without the fallback on appearances and perceptions, defenders of campaign finance reforms would be left with the difficult job of proving that (1) campaign contributions have actually corrupted representatives, and (2) antibribery laws are insufficient to combat actual corruption. The existence of the fallback state interest of preventing appearances allows judges to say that, while they think examples of actual corruption justify the given reform, the existence of widespread appearances of corruption removes all doubt.⁴⁸

Professors Persily and Kammie’s points are valid across the spectrum of democracy. Yet, as Parts II and III of this Article explain in more detail, proving corruption is more of a hurdle in the judiciary than in any other political arena. Furthermore, the judiciary is precisely where even the most marginal and subconscious influence of financial support can jeopardize both systemic integrity and individual rights.

In addition to the dynamic of the “unlimited trips to the buffet” with the “thimble-sized spoon”⁴⁹ and the other unintended consequences described above, critics of campaign finance restrictions point to another significant worry: if “reform advocates had their way, they would discover what the Corps of Engineers learned over the years in trying to redirect the Mississippi. Money, like water, will seek its own level” and, consequently, “[t]he price of containment may be uncontrolled flood damage elsewhere.”⁵⁰

C. *The Rising Tide of Independent Expenditures*

Whether the money-like-water dynamic is categorically unavoidable or merely a function of the *Buckley* jurisprudence is a debate beyond the scope of this article. One product of the “uncontrolled flood damage,” however, is very much on point—the

47. *Id.* at 133.

48. *Id.* at 134.

49. Issacharoff & Karlan, *supra* note 40, at 1711.

50. *Id.* at 1713.

extent to which the real-world result of *Buckley*'s sharp line between regulable contributions and non-regulable expenditures is the rise of contemporary big-dollar independent expenditure campaigns.

Writing nearly 30 years ago, Judge Skelly Wright⁵¹ lamented even then how, “in recent years, [Political Action Committees (PACs)]—especially those connected with corporations and trade associations—have grown explosively in numbers and influence”⁵² and, more tellingly, that FECA, along with *Buckley* “facilitated their formation and operation.”⁵³ PAC contributions to candidates and independent expenditures are apples and oranges.⁵⁴ Yet, Judge Wright's description included, as foreshadowing, a reflection that “PAC's, especially those with ideological or single-issue orientations, spend large sums of money on highly effective campaigns against ‘hit listed’ activities for or against particular candidates.”⁵⁵ In 1980 alone, PACs had “devoted more than \$14 million” to independent expenditure campaigns.⁵⁶ In a comment that would prove only more prescient over time, Judge Wright stated simply, “Because *Buckley v. Valeo* constitutionally invalidated ceilings on ‘independent’ expenditures, this type of spending may theoretically be unlimited.”⁵⁷ Judge Wright found it telling that the biggest injections of wealth “into the political bloodstream” came from “the more highly regulated industries, such as oil, transportation, utilities, drugs, health care, and government contracting.”⁵⁸ Looking at Judge Wright's list three decades later—especially in the context of

51. Judge Wright was on the panel of the D.C. Circuit Court of Appeals that was reversed by the Supreme Court in *Buckley*. See *Buckley v. Valeo*, 519 F.2d 821, 831 (D.C. Cir. 1975), *aff'd in part, rev'd in part*, 424 U.S. 1 (1976).

52. J. Skelly Wright, *Money and the Pollution of Politics: Is the First Amendment an Obstacle to Political Equality?*, 82 COLUM. L. REV. 609, 614–15 (1982) (noting that between 1974 and 1982, number of corporate PACs went from 89 to 1327 and that “[b]etween 1976 and 1980 PAC's more than doubled the amounts of money they poured into House and Senate campaigns—\$22.6 million in 1976, \$35.2 million in 1978, and \$55.3 million in 1980”).

53. *Id.* at 614.

54. For a clear and insightful overview of the tax and campaign finance distinctions between PACs, 527s, and the consequences of being defined, or not defined, as a “political committee” under the law, see Paul S. Ryan, *527s in 2008: The Past, Present, and Future of 527 Organization Political Activity Regulation*, 45 HARV. J. ON LEGIS. 471 (2008).

55. Wright, *supra* note 52, at 616.

56. *Id.*

57. *Id.*

58. *Id.*

judicial elections—one is hard pressed not to think of the French epigram, *plus ça change, plus c'est la même chose*.⁵⁹

Largely because the rhetoric concerning campaign finance involves the intersection of the First Amendment and politics, it is marked as much by tropes, anecdotes, and shoehorning as by meritorious dialectic. Professors Issacharoff and Karlan treat reformers' less-than-credible overreliance on corruption arguments with appropriate skepticism.⁶⁰ Similarly, the most insistent opponents of campaign regulation overemphasize homespun stories that—were they the rule rather than the rarest of exceptions—would indicate that the real aim of campaign regulation is neither to prevent corruption nor to reduce political inequality. Instead, campaign regulation opponents attempt to create the impression that campaign regulations seek to prevent Mom & Pop from ever expressing a political viewpoint.⁶¹ Indisputably, true stories of the unwittingly ensnared exist in campaign finance just as in other regulatory schemes. Regarding independent expenditures in political campaigns, however, the numbers tell a very different story; the vanguard of the genre is not Mom and Pop—at least not unless Mom and Pop spend tens of millions through a “527.”

59. “The more things change, the more they stay the same.”

60. See Issacharoff & Karlan, *supra* note 40, at 1708 (addressing paradoxes inherent within reformers' assertions that the election process is corrupted by the influx of money).

61. Former FEC Commissioner Bradley A. Smith employs this type of advocacy particularly effectively. Below is a typical example of this technique, as well as one of Smith's favorite anecdotes:

In the autumn of 2000, Harvey Bass, the owner of Harvey Bass Furniture and Appliance in the west Texas town of Muleshoe, painted “Save our Nation: vote Democrat; Al Gore for President” on the side of a leftover refrigerator shipping box. Bass left the homemade sign on the porch of his business on State Highway 214, to be seen by passersby. Soon another local resident, Bill Liles, “got tired of looking at it, especially the ‘Save our Nation’ part.” Liles and a friend, one Mark Morton, decided to make a sign supporting Gore's rival, Texas Governor George W. Bush. The two decided that their sign “should be bigger and better.” They obtained a large plywood board, hired a professional sign painter, and mounted the finished product on the side of a cotton trailer obtained from another local resident, Don Bryant. They parked the trailer, with its sign, across the street from Mr. Bass's store. “As word spread the sign became a topic of conversation at local gathering places. Mostly the Spudnut Shop on Main Street and the Dinner Bell Café on Highway 84. People started coming by and donating to help pay for the cost of the sign.”

This spontaneous burst of political activity ended when another Muleshoe resident, Don Dyer, filed a complaint with the Federal Election Commission against Liles, Morton, Bryant, and one of their “contributors.”

Bradley A. Smith, *Campaign Finance Reform: Searching for Corruption in All the Wrong Places*, 2003 CATO SUP. CT. REV. 187 (2003).

Named for the section of the Internal Revenue Code under which they are organized, 527s frequently make independent expenditures to support or oppose political candidates.⁶² “In the world of campaign finance, 2004 was without a doubt the year of the 527 organization.”⁶³ Estimates of spending by 527s in the groundbreaking 2004 federal election “amounted to at least \$405 million, accounting for more than one-tenth of total federal election spending and perhaps twenty to twenty-five percent of spending in the presidential campaign.”⁶⁴ The year 2004 also produced what remains the best-known 527 organization ever: the Swift Boat Veterans for Truth,⁶⁵ whose campaign against John Kerry generated what pollsters described as a “staggering” awareness among nearly three out of four voters in 2004.⁶⁶ When voters in that election were asked which 527 group had the most impact on the 2004 race, “almost as many voters cited Swift Boat Veterans as cited all the other 527’s combined.”⁶⁷

Even the strategic deployment of 527s, through which Democratic-leaning 527s outspent their Republican-leaning adversaries by nearly a 3:1 spending margin,⁶⁸ was a lesson in unintended consequences: “The Democrats built these new shadow-party advocacy groups to attack the president early in the campaign season and build voter-turnout machines. Then they watched Bush partisans

62. 26 U.S.C. § 527 (2006).

63. Richard Briffault, *The 527 Problem . . . and the Buckley Problem*, 73 GEO. WASH. L. REV. 949, 949 (2005).

64. *Id.* (citing Stephen R. Weissman & Ruth Hassan, *BCRA and the 527 Groups*, in *THE ELECTION AFTER REFORM: MONEY, POLITICS, AND THE BIPARTISAN CAMPAIGN REFORM ACT 79* (Michael Malbin ed., 2006)).

65. See John J. Miller, *What the Swifties Wrought*, NAT’L REV., Nov. 29, 2004, at 18 (asserting that the Swift Boat Campaign “will go down in history for its stunning effectiveness”); Briffault, *supra* note 63.

66. Public Release of National Survey Results, Fabrizio, McLaughlin & Associates, *527’s Matter and Swift Boat Vets Top the Heap – Best Known, Most Impact and Highest Ad Recall* (Nov. 4, 2004), available at <http://www.fabmac.com/FMA-2004-11-04-527-Effects.pdf>.

67. *Id.* The best known of the 527s supporting Senator Kerry was MoveOn.org. See *id.* MoveOn.org was supported with millions of dollars from sources including George Soros and Peter Lewis. For a chart detailing MoveOn’s top contributors during the 2004 cycle *MoveOn.org: Top Contributors, 2004 Cycle*, CENTER FOR RESPONSIVE POLITICS (2010), http://www.opensecrets.org/527s/527cmtedetail_contribs.php?ein=200234065&cycle=2004.

68. Editorial, *The Soft Money Boomerang*, N.Y. TIMES, Dec. 29, 2004, at A20, available at <http://query.nytimes.com/gst/fullpage.html?res=9B03E2D71639F93AA15751C1A9629C8B63>.

adapt the same financing device to float the campaign's most notorious and devastating attack ads, the Swift Boat assault"⁶⁹

At least in terms of public expression—with private and strategic thinking being inherently harder to measure—527s were broadly scorned, even by their biggest beneficiaries. Former President George W. Bush, for example, stated, “I don’t think we ought to have 527’s [sic] I can’t be more plain about it and I wish—I hope my opponents join me in saying—condemning these activities of 527’s [sic]. I think they are bad for the system.”⁷⁰ It is not often that a question of campaign finance generates agreement so broad that it encompasses President Bush and Wisconsin Senator Russell Feingold. Senator Feingold argues that 527s allow “wealthy individuals to drown out the voices of average citizens.”⁷¹ Senator Feingold’s critique was echoed, most notably, by his frequent collaborator in regulating campaign finance, Senator John McCain. Senator McCain asserted that 527s must be limited so that “political power in this country does not lie *solely* in the hands of big corporations, labor unions, and the wealthiest of the wealthy.”⁷²

The ideological diversity represented by these three political figures would, at least superficially, seem to bode well for those favoring limits on independent expenditures. However, on closer examination, one realizes quickly that not one of these arguments references “corruption.” Moreover, Senators Feingold and McCain expressly invoke the prohibited leveling rationale, while President Bush’s “bad for the system” remark seems to strongly indicate a non-corruption concern. Consequently, even if modest registration requirements, disclosure provisions, and other regulations are imposed on 527 organizations, fundamental reforms aimed at limiting their reach run headlong in the *Buckley* bulwark, which makes corruption the sole justification for regulation of campaign financing. A circular relationship thus connects the rhetoric with the *Buckley*-based reality. First, this reality requires that those who want to limit

69. *Id.*

70. Elizabeth Bumiller & Kate Zernike, *President Urges Outside Groups to Halt All Ads*, N.Y. TIMES, Aug. 24, 2004, at A1, available at <http://www.nytimes.com/2004/08/24/politics/campaign/24swift.html>.

71. Briffault, *supra* note 63, at 995 (citations omitted) (quoting *Hearing to Examine and Discuss S. 271, a Bill Which Reforms the Regulatory and Reporting Structure of Organizations Registered Under Section 527 of the Internal Revenue Code*, 109th Cong. (2005) (statement of Sen. Russell Feingold)).

72. *Id.* (emphasis added) (citations omitted) (quoting *Hearing to Examine and Discuss S. 271, a Bill Which Reforms the Regulatory and Reporting Structure of Organizations Registered Under Section 527 of the Internal Revenue Code*, 109th Cong. (2005) (statement of Sen. John McCain)).

expenditures must emphasize an anti-corruption nexus that is only tangentially connected to their real democratic concerns, and must focus on limiting contributions rather than expenditures. Second, those who oppose or are at least less sanguine about expenditure limitations can decry massive special interest capture without any risk that the money will stop flowing.

The combination of the *Buckley* bulwark against regulation of 527s and the high-profile success of the Swift Boat campaign in the 2004 presidential election⁷³ rendered imitation and expansion of the independent expenditure model virtually inevitable, even at the state level. But where did the money come from? In some respects, the role of 527s in the 2004 election dovetails with the hydraulics theory.⁷⁴ As Professor Briffault notes, “many observers contend[ed] that contributions to and expenditures by 527s were little more than evasions of the recently enacted Bipartisan Campaign Reform Act of 2002 (“BCRA”) and re-creations of the soft money problem that BCRA was supposed to have eliminated.”⁷⁵ However, in analyzing a study conducted by Steve Weissman and Ruth Hassan, Professor Briffault describes a dynamic less marked by hydraulics. Instead, the ability of singular parties contributes to the dynamic; Briffault describes these parties as “megadonors” who play, in essence, a mega-role:

[I]t was not simply a matter of old money finding new venues. The real source of the explosion of 527 activity was contributions from individuals, especially very large contributions by individuals. Most of this money had not been in the pre-BCRA soft money system at all, but instead entered the electoral process in 2004 for the first time. A total of 1882 individuals made contributions to 527s of \$5000 or more apiece, and they gave an aggregate of \$256 million. Although the median donation was only \$12,000, the mean contribution was \$135,805—twenty-seven times the statutory ceiling on individual donations to FECA political committees—because of the dominant role of a small number of extremely large donors. Indeed, there were twenty-four individuals or couples who each gave \$2 million or more. This handful of megadonors collectively gave a staggering \$142.5 million. To put this number in context, it is

73. Richard Briffault notes that “the activities of the pro-Republican 527s, particularly the Swift Boat Veterans, demonstrate that 527s can give powerful support to a candidate while still operating independently of the candidate’s campaign.” *Id.* at 969.

74. See *supra* note 40 and accompanying text.

75. Briffault, *supra* note 63, at 950.

roughly equal to the aggregate of \$149.2 million in public funds provided to the two presidential nominees for the general election campaign. *The \$142.5 million in 527 funds, however, came from just two dozen megadonors, while the \$149.2 million in public funds came from the entire American public.*⁷⁶

A similar “megadonor” dynamic is increasingly coming to the fore in state judicial elections. While unregulated political activity by 527 organizations has not affected all or even most state and local governments,⁷⁷ isolated state contests have turned into flashpoints for large independent expenditures. And as Part II explains, as a proportional matter, “megadonors” play an even more outsized role in judicial contests than the role Professor Briffault describes above in the presidential context.

II. CAMPAIGN CASH AND THE NEW POLITICS OF JUDICIAL ELECTIONS

Judicial elections are particularly susceptible to being decided by large contributions or expenditures by particular individuals or organizations for multiple reasons. Judicial elections are low-salience affairs, making contributions from the so-called “average” voter hard to find.⁷⁸ At the same time, for the “exceptional” voter who has a high-value lawsuit pending or routinely asks the state courts to vindicate their economic interests, the concentrated benefit could scarcely be higher. Thus, for the “exceptional” interest, who is a major and/or repeat player in the courts, spending even millions of dollars on a candidate for the state supreme court quickly yields a favorable cost-benefit analysis. On the other hand, for the vast majority of citizens—with hugely important but highly diffuse interests in the caliber and fairness of their state courts—equivalent “major” spending is irrational. Not only is it a financial impossibility, but the cost-benefit analysis of even *minimal* spending on a judicial campaign is abysmal. There is not currently, nor is there likely to be, a “small-donor revolution” in judicial elections any time soon. Three million dollars in West Virginia or Alabama or Wisconsin—or even in more expensive markets like Washington, Illinois, Michigan or Ohio—can buy a whole lot of misinformation.

76. Briffault, *supra* note 63, at 964 (emphasis added) (internal citations omitted).

77. Ryan, *supra* note 54, at 505.

78. *See, e.g., infra* Part II.D (showing clear examples of “exceptional” spenders).

And as campaign spending has skyrocketed overall in the past few decades, judicial campaign spending, specifically, has gone into orbit, at least by historical standards.

A. *Everything's Bigger in Texas*

In the modern era, the specter of massive judicial campaign support from parties and lawyers with specific interests before the court is traceable to the 1982 Texas Democratic Primary. In that primary, a controversial repeat litigant, described as a “South Texas rancher and oil man,” contributed \$200,000 to an unsuccessful candidate for the Texas Supreme Court—a sum that represented more than 90% of the candidate’s campaign support.⁷⁹

Three years later, Texas once again pioneered today’s brave new world of judicial elections in a highly charged case between energy giants Pennzoil and Texaco. Two days after filing its answer, Pennzoil contributed \$10,000 to the pre-trial judge through one of its lawyers⁸⁰ as well as another \$10,000 to the judge in charge of assigning trial judges.⁸¹

Texaco’s attorneys, who demanded a new trial on the basis of the \$10,000 contribution by their adversary to the trial judge, then responded by contributing “\$72,700 to seven Texas Supreme Court justices who were expected to make the final ruling in the case.”⁸² The seven justices to whom Texaco contributed included three justices not even up for re-election.⁸³ Ultimately, Pennzoil and its lawyers contributed \$315,000 to judges involved in the litigation, outspending Texaco’s \$72,700 by a factor of four.⁸⁴ “Perhaps coincidentally, the courts ultimately ruled in favor of the highest bidder. Although it is clearly plausible that Pennzoil prevailed solely on the merits of its arguments, the attack on judicial independence and integrity that stemmed from the contributions tainted the outcome.”⁸⁵

79. See Anthony Champagne, *The Selection and Retention of Judges in Texas*, 40 Sw. L.J. (Special Issue) 53, 84 (1986).

80. See *Texaco, Inc. v. Pennzoil Co.*, 729 S.W.2d 768, 842 (Tex. App. 1987).

81. See Mark Andrew Grannis, Note, *Safeguarding the Litigant’s Constitutional Right to a Fair and Impartial Forum: A Due Process Approach to Improprieties Arising From Judicial Campaign Contributions from Lawyers*, 86 MICH. L. REV. 382, 404 (1987).

82. Jason Miles Levien & Stacie L. Fatka, *Cleaning Up Judicial Elections: Examining the First Amendment Limitations on Judicial Campaign Regulation*, 2 MICH. L. & POL’Y REV. 71 (1997).

83. See Kathryn Abrams, *Some Realism About Electoralism: Rethinking Judicial Campaign Finance*, 72 S. CAL. L. REV. 505, 517 (1999).

84. Levien & Fatka, *supra* note 82, at 71.

85. *Id.* at 71–72.

While the facts and figures in Pennzoil were extreme, they also reflected a near-standard operating procedure within the state's courts for decades. In a study of the Texas State Bar regarding the state's courts, 79% of attorneys surveyed indicated their belief that campaign contributions significantly influence a judge's decision.⁸⁶

B. 2000–2009: *Cash in the Courtroom Goes National*

The national trend of big-money judicial campaigns began in the late 1990s. In the 1999–2000 election cycle, state supreme court candidates raised nearly \$46 million—a 62% increase over 1998.⁸⁷ Alabama's 2000 election for a seat on the state high court set a short-lived judicial-race record of \$4.8 million.⁸⁸ That same year, “national and business media were reporting on how the business sector was fighting to shift the balance on state courts back from what many considered a pro-plaintiff bias.”⁸⁹ For example, in 2000, the U.S. Chamber of Commerce and its state chapters announced that they were going to prioritize the funding of high court campaigns. They were not kidding. Commenting on the 62% jump in overall spending from 1998 to 2000, Roy Schotland wrote: “That increase was no ‘sport,’ but a new peak in a clear, dramatic trend whether viewed in terms of total dollars, or dollars-per-seat, or percentages.”⁹⁰ In the decade since the 2000 watershed, big money has become a prerequisite for a viable high-court campaign, and for many lower state courts as well. The speed of the geographically dispersed transformation is remarkable and measurable both empirically and anecdotally.

Perhaps the best way to understand the rise in judicial spending starting in 2000 is to consider the following graph. This graph illustrates, at two-year intervals, the percentage of states with high-court elections in which the campaigns included television advertising, which is the biggest driver in campaign costs from 2000-06:

86. See Alexander Wohl, *Justice for Rent*, AM. PROSPECT, Nov. 30, 2002, available at http://www.prospect.org/cs/articles?article=justice_for_rent.

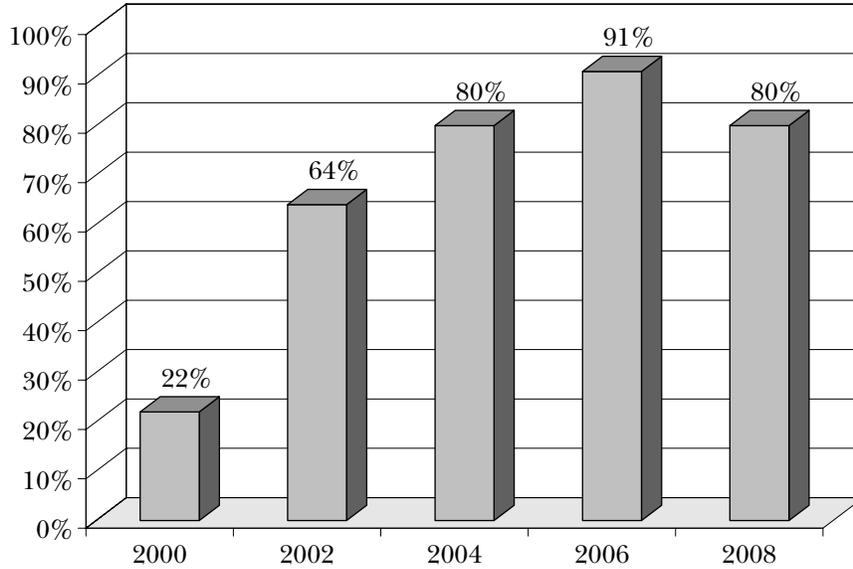
87. See SAMPLE ET AL., *supra* note 11, at 5.

88. See DEBORAH GOLDBERG ET AL., THE NEW POLITICS OF JUDICIAL ELECTIONS 2004: HOW SPECIAL INTEREST PRESSURE ON OUR COURTS HAS REACHED A “TIPPING POINT”—AND HOW TO KEEP OUR COURTS FAIR AND IMPARTIAL 32 (2005), available at http://brennan.3cdn.net/dd00e9b682e3ca2f17_xdm6io68k.pdf.

89. See SAMPLE ET AL., *supra* note 11, at 39.

90. Roy A. Schotland, *Financing Judicial Elections, 2000: Change and Challenge*, 2001 L. REV. MICH. ST. U.-DETROIT C.L. 849, 862.

Percentage of States With Contested Supreme Court Elections Featuring TV Advertising, 2000-2008⁹¹



The chart verifies the highly compressed chronology of the transformation in judicial campaigns. In just eight years, television advertising in state high-court races went from a one-in-five exception to a prerequisite in all but the least seriously contested races. Moreover, where significant television advertising exists, massive spending—in the form of either contributions or outside expenditures, or both—is guaranteed to exist as well.

Across the country, judicial campaign spending records have been broken repeatedly throughout the past decade. Forty percent of the states holding contested supreme court elections (nine of twenty-two) broke aggregate candidate fundraising records in the 2003–2004 election cycle.⁹² In the 2005–2006 cycle, “half of the states that held entirely privately-financed, contested supreme court elections (five of ten) broke state fundraising records.”⁹³ In 2007–2008, state supreme court candidates raised \$45.6 million;

91. See SAMPLE ET AL., *supra* note 11, at 24. Counting a bitterly contested and expensive 2007 Wisconsin campaign, the 2007–08 cycle featured sixteen states with contested supreme court races, with television ads running in fourteen of them (more than eighty-five percent). *Id.*

92. GOLDBERG ET AL., *supra* note 88, at 13.

93. JAMES SAMPLE ET AL., JUSTICE AT STAKE CAMPAIGN, THE NEW POLITICS OF JUDICIAL ELECTIONS 2006 15 (2006), available at http://www.brennancenter.org/content/resource/the_new_politics_of_judicial_elections_2006/.

this sum was seven times the 1989–90 total.⁹⁴ The figure marked the third time in the last five two-year cycles that, nationally, state supreme court candidates raised more than \$45 million.⁹⁵ Alabama offers a good illustration. In the 2005–2006 election cycle alone, state supreme court candidates in Alabama raised \$13.4 million, surpassing the previous state record by more than \$1 million.⁹⁶ The three candidates for Chief Justice raised a combined \$8.2 million, making it the most expensive judicial race in state history and the second-most expensive judicial campaign in American history.⁹⁷

During the decade as a whole, “[twenty] of the [twenty-two] states that elect [s]upreme [c]ourt judges set spending records; only Texas and North Dakota had their highest-spending elections in the 1990s.”⁹⁸ In the 1990s, only three states—Alabama, Pennsylvania, and Texas—had judicial campaigns in which over \$1 million was raised. From 2000 to 2009, sixty-six campaigns raised \$1 million or more in twelve states—over half of the twenty-two states that hold judicial elections for their highest courts.⁹⁹ The numbers are even more dramatic when viewed from a broader perspective. From “2000–09, Supreme Court candidates raised \$206.9 million nationally, more than double the \$83.3 million raised from 1990–1999.”¹⁰⁰ That increase of more than 250% is *ten times* the increase in the consumer price index during the same period.¹⁰¹

The judicial candidate with the most funds in a race generally wins the election. In 2006, the candidate who raised the most money in state high court races won 68% of the time.¹⁰² In 2004, that figure was 85%.¹⁰³ In itself, that proves little—the best candidates are also often the best at fundraising. The dynamic, however, exists in tandem with the fact that a substantial proportion of the contributions come from those with specific interests before the courts. For example, research by the National Institute on Money in

94. See SAMPLE ET AL., *supra* note 11, at 8.

95. See *id.*

96. SAMPLE ET AL., *supra* note 93, at 15.

97. *Id.* Former Alabama State Bar President Mark White is a no-nonsense crusader to whom those committed to impartial courts owe substantial gratitude. In 2006, he noted that candidates for judicial office in Alabama spent more than twice Alabama’s total annual spending on civil legal services for the poor. Editorial, *Justice at any price*, BIRMINGHAM NEWS, July 23, 2008, at A8, available at 2008 WLNR 13791071.

98. SAMPLE ET AL., *supra* note 11, at 8.

99. See *id.*

100. *Id.*

101. See *id.*

102. *Id.* at 31.

103. See *id.*

State Politics identified and disaggregated 84% of directly contributed funds in 2005–2006 state supreme court races.¹⁰⁴ Business interests represented 44% of contributed funds, with lawyers making up the second-largest source at 21% of all contributed funds.¹⁰⁵

C. “*There is no truth. There is only perception.*”¹⁰⁶

It is often remarked in varying forms that “without public confidence, the judicial branch could not function.”¹⁰⁷ The flood of cash into judicial races is, predictably, jeopardizing that confidence. According to a 2004 public poll, more than 70% of Americans believe that judicial campaign contributions have at least some influence on judges’ decisions in the courtroom.¹⁰⁸

Staggeringly, nearly half of judges themselves agree with the public’s perception. In a written survey of 2428 state lower, appellate, and supreme court judges, almost half (46%) of the judges surveyed indicated a belief that campaign contributions influence judges’ decisions.¹⁰⁹ More than 70% of surveyed judges expressed concern regarding the fact that “[i]n some states, nearly half of all supreme court cases involve someone who has given money to one or more of the judges hearing the case.”¹¹⁰

Referring to the *Avory* case described below,¹¹¹ Justice O’Connor, put the matter in provocative terms:

104. *Id.* at 18. This yeoman’s work was made more challenging by lax campaign disclosure requirements, which made it impossible to identify and disaggregate 16% of the contributed funds nationwide, and a higher amount in particular states. *Id.* at fig.11.

105. *Id.* at 18 fig.11.

106. In the original French, “Il n’y a pas de Vrai! Il n’y a que des manières de voir.” Letter from Gustave Flaubert to M. Léon Hennique, (Feb. 2-3, 1880), in CORRESPONDANCE: ANNÉE 1880, available at <http://flaubert.univ-rouen.fr/correspondance/conard/outils/1880.htm>.

107. *In re Raab*, 793 N.E.2d 1287, 1292 (N.Y. 2003).

108. JUSTICE AT STAKE CAMPAIGN, MARCH 2004 SURVEY HIGHLIGHTS: AMERICANS SPEAK OUT ON JUDICIAL ELECTIONS (2004), <http://faircourts.org/files/Zogby-PollFactSheet.pdf>. These 2004 poll results are consistent with a 2001 nationwide poll, in which 76% of those surveyed stated their belief that campaign contributions influence judges’ decisions. GREENBERG QUINLAN ROSNER RESEARCH INC. & AM. VIEWPOINT, JUSTICE AT STAKE FREQUENCY QUESTIONNAIRE 4 (2001), http://www.gqrr.com/articles/1617/1412_JAS_ntlsurvey.pdf.

109. GREENBERG QUINLAN ROSNER RESEARCH INC. & AM. VIEWPOINT, JUSTICE AT STAKE – STATE JUDGES FREQUENCY QUESTIONNAIRE 5 (2002), http://www.gqrr.com/articles/1617/1411_JAS_judges.pdf.

110. *Id.* at 9.

111. See *infra* Part II.E.

In 2004, there was a race for the Illinois Supreme Court, right here. It cost just over \$9 million for that race. As you might have guessed, the winner of that race got his biggest contributions from a company that had an appeal pending before the Illinois Supreme Court. You like that? . . . Sounds a lot like the *Caperton* case, doesn't it?¹¹²

A recent study of rulings by the Ohio Supreme Court by Adam Liptak and Janet Roberts added a layer of empirical support for these perceptions.¹¹³ The study found that Ohio justices routinely sat on cases after having received campaign contributions from the parties involved, and that they then voted in favor of those contributors 70% of the time. One justice voted in favor of his contributors 91% of the time.¹¹⁴

The perception that campaign contributions buy influence on the bench in pending or imminent cases is so strong that litigants and lawyers give even when their candidate cannot lose. A recent *Los Angeles Times* study found that Nevada judges running unopposed collected hundreds of thousands of campaign dollars from litigants and lawyers, frequently "within days of when a judge took action in the contributor's case"¹¹⁵

D. Super Spenders

It is useful to recall Professor Briffault's analysis of 527 "megadonors" in the federal context. Briffault noted, most strikingly, that "\$142.5 million in 527 funds . . . came from just two dozen megadonors" in the 2004 presidential election while, by comparison, "\$149.2 million in public funds came from the entire American public."¹¹⁶ Briffault's eye-popping analysis is consistent with figures pointing to the outsized influence of a very few big spenders in state judicial elections. Super spenders fueled much of the cash boom in judicial elections in the past decade. These super spenders, who are stakeholders in litigation, can dominate election spending in either of two ways: (1) by writing huge campaign

112. Abdon M. Pallasch, *O'Connor Urges Illinois to Select Judges by Merit*, CHI. SUN-TIMES, May 20, 2010, at 7.

113. Adam Liptak & Janet Roberts, *Campaign Cash Mirrors a High Court's Rulings*, N.Y. TIMES, Oct. 1, 2006, at A1, available at http://www.nytimes.com/2006/10/01/us/01judges.html?_r=1&scp=1&sq=Campaign%20Cash%20Mirrors%20a%20High%20Court's%20Rulings&st=cse.

114. *See id.*

115. Michael J. Goodman & William C. Rempel, *In Las Vegas, They're Playing with a Stacked Judicial Deck*, L.A. TIMES, June 8, 2006, at A1, available at <http://www.corpwatch.org/article.php?id=13692>.

116. Briffault, *supra* note 63, at 964.

checks to candidates or (2) by going outside the system and spending millions on their own independent TV ads and other election communications, seeking to help or hurt a candidate on the ballot.¹¹⁷

“A review of ten states with the highest judicial campaign spending shows two separate worlds—a small coterie of organized super spenders who dominate election financing, and a large number of small contributors who simply cannot keep up.”¹¹⁸ As a consequence, a large number of justices in those states owe their elections to a few key benefactors. The gap is best shown by twenty-nine elections held from 2000–2009 in these ten states.¹¹⁹ In each of these elections, at least one of the candidates benefited from \$1 million or more in other people’s money—either in direct contributions or through independent election spending by other groups.¹²⁰

The top five super spenders from each of the twenty-nine elections—a total of 145 super spenders—spent an average of \$473,679 apiece.¹²¹ By contrast, the remaining 116,000 donors averaged \$850.¹²² One finds relative super spenders even within the super spender category. Half of the money from those 145 spenders came from just twenty sources even while excluding money from self-financing candidates.¹²³ In all, the 145 super spenders accounted for slightly over 40% of all campaign cash in the twenty-nine elections.¹²⁴ Moreover, the disparity was widespread and not just the result of a few outlier contests. In twenty-two of twenty-nine elections, the top five spenders averaged more than \$200,000 apiece—and in twelve elections, they averaged more than \$500,000.¹²⁵ In twenty-one of twenty-nine elections, a mere five spenders accounted

117. SAMPLE ET AL., *supra* note 11, at 9.

118. *Id.* at 9.

119. These ten states are Alabama, Ohio, Pennsylvania, Illinois, Texas, Michigan, Mississippi, Louisiana, Nevada, and West Virginia. *See id.* at 6 fig.2.

120. *See id.* at 11.

121. *See id.* at 10–11.

122. *Id.* at 10.

123. *Id.* This figure is based on the Author’s analysis of the raw data underlying the super spender figures used in the *New Politics* report. The raw data are on file with the Author.

124. *Id.* at 9. This further breakdown of the top twenty spenders within the 145 super spenders does not appear within the *New Politics* report itself, but is a further calculation using the same raw spending data that my Co-Authors and I used for the report. Likewise, the median for expenditures by the 145 super spenders is \$198,016. The raw data are on file with the Author.

125. *Id.*

for at least 25% of all campaign funding.¹²⁶ In nine elections, five super spenders accounted for more than 50%, exceeding thousands of contributors combined.¹²⁷ Finally, of the super spenders that laid out more than \$100,000, only one was an individual: Don Blankenship, whose \$3 million in expenditures in the 2004 West Virginia election led to *Caperton*.¹²⁸

Will this super spender trend continue? If so, it will inevitably lead to courtrooms across the country in which the physical separation of the litigation aisle demarks dramatic asymmetries in financial support. Clearly, no one can say for certain, but there is every reason to believe the trends will not only continue, but worsen. Perhaps most tellingly, consider that the above-described sea changes in judicial elections occurred in the pre-*Citizens United* era, an era in which many states had laws equivalent to the federal bans on spending corporate and union general treasury funds (as opposed to via PACs) on state campaigns.¹²⁹ No such restrictions survive *Citizens United*, a stark reality that, as Justice John Paul Stevens recognized in dissent, is likely to have a particularly pronounced effect on court campaigns: “At a time when concerns about the conduct of judicial elections have reached a fever pitch, the Court today unleashes the floodgates of corporate and union general treasury spending in these races.”¹³⁰

E. Case Study: Avery v. State Farm

Behind these figures are the individual races that drive those numbers. In 2004—the year of the West Virginia campaign that culminated in *Caperton*—Illinois arguably best exemplified the nexus of pending high-stakes litigation and big-money campaigns. In a race for a seat on the Illinois Supreme Court, two candidates combined to raise more than \$9.3 million, a figure that was nearly double the previous national record for a state judicial election.¹³¹ To put the sum in further perspective, \$9.3 million was also more than was spent in eighteen U.S. Senate races that year.¹³² Moreover, the sums were not necessitated by a costly statewide campaign akin

126. *Id.*

127. *Id.*

128. *Id.*

129. *Life After Citizens United*, NATIONAL CONFERENCE OF STATE LEGISLATURES, <http://www.ncsl.org/default.aspx?tabid=19607#laws> (last visited Dec. 28, 2010).

130. *Citizens United v. FEC*, 130 S. Ct. 876, 968 (2010) (citation omitted) (Stevens, J., concurring in part and dissenting in part).

131. See James Sample, *The Campaign Trial: The True Cost of Expensive Court Seats*, SLATE, Mar. 6, 2006, <http://www.slate.com/id/2137529>.

132. GOLDBERG ET AL., *supra* note 88, at 18.

to a U.S. Senate race. Instead, the Illinois Supreme Court seat in question was elected by a single district—and a rural one at that.¹³³ The record sums were less a natural progression than they were a function of litigation timing.

In May 2003, the Supreme Court of Illinois heard oral arguments in the appeal of *Avery v. State Farm Mutual Automobile Insurance Company*.¹³⁴ “The dispute involved a class action against State Farm on behalf of 4.7 million policyholders in forty-eight states.”¹³⁵ Although the argument was heard in May 2003, the appeal would not be decided until seventeen months later, after the November 2004 election. In other words, the appeal was pending before the Supreme Court of Illinois—and had been for a year—by the time the 2004 campaign began.¹³⁶ The stakes in *Avery* were huge. State Farm’s appeal sought to overturn a \$1 billion lower-court verdict against the insurer, including \$456 million in contractual damages.¹³⁷

Lloyd Karmeier, the victorious candidate in the race, was supported by \$350,000 in direct contributions from employees, lawyers, and others (e.g., amici) directly involved with State Farm and/or its pending appeal, and by an additional \$1 million from larger groups of which State Farm was a member or to which it contributed.¹³⁸

133. *Id.*

134. 835 N.E.2d 801 (Ill. 2005), *cert. denied*, 547 U.S. 1003 (2006).

135. Sample, *supra* note 131.

136. *Id.*

137. *Avery*, 835 N.E.2d at 817.

138. For example, Justice Karmeier received \$1.9 million in contributions from the Illinois Republican Party, which received over \$2 million from the U.S. Chamber of Commerce. See GOLDBERG ET AL., *supra* note 88, at 26 fig.17. Employees at State Farm were directors of the U.S. Chamber of Commerce. Justice Karmeier also received nearly \$1.2 million in contributions from the Illinois Civil Justice League’s political committee, JUSTPAC. Sample, *Campaign Trial*, *supra* note 131. The largest contributors to JUSTPAC included the American Tort Reform Association (\$415,000), the U.S. Chamber and the Illinois Chamber of Commerce (\$200,000), and the Coalition for Jobs, Growth and Prosperity (\$150,000). See RACHEL WEISS, THE INSTITUTE ON MONEY IN STATE POLITICS, FRINGE TACTICS: SPECIAL INTEREST GROUPS TARGET JUDICIAL RACES 11 (2005), available at <http://www.followthemoney.org/press/Reports/200508251.pdf>. State Farm was a member of and contributor to both the American Tort Reform Association and the Illinois Coalition for Jobs, Growth and Prosperity. See *Petition for Writ of Certiorari* at 8 n.3, *Avery v. State Farm Mut. Auto. Ins. Co.*, 547 U.S. 1003 (2006) (No. 05-842), 2005 WL 3662258, at *8 [hereinafter “*Avery Petition*”].

Justice Karmeier’s opponent, Judge Gordon Maag, received \$2.8 million in contributions from the Illinois Democratic Party, \$1.2 million from Justice For All PAC, and more than \$50,000 from the Illinois State Federation of Labor. See GOLDBERG ET AL., *supra* note 88, at 19; *Avery Petition* at 7. The Illinois Democratic

Immediately upon winning the race, Karmeier called the fundraising in the race “obscene for a judicial race” and queried, some might say deliciously: “What does it gain the people? How can people have faith in the system?”¹³⁹ Whatever the answers to the questions were at the time Karmeier posed them, the people’s “gain[s]” and “faith in the system” were unlikely to increase when, almost immediately upon taking the bench, Karmeier cast a tie-breaking vote nixing the \$456 million claim against State Farm.¹⁴⁰

An editorial in the *St. Louis Post-Dispatch* reflecting on *Avery* explained that Justice Karmeier also cast the deciding vote in reversing a \$10.1 billion judgment against Philip Morris USA.¹⁴¹ Reportedly, Philip Morris, along with a business lobbying group backing it, spent more than \$1 million supporting Karmeier in the 2004 election.¹⁴² Lawyers, judges, and press around the state and country expressed concerns summed up by the *Post-Dispatch*’s editorial:

The juxtaposition of gigantic campaign contributions and favorable judgments for contributors creates a haze of suspicion over the highest court in Illinois. . . . Although Mr. Karmeier is an intelligent and no doubt honest man, the manner of his election will cast doubt over every vote he casts in a business case. This shakes public respect for the courts and the law—which is a foundation of our democracy.¹⁴³

The *Avery* plaintiffs sought review on due process grounds, and the Supreme Court denied certiorari.¹⁴⁴ When I analyzed the chances for certiorari at the time, I believed that, short of express criminal bribery, the following circumstances could scarcely get more compelling: (1) the race was the nation’s most expensive judicial race ever; (2) the election happened while an appeal was pending before the court—and had been for seventeen months

Party received almost \$2 million from lawyers and law firms. Justice For All PAC’s \$1.2 million in contributions to Judge Maag constituted the entirety of its political expenditures in 2004. See GOLDBERG ET AL., *supra* note 88, at 27 fig.18. The organization received a combined \$670,000 from an Illinois law firm, Simmons Firm LLC, and one of its attorneys, Randall A. Bono. *Id.* Justice for All PAC also received more than \$90,000 in contributions from attorney Stephen N. Tillery, and a \$50,000 contribution from attorney Barry Johan. See *id.*

139. GOLDBERG ET AL., *supra* note 88, at 19.

140. Sample, *supra* note 131.

141. Editorial, *Buying Justice?*, ST. LOUIS POST-DISPATCH, Dec. 20, 2005, at B8.

142. *Id.*

143. *Id.*

144. *Avery v. State Farm Mut. Auto. Ins. Co.*, 547 U.S. 1003 (2006). By way of disclosure, I served as counsel on an amicus brief in support of certiorari in *Avery*.

including the entirety of the campaign; (3) the principal funders were interested parties in the litigation; and (4) the winning candidate then cast an outcome-determinative vote.¹⁴⁵ I thus believed these circumstances were enough for the Supreme Court to consider whether due process could be jeopardized by outsized judicial campaign support. Therefore, once the Court denied certiorari,¹⁴⁶ I believed the Court would almost certainly continue to decline to weigh in on the due process question. Fortunately, as discussed at length in Part III, my analysis proved to be wrong—not only with respect to the question of certiorari on the issue, but also with respect to whether the brave new world of judicial elections would produce an even more egregious scenario.

It is worth noting briefly here that one of the most interesting (and essentially unnoticed) distinctions between *Avery* and *Caperton* pertains specifically to the issue of the *Buckley* contribution–expenditure dichotomy. In *Avery*, even the large contributions, while funneled through a variety of political action committees, trade groups, and umbrella organizations, ultimately made their way to the candidates as direct contributions. This was made possible by the fact that Illinois had no state contribution limits whatsoever, symptomatic perhaps of a certain kind of state politics. On the contrary, under West Virginia law, an individual may give no more than \$1000 per election cycle to a campaign.¹⁴⁷ Thus, in 2004, Don Blankenship contributed the \$1000 maximum to his favored candidate’s campaign¹⁴⁸ and then spent \$3 million to support that candidate the only ways he legally could—by personal independent expenditures and via a 527.¹⁴⁹ By way of contrast, equivalent sums were given to the candidates in *Avery* directly.¹⁵⁰

Comparing *Avery* and *Caperton* thus highlights that from a *Buckley* standpoint, for all their similarities, significant differences exist regarding the monetary sources and forms at issue. In *Avery*, the pertinent spending came in the form of direct contributions from concentrated, though still comparatively dispersed sources. In

145. See Brief for 12 Organizations Concerned About the Influence of Money on Judicial Integrity, Impartiality, and Independence as Amici Curiae Supporting Petitioners at 3, 9, 19–20, *Avery*, 547 U.S. 1003 (No. 05-842), 2006 WL 295175, at *3, *9, *19–20; see also Sample, *supra* note 131.

146. *Avery v. United States*, 130 S. Ct. 766 (2009).

147. W. VA. CODE ANN. § 3-8-12(f) (West 2010).

148. *Caperton v. A.T. Massey Coal Co.*, 129 S. Ct. 2252, 2257 (2009).

149. See *infra* Part III (detailing legal categorizations of Mr. Blankenship’s expenditures).

150. See *supra* note 138.

Caperton, the pertinent support came exclusively in the form of independent expenditures, but entirely from one individual.

This approach would leave a wise but not wealthy candidate for judicial office who faces a substantial independent expenditure campaign with essentially three options: (1) subtly and sometimes not so subtly signaling a favorable disposition towards other organized, concentrated interests—a clear abdication of the ideals of the judicial role;¹⁵¹ (2) the daunting and statistically almost impossible specter of raising enough money to meaningfully counter the funded speech—even inaccurate funded speech without engaging in option (1); or (3) surrender—often before ever entering a race at all, thereby creating an adverse selection problem that degrades the quality of the courts, even apart from concerns about their impartiality.

III.

“CAPERTON CONTRIBUTIONS” AND DUE PROCESS

Part III of this Article coins the term “*Caperton* contributions” to describe the equation of independent expenditures with contributions in the judicial elections context. As Justice Stevens articulates in dissent in *Citizens United*, *Caperton* “underscores the old insight that, on account of the extreme difficulty of proving corruption, ‘prophylactic measures, reaching some campaign spending not corrupt in purpose or effect, may be nonetheless required to guard against corruption.’”¹⁵² Justice Stevens, of course, is using *Caperton* to argue for such prophylactic measures in the non-judicial elections context, where there is no commitment to impartiality, and where there is an entirely different democratic expectation—namely, of constituent influence, including moderate influence via contribution-limited financial support. In the context of the courts, financial influence itself is a concern, as is the perception thereof. Part III of this Article thus argues that just as the difficulty of prov-

151. Professor Karlan recently described this dynamic as follows:

Money can play a critical role in judicial elections. Especially because many judicial elections are low-salience, down-ballot races, political spending often serves as the major source of information to voters. Just as judicial candidates may face a temptation to shade their decisions to attract voters’ support, so too they may face the temptation to shade their decisions to attract the financial support that enables them to appeal to voters.

Pamela S. Karlan, *Electing Judges, Judging Elections, and the Lessons of Caperton*, 123 HARV. L. REV. 80, 90 (2009) (citation omitted).

152. *Citizens United v. FEC*, 130 S. Ct. 876, 968 (2010) (Stevens, J., concurring in part and dissenting in part) (quoting *Buckley v. Valeo*, 424 U.S. 1, 30 (1976)).

ing corruption justifies prophylactic measures, the differences in democratic expectations in the courts as compared to the constituent branches justify differential treatment of expenditures in a judicial elections context. Given the controversial and compelling equation of expenditures and contributions in *Caperton*, this Article coins the term “*Caperton* contributions” to provide a prospective shorthand for just such differential treatment.

A. *Facts That Push Formalism Beyond Its Limits*

In August 2002, a West Virginia jury returned a verdict that found A.T. Massey Coal Co. and its affiliates liable for fraudulent misrepresentation, concealment, and tortious interference with existing contractual relations.¹⁵³ The jury awarded plaintiffs Hugh Caperton, Harman Development Corp., Harman Mining Corp., and Sovereign Coal Sales the sum of \$50 million in compensatory and punitive damages.¹⁵⁴ During the consideration of post-verdict motions and just prior to appealing the verdict to West Virginia’s sole appellate court, Massey’s chairman, chief executive officer, and president, Don Blankenship, spent approximately \$3 million in 2004 supporting attorney Brent Benjamin in a campaign against an incumbent of that court, West Virginia Supreme Court of Appeals Justice Warren McGraw.¹⁵⁵

Blankenship contributed \$1000, the statutory maximum under West Virginia law, to Benjamin’s campaign.¹⁵⁶ Additionally, he spent \$517,707.53 in his own name “to support” Justice Benjamin’s candidacy, including via radio and newspaper advertisements, campaign fliers, and telephone calls to registered voters.¹⁵⁷ Most substantially, Blankenship donated almost \$2.5 million to a 527 organization named “And For The Sake Of The Kids,” which supported Benjamin and/or attacked McGraw.¹⁵⁸

Finally, though it is much less easily quantified both by nature and intent, Blankenship—as the CEO of one of the largest compa-

153. *Caperton*, 129 S. Ct. at 2254.

154. *Id.* at 2257.

155. *Id.*

156. *Id.*

157. *Id.* (quoting Blankenship’s state campaign financial disclosure filings). Mr. Blankenship’s \$3 million in expenditures were, by legal definition, for the purpose of influencing the election. Mr. Blankenship spent \$517,707 that he reported to the state as “expenditures” on W. Va. Official Form F-7B. Form F-7B is used only to report an expenditure “for a communication which expressly advocates the election or defeat of a clearly identified candidate” See W. VA. CODE ANN. §3-8-1a(15) (West 2010) (defining “independent expenditure”).

158. See *Caperton*, 129 S. Ct. at 2257.

nies in the state—was instrumental in helping to raise additional money contributed to the 527, as well as a significant portion of the roughly \$800,000 amassed by the “official” Benjamin campaign. In little noticed but startling comments—particularly from a jurist—just prior to oral argument in the U.S. Supreme Court, former West Virginia Chief Justice Richard Neely described Blankenship’s less quantifiable efforts:

Just factually, Massey Coal buys a lot of stuff. It buys mine gear. It buys . . . equipment. It buys Caterpillars. Blankenship knocked down every person Massey did business with, for the purpose of raising money to go into this 527. The 527, as a matter of fact, was supremely well coordinated with Justice Benjamin’s campaign. And Massey saw to it that clean, reportable contributions from all kinds of people that nobody ever heard of—little old ladies in tennis shoes, right—who were the mothers of people who owned equipment suppliers, right, went to Benjamin himself so that he could put up huge billboards that said, “Who is Brent Benjamin?”—Just, name recognition, to get the campaign off the ground.¹⁵⁹

Even completely setting aside money he might have raised through these less than salutary methods, however, Blankenship’s expenditures were substantial. Blankenship’s spending accounted for 60% of all expenditures in support of Benjamin’s candidacy. Blankenship was thus very much a “super spender.” Indeed, Blankenship was the principal financial force behind Benjamin’s campaign. Benjamin defeated the incumbent McGraw; he then repeatedly refused to disqualify himself from the multimillion-dollar Massey case, although West Virginia’s Code of Judicial Conduct required Justice Benjamin’s disqualification whenever his impartiality “might reasonably be questioned.”¹⁶⁰

159. Press Release, American Constitution Society, Judicial Elections and Due Process: On Caperton v. A.T. Massey Coal Company, et al. (Feb. 26, 2009), *available at* <http://www.acslaw.org/node/8257>.

160. *See* W. VA. CODE OF JUDICIAL CONDUCT CANON § 3(E)(1) (2010) (requiring a judge to “disqualify himself or herself in a proceeding in which the judge’s impartiality might reasonably be questioned”); *see also* JAMES SAMPLE, DAVID POZEN & MICHAEL YOUNG, FAIR COURTS: SETTING RECUSAL STANDARDS 17–18 (Brennan Center for Justice 2008), *available at* http://brennancenter.org/content/resource/fair_courts_setting_recusal_standards/ (noting near universality of this standard in the states). I have asserted the relatively non-controversial proposition that, while the “myriad extreme facts in the case presented the Court with a landmark question of constitutional law . . . it was an easy state question that gave birth to the difficult federal issue.” James Sample, Caperton: *Correct Today, Compelling Tomorrow*, 60 SYRACUSE L. REV. 293, 294 (2010).

Instead, in November 2007, Benjamin voted in a 3–2 majority to overturn the jury’s verdict against Massey; after the decision, he had unusually plentiful opportunities to reconsider his refusal to disqualify himself. Previously undisclosed photographs were uncovered that showed Blankenship’s coziness with the court extended beyond Benjamin himself. In January 2008, photos surfaced depicting Blankenship and another member of the 3–2 majority, then-Chief Justice Elliot Maynard, vacationing together in the French Riviera during the pendency of Blankenship’s company’s appeal.¹⁶¹ After these photos were discovered and published on front pages of newspapers across the country,¹⁶² Maynard stepped down from the case.¹⁶³ He nonetheless asserted he had remained impartial all along. Blankenship’s cozy relationship with Maynard and his separate support of Benjamin led still a third Justice to take the unusual step of criticizing Blankenship—the CEO of a litigant before the court—and his own colleagues. Larry Starcher, one of the original dissenting justices in that 3–2 decision, decried what he termed the “cancer” of moneyed influence in his court. He further asserted that “John Grisham got it right when he said that he simply had to read the *Charleston Gazette* to get an idea for his next novel.”¹⁶⁴ He

161. *Caperton*, 129 S. Ct. at 2258.

162. See, e.g., Adam Liptak, *Motion Ties W. Virginia Justice to Coal Executive*, N.Y. TIMES, Jan. 15, 2008, available at <http://www.nytimes.com/2008/01/15/us/15court.html>.

163. *Caperton*, 129 S. Ct. at 2258.

164. *Id.* The details of each of the recusal motions are beyond the scope of this article, which is focused prospectively on the issue of independent expenditures as opposed to a retrospective just on *Caperton* itself. It is worth noting, however, that Benjamin’s denials were accompanied by memoranda reflecting a complete lack of understanding of the rule-based, as opposed to constitutionally based disqualification rules. Jeffrey Stempel recently wrote an excellent analysis to this end; Stempel’s analysis focused on Benjamin’s assertion that “no objective information” had been advanced demonstrating that Benjamin would actually be biased. Stempel incisively notes the following:

[Benjamin’s] memorandum establishes that he misunderstood—or did not want to understand—what is meant by the objective test for recusal. Instead of focusing on what outside observers would think based on their observations (for example, \$3 million in campaign support from interested de facto party for judge deciding case important to benefactor), Justice Benjamin instead seems to view the objective test as a matter of whether he personally is persuaded by the motion for recusal, notwithstanding whatever number of reporters, editorialists, or commentators may disagree. In addition, Justice Benjamin appears to focus primarily on whether he is biased or prejudiced, giving implicit short shrift to the correct standard of impartiality. Even when nodding in this direction, Justice Benjamin mangles the concept. He states that recusal is not required because there is insufficient evidence to prove that he would fail to be fair and impartial—but the correct inquiry is whether ob-

then disqualified himself from the case, urging Benjamin to do likewise.¹⁶⁵

Subsequently, the West Virginia Supreme Court, with two substitute judges replacing Starcher and Maynard, re-heard the appeal. Benjamin once again voted in a 3–2 majority to overturn the verdict against Massey. Hugh Caperton and Harman Mining successfully sought certiorari in the U.S. Supreme Court, arguing that Benjamin’s failure to recuse himself violated the Fourteenth Amendment’s guarantee of Due Process.¹⁶⁶

B. *Caperton and the Buckley Conundrum*

1. Framing the Issue Before the Court

The petitioners in *Caperton*, led by Ted Olson, clearly and correctly characterized Blankenship’s support for Benjamin as independent expenditures, i.e., they were on the wrong side of the *Buckley* dichotomy. For example, in their merits stage brief before the Court, Petitioners stated that “Mr. Blankenship contributed the maximum amount permitted by West Virginia law to Benjamin’s campaign committee. He then spent 3,000 times that amount—some \$3 million—to underwrite independent advertisements supporting Benjamin, while publicly urging others to make additional donations to the campaign.”¹⁶⁷ Then, in attempting to refute Benjamin’s own contention that his “campaign was completely indepen-

servers would harbor doubts about his impartiality. Instead of playing by the rules and correctly applying the proper standard, Justice Benjamin imposes a burden not required by law—the burden to prove that he cannot be impartial.

Jeffrey W. Stempel, *Impeach Brent Benjamin Now!? Giving Adequate Attention to Failings of Judicial Impartiality*, 47 SAN DIEGO L. REV. 1, 38 (2010) (internal citations omitted). Under the circumstances, the rule-based necessity of Benjamin’s disqualification was apparent to virtually all involved—except Benjamin himself. In dissent on the constitutional ruling requiring Benjamin’s disqualification, Justice Scalia, for example, nonetheless noted that that it is “[u]ndoubtedly” clear that “[i]n the best of all possible worlds” judges should “sometimes recuse even where the clear commands of . . . due process law do not require it. . . .” See *Caperton*, 129 S. Ct. at 2275 (Scalia, J., dissenting).

165. See *Voluntary Disqualification of Justice Larry V. Starcher, A.T. Massey Coal Co. v. Caperton*, 679 S.E.2d 223, (W. Va. 2008) (No. 33350), available at <http://www.state.wv.us/wvscap/press/caperton.pdf>.

166. See *infra* Part IV.A (addressing in depth close interrelationship that exists between due process clause, judicial elections, and public’s strong desire for fair and impartial judiciary).

167. Brief for Petitioners at 2, *Caperton v. A.T. Massey Coal Co.*, 129 S. Ct. 2252 (2008), 2008 WL 5433361.

dent of any independent expenditure group,”¹⁶⁸ the petitioners contended further that the “debt of gratitude in the case [was] not diminished” merely by the fact that Blankenship’s support involved independent expenditures because the “end result was the same: Mr. Blankenship’s expenditures were directly responsible for hundreds of pro-Benjamin and anti-McGraw campaign advertisements that unquestionably helped Justice Benjamin—a previously unknown and underfunded candidate—prevail in his sharply contested race.”¹⁶⁹ Amici in support of Caperton also emphasized the fact that the majority of Blankenship’s support for Benjamin came in the form of independent spending.¹⁷⁰

For Massey, however, *Buckley* formalism trumped the facts. Accordingly, Massey argued to the Court that “Blankenship contributed only \$1,000 directly to Benjamin’s campaign. Apart from a \$1,000 contribution by A.T. Massey Coal Company’s PAC neither Massey nor any of its subsidiaries contributed money to Benjamin’s campaign.”¹⁷¹ Phrased differently, arguing from *Buckley*, the focus of the case was a garden variety \$1000 campaign contribution, while the elephant in the room—three-million dollars in independent support—was but an inconsequential footnote. If this proposition strikes one as absurd, it is all the more remarkable that, as noted in the introduction to this Article, precisely such a result would have obtained were it not for Justice Kennedy’s vote.¹⁷² In fact, Chief Justice Roberts’s dissent fully embraces and endorses the Alice in Wonderland qualities of a rigidly-applied *Buckley* contributions–expenditures distinction, such that the operative expenditure was \$1000 rather than \$3 million as follows: “It is true that Don Blankenship spent a large amount of money in connection with this election. But this point cannot be emphasized strongly enough: Other than a \$1000 direct contribution from Blankenship, *Justice Benjamin and his campaign had no control over how this money was spent.*”¹⁷³ Further playing the independent and uncoordinated

168. *Id.* at 34.

169. *Id.* at 17–18.

170. *See, e.g.*, Brief of the Brennan Center for Justice at NYU School of Law et al. as Amici Curiae Supporting Petitioners at 23, *Caperton*, 129 S. Ct. 2252 (No. 08-22), 2009 WL 45972 (asserting that “[a]s a practical matter, distinctions between contributions and expenditures have only marginal salience when it comes to the fundamental fairness concerns at the core of due process”).

171. Brief for Respondents at 4, *Caperton*, 129 S. Ct. 2252 (No. 08-22), 2009 WL 216165.

172. *Caperton*, 129 S. Ct. at 2256 (2009).

173. *Id.* at 2273 (Roberts, J., dissenting). Legal ethics expert Keith Swisher recently wrote an excellent piece in which he “respectfully dissent[ed] from the

card, Massey argued that Blankenship “did not and does not have any friendship or other personal connection with Benjamin, and [that Blankenship’s] support was not solicited by [Benjamin].”¹⁷⁴ As it happened, this latter assertion proved to be an instance of extraordinarily bad lawyering; the contention was not only demonstrably false, but was directly contradicted in the *New York Times* by, of all people, Blankenship himself, who said that that he and Benjamin had indeed met privately before the election; discussed the campaign; and specifically discussed, inter alia, “raising money.”¹⁷⁵ Worse, the false assertion in Massey’s brief was then repeated at oral argument.¹⁷⁶ The Massey brief went so far as to deny that Blankenship and Benjamin “even knew one another, before or after the election.”¹⁷⁷

Massey’s lawyers’ stunning misrepresentations about their own client in this regard were compounded by an equally bizarre strategic error. Immediately following oral argument in the U.S. Supreme Court, Massey’s lawyers sought leave to submit a supplemental brief to bring to the Court’s attention a transparently self-serving press release from the West Virginia Supreme Court of Appeals’ public information office—an office which was subject to oversight by Benjamin in his new capacity as Chief Justice.¹⁷⁸

dissent” stating that “every question, save one or two, can be answered (and the ones that cannot seem to reflect more poorly on the questioner’s drafting than the majority’s analysis).” Posting of Keith Swisher to Judicial Ethics Forum, Caperton: *Answers to the Chief Justice’s “Twenty Questions” Times Two* (June 15, 2009), <http://judicialethicsforum.com/2009/06/15/caperton-answers-to-chief-justice-roberts-twenty-questions-times-two>. Swisher also noted that “an umpire who merely calls balls and strikes should be less concerned with questions not before the court, and indeed, [that] every case could spawn a multitude of forward-looking questions not raised by the facts at hand. . . .” *Id.* Swisher’s article goes on to propose answers to each of the 40 questions, and should be considered a must-read for those interested in the case.

174. *Id.* at 3.

175. Adam Liptak, *Justices Hear Arguments on Money-Court Nexus*, N.Y. TIMES, Mar. 4, 2009, at A18, available at http://www.nytimes.com/2009/03/04/washington/04scotus.html?_r=1.

176. Transcript of Oral Argument at 30, *Caperton v. A.T. Massey Coal Co.*, 129 S. Ct. 2252 (2009) (No. 08-22), 2009 WL 527723 [hereinafter “*Caperton* Transcript”].

177. Brief for Respondents at 55, *Caperton v. A.T. Massey Coal Co.*, 129 S. Ct. 2252 (2009) (No. 08-22), 2009 WL 216165.

178. Supplemental Brief for Respondents at 1, *Caperton v. A.T. Massey Coal Co.*, 129 S. Ct. 2252 (2009) (No. 08-22) (citing Press Release, Supreme Court of Appeals, State of West Virginia, Summary of Chief Justice Benjamin’s Dispositive Voting Record Regarding Massey Energy Cases from 01/01/2005 to 12/31/2008 (Mar. 2, 2009), available at http://www.state.wv.us/wvsca/press/march2_09.htm),

Massey's gambit opened the door for Olson and Caperton to oppose the motion on the ground that the press release, consisting as it did of merely recycled arguments, failed to meet the standard for "new matter" deserving of the Court's attention.¹⁷⁹ Olson adroitly noted that, in contrast to the recycled press release, Blankenship's interview with the *Times*—conducted weeks earlier—was effectively made new all over again because Andrew Frey, Massey's lawyer, expressly denied before the U.S. Supreme Court a fact that his own client had expressly admitted to the press. Frey's denial was perhaps unwitting or perhaps strategic. Either way, it was clearly unwise. Indeed, the exchange was remarkable enough to those paying close attention that Liptak, acting in his capacity as the Supreme Court reporter for the *Times*, included the following passage in his story the day after the oral argument:

Justice Scalia explored the nature of the relationship between Mr. Blankenship and Justice Benjamin.

"This contributor never even met the judge, did he?" Justice Scalia asked Mr. Olson, who said the answer was not clear.

Justice Scalia's question appeared to be based on an assertion in a brief Massey filed with the court in January. It said "there is no indication that Blankenship and Justice Benjamin even knew one another, before or after the election."

But Mr. Blankenship said in an interview last month that he had met with Mr. Benjamin before the election. Mr. Blankenship said that his spending had been mainly intended to oust the incumbent justice, Warren McGraw, but that he nonetheless wanted to meet the justice's opponent.

"I thought, if I want to beat this guy I ought to know who's running against him," Mr. Blankenship said, adding that the meeting did not go well.

"When he got through talking to me, I said, 'Mr. Benjamin, I don't know who you are, but if you go around talking to business people about raising money, you need to do more listening than you do talking'," Mr. Blankenship recalled.¹⁸⁰

The above sequence, especially when viewed through the lens of former Chief Justice Neely's comments, reflects much more than

available at http://www.abanet.org/publiced/preview/briefs/pdfs/07-08/08-22_RespondentSuppmotion.pdf.

179. Opposition to Respondents' Motion for Leave to File a Supplemental Brief at 1, *Caperton v. A.T. Massey Coal Co.*, 129 S. Ct. 2252 (No. 08-22), available at http://www.abanet.org/publiced/preview/briefs/pdfs/07-08/08-22_PetitionerOpptoMotionSupp.pdf.

180. Liptak, *supra* note 175.

just the inside baseball of one particular case. Indeed, the sequence raises a common sense question germane to independent campaigns more generally. As a practical matter, just how surprising is it that a de jure uncoordinated \$3 million “independent” campaign would be coordinated de facto? Even the fabulously wealthy are unlikely to spend millions on the metaphorical complete mystery behind door number two. Suffice it to say that the only distinctions making the *Caperton* scenario unique in this respect are that usually, the person or entity responsible for spending the millions is hardly inclined to admit coordination (or near coordination) to the *New York Times*. Moreover, if they were to do so, it is even less likely that their lawyer (with honest ignorance being the best-case explanation) would contradict the client in briefs and oral presentations to the U.S. Supreme Court.

Speaking on the same panel as former West Virginia Chief Justice Richard Neely, judicial ethics scholar Amanda Frost adeptly framed Blankenship’s activities relationship to the dynamics of judicial campaigns more generally:

Blankenship recently gave an interview to the *New York Times* where he admitted . . . [that] “w[e] set it up to make Justice McGraw look bad on issues involving children. We found a case where he released someone who was accused of being a pedophile, we publicized that case, and he said, “but that was not why I started this organization. I started And For the Sake of Kids to get McGraw off the court because I wanted someone who could protect my corporate interests.”

That’s I think the most disturbing thing about this case because it’s happening not just in West Virginia, it’s happening in many, many states. People take a hot-button issue—like releasing somebody early from a sentence, or some pedophile getting out of jail . . . and use that as a means to try to get someone into office who they think will benefit them financially.¹⁸¹

To a certain extent, the portion of Frost’s second paragraph relating to the content of the campaigns accurately details a mix of the unfortunate and the inevitable—and a reality about which little

181. Amanda Frost, Press Briefing on *Caperton v. A.T. Massey Coal Company, et al.*, American Constitution Society: Judicial Elections and Due Process, at 56:20 (Feb. 26, 2009), available at <http://www.acslaw.org/node/8257>. Professor Keith Swisher has noted that “the tough-on-crime message, or some derivation thereof, is among the most, if not *the* most, prevalent in judicial campaigns.” Keith Swisher, *Pro-Prosecution Judges: “Tough on Crime,” Soft on Strategy, Ripe for Disqualification*, 52 ARIZ. L. REV. 317, 327 (2010).

can (or should) be done from a regulatory standpoint. Judicial selection, regardless of whether via appointment or election, involves politics, including aspects of the mess of politics. Too often, critics of judicial elections—and they have much to criticize—argue from the false premise that elections have cornered the judicial selection market for unseemliness, a lack of transparency, and misdirection. Would that it were so.¹⁸² Still, the distortional dynamics that Frost correctly describes are greatly exacerbated by big money election expenditures. Which is to say that it is a simple fact that, absent content regulation (a “cure” worse than the disease), misinformation is inevitable. Yet regardless of whether the information is accurate or misleading, two facts are indisputable: (1) unlimited millions of dollars can disseminate a whole lot more of it than can, for example, thousands of dollars; and (2) the candidates, against whom big-money independent expenditure campaigns are waged, have limited places to turn for the money required to respond, with key stakeholders before the bench being by far the most fertile—and independence threatening—sources.

2. Sub Silentio *Genius*

A rigid, formalistic application of *Buckley’s* contribution–expenditure distinction in *Caperton* would have resulted in exactly the position that Massey advocated—Blankenship’s trivial \$1000 contribution would not have triggered recusal even under the ethics rules, much less under the Due Process Clause. The \$3 million in expenditures would simply not have been relevant. Opening the Pandora’s Box of directly challenging *Buckley’s* contribution–expenditure distinction, however, was not an option in *Caperton*—and certainly would not have been a winning option had it been pressed.¹⁸³ Moreover, *Caperton* occurred just seven years after the Court’s decision in *Republican Party of Minnesota v. White*.¹⁸⁴ In *White*, the Court held that Minnesota’s “announce clause,” which prohibited a “candidate for a judicial office” from “announc[ing]

182. See Brennan Center for Justice at New York University School of Law, *Testimony of Adam Skaggs on MD Judicial Elections and Senate Bill 833* (Mar. 9, 2010), http://www.brennancenter.org/content/resource/testimony_of_adam_skaggs_on_md_judicial_elections_and_senate_bill_833/ (noting that merit systems do not of themselves provide guarantee of transparent, independent judiciary).

183. Note, in this regard, that even the far narrower ground that was ultimately pursued produced only a 5–4 victory for the Petitioners. *Caperton*, 129 S. Ct. at 2252 (2009).

184. 536 U.S. 765 (2002).

his or her views on disputed legal or political issues,”¹⁸⁵ violated the First Amendment.¹⁸⁶ The Court left little doubt that, despite the due process interest, much speech surrounding judicial campaigns was identical to that surrounding legislative and executive elections, at least inasmuch as it was protected by the First Amendment. Most notably, for our purposes of analyzing *Caperton*, Justice Kennedy concurred in *White*, stating:

[Minnesota] may strive to define those characteristics that exemplify judicial excellence. It may enshrine its definitions in a code of judicial conduct. It may adopt recusal standards more rigorous than due process requires, and censure judges who violate these standards. What Minnesota may not do, however, is censor what the people hear as they undertake to decide for themselves which candidate is most likely to be an exemplary judicial officer. Deciding the relevance of candidate speech is the right of the voters, not the State.¹⁸⁷

White made it clear that in judicial elections—just as in legislative or executive contests—“[d]ebate on the qualification of candidates” lies “at the core . . . of First Amendment freedoms, not at the edges.”¹⁸⁸ *White*, however, dealt with speech *qua* speech as opposed to money *qua* speech. Justice Scalia’s opinion for the Court in *White* stated, *inter alia*, that: “We neither assert nor imply that the First Amendment requires campaigns for judicial office to sound the same as those for legislative office.”¹⁸⁹ Still, *White* reinforced the notion that the *Buckley* split would apply to judicial and non-judicial contexts in exactly the same manner.

In *Caperton*, however, Justice Kennedy proved unwilling to push *Buckley* formalism to what would have been an illogical and untenable result. Rather than confronting the *Buckley* conundrum, or delving into the thicket of just how different judicial elections should or should not be, the Court narrowly held (5–4) that Benjamin’s participation violated the Due Process Clause of the Fourteenth Amendment. For present purposes, the italicized text is intended to highlight the Court’s treatment of the *Buckley* distinction vis-à-vis the holding in *Caperton*:

To provide some perspective, Blankenship’s \$3 million in *contributions* were more than the total amount spent by all other

185. *Id.* at 768 (quoting 52 MINN. CODE OF JUDICIAL CONDUCT CANON 5(A)(3)(d)(i) (2000)).

186. *Id.* at 774–84.

187. *Id.* at 794 (Kennedy, J., concurring).

188. *Id.* at 781 (internal quotations omitted).

189. *Id.* at 783.

Benjamin supporters and three times the amount spent by Benjamin's own committee.¹⁹⁰

...

There is a serious risk of actual bias—based on objective and reasonable perceptions—when a person with a personal stake in a particular case had a significant and disproportionate influence in placing the judge on the case by raising funds or directing the judge's election campaign when the case was pending or imminent. The inquiry centers on the *contribution's* relative size in comparison to the total amount of money *contributed* to the campaign, the total amount spent in the election, and the apparent effect such *contribution* had on the outcome of the election.¹⁹¹

...

Just as no man is allowed to be a judge in his own cause, similar fears of bias can arise when—without the consent of the other parties—a man chooses the judge in his own cause. And applying this principle to the judicial election process, there was here a serious, objective risk of actual bias that required Justice Benjamin's recusal.¹⁹²

To say that it did not take long for campaign finance experts to focus on *Caperton's* treatment of Blankenship's nominal independent expenditure as a contribution would be an understatement. Perhaps the best, and certainly the quickest example, is the post by Rick Hasen described at this start of the Article, suggesting that the "equating" of independent expenditures and campaign contributions in the *Caperton* opinion was "inadvertent."¹⁹³ On the contrary, Justice Kennedy's equation of Blankenship's expenditures with contributions was clearly intentional and enabled a victory for fact over formalism, a first-of-its-kind acknowledgment that in the judicial elections context, big money spending—whether in the form of contributions or expenditures—implicates not only the First Amendment but due process as well.

Critics pounced. Ronald Rotunda decried that Justice Kennedy "never explains why he blurred the distinction between contributions and expenditures. All we know is that Kennedy acknowledges (only once) that Blankenship engaged in "independent expenditure." But then, a dozen times he repeatedly re-labels these "inde-

190. *Caperton v. A.T. Massey Coal Co.*, 129 S. Ct. 2252, 2257 (2009) (emphasis added).

191. *Id.* at 2263–64 (emphasis added).

192. *Id.* at 2265 (emphasis added).

193. Hasen, *supra* note 3.

pendent expenditures” as “contributions.”¹⁹⁴ Rotunda’s frustration is that Justice Kennedy “treats the two words as synonyms and never explains why.”¹⁹⁵ Fair enough, but Rotunda goes further, asserting that “[o]ne can understand the concern . . . that parties or their lawyers may try to ‘buy’ a judge with their campaign *contributions* to their favorite candidate, but that could not have occurred in this case because no party, no litigant, no lawyer, and no one else connected to the Caperton case gave more than \$1,000, the statutory maximum.”¹⁹⁶ Translation: pay no attention to the \$3 million behind door number two.

Likewise, ardent campaign finance regulation opponent Brad Smith¹⁹⁷ notes that “[b]efore *Caperton*, the U.S. Supreme Court had explicitly rejected the argument that Congress may restrict the funding of independent activity that merely ‘benefits’ a candidate.”¹⁹⁸ While this is a valid characterization of the Court’s jurisprudence, like Rotunda, Smith goes further. He asserts the following:

Kennedy goes along with the popular press descriptions of the facts and some of the briefs supporting *Caperton*. These descriptions . . . routinely describe Blankenship’s activity in terms of “contributions,” “contributed,” etc., and Justice Kennedy’s opinion likewise calls Blankenship’s expenditures “contributions.” Collapsing the distinction may reflect poor draftsmanship or even a poor understanding of the facts. It may even reflect a willingness to abandon the contribution/expenditure distinction at the center of post-*Buckley* campaign finance law.¹⁹⁹

Note the strategic framing at work here. Smith presents the issue as if Justice Kennedy—not only a Supreme Court Justice, but a First Amendment stalwart, particularly in his opposition to cam-

194. Ronald D. Rotunda, *Judicial Disqualification in the Aftermath of Caperton v. A.T. Massey Coal Co.*, 60 SYRACUSE L. REV. 247, 262 (2010) (citing *Caperton*, 129 S. Ct. at 2256).

195. *Id.* at 263.

196. *Id.* at 258 (emphasis added).

197. See Bradley A. Smith, *Faulty Assumptions and Undemocratic Consequences of Campaign Finance Reform*, 105 YALE L.J. 1049 (1996).

198. Stephen M. Hoersting & Bradley A. Smith, *Speech and Elections: The Caperton Caper and the Kennedy Conundrum*, 2008-09 CATO SUP. CT. REV. 319, 335 (2008). Smith’s paper is co-authored with Stephen Hoersting, who works for the Center for Competitive Politics, an organization founded and chaired by Smith to advocate against campaign finance regulations. Center for Competitive Politics, About Us, http://www.campaignfreedom.org/about_us/ (last visited Jan. 1, 2011).

199. Hoersting & Smith, *supra* note 198, at 344.

paigned finance regulation—is merely passively “going along with the popular press.” For Smith and those who share his perspective, Justice Kennedy’s “use of [contribution] nomenclature may allow the Court to escape the logical problems for the protection of independent expenditures that seem to have been created by the ruling in *Caperton*.”²⁰⁰ Reasonable people can and do disagree about many aspects of campaign finance jurisprudence. However, it is worth pausing to query whether, from a satellite perspective, the “logical problems” that “have been created” have *Caperton* (which treats \$3 million in campaign support as consequential) as their creator, or Smith’s reading of *Buckley* (which treats \$3 million as inconsequential)?

Similarly, James Bopp, who successfully argued *White* in the Supreme Court, argues that “[c]onstruing *Caperton* beyond its legitimate scope” will lead to a circumstance in which “third party political speech during judicial campaigns will be chilled because those wishing to contribute or spend their own money to support a candidate . . . will refrain from doing so because it may somehow interfere with a judicial candidate’s ability to serve as a judge.”²⁰¹ For this argument, only one Amendment exists—the First. Further, only one interest—anti-corruption—is implicated by independent expenditures in the judicial elections context. Bopp states: “Corruption is a serious charge and serious charges demand serious evidence, particularly where one must overcome the presumption of impartiality accorded to judges. The government has the affirmative burden of demonstrating that its regulation of campaign finance minimizes corruption. Yet it has failed to meet this burden.”²⁰² *Caperton* properly leaves many questions unanswered. However, among the matters that *Caperton* expressly does settle is that both of the above propositions are now wrong as a matter of settled law.

First, if a third party is indeed chilled from engaging in substantial, disproportionate spending, it will not be because that spending somehow interferes with a judicial candidate’s ability to serve as a judge generally—but rather, only because the would-be spender recognizes that such financial support specifically interferes with the judge’s fitness to rule in the benefactor’s own case. Second, *Caperton* makes clear that corruption is not the only inter-

200. *Id.*

201. James Bopp, Jr. & Anita Y. Woudenberg, *Extreme Facts, Extraordinary Case: The Sui Generis Recusal Test of Caperton v. Massey*, 60 SYRACUSE L. REV. 305, 305, 335 (2010).

202. *Id.* at 313.

est that justifies balancing the First Amendment against other constitutional concerns. This reality, evident throughout the opinion, was perhaps most prominently foreshadowed by Justice Kennedy's statement to Massey's counsel at oral argument:

I want you to be able to elaborate your full theory of the case, but just so you know, it—it does seem to me that the appearance standard has—has much to recommend it. In part it means that you don't have to inquire into the actual bias; it's—it's more objective. Now, of course it has to be controlled, it has to be precise. But I just thought that you know that I—I do have that inclination.²⁰³

When Massey's counsel argued in response that Due Process cannot rest on appearances, Justice Kennedy replied: "But our whole system is designed to ensure confidence in our judgments. . . . And it seems—it seems to me litigants have an entitlement to that under the Due Process Clause."²⁰⁴

Drawing on this exchange from oral argument, Jed Shugerman points out that "[f]rom these unambiguous signals, it appears that an 'appearance of bias' standard had five votes as of March 3, 2009."²⁰⁵ Shugerman notes that while the Court's holding somewhat backed away from this reliance on appearances for purposes of Due Process analysis in favor of a "probability of actual bias" standard, this switch has the attribute of "seem[ing] to address real harm without having to prove something so inherently subjective."²⁰⁶ The shift, however, may have been more significant given that Justice Kennedy frames the issue being decided as whether there is a "serious risk of actual bias—based on objective and reasonable perceptions."²⁰⁷ In any event, the notion that "corruption" remains the only predicate for the consideration of interests other than the First Amendment is, after *Caperton*, demonstrably false.

Despite the isolated criticism from advocates of unlimited monetary influence in campaigns such as Smith and Bopp, the *Caperton* decision—including its treatment of massive independent expenditures—closely tracked the analysis suggested by, among others, Caperton's counsel Ted Olson, whose credentials as a con-

203. *Caperton* Transcript, *supra* note 176, at 33.

204. *Id.* at 37.

205. Jed Handelsman Shugerman, *In Defense of Appearances: What Caperton v. Massey Should Have Said*, 59 DEPAUL L. REV. 529, 541 (2010).

206. *Id.*

207. *Caperton v. A.T. Massey Coal Co.*, 129 S. Ct., 2262, 2263 (2009).

servative lawyer are unrivaled.²⁰⁸ Moreover, the counsel of the Conference of Chief Justices submitted the most remarkable brief in the case. The brief on behalf of the Conference of Chief Justices stated, in essence, four things: (1) “the Constitution may require the disqualification of a judge in a particular matter because of extraordinarily out-of-line campaign support from a source that has a substantial stake in the proceedings”;²⁰⁹ (2) “[b]ecause the applicability of the Due Process Clause in the campaign spending context depends on the particular facts of each case, no bright-line rule can or should be attempted”;²¹⁰ (3) fears that a ruling constitutionally requiring disqualification would “open the floodgates”—such as those ultimately articulated by Chief Justice Roberts and Justice Scalia—were “unfounded”;²¹¹ and (4) “due process review . . . would be limited to cases of extraordinary support.”²¹²

The CCJ’s brief marked the first time in the Conference’s six-decade history as an organized entity that it had filed an amicus brief in a case involving review of a state court judgment—that is to say, in effect, one of its own. Consequently, given that judicial pedigree, the Supreme Court’s ultimate decision itself, which closely tracks the CCJ’s perspective, has been particularly well received among jurists.²¹³ Likewise, for every critic there are advocates who describe *Caperton* as “a welcome result for anyone interested in en-

208. N.Y. Times, *People: Theodore B. Olson*, http://topics.nytimes.com/top/reference/timestopics/people/o/theodore_b_olson/index.html (last updated Oct. 19, 2009) (noting, inter alia, that “[e]ven before *Bush v. Gore*, Mr. Olson long served as the leading appellate litigator of the Republican establishment”).

209. Brief of the Conference of Chief Justices as Amicus Curiae Supporting Neither Party at 4, *Caperton v. A.T. Massey Coal Co.*, 129 S. Ct. 2252 (2009) (No. 08-22), 2009 WL 45973 [hereinafter “CCJ Brief”].

210. *Id.* at 22.

211. *Id.* at 23.

212. *Id.*

213. See John Schwartz, *Uncertainty in Law Circles Over New Rules for Judges*, N.Y. TIMES, June 10, 2009, at A20 (citing Alabama Chief Justice Sue Bell Cobb’s view that *Caperton* is a “good thing” because it will push judges to be more careful”); Nathan Koppel, *Massey Coal Ruling Getting Thumbs Up in Judicial Circles*, WALL ST. J. LAW BLOG (June 8, 2009, 2:47 PM) <http://blogs.wsj.com/law/2009/06/08/massey-coal-ruling-getting-thumbs-up-in-judicial-circles> (noting, among others, Indiana Justice Randall Shepard’s view that “it was wise of the majority to focus not just on the amount of a particular contribution . . . but its size relative to the total amount of contributions”); see also Tony Mauro, *Coping with Caperton*, THE BLT: THE BLOG OF LEGAL TIMES (June 10, 2009, 4:21 PM) <http://legaltimes.typepad.com/blt/2009/06/coping-with-caperton-a-conversation-with-tom-phillips.html> (quoting former Texas Chief Justice Tom Phillips’ view that “*Caperton* established a principle that is really important: There are constitutional concerns with a judge sitting in judgment of a case where a party is a significant donor”).

sureing the fairness and integrity of the judicial decision-making process.”²¹⁴ This is because, in large part, “the vast majority of the expenditures in *Caperton* were not direct contributions to the judge but rather were made by an independent group on the judge’s behalf”²¹⁵ and yet, the court majority clearly—if ever so quietly—recognized as much.

There are, of course, salient differences between the ex post remedy of constitutionally mandated recusal due to independent expenditures, and upholding ex ante limits on the expenditures themselves. Writing for the Court in *Citizens United*, Justice Kennedy zeroed in on the distinction by noting that “*Caperton*’s holding was limited to the rule that the judge must be recused, not that the litigant’s political speech could be banned.”²¹⁶ On its face, the statement is indisputably correct as a characterization of what *Caperton* did and did not address. It would, however, have been patently inappropriate for the Court to address the constitutionality of ex ante expenditure limits in *Caperton*—much less in state judicial elections generally—for numerous reasons, the most important of which was that no such limits were even remotely at issue in the case. Second, Justice Kennedy’s reference to a “ban”—as opposed to contribution-like limits—is uniquely suited to *Citizens United*, which did involve a ban on expenditures, albeit one applicable only to expenditures from a very particular source—corporate and union general treasuries.²¹⁷ By definition, expenditure limits allow more

214. J. Gerald Hebert, *A Victory for Judicial Integrity from High Court*, CAMPAIGN LEGAL CENTER BLOG (June 8, 2009), http://www.clcblog.org/blog_item-289.html (emphasis added).

215. *Id.*

216. *Citizens United v. FEC*, 130 S. Ct. 867, 910 (2010).

217. See, for example, *id.*, in which the opening sentence of the opinion for the Court notes that “federal law prohibits corporations and unions from using general treasury funds to make independent expenditures.” Many post-*Citizens United* discussions have involved more heat than light. Recent papers by Rick Hasen and Justin Levitt not only go against that grain, but offer compelling analyses of the decision and what can be reasonably anticipated to change in its aftermath. See Rick Hasen, *Citizens United and the Illusion of Coherence*, 109 MICH. L. REV. 581, 584 (2011) (noting that unusually broad language in *Citizens United* “will force the Court either to adopt a view that no limits on money in politics are ever constitutional or, more likely, to vote to sustain some limits [such as contribution limits] . . . through doctrinal incoherence”). Of particular pertinence to the thesis of this particular article, Hasen notes that due to the interplay of *Citizens United* and *Caperton*, the “Court’s new doctrine is already incoherent” in that “[t]he *Citizens United* majority did not satisfactorily explain how independent expenditures, which apparently cannot corrupt, were so corruptive, apparently corruptive, or distorting of a judicial election” as to lead to the result in *Caperton*. *Id.* at 584; see also Justin Levitt, *Confronting the Impact of Citizens United*, 29 YALE L. & POL’Y REV.

space for First Amendment interests than bans. In the judicial elections context, such limits also offer a reasonable *ex ante* mechanism for balancing the First Amendment against due process interests and the integrity of the courts. Far from undermining arguments for differential treatment of judicial elections, *Citizens United* underscores their very differences. The fact that Justice Kennedy writes the opinion for the Court in *Citizens United* as well as *Caperton*—reflects, among other things, what was obvious all along—Justice Kennedy knows, very well, the difference, under *Buckley*, between expenditures and contributions, and found it appropriate to equate the concepts only in a case involving court elections.²¹⁸

Finally, as the Conference of Chief Justices themselves noted in inviting the constitutional floor that the Court ultimately set in *Caperton*, going beyond that floor is part and parcel of “the state courts’ legitimate constitutional role in protecting the due process rights of all litigants to fair and impartial trials in elective judicial systems.”²¹⁹ Accordingly, there could scarcely be a more appropriate arena for a classical state-as-laboratory Brandeis approach—including via the adoption of the “*Caperton* contribution” concept of reasonable expenditure limits in judicial elections.

(forthcoming 2010), available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1676108 (noting, most pertinently, that *Citizens United* “flatly, and without empirical evidence, discounted [independent expenditure-based *quid pro quo*] exchanges as factual impossibilities,” leading commentators to correctly note that “the Court’s assessment contradicts its evaluation of very similar problems in *Caperton*”).

218. Rick Hasen asserts that the *Citizens United* “majority’s attempt to distinguish *Caperton* as a recusal case is unpersuasive” and posits, consistent with the thesis of this article, that he “suspect[s]” that Justice Kennedy believes that “independent spending does have the potential to corrupt, but *that outside the context of judicial elections* . . . the state’s interest in preventing such corruption is outweighed by the considerable First Amendment costs of limiting such spending.” Hasen, *supra* note 217, at 612 & n.231. Professor Hasen, commenting in that article on a draft of this one, notes that while he agrees with the result in *Caperton*, he thinks “relabeling independent spending as a ‘contribution’ . . . will only increase the incoherence of existing law,” but on the substance of the arguments, we come down in exactly the same place. *Id.* at n.231. The terminology of “*Caperton* contributions” has, in my view, the attribute of capturing the concept in just two words, but Hasen is also rightly concerned with labels, given the Court’s increasingly confounding campaign finance jurisprudence. *See id.*

219. CCJ Brief, *supra* note 209, at 14.

IV.
JUDICIAL CAMPAIGN EXPENDITURE LIMITS: AN
EXCEPTION BORN OF EXCEPTIONALISM

A. *The Due Process Interest in Limiting Judicial
Campaign Expenditures*

In his 2008 year-end report on the federal courts, Chief Justice Roberts, writing in the context of making the case for salary increases for federal judges, unintentionally underscored the core concern at issue in *Caperton* and in any case in which a significant stakeholder is also a super spender—regardless of whether that spending is nominally on the expenditure side of the *Buckley* divide. According to Chief Justice Roberts, America’s courts “guarantee that those who seek justice have access to a fair forum where all enter as equals and disputes are resolved impartially under the rule of law.”²²⁰ A judge generally has only two litigants before him or her. Consequently, from the standpoint of individual litigants—even in cases involving larger structural questions—what follows is almost always a zero-sum game: one litigant wins, and one loses.

A few commentators, including most notably Professor Karlan, have lamented that “fundamentally, *Caperton* continues the Court’s problematic insistence on addressing structural problems through the lens of protecting individual rights.”²²¹ Similarly, Bob Bauer argues that Justice Kennedy used “the only constitutional tool in the kit, Due Process,” and consequently “disguised a large public question as a private wrong, moving the resolution from any public political forum to the federal courthouse.”²²² The flip side of the allegedly “problematic insistence” is, of course, a necessary and healthy byproduct of the case or controversy requirement. The case arose in the context of the individual rights of Hugh Caperton and Harman Mining Company and the Court addressed it accordingly. Further, unfair adjudications, even for specific individuals or entities, pose a systemic risk to the legal system beyond the case at hand. If the public perceives, as it does, that justice is for sale, even if that is not true in all but the rarest of instances, then not only are

220. Chief Justice John G. Roberts, Jr., 2008 Year-End Report on the Federal Judiciary 7 (Dec. 31, 2008), available at <http://www.supremecourt.gov/publicinfo/year-end/2008year-endreport.pdf>.

221. Pamela S. Karlan, *Electing Judges, Judging Elections and the Lessons of Caperton*, 123 HARV. L. REV. 80, 81 (2009).

222. Bob Bauer, *Kennedy’s Problem in Caperton v. Massey and the Unfortunate Solution He Chose*, MORE SOFT MONEY HARD LAW WEB UPDATES, (June 12, 2009) http://www.moresoftmoneyhardlaw.com/updates/the_supreme_court.html?AID=1453.

individual litigants aggrieved, but the public's faith in the rule of law is severely undermined.

Structurally, the core role of the courts, as distinct from the legislative and executive branches, is to manifest the one branch of government where all—spenders and non-spenders alike—enter as equals under law, and where the public perceives as much. There can be no doubt, however, that monetary influence, in addition to its systemic challenges, also undermines the due process rights of individual litigants. Consequently, where states fail to include ex ante safeguards to address the concern systemically, ex post remedies like recusal are inherently and wisely focused—as *Caperton* was—on individual rights.

Just as importantly though, the individual rights frame is what gives judicial campaign finance its special flavor. Inherently, the protection of individual rights is a decidedly non-trivial matter to the aggrieved individuals who may “enter as equals,” but who enter the courtroom to face an adversary who spent millions supporting the campaign of the jurist who will decide their case. Moreover, in the courts context, the protection of a structural procedural right—to a hearing before an arbiter who is and objectively appears to be impartial—is patently absent in the constituent-based branches.

It is this individual-right-as-structural-right dynamic that Ted Olson captured in rebuttal during the oral argument before the Court. Massey's counsel Andy Frey implicitly encouraged the Justices to empathize with Justice Benjamin's perspective; correspondingly, these Justices would see his refusal to disqualify through the filter of their own faiths in their own respective commitments to impartiality. Frey asked the Justices the following rhetorical question during the argument: “[I]f you were in Justice Benjamin's situation, do you really think you would be incapable of rendering an impartial decision in a case involving Massey? Because if the answer to that is no . . . then there's no justification for saying that Justice Benjamin would.”²²³ Moments later during rebuttal, Olson deftly reframed the issue, encouraging the Justices to see the issue from the perspective of the individual litigant. In so doing, Olson was also implicitly drawing on the notion that the right the individual is seeking is, in fact, the most fundamental structural characteristic distinguishing the courts from the legislative and executive branches:

Would a detached observer conclude that a fair and impartial hearing would be possible? So instead of the question that Mr.

223. *Caperton* Transcript, *supra* note 176, at 41.

Frey was asking . . . I would like to ask you to ask this question. If this was going to be the judge in your case, would you think it would be fair and would it be a fair tribunal if the judge in your case was selected with a \$3 million subsidy by your opponent?²²⁴

Speaking at a September 2009 conference on state courts, hosted by Justice O'Connor, First Amendment scholar Kathleen Sullivan opened her remarks by indicating that, in her view, "the [*Caperton*] case was . . . a no-brainer . . . [and that] given all the particular facts and optics in the case, Justice Benjamin should have recused himself under the West Virginia standard or the ABA standard and it should have been an easy matter that didn't have to go to a Due Process ruling."²²⁵ Like Olson, Sullivan is generally opposed to campaign finance regulation,²²⁶ making her strong support of the decision in *Caperton* further evidence that judicial election restrictions are different. For Olson and Sullivan, opposition to most campaign limitations on First Amendment grounds, but support for broader limits in the judicial election context, is a principled juxtaposition. The position embraces the fundamental premise—disputed by those on the fringes—that judicial elections really are and ought to be different from legislative and executive races in constitutionally meaningful respects. To that end, it is worth noting that Olson argued both *Caperton* and *Citizens United* before the Supreme Court. Moreover, Justice Kennedy, who wrote opinions for the majority in both cases, agreed with his perspective on each occasion.²²⁷

224. *Id.* at 55–56.

225. Kathleen Sullivan, Professor, Stanford Law School, *One Symptom of a Serious Problem—Caperton v. Massey*, Address at the Seattle University School of Law (Sept. 14, 2009), available at <http://medialaw.seattleu.edu/events/judicial/>; see also Sample, *supra* note 12, at 294. While Benjamin's failure to disqualify himself under the rules proved so egregious as to become a constitutional case, less compelling scenarios put judges in a tough position. Indiana Professor Charlie Geyh, who has ably spearheaded an ABA project on recusal, recently testified before Congress that "judges who are deeply committed to the appearance and reality of impartial justice are called upon to acknowledge, in the context of specific cases, that despite their best efforts to preserve their impartiality, they are partial or appear to be so." *Examining the State of Judicial Recusals After Caperton v. A.T. Massey: Hearing Before the Subcomm. on Courts and Competition of the H. Comm. on the Judiciary*, 111th Cong. 2 (2009) (statement of Charles G. Geyh, Associate Dean of Research, John F. Kimberling Chair in Law, Indiana University Maurer School of Law at Bloomington), available at <http://judiciary.house.gov/hearings/pdf/Geyh091210.pdf>.

226. See *infra* note 228 and accompanying text.

227. See *Caperton v. A.T. Massey Coal Co.*, 129 S. Ct. 2252 (2009); *Citizens United v. FEC*, 130 S. Ct. 867 (2010).

The significance of Professor Sullivan's comments was largely lost on the national legal establishment. However, the comments raised the eyebrows of at least one campaign finance scholar, Rick Hasen, who was a fellow panelist with her in Seattle. Blogging about the conference, Hasen noted that "[i]t was a somewhat *surreal* experience to sit next to Kathleen (usually an ardent opponent of campaign regulations aside from disclosure) [and] to hear Kathleen call for consideration of a ban for everyone on independent spending in judicial campaigns"²²⁸ Most pertinently for the purposes of this article, Sullivan's "call" included the following:

The question, I think, that's relevant for us to think about . . . is what latitude is there for trying to impose more *ex ante* limits on the kind of judge buying that occurred, that led to the *Caperton* ruling. Now, I strongly support First Amendment limits on expenditure limitations in the political campaign context, but I think there is no reason to think they apply exactly the same way in judicial election contexts.²²⁹

Long prior to *Caperton*, the Court had recognized that "[a] fair trial in a fair tribunal is a basic requirement of due process."²³⁰ More than half a century ago, the Court articulated what it described as the "stringent rule" that to "perform its high function in the best way justice must satisfy the appearance of justice" even though such a rule "may sometimes bar trial by judges who have no actual bias and who would do their very best to weigh the scales of justice equally between contending parties."²³¹ Taken at face value, such a proposition is hardly controversial. Applying it is a much more difficult task. Accordingly, the Court had also recognized the "fundamental tension between the ideal character of the judicial office and the real world of electoral politics[.]"²³²

Testing this "tension" in *White*, Justice Scalia's opinion for the Court took pains to include the caveat that the Court was neither asserting nor implying that the "First Amendment requires campaigns for judicial office to sound the same as those for legislative office."²³³ The Scalia opinion included a significant exposition on due process, in which it found that while candidate announcements

228. Rick Hasen, *Sandra Day O'Connor: End Judge Elections in Wash.*, ELECTION LAW BLOG (Sept. 15, 2009, 8:17 AM), http://electionlawblog.org/archives/2009_09.html (emphasis added).

229. Sullivan, *supra* note 225.

230. *In re Murchison*, 349 U.S. 133, 136 (1955).

231. *Id.* (internal quotations omitted).

232. *Chisom v. Roemer*, 501 U.S. 380, 400 (1991).

233. *Republican Party of Minn. v. White*, 536 U.S. 765, 783 (2001).

relevant to issues were protected speech, preventing bias for or against particular parties is essential under the Due Process Clause.²³⁴

Seven years after *White*, for all the faldederal over its sensational and extreme facts, *Caperton* (despite its somewhat narrow holding) gave teeth to the protection of the specific aspect of due process that *White* endorsed. Whatever First Amendment interest a litigant or counsel has in contributing or spending in a campaign, she has no constitutionally protected interest in gaining an unfair litigation advantage on that basis. In so holding, *Caperton* adds muscle to the Court's earlier articulations of judicial independence boilerplate. Moreover, the fact that the Court's action came in a circumstance where, indisputably, the spending that created the "serious risk of actual bias-based on objective and reasonable perceptions"²³⁵ was almost entirely expenditure-based rather than contribution-based is all the more remarkable.

B. *Expenditure Limits and the Equality Rationale*

Professor Roy Schotland recently captured the essence of the decision's broader ramifications, asserting simply that "*Caperton*'s fundamental holding is that judicial elections are different. One cannot conceive of a court holding that a legislator (or executive) would be barred from acting in X matter because a campaign supporter was involved."²³⁶ Schotland is correct on both counts. Viewed through a campaign finance lens, "*Caperton* contributions" embody these differences.

The case for regulating both contributions and expenditures in judicial campaigns is not only compelling but also urgently necessary to prevent a developing constitutional crisis. First, the core "Schotland" difference distinguishing even elected judiciaries from the constituent branches is one such compelling indicator of the need for regulation. Second, the Supreme Court has never specifically addressed the constitutionality of limits on independent expenditures in judicial elections. Third, judicial campaign spending has exploded since *Buckley* and particularly in the last decade and with respect to independent spending and the "super-spender" dynamic. Fourth, as Erwin Chemerinsky has noted, *Buckley* "did not create an absolute bar to government regulation of expenditures"

234. *See id.* at 775–77.

235. *Caperton v. A.T. Massey Coal Co.*, 129 S. Ct. 2252, 2263 (2009).

236. Roy A. Schotland, *Caperton Capers: Comment on Four of the Articles*, 60 SYRACUSE L. REV. 337, 344 (2010).

but “rather, it imposed strict scrutiny as the test that must be met.”²³⁷ Fifth, as Kathleen Sullivan asserts, *White* is “a very narrow decision” and “lower courts that have interpreted *White* broadly have gotten *White* wrong” because they “ignore the countervailing interests of a constitutional magnitude that are implicated by limiting the *expenditure* or the contribution of funds and the use of those funds in judicial campaigns.”²³⁸ Sixth, Sullivan correctly recognizes that “nothing in *White* ruled out the possibility that those countervailing constitutional interests could trump whatever free speech interest there might be in a particular case.”²³⁹ Given this combination of factors, Justice Kennedy’s silent but clear equation of independent expenditures and contributions in the judicial elections context, leading to the concept that this Article coins as “*Caperton* contributions,” guides the way forward for one of the best available *ex ante* means of preserving the actual and perceived integrity of the nation’s state courts.

Judge Guido Calabresi, concurring in the Second Circuit’s 2005 denial of rehearing *en banc* in *Landell v. Sorrell*, stated bluntly that the failure to address inequality concerns “has been, since *Buckley*, the huge elephant—and donkey—in the living room in all discussions of campaign finance reform.”²⁴⁰

In his admirable—and unsuccessful—push for rhetorical realism, Judge Calabresi’s words echo those of Judge Skelly Wright. As part of the ultimately overturned D.C. Circuit panel in *Buckley* itself, Judge Wright later lamented that “the dominance of wealth in the political process is inconsistent with both the philosophical meaning and the practical exercise of political equality.”²⁴¹ Judge Wright buttressed his perspective with that of no less than political philosopher John Rawls, who asserted simply that the “liberties protected by the principle of participation lose much of their value whenever those who have greater private means are permitted to use their advantages to control the course of public debate.”²⁴² Ronald Dwor-

237. Chemerinsky, *supra* note 39, at 134.

238. See Sullivan, *supra* note 225 (detailing three such “major” countervailing interests: (1) the rule of law; (2) separation of powers; and (3) the individual litigant’s interest in due process) (emphasis added).

239. *Id.* (noting other countervailing interests that have been found compelling enough to “overwhelm” the speech interest in the Hatch Act cases).

240. *Landell v. Sorrell*, 406 F.3d 159, 162 (2d Cir. 2005) (Calabresi, J., concurring in the denial of rehearing *en banc*), *rev’d sub nom.*, *Randall v. Sorrell*, 548 U.S. 230 (2005). The Second Circuit initially upheld, but then the Supreme Court struck down several provisions of Vermont campaign finance law.

241. Wright, *supra* note 52, at 629.

242. *Id.* at 630 (citing JOHN RAWLS, A THEORY OF JUSTICE 225 (1971)).

kin is more direct, asserting that “[o]ur politics are a disgrace and money is the root of the problem.”²⁴³

Calabresi, Wright, Mills, and Dworkin are, of course, contemplating rules applicable to non-judicial campaigns. Suffice it to say, their sentiments apply a fortiori to campaigns for offices that, unlike the constituent-based branches, manifest the highest ideals of equality under law.

V. CONCLUSION

In a short essay published in 1998, just prior to the proliferation of judicial campaign spending described in this Article, Dean Chemerinsky asserted that “only limits on both contributions and expenditures can succeed” in the task of preserving “an independent judiciary and public confidence that one exists.”²⁴⁴ Acknowledging that “expenditure restrictions limit political speech,” Chemerinsky asserted that “this is constitutional because it is the only apparent way to achieve an undoubtedly compelling interest: preserving an independent judiciary in the face of the corrosive effects of ever larger spending in judicial elections.”²⁴⁵

Paradoxically, it is both regrettable and fortunate that Dean Chemerinsky’s assertion is exponentially more powerful today than it was barely more than a decade ago. Regrettable because in terms of the sums at issue, the threats to impartial justice, and the public’s growing perceptions of same, the differences of a decade are differences not of degree, but in kind. Fortunate too though, because *Caperton* makes clear that judicial elections are manifestly—rather than merely rhetorically—different from legislative and executive campaigns in constitutionally meaningful respects. Judicial elections—for some a bright spot, for others a backwater of American exceptionalism—are for better or worse, here to stay. *Caperton* points a way forward. It demonstrates that the rejection of strict path dependency—based on campaign finance jurisprudence pertaining solely to the political branches—is now not only settled constitutional law, but also sound public policy.

This Article saves for future discussions, and hopefully for the states as laboratories, debates regarding the merits of particular

243. Ronald Dworkin, *Free Speech and the Dimensions of Democracy*, in *If Buckley Fell: A First Amendment Blueprint for Regulating Money in Politics* 63 (Josh Rosenkranz ed., Century Foundation Press 2000).

244. Chemerinsky, *supra* note 39, at 149.

245. *Id.*

proposals involving spending—as opposed to merely contribution—limits in judicial elections. To be sure, all such proposals will involve tradeoffs and “hydraulics”²⁴⁶ of their own. Also to be sure, it is clear—in part because of campaign hydraulics—that spending limits are just one of the multi-prong measures needed to protect the courts from what Justice Sandra Day O’Connor bluntly calls the “wrong” of “cash in the courtroom.”²⁴⁷ In addition to limits on expenditures—which, like contribution limits are prohibitions only in excess of trigger thresholds—other *ex ante* measures such as voluntary judicial public financing systems promote a vibrant political discourse while also respecting the “countervailing concerns of constitutional dimension”²⁴⁸ that are applicable to the courts.

Inevitably, even the best *ex ante* measures, including spending limits, will be imperfect. Like any preemptive effort, they will invariably fail to anticipate every scenario that circumstance, chance and intent can conceive. Fortunately, because the vast majority of state court jurists are deeply committed to preserving the integrity of their courts, only the rarest exceptions ever morph into a one-person Constitutional crisis on the scale of Justice Benjamin.

When, however, a jurist acts unilaterally and subjectively against the clear weight of objective perceptions, *ex post* measures such as strengthened procedures to enforce non-constitutional recusal rules are an effective and important fairness protection.²⁴⁹ Spurred by *Caperton*, in 2010, the state of Michigan enacted just such a process—one that, if it had applied in West Virginia, would have prevented *Caperton* from ever turning into a federal constitutional case. Michigan Chief Justice Marilyn Kelly spearheaded the court’s adoption of the new process in large part because, in her words, it “permits a justice’s recusal where that justice is unable to

246. See Issacharoff & Karlan, *supra* notes 40.

247. Dorothy Samuels, Op-Ed., *The Selling of the Judiciary: Campaign Cash ‘in the Courtroom’*, N.Y. TIMES, Apr. 15, 2008, at A22.

248. Sullivan, *supra* note 225; see also North Carolina Right to Life Comm. Fund for Indep. Political Expenditures v. Leake, 524 F.3d 427, 441 (4th Cir. 2008) (upholding North Carolina’s judicial public financing system and stating that “the concern for promoting and protecting the impartiality and independence of the judiciary is not a new one; it dates back at least to our nation’s founding . . . [T]he provisions [of North Carolina’s public financing system for judicial campaigns] challenged today, which embody North Carolina’s effort to protect this vital interest in an independent judiciary, are within the limits placed on the state by the First Amendment”).

249. *Caperton* is already spurring halting but positive progress in the area of court reform, particularly in the area of recusal, a dynamic I describe at length in James Sample, *Court Reform Enters the Post-Caperton Era*, 58 DRAKE L. REV. 787 (2010).

render an unbiased decision *and* unable or unwilling to acknowledge that fact.”²⁵⁰ Chief Justice Kelly noted further that the “justice system and this Court can only be stronger for it.”²⁵¹

Scholars have long debated whether and to what degree overturning or modifying *Buckley*’s contribution–expenditure distinction would improve the quality, and perhaps even the quantity, of American political discourse. While I concur with Professor Neuborne’s belief that free speech values “would be advanced, not endangered, by the placement of generous ceilings on campaign spending,”²⁵² in judicial and non-judicial elections alike, that, manifestly, is a controversial question on which principled people reasonably disagree. This Article posits a more modest and potentially more broadly acceptable proposition: given compelling state interests in an impartial and independent judiciary, expenditure limits in judicial elections meet strict scrutiny, and appropriately balance the First Amendment with due process and structural concerns of equally important constitutional magnitude.

Even those members of the legal community—ranging from Ted Olson and Kathleen Sullivan to Justice Kennedy—who traditionally believe that campaign regulation in the constituent branches unconstitutionally infringes on the First Amendment recognize the compelling countervailing interests applicable in judicial elections. Notably, Justice Kennedy’s equation of contributions and expenditures is a refreshing rejection of *Buckley* formalism in the manifestly unique context of campaigns for the bench. *Caperton* provides a solid constitutional foundation for expenditure limits in judicial elections. If adopted, particularly in a post-*Citizens United* era, such limits would substantially reduce the risk of the “mutually assured destruction” in our courts of which Justice O’Connor so wisely warns.

250. *Id.* at 807 (citation omitted).

251. *Id.*

252. Burt Neuborne, *Soft Landings*, in *IF BUCKLEY FELL: A FIRST AMENDMENT BLUEPRINT FOR REGULATING MONEY IN POLITICS* 184 (Josh Rosenkranz ed., Century Foundation Press 2000); see also NEUBORNE, *supra* note 14, at 14 (“It remains to be seen whether the same assumption—that all limits gravely injure the quantity of political speech—would be justified in the context of more generous spending ceilings.”).

PRICE INCLUDES TAX: PROTECTING CONSUMERS FROM TAX-EXCLUSIVE PRICING

HAYES HOLDERNESS*

INTRODUCTION

Millions of Americans are suffering economic harm every day, and not just as a result of the current economic crisis. Savings are being lost, people are sacrificing leisure time to make ends meet, debt is rising, and consumers are finding that they are unable to afford the goods they prefer. The cause of this harm is a seemingly benign practice, a practice so widespread that many people probably have not thought twice about it. That practice is tax-exclusive pricing, whereby prices are presented without tax included.

The European Community requires that prices on consumer goods include tax.¹ Indeed, other countries have embraced this idea and also require tax-inclusive pricing.² This practice is thought to protect consumers from becoming misled as to the total cost of the products they purchase. Tax-exclusive pricing presents a problem because it lowers the salience of taxes, causing them to become hidden from consumers. Recent studies and economic analyses of behavior show that consumers will systematically undervalue the cost of hidden taxes and, as a result, will over-consume.³ Overconsumption can lead to decreased savings, overworked individuals, and increased amounts of debts or cuts in consumption of other desired products. To prevent these effects, consumers must be given the total cost of a product before they decide to buy it, not after.

* J.D. 2011, New York University School of Law; B.A. Political Science 2006, The University of North Carolina at Chapel Hill. Articles Editor for the *New York University Annual Survey of American Law* in 2010–11. I would like to thank Professor Lily Batchelder and my colleagues in her Tax and Social Policy Seminar for their insight and support.

1. Council Directive 98/6, art. 2, 1998 O.J. (L 80) 27 (EC).

2. See Value Added Tax Law, 5736-1975, 30 LSI 46 (1975-76) (Isr.); ALAN A. TAIT, VALUE ADDED TAX: INTERNATIONAL PRACTICE AND PROBLEMS 357 (1988) (providing that New Zealand requires tax-inclusive pricing).

3. See generally Raj Chetty, Adam Looney & Kory Kroft, *Salience and Taxation: Theory and Evidence*, 99 AM. ECON. REV. 1145 (2009) (studying hidden taxes in grocery stores and in alcoholic beverage sales); Amy Finkelstein, *E-ZTax: Tax Salience and Tax Rates*, 124 Q. J. ECON. 969 (2009) (studying hidden charges at toll plazas).

The harms resulting from the undervaluation of hidden taxes have received scholarly attention, but this Note approaches them from a different angle, analyzing them through the lens of United States consumer protection law. This analysis concludes that the Federal Trade Commission (FTC or Commission) should use its rulemaking authority to mandate tax-inclusive pricing for consumer goods. For the purposes of this Note, “consumer goods” are understood to be those products and services purchased by “consumers”; that is, unsophisticated end users.⁴ Therefore, business purchases and distributor purchases are excluded from this analysis for the most part. A mandate requiring tax-inclusive pricing would ensure that the costs of taxes are not hidden from consumers by presenting them with the information they need to make accurate purchasing decisions. There are three general forms this mandate could take: it could require that tax and price both be stated separately but not added together, that tax and price both be stated separately and added together, or that fully tax-inclusive prices with no breakdown be presented. As will be developed, the FTC should require that tax and price both be stated separately and added together. There are two ways to accomplish this scheme: having both the tax and the tax-inclusive price presented on the price tag,⁵ or having only the tax-inclusive price on the price tag along with a breakdown of tax and price on the receipt. Either way should suffice to prevent consumer harm.

In a country that prides itself on demand-driven markets, consumers should be given as much information relevant to their purchasing decisions as is reasonably necessary and possible.⁶ What information is reasonably necessary is debatable, but cost is a necessary piece of information for any consumer-driven market. Consumers must see the total cost of a product on the price tag, or they will continue to make imperfect decisions. Some taxes are already

4. See BLACK'S LAW DICTIONARY 359 (9th ed. 2009) (defining consumer product as “[a]n item of personal property that is distributed in commerce and is normally used for personal, family, or household purposes”); accord FTC Cooling-Off Rule, 16 C.F.R. § 429.0(b) (2010). See BLACK'S LAW DICTIONARY, *supra*, at 358 (defining consumer as “[a] person who buys goods or services for personal, family, or household use, with no intention of resale; a natural person who uses products for personal rather than business purposes”).

5. “Price tag” should be taken to include both the posted price in-store and online (as the case may be) and advertised prices.

6. Neil W. Averitt & Robert H. Lande, *Consumer Sovereignty: A Unified Theory of Antitrust and Consumer Protection Law*, 65 ANTITRUST L.J. 713, 722–23 (1997) (discussing how failures “inside the head” of consumers, such as making decisions on incomplete information, can cause markets to operate inefficiently).

included in price in the United States, most notably excise taxes. Excise taxes are taxes imposed on specific goods—unlike sales taxes, which are imposed on purchases generally.⁷ However, there are still many taxes that are not, such as state sales taxes, and consumers are being harmed by the exclusion of those taxes. Almost every consumer, acting rationally or irrationally, has probably ignored taxes in his or her life, and has over-consumed as a result. While the harm from each individual purchase is relatively small, the aggregate harm to consumers is monumental. Mandating that consumers be presented with a tax-inclusive price will ensure that these harms are avoided.

This Note is divided into five parts. Part I describes the current tax disclosure scheme for consumer goods and also how consumers make their purchasing decisions. This Part provides background information necessary to frame the remaining discussions, describing the current tax disclosure regime and illustrating the consumer purchase decisionmaking process that leads to the undervaluation of undisclosed taxes. Because current law does not require the disclosure of all taxes on the price tag and because consumers make their purchasing decisions based on price tags and not total cost, the current tax disclosure regime can be expected to negatively affect consumer purchasing decisions.

Part II considers the effects of tax salience on consumers, concluding that hidden taxes should be expected and that consumers undervalue those taxes. Both rational actor theory and behavioral economics analyses are used to reach this conclusion. Rational actor theory shows that consumers should be expected to ignore taxes on very low priced items, calculate taxes on highly priced items, and approximate taxes on the remainder of their purchases. Behavioral economics enhances this showing by explaining why individuals, suffering from multiple cognitive biases, will either ignore or approximate the taxes on most items; and when they approximate, they will systematically undervalue the cost of those taxes.

Part III discusses the welfare implications of consumers' undervaluation of taxes. This Part finds that hidden taxes likely harm social welfare by fostering over-taxation and by operating in a regressive manner. Individuals suffer from the income effects of their undervaluations, causing them to decrease their savings, work more than they would prefer, take on increased amounts of debt, or cut consumption of other desired products. Given that hidden

7. Compare BLACK'S LAW DICTIONARY, *supra* note 4, at 646 (defining excise tax), with *id.* at 1597 (defining sales tax).

taxes likely cause social harm, there do not appear to be any benefits to individuals that offset the harms of overconsumption.

Part IV considers the policy implications of these harms. This Part establishes that the FTC is the best actor to solve the problem of these harms, and that it has the authority to act. By analyzing the harms through the FTC's unfair trade practice jurisprudence, it is shown that tax-exclusive pricing is an unfair trade practice. The strengths and weaknesses of the different tax disclosure schemes are then considered, and it is shown why requiring that both the tax-inclusive price and the tax on an item be presented is the best way to prevent consumer harms. Part V concludes.

I. BACKGROUND

This Part will briefly discuss the current tax disclosure scheme and the process of consumer decisionmaking to help frame the rest of the Note. Recall that consumer goods are those products and services purchased by unsophisticated end users, and that consumers are understood to be those end users.

A. *The Current Tax Disclosure Scheme*

Currently, state and local governments are the main entities imposing taxes on consumer goods. To date, the federal government does not levy taxes on consumer goods as a whole, but does impose taxes on some discrete categories of goods.⁸ However, some academics and policy makers have advocated for a value added tax that would tax general consumption.⁹ If such a tax were imposed, then the issues discussed in this Note would become relevant to such a tax.

Taxation schemes differ from state to state. There are a plethora of taxes that could be levied on consumer goods.¹⁰ For simplicity's sake, this Note will focus on state sales taxes, which are probably the most prominent type of taxes that can become hid-

8. *See, e.g.*, I.R.C. § 5701 (federal tax on tobacco products). Federal taxes on consumer goods are found in Subtitles D and E of the Internal Revenue Code, I.R.C. §§ 4000–5891.

9. *See generally* Paul R. McDaniel, *A Value Added Tax for the United States? Some Preliminary Reflections*, 6 J. CORP. L. 15 (1981) (discussing issues surrounding adoption of value added tax in United States).

10. *See, e.g.*, KY. REV. STAT. ANN. §139.200 (West 2010) (levying general sales tax, hotel tax, tax on sewer services, tax on prepaid calling services, tax on communications services, and tax on distribution services for natural gas).

den. Forty-five of the fifty states impose a sales tax, which is not typically included in the good's posted price.¹¹

While this Note will only make reference to sales taxes, other taxes on consumer goods should not be ignored when considering the welfare and policy implications of the current taxation scheme. For now it suffices to say that consumer goods are taxed differently by each state, and not all of these taxes are included in posted prices.

While posted prices across the United States do not typically include all taxes, the full price of products is revealed at the register before the consumer ultimately pays. Therefore, taxes are technically disclosed to consumers before they make their purchase. However, this form of disclosure has little impact on consumers, as a brief exploration of the process of consumer purchasing decisions will show.

B. Consumer Purchasing Decisions

The previous Section gave a brief background of the current state of tax disclosure laws, showing that not all taxes on consumer goods are included in price. This Section shows that consumers make their purchasing decisions before they reach the register, and they are very reluctant to change those decisions, even when presented with full price information at the register. As such, when coupled with the fact that not all taxes are included in price, it can be expected that undisclosed taxes will affect consumer perceptions of the cost of a product. While direct research on this particular behavior is somewhat sparse, research into consumer choice supports the conclusion that purchasing decisions are made before the register, and rational and behavioral analyses explain why consumers are reluctant to change their minds.¹²

Consumers take a number of different approaches to their purchasing decisions.¹³ These approaches may be complex and

11. Alaska, Delaware, Montana, New Hampshire, and Oregon are the five states without any general sales taxes. See 2 JEROME R. HELLERSTEIN, WALTER HELLERSTEIN & JOHN A. SWAIN, *STATE TAXATION* ¶ 12.02 (3d ed. 2010). However, some municipalities in Alaska do charge a local sales tax. *Id.*

12. Under rational actor theory, individuals weigh the costs of choices and select the least costly alternative. See *infra* note 30 and accompanying text. Behavioral analyses look to actual human behavior to formulate predictions of future human behavior. See Christine Jolls, Cass R. Sunstein & Richard Thaler, *A Behavioral Approach to Law and Economics*, 50 *STAN. L. REV.* 1471, 1476–77 (1998).

13. See Michele D. Bunn, *Taxonomy of Buying Decision Approaches*, 57 *J. MARKETING* 38, 38, 46–50 (1993) (offering detailed analysis of buyer decision approaches for different types of purchases).

multi-stepped, but what is important for this Note is that consumers tend to make their decision to purchase before they reach the register.¹⁴ Price is an important dimension in consumer purchasing decisions, and posted prices carry heavy weight because consumers make their purchasing decisions based on the information that is readily accessible to them at the time they make their decision. This occurs when they select the item, not when they pay for it. Information that is given later, such as the fully tax-inclusive price, has a significantly reduced impact on the consumer.¹⁵ This reduced impact creates the potential for valuation mistakes by consumers.

Further, consumers should not be expected to change their decisions once they are presented with full price information at the register. Rational actor theory can explain this behavior if the cost of turning back and finding different products outweighs the additional disclosed cost. It seems plausible that it would be less costly to simply buy the product, since in many cases the additional cost from the tax will be relatively small, and the hassle of finding new products can be time-consuming. Even if rational actor theory would imply turning back, the endowment effect is a bias in individuals that causes them to attach above market value to items they already possess.¹⁶ If a consumer has already decided to purchase a product, then he will value it higher than its economic value and thus be reluctant to swap it out for another item upon disclosure of the true price.

Given this model of consumer decisionmaking, the practice of tax-exclusive pricing can be expected to affect consumer percep-

14. Richard M. Bird, *Policy Forum: Visibility and Accountability—Is Tax-Inclusive Pricing a Good Thing?*, 58 CAN. TAX J. 63, 68–69 (2010) (“[M]ost consumer purchase decisions are made in the aisles, not at the cash register.”); see Chetty et al., *supra* note 3, at 1149 (noting that in their model explaining the effects of tax salience, “[t]he price that consumers see when deciding what to purchase is [the posted price]; the sales tax is not included in the posted price”).

15. See Stephanie Stern, *Temporal Dynamics of Disclosure: The Example of Residential Real Estate Conveyancing*, 2005 UTAH L. REV. 57, 73–81 (2005). Stern discusses the “considerable body of research establishing that individuals have difficulty changing course once they have made an overt commitment to a certain path or action.” *Id.* at 73.

16. Richard Thaler, *Toward a Positive Theory of Consumer Choice*, 1 J. ECON. BEHAV. & ORG. 39, 44 (1980) (“[A] certain degree of inertia is introduced into the consumer choice process since goods that are included in the individual’s endowment will be more highly valued than those not held in the endowment.”). See generally Daniel Kahneman, Jack L. Knetsch & Richard H. Thaler, *Experimental Tests of the Endowment Effect and the Coase Theorem*, 98 J. POL. ECON. 1325 (1990) (discussing experiments observing endowment effect).

tion of the total cost of a product. Those effects are considered in the following Part II.

II. THE EFFECTS OF THE CURRENT TAX DISCLOSURE SCHEME

The current tax disclosure scheme for consumer goods leads to interesting and perhaps unexpected effects on consumers. Taxes that are not presented in the posted price, such as sales taxes, are so weakly salient that they effectively become hidden from consumers. Consumers have been shown to undervalue the cost of such hidden taxes, and indeed, rational and behavioral actor models show that consumers should be expected to undervalue hidden taxes. As will be seen, rational actor theory shows that consumers will approximate the cost of hidden taxes in some situations, and behavioral analysis builds on the rational actor discovery to show that consumers frequently approximate, and subsequently underestimate, the cost of taxes. This Part explores these ideas—first establishing that taxes on consumer goods which are not disclosed at the time the decision to purchase is made should be expected to become hidden, then showing why consumers will undervalue those taxes once they are hidden.

A. *Tax Salience and Hidden Taxes*

This Section takes a look at the background literature surrounding tax salience and its effect on the perceived cost of taxes. Tax salience refers to how noticeable or easy to process taxes are.¹⁷ Reducing either the ability to notice the tax or the ability to figure the tax out will result in lower salience. Some taxes are so weakly salient that they can be aptly described as hidden, in that they “collect revenue or redistribute wealth without also affecting decisions about whether or where to earn or spend.”¹⁸ In essence, the idea behind hidden taxes is that they affect people’s behavior less than they predictably should under traditional economic models of anal-

17. Brian Galle, *Hidden Taxes*, 87 WASH. U. L. REV. 59, 62 (2009) (citing Edward J. McCaffery & Jonathan Baron, *Isolation Effects and the Neglect of Indirect Effects of Fiscal Policies*, 19 J. BEHAV. DEC. MAKING 289, 289 (2006) for support of this definition of “salient” taxes). Chetty et al. offer an alternate definition by defining tax salience as “the visibility of the tax-inclusive price.” Chetty et al., *supra* note 3, at 1146 n.2. This definition removes considerations of the tax rate. As this Note will assume that consumers are aware of tax rates, whichever definition the reader finds more accessible will suffice.

18. Galle, *supra* note 17, at 61.

ysis. Hidden taxes could conceivably arise from many different practices,¹⁹ but this Note focuses on taxes that become hidden because of their low salience levels.

There are two seminal studies in the low salience hidden-tax literature. The first study concerns tolls paid on roads.²⁰ The author collected data from 123 toll plazas around the country that had been in operation since at least 1985. Many of these toll plazas have electronic toll collection devices which allow drivers to pass through toll collection sites without having to stop to manually pay the toll on the road. After comparing the data while controlling for certain differences, the author found that consumers were less sensitive to toll increases when they paid through a debit from an electronic toll collection device, such as E-Z Pass, than when they paid in cash.²¹ The author suggests that because of the passive nature of electronic toll collection, the tolls and toll increases were lowly salient to electronic toll payers.²² The study found that tolls collected at facilities with electronic toll collection devices rose on average 20–40% more than facilities without those devices, and it attributed this finding to the low salience of the tolls.²³ Since the author did not find a comparative decline in usage with the rise of tolls, it is possible that electronic toll payers were undervaluing or ignoring the tolls, as they did not adjust their behavior relative to the increase in toll prices.

The second study had two findings. The first part concerned taxes on consumer goods in grocery stores, and the second observed the effects of more and less salient taxes on alcoholic beverage sales.²⁴ For the first part of the study, the authors presented tax-inclusive pricing on selected types of consumer goods subject to sales taxes in a chain of grocery stores for a period of three weeks

19. See generally Aradhna Krishna & Joel Slemrod, *Behavioral Public Finance: Tax Design as Price Presentation*, 10 INT'L TAX & PUB. FIN. 189 (2003) (discussing various pricing techniques such as framing, disaggregation, and timing issues that could also possibly cause taxes to become hidden). The authors describe how various practices can affect price perception, with the idea that businesses would like for consumers to perceive prices to be lower than they are, to encourage extra consumption. *Id.* at 190–98. They then predict that such techniques might also be able to be used by the government to lower individuals' perception of tax burdens, focusing mainly on the U.S. income tax. *Id.*

20. Amy Finkelstein, *supra* note 3, at 969–71.

21. *Id.* at 970, 983–86 (discussing methodology and data).

22. *Id.* at 970–971.

23. *Id.* at 971.

24. Chetty et al., *supra* note 3, at 1146.

and monitored purchasing trends through register scanner data.²⁵ They found that when sales taxes were presented with the price on the shelf, rather than only at the register, consumers were more responsive to the taxes. Consumers were shown to buy less of a product once its total cost was disclosed by the inclusion of tax.²⁶ This is direct evidence that the salience of taxes affects the purchasing decisions of consumers because when the taxes were made salient by increasing their noticeability, consumers altered the expression of their preferences.

The results from the grocery store part of the study were supported by a study of taxes on alcoholic beverages. In this study, the authors compared data on aggregate annual beer consumption by state, compiled from administrative state tax records, with data on excise and sales tax rates on beer by state.²⁷ The authors reported finding that increases in excise taxes on alcoholic beverages, which were included in price and thus highly salient to consumers, caused a greater decline in consumption than increases in sales taxes, which were not as salient.²⁸ The authors concluded that this reaction by consumers shows that the salience level of taxes directly impacts consumer purchasing decisions.²⁹

These two studies suggest that consumers undervalue taxes that are not above a certain level of salience. The remaining sections of this Part II ask if traditional rational actor and behavioral analyses can explain this behavior, concluding that they can and that undervaluation should be expected for a majority of purchases.

B. A Rational Explanation: Rational Ignorance and Approximation

This Section establishes a basic understanding of consumer behavior using rational actor theory, which will subsequently be enhanced with behavioral analyses. Rational actor theory shows that consumers can be expected to approximate the cost of hidden taxes, and behavioral economics then shows that approximation

25. *Id.* at 1150–53 (describing methodology of grocery store experiment). One important part of this experiment was that all the similar consumer products studied had their taxes disclosed. As such, potential confusion between tax-inclusive and tax-exclusive pricing was controlled for. Additionally, the tax-inclusive price was clearly visible.

26. *Id.* at 1153–58 (describing results of grocery store experiment). Particularly telling from the results of the experiment is the observation that making a 7.375% sales tax salient decreased demand by 7.6%. *Id.* at 1155.

27. *Id.* at 1158–59 (describing methodology of alcohol sales study).

28. *Id.* at 1160–64 (describing results of alcohol sales study).

29. *Id.* at 1164 (“This finding supports the claim that the excise tax has a larger effect on demand than the sales tax because it is fully salient.”).

should occur most of the time and that undervaluation of total costs should be expected when consumers approximate.

Traditional rational actor economic analyses of human behavior assume that “all human behavior can be viewed as involving participants who maximize their utility from a stable set of preferences and accumulate an optimal amount of information and other inputs in a variety of markets.”³⁰ Under the most basic of these models—in which actors are perfectly rational and have perfect information—taxes could not become hidden. Consumers would have perfect information about the tax, and even if it was not salient, the consumer would still consider its cost when making a purchasing decision. While this basic model is clearly an idealized view of the world, it offers a good starting place to consider the effects of hidden taxes. Removing the assumption of perfect information provides a much more realistic view of the world, and under it the rational actor model can explain why some hidden taxes are not calculated.

A quick note here is important. The analysis in this Section will focus on how consumers evaluate the imperfect information they receive regarding taxes. It is important to understand that perfect rationality will be assumed throughout this analysis in the sense that consumers will be assumed to make rational decisions based on the information that they have. For example, if Abe’s rational preference would be for a candy bar over all other things when the price is \$0.99, it will be assumed that Abe will buy the candy bar and not something else. This assumption is made because consumer protection law does not aim to force consumers to make rational choices, but rather aims to give them the information they need to make such choices.³¹ However, assuming perfect rationality in the ultimate decision in the presence of imperfect information does not mean that perfect rationality is used when evaluating imperfect information.³² The important point of this aside is that when rational-

30. GARY S. BECKER, *THE ECONOMIC APPROACH TO HUMAN BEHAVIOR* 14 (1976).

31. *See infra* text accompanying notes 108–111. The author makes no judgment about this assumption of rationality.

32. Individuals who do not use perfect rationality to evaluate imperfect information are said to use “bounded rationality.” Bounded rationality occurs because “[t]he capacity of the human mind for formulating and solving complex problems is very small compared with the size of the problems whose solution is required for objectively rational behavior in the real world—or even for a reasonable approximation to such objective rationality.” Thaler, *supra* note 16, at 40 (quoting HERBERT A. SIMON, *MODELS OF MAN* 198 (1957)). Individuals are unable to process all of the information they are given so they use shortcuts, or heuristics, to come to

ity is discussed, it refers to the evaluation of information, not to the making of the final purchasing decision.

1. Rational Ignorance

When faced with imperfect information, the rational actor will weigh the opportunity cost of discovering the information against the cost of acting with the imperfect information.³³ If there is greater benefit to acting on full information than on incomplete information, the missing information will be discovered. The imperfect information in the hidden tax scenario is the actual costs of the taxes, which are not known because they are not disclosed at the time the consumer makes her purchasing decisions. The cost of calculation of the cost of the taxes is the sum of the costs of discovering and of applying the tax rate to the price of a product.³⁴ The cost of ignorance is the cost of the tax minus the cost of time saved from not calculating. If the cost of calculation is larger than the cost of ignorance, consumers will intentionally ignore the cost of the tax.³⁵ For example, Betty may be a lawyer making \$180 an hour who wants to buy a \$39.99 shirt subject to a 4.75% sales tax that is not included in the price. Assume that Betty can work whenever she

their final decisions. These decisions may appear irrational when an “incorrect” choice is made, but technically they are rational on the basis of the information gained from the shortcuts the individual has used. The use of the shortcuts skews information and thus can lead to a rationally made irrational choice. See Jolls et al., *supra* note 12, at 1477–78. In a companion piece to their published article, Chetty et al., *supra* note 3, Chetty et al. propose a theory of bounded rationality to explain the effects of low tax salience. Raj Chetty, Adam Looney & Kory Kroft, *Salience And Taxation: Theory and Evidence* 28–35 (Nat'l Bureau of Econ. Research, Working Paper No. 13,330, 2007) [hereinafter Chetty et al., Working Paper], available at <http://www.nber.org/papers/w13330>.

33. BECKER, *supra* note 30, at 6–7.

34. For the sake of the analysis in this Note, consumers will be assumed to know the taxes that exist on consumer goods. Chetty et al. make a similar assumption in their article and also provide some anecdotal evidence that consumers are indeed aware of taxes on consumer goods and the rates of those taxes. See Chetty et al., *supra* note 3, at 1165. Whether it is unrealistic or not to assume that individuals are aware of all the taxes on consumer goods, individuals' knowledge of taxes, while closely related to this Note, is beyond the scope of this Note.

35. Chetty et al., Working Paper, *supra* note 32, at 30–32, and Finkelstein, *supra* note 3, at 973–75, 980–83, both propose this idea as a reason for consumer ignorance of taxes. One can imagine the costs of calculating the cost of the tax would be quite high when compared to the cost of acting without full information. Assuming that the existence of all of the taxes on the product is known, the costs of calculation arise from the need to discover the tax rate and to apply it to the stated price. The cost of acting without full information is the cost of the tax assuming an all-or-nothing approach, but may be smaller than the full cost of the tax when approximation is taken into account, as noted below.

likes, for as long as she likes, so that her time is presumptively worth \$0.05 a second. The cost of calculation, assuming it always takes her 30 seconds to figure out the cost of the tax and add it to the price, is \$1.50. The cost of ignorance would be \$1.89 – \$1.50, or \$0.39. Because the cost of calculation is greater than the cost of ignorance, Betty will ignore the tax. If Betty were purchasing a \$3999.99 high definition television instead of the shirt, she would calculate the cost of the tax, since the cost of calculation would remain \$1.50, but the cost of ignorance would jump to \$190 – \$1.50, or \$188.50, assuming the same 4.75% tax rate.³⁶

The rational ignorance model shows that consumers are more likely to ignore the cost of taxes when prices are relatively low and more likely to calculate the cost of taxes when prices rise. However, calculation and ignorance are not the only two options available to consumers when considering the cost of hidden taxes. Approximation of the cost could also occur.

2. Approximation

The rational ignorance model does not have to entail all-or-nothing ignorance of taxes. Consumers are also able to approximate the cost of hidden taxes. The cost of approximation is the sum of the cost of figuring out an approximate cost and the difference between that approximation and the actual cost, minus the cost of the time saved from not calculating.³⁷ When the cost of approximation is less than the costs of calculation and of ignorance, rational consumers will approximate.

The three situations—ignorance, approximation, and calculation—should be expected to occur under different circumstances. Ignorance should be expected when prices are very low, such as when buying a \$1.00 cup of coffee. In such circumstances, the time needed to form an approximation or to calculate the actual cost of the tax will cost more to the consumer than the tax itself does. When prices—and thus the cost of taxes on the product—rise above a certain level, which can still be relatively low, the cost of the tax becomes greater than the cost of the time needed to form an approximation, and approximation will ensue. For example, Chuck buys a \$3.00 cappuccino instead of the \$1.00 cup of coffee in a state

36. This example operates under a number of simplifying assumptions, one of the biggest being that Betty would be able to become aware of the costs of ignorance and of calculation, but should remain illustrative nonetheless.

37. An approximation could be achieved in a number of ways. For example, one could approximate using a simple tax rate, such as 10%, or one could simply add a few dollars to the price.

where the sales tax is 4.75%. Assume Chuck always approximates by adding \$0.04 for each dollar he spends, and also that his time is worth \$0.01 a second and it takes him five seconds to approximate rather than ten seconds to calculate. Chuck's cost of approximation is -\$0.03.³⁸ The cost of ignorance is \$0.04.³⁹ The cost of calculation is \$0.10.⁴⁰ The cost of approximation is less than the costs of ignorance and calculation, so Chuck will approximate.

As prices rise, the costs of approximation and of ignorance also rise, but the cost of calculation remains constant.⁴¹ This is because the cost of calculation is not tied to price, whereas the costs of approximation and of ignorance are. Therefore, even with approximation as an option, rational consumers will be more likely to calculate the costs of taxes as prices rise.⁴²

In all likelihood, rational consumers will find approximation to be valuable for a significant number of goods purchased. Because the practice of approximation suggests that consumers might undervalue some taxes while overvaluing others, hidden taxes might not necessarily lead to undervaluation by consumers of their total tax burden. One might expect an averaging out of overestimation and underestimation of tax burdens such that the end result comes close to the actual tax burden, resulting in no significant net consumer harm. Rational actors should also be able to learn from past mistakes and become increasingly accurate in their approximations, reducing each individual instance of harm.⁴³ Therefore, traditional rational actor models cannot fully explain the undervaluation of hidden taxes and likely assume too much about the cogni-

38. Cost of approximation = $-\$0.03 = \0.05 [cost of figuring out approximation] + $\$0.02$ [cost of tax ($\$0.14$) - approximated cost ($\0.12)] - $\$0.10$ [cost of time saved from not calculating]. In this case, Chuck receives a benefit from approximating, seen from the fact that the cost is negative.

39. Cost of ignorance = $\$0.04 = \0.14 [cost of tax] - $\$0.10$ [cost of time saved from not calculating].

40. Cost of calculation = $\$0.10 = \0.01 [cost of time] * 10 [time to calculate].

41. Chetty et al., Working Paper, *supra* note 32, at 34. ("[T]he cognitive cost is a fixed amount paid at each transaction, whereas the benefit of computing [the tax-inclusive price] scales up with expenditure on the good.")

42. This observation assumes that the marginal utility of money remains constant. Otherwise, if individuals making expensive purchases have a decreasing marginal utility of money (perhaps because they are wealthy), then the cost of ignorance will decrease as the utility of each dollar lost decreases.

43. See Lawrence E. Blume & David Easley, *Rational Expectations and Rational Learning*, in ORGANIZATIONS WITH INCOMPLETE INFORMATION: ESSAYS IN ECONOMIC ANALYSIS: A TRIBUTE TO ROY RADNER 61, 65-76 (Mukul Majumdar ed., 1998) (discussing the dynamics of learning in the context of single-actor decisionmaking).

tive capabilities of consumers.⁴⁴ Further development is needed. Removing the assumption thus far of perfect rationality opens the door to cognitive explanations that explain why undervaluation of hidden taxes should be expected in almost every case.

*C. A Behavioral Explanation: Behavioral Ignorance
and Undervaluation*

This Section will use behavioral analysis to supplement the rational actor model of consumer behavior in response to hidden taxes developed in the previous section. In particular, cognitive biases explain both why consumers should not be expected to calculate the costs of hidden taxes and why approximation will lead to undervaluation in almost all cases. These cognitive biases raise the cost of information, making it unlikely that consumers will find it worthwhile to calculate the cost of hidden taxes for most of their purchases.

1. Behavioral Ignorance

The rational actor model would have consumers calculate the cost of taxes when the cost of calculation is less than the cost of acting with imperfect information. This Section explores cognitive biases and processes that raise the cost of calculation in the rational actor model.⁴⁵ The cost of calculation can become so high that consumers will likely never calculate the cost of hidden taxes. Some of the biases and errors most relevant to this idea include a lack of computational skills needed to figure out the tax burden,⁴⁶ the possibility of hyperbolic discounters who do not correctly value future

44. See Galle, *supra* note 17, at 83 (noting evidence that consumers' choices are not always the result of considered reflection).

45. See *id.* at 83–84 (surveying cognitive processes that could lead to ignorance of taxes).

46. See B. Douglas Bernheim & Antonio Rangel, *Behavioral Public Economics: Welfare and Policy Analysis with Non-Standard Decision Makers*, in *BEHAVIORAL ECONOMICS AND ITS APPLICATIONS* 7, 24–25 (Peter Diamond & Hannu Vartiainen eds., 2007); Daniel Kahneman, *Maps of Bounded Rationality: Psychology for Behavioral Economics*, 93 *AM. ECON. REV.* 1449, 1453, 1459, 1464 (2003); Daniel Read, George Loewenstein & Matthew Rabin, *Choice Bracketing*, 19 *J. RISK & UNCERTAINTY* 171, 187 (1999).

events,⁴⁷ and the possibility that people have trouble combining the cost of the tax with the price even when it is figured out.⁴⁸

For example, Dave might lack the ability to multiply a \$3.99 loaf of bread by a 4.75% sales tax rate, raising the cost of calculation to a relatively high level—perhaps the cost of calculator—which would outweigh the cost of approximation, probably just a few cents in this case. Dave should be expected to approximate or ignore the tax. Emily might be able to make the calculation of the tax on the loaf of bread, but is a hyperbolic discounter who incorrectly undervalues the future benefit of calculating the tax and thus ignores the cost of the tax. She may think that the value of her current time saved by not calculating the cost of tax is worth more than the future \$0.19 of tax, only to find at the register that she was wrong and should not have ignored the cost of the tax.⁴⁹ Finally, Frank might have the ability to calculate the tax, but experiences frustration and confusion when trying to add it to the price, adding to his cost of calculation, and opts to avoid this burden by approximating or ignoring the tax. It might seem counterintuitive that Frank could multiply but not add, but if the sales tax were a simple 10% and the price was \$3.99, there could conceivably be some initial confusion and frustration at adding \$0.39 to \$3.99, enough that Frank might prefer just to approximate the total price.

These behaviors can explain why actors would not calculate the cost of hidden taxes within the rational actor model.⁵⁰ While one could argue that these are irrational behaviors, where the consumer does not weigh the cost of calculation against the cost of acting with

47. See R. H. Strotz, *Myopia and Inconsistency in Dynamic Utility Maximization*, 23 REV. ECON. STUD. 165, 170–71, 177–78 (1955) (describing effects of time on a consumer's optimal demand curve); see generally Shane Frederick, George Loewenstein & Ted O'Donoghue, *Time Discounting and Time Preference: A Critical Review*, 40 J. ECON. LITERATURE 351 (2002) (reviewing literature on myopia). Hyperbolic discounting is also referred to as "myopia" and is discussed again below, *infra* text accompanying note 57.

48. See Kahneman, *supra* note 46, at 1459.

49. Procrastination is a common expression of myopia. The pain of doing something now feels much greater than of doing that same thing in the future, though technically the pain should be the same (barring changed circumstances).

50. In addition, taxation schemes vary from state to state and even from good to good within each state. For the purpose of simplicity, this Note assumes that consumers are aware of the taxes on consumer goods, see *supra* note 34, but to the extent that they actually are not aware of the taxes, these two forms of variations could significantly increase the cognitive costs of calculating taxes to a point where consumers give up on calculation in most cases.

imperfect information at all,⁵¹ it is more accurate to consider them as supplementing the understanding obtained through the rational actor model by increasing the cost of calculation.⁵² In addition to explaining why calculation should not be expected, behavioral analysis further shows that consumers will undervalue costs when they approximate.

2. Undervaluation

The rational actor model showed that when consumers do not calculate the costs of taxes, they will approximate those costs unless price is very low. Consumers can approximate the costs of taxes directly by estimating the costs and adding them to the posted price of a product or indirectly by simply estimating the total cost of a product. Regardless of the means of approximation, behavioral analysis shows that approximation leads to systematic undervaluation of taxes. Given the results of the studies, undervaluation may be expected. Even so, it could be expected that consumers would overestimate the cost of taxes. Therefore, it is important to establish why consumers will undervalue the costs of hidden taxes when they do elect to approximate.

The universe of behavioral economics is quite large, and this Note will address cognitive biases that are some of the strongest indicators of consumers' tendency to underestimate the cost of taxes. The discussion here is not meant to be exhaustive. It examines the anchoring bias, the disaggregation bias, the optimism bias, and consumer myopia. These biases do not operate in a vacuum and a consumer may suffer from multiple biases, however, for ease, their effects are explained individually.

The anchoring bias causes individuals to gravitate towards a starting point when making estimations.⁵³ The bias can take hold even if the starting point is not an accurate value for whatever is being considered, although starting points that are close to being

51. Galle, *supra* note 17, at 83–84 (providing examples of situations where consumers do not calculate taxes regardless of the cost of avoiding taxation).

52. Jolls et al. note that “[t]he project of behavioral law and economics . . . is to take the core insights and successes of economics and build upon them by making more realistic assumptions about human behavior.” Jolls et al., *supra* note 12, at 1487.

53. See, e.g., Amos Tversky & Daniel Kahneman, *Judgment under Uncertainty: Heuristics and Biases*, 185 Sci. 1124, 1128 (1974); see also Stern, *supra* note 15, at 78 (“Anchoring is a pervasive judgment bias in which starting points or initial values systematically and disproportionately influence decision makers.”).

accurate create a stronger bias.⁵⁴ Anchoring affects consumers who engage in indirect approximation of the cost of the tax. If consumers see the presented tax-exclusive price as the starting point for estimating the total cost of the product, then they are likely to underestimate the total cost, and implicitly the cost of the tax, by staying too close to the posted price. Returning to the \$3.99 loaf of bread, with the anchoring bias Greg might approximate the total cost of the bread, but stay too close to the posted price, even if he understood abstractly that the cost of the tax should add more to the price. Under a 4.75% sales tax, Greg might approximate the total cost at \$4.10 when the actual cost is \$4.18.

Closely related to the anchoring bias is the disaggregation bias. The disaggregation bias causes individuals to underestimate the total cost of a product when prices are presented as broken down in smaller parts.⁵⁵ This undervaluation occurs because individuals have difficulty adding the parts together and will thus gravitate toward the most salient parts of the price and undervalue the less salient parts of the price. It stands to reason therefore, that by making prices tax-exclusive, the total cost of the product will be underestimated because the tax will be undervalued as the smaller, less salient part of the total cost, even if the initial estimation is fairly accurate. For example, Henry, suffering from the disaggregation bias, might have trouble adding together the two price components of his bread purchase—the \$3.99 posted price and \$0.19 of tax—and thus focus on the larger and more salient \$3.99, finding it easier to do so while adding a few cents on the top to account for the tax he knows is there.

Further, the optimism bias, which basically states that individuals do not rationally perceive negative events in their lives and thus undervalue their impact or probability, suggests that consumers who were not certain of their tax liability would tend to undervalue

54. Stern, *supra* note 15, at 79 (noting that “[r]esearchers have found anchoring effects even when the anchor presented is uninformative or extreme,” but also that “the anchoring effect is strongest when the anchor is moderate rather than extreme”).

55. See Krishna & Slemrod, *supra* note 19, at 192–94 (citing various pieces of experimental data pointing to disaggregation bias phenomenon); see also John M. Clark & Sidne G. Ward, *Consumer Behavior in Online Auctions: An Examination of Partitioned Prices on eBay*, 16 J. MARKETING THEORY & PRAC. 57, 57–66 (2008); Tanjim Hossain & John Morgan, *Plus Shipping and Handling: Revenue (Non) Equivalence in Field Experiments on eBay*, 6 ADVANCES ECON. ANALYSIS & POL’Y, no. 2, art. 2, 2006, at 1–4. See generally Vicki G. Morwitz, Eric A. Greenleaf & Eric J. Johnson, *Divide and Prosper: Consumers’ Reactions to Partitioned Prices*, 35 J. MARKETING RES. 453, 453–63 (1998).

its negative impact on them: they would assume the tax is not as big as it is.⁵⁶ Isabelle might see a 4.75% sales tax on the \$3.99 loaf of bread and assume that there is no way its cost could be as high as \$0.19. Thus, she will underestimate its cost, perhaps as \$0.10.

Consumer myopia refers to the tendency in individuals to undervalue future expenses and overvalue current expenses.⁵⁷ Myopia causes consumers to give undue weight to the posted price and undervalue the impact of the hidden tax, which can be considered a future expense since it is not disclosed at the time the posted price is disclosed, but figured out or approximated sometime thereafter. Jenny might be aware that a tax exists on a \$3.99 loaf of bread, but be unsure what it will be. Since myopia causes future harms to seem less painful, Jenny will assume the cost of the tax to be less than what it will actually be.

These biases explain why consumers will undervalue the costs of hidden taxes. When combined together they can create powerful cognitive hurdles to accurate approximation. However, there may be some cognitive processes that would tend to push consumers in the other direction, processes that would diminish the harm from undervaluing hidden taxes. The next Section explores those processes and finds that consumers should not be able to significantly mitigate the harms from undervaluing taxes.

3. Can Consumers Mitigate the Effects of Hidden Taxes?

Processes of bias removal and of preference ordering imply that the harms resulting from the undervaluation of hidden taxes should be short lived or minimal in the long run.⁵⁸ Consumers suffering harm from their biases are thus incentivized to correct those biases.⁵⁹ Individuals can learn of the taxes themselves or third par-

56. See generally David A. Armor & Shelley E. Taylor, *When Predictions Fail: The Dilemma of Unrealistic Optimism*, in HEURISTICS AND BIASES: THE PSYCHOLOGY OF INTUITIVE JUDGMENT 334 (Thomas Gilovich, Dale Griffin & Daniel Kahneman eds., 2002) (reviewing and discussing optimism bias); see also J.D. Trout, *Paternalism and Cognitive Bias*, 24 L. & PHIL. 393, 403–04 (2005) (discussing how overconfidence can make individuals overly optimistic about their abilities to evaluate information).

57. For literature regarding myopia, see *supra* note 47.

58. Galle, *supra* note 17, at 85. As will be seen in Part IV, bias removal is exactly what this Note argues for. However, the process referred to here is individuals' ability to remove biases without government action.

59. *Id.* at 89 (citing Jonathan Klick & Gregory Mitchell, *Government Regulation of Irrationality: Moral and Cognitive Hazards*, 90 MINN. L. REV. 1620, 1647–48 (2006)).

ties can intervene and inform the individuals. However, problems abound with processes of bias removal.⁶⁰

Initially, it will be hard for irrational individuals to remove their own biases. They would have to be aware that they were constantly underestimating. As the harm to each individual is not substantial, the injury causing the harm may be hard to locate. Assume that an individual overspends by \$10 a month due to undervaluing hidden taxes. Only the most pious of bookkeepers would likely be able to discover that the \$10 of harm resulted from undervaluing taxes and not some other expenditure or mistake, such as simply dropping a ten dollar bill. Also, if individuals do not check their estimations each time, they will not be able to correct themselves very easily.

Even third parties that attempt to alert individuals to the hidden taxes may be met with some resistance from consumers if they are not offering simple disclosures of total price, as advocated for in this Note.⁶¹ Notably, consumers are the ones who actually go to the store and see what they buy and how much it costs. Because consumers have first-hand information on their own purchases, they may simply not believe the others who do not have that information. Furthermore, as the harm on each individual purchase is likely to be relatively small, the consumers may just not care about it once it is revealed to them, a reaction that harkens back to disaggregation theory. Consumers may thus have a hard time realizing the aggregate harm they suffer. Optimism bias would cause consumers to believe that the harm that the groups reveal is not happening to them, but to other consumers. These and other cognitive biases and errors will lead consumers to dismiss the watchdog groups' claims of serious harm.

It is also argued that the effects of undervaluing hidden taxes are minimal because of consumer ordering of preferences, which is the idea that consumers will anticipate that some unknown income effects will adversely affect their budgets and thus they will purchase the goods they prefer the most before purchasing the goods they prefer the least.⁶² Therefore, because the consumers are only de-

60. *See id.* at 89–93 (discussing complications with effectiveness of bias removal).

61. *See id.* at 90–92 (discussing potential problems of external actors attempting to educate taxpayers on their tax burdens including noisy signals and credibility issues).

62. *See id.* at 79–80; *see also* Chetty et al. Working Paper, *supra* note 32, at 41–46 (describing various possibilities of how consumer may order his preferences).

nied their least preferred goods, the harm incurred is presumably minimized. It is unclear though that individuals order preferences in such a way, and the anticipation scenario rests on rather unstable ground.⁶³ Regardless, this idea admits undervaluation and the harm, but attempts to justify them by claiming that the harm is not that great. However, income effects of taxes are real, and where income is lost, it should not matter if it is on the front or back end of consumer preferences. Consumers will have lost the ability to spend their money and express their preferences accurately. On balance then, processes of bias removal and preference ordering are not likely to counterbalance the effects of the anchoring bias, the disaggregation bias, the optimism bias, and consumer myopia.

In conclusion, Part II has examined how consumers react to hidden taxes created by low salience levels. Consumers can process these hidden taxes in three ways: by ignoring them, approximating them, or calculating them. Given consumers' imperfect rationality when considering hidden taxes, it should be expected that they will not calculate the costs of the taxes for the majority of their purchases, but rather will approximate or ignore them, leading to undervaluation of those costs.

III. WELFARE IMPLICATIONS

Part II established that consumers should be expected to undervalue the cost of hidden taxes in almost all situations. Part III will examine the impact of undervaluation on both social and individual consumer welfare. As will be seen, the social impacts are somewhat speculative, but can be described in general terms. It is likely that hidden taxes lead to over-taxation and that they are regressive. Taxes are regressive when they impact poorer consumers more than wealthy consumers.⁶⁴ Regressive taxes impose less of an effective burden on individuals as their wealth increases.⁶⁵ Regres-

63. Galle, *supra* note 17, at 80, 87–89 (noting that it is unclear if consumers actually calculate taxes, and if they do, whether they act rationally with that information); *see also* Marianne Bertrand, Sendhil Mullainathan & Eldar Shafir, *Behavioral Economics and Marketing in Aid of Decision Making Among the Poor*, 25 J. PUB. POL'Y & MARKETING 8, 10 (2006).

64. *See* BLACK'S LAW DICTIONARY 1597 (9th ed. 2009) (defining regressive tax); *accord* Joseph Bankman & Thomas Griffith, *Social Welfare and the Rate Structure: A New Look at Progressive Taxation*, 75 CAL. L. REV. 1905, 1908 (1987) ("Under a regressive tax, the percentage of income paid to the government falls as income rises, although the absolute amount paid to the government may rise, fall, or remain constant.").

65. Galle, *supra* note 17, at 69, 102.

sive taxes create welfare concerns because they burden one class of individuals—the poor—more than another—the rich—even though it is thought that the rich have the better ability to pay. The impact on individuals is much clearer. Undervaluation causes individuals to over-consume,⁶⁶ resulting in individual welfare losses.⁶⁷

A. Society

Section A explains the effects of hidden taxes on social welfare. The effects that the undervaluation of hidden taxes has on social welfare are not entirely clear, but can be discussed generally. On the one hand, some argue that undervaluation will lead to over-taxation and that hidden taxes are particularly regressive.⁶⁸ On the other hand, some argue that undervaluation can increase social welfare by reducing the deadweight losses associated with the taxes.⁶⁹

Finkelstein's study addresses the over-taxation issue and focuses on fiscal illusion.⁷⁰ This bias is said to cause individuals to undervalue the cost of government when they are unaware of government revenues.⁷¹ If individuals are unaware of taxes—whether from fiscal illusion or something else⁷²—then opportunists in government may use this ignorance to tax people more heavily than is

66. See *supra* Part II.A.; *infra* Appendix, Figure 2.

67. See William Congdon, Jeffrey R. Kling & Sendhil Mullainathan, *Behavioral Economics and Tax Policy* 8 (Nat'l Bureau of Econ. Research, Working Paper No. 15,328, 2009), available at <http://www.nber.org/papers/w15328> (“[W]ith a binding budget constraint, spending too much on the good with a hidden tax will leave less income for subsequent purchases—distorting individual consumption and decreasing welfare.”).

68. Galle, *supra* note 17, at 93–104 (suggesting that hidden taxes could lead to over-taxation and to regressive taxation).

69. *Id.* at 93; see also Congdon et al., *supra* note 67, at 10.

70. Finkelstein, *supra* note 3, at 970.

71. See generally Wallace E. Oates, *On the Nature and Measurement of the Fiscal Illusion: A Survey*, in *TAXATION AND FISCAL FEDERALISM: ESSAYS IN HONOUR OF RUSSELL MATHEWS* 65 (Geoffrey Brennan et al. eds., 1988) (discussing different versions of fiscal illusion hypothesis and examining empirical studies); Brian E. Dollery & Andrew C. Worthington, *The Empirical Analysis of Fiscal Illusion*, 10 *J. ECON. SURVS.* 261 (1996) (examining and evaluating empirical research on five main hypotheses within fiscal illusion).

72. Oates finds that a persuasive case for the fiscal illusion has not yet been made. See Oates, *supra* note 71, at 65–66. However, to the extent that it does exist, it only mandates that individuals be shown the revenues the government brings in. In the case of consumer goods, the cost of taxes is presented on receipts, so consumers have the opportunity to see what the government is taking in from them. For the purposes of this Note's discussion of over-taxation, it is not particularly important why individuals are ignoring their taxes, only that they are ignored.

necessary. It may also simply be the case that when taxes are not properly evaluated because they are hidden, they may not be set at an optimal level.⁷³ Taxpayers might accept over-taxation because they perceive the tax rate to be at an acceptable level. No bad actors are necessary in this scenario. Finkelstein's study provides some robust evidence that hidden taxes can lead to non-optimal taxation; her study found evidence of over-taxation.⁷⁴ Indeed, over-taxation has been a concern since at least the time of John Stuart Mill,⁷⁵ and such over-taxation seems a very likely result of the undervaluation of hidden taxes.

As to the regressivity concern, given the nature of hidden taxes and the processes that cause people to ignore them, poorer consumers will likely be disproportionately affected by the hidden taxes. Such consumers' cognitive abilities are likely not as developed as richer consumers, which would predictably lead to less accurate estimations of the cost of hidden taxes.⁷⁶ In addition, poorer consumers are likely to spend a higher proportion of their budgets than rich consumers on inexpensive items, which should lead to more approximation or ignorance of tax costs, especially where cognitive costs are higher for poorer consumers. If poor consumers are disproportionately affected by hidden taxes, then hidden taxes may be said to be regressive.⁷⁷ Finally, sales taxes, which are some of the most prominent hidden taxes, are widely recognized as regres-

73. The widely accepted and used basic model for determining what taxation scheme is optimal is found in James Mirrlees's seminal 1971 article *An Exploration in the Theory of Optimum Income Taxation*. Under the Mirrlees model, the goal of government is to choose a tax structure that maximizes the welfare of society as defined by a social welfare function. See J. A. Mirrlees, *An Exploration in the Theory of Optimum Income Taxation*, 38 REV. ECON. STUD. 175, 178 (1971).

74. See Finkelstein, *supra* note 3, at 991–1005 (describing impact of electronic toll collection on political behavior).

75. See 2 JOHN STUART MILL, PRINCIPLES OF POLITICAL ECONOMY WITH SOME OF THEIR APPLICATIONS TO SOCIAL PHILOSOPHY 491–93 (New York, D. Appleton & Co. 1892) (1848).

76. Galle, *supra* note 17, at 102–03 (“[A] regressive influence is the likelihood that the disutility of engaging in calculations diminishes as wealth increases. More precisely, it seems likely that the difficulty of carrying out mathematical operations declines with education, and education correlates with wealth. Wealthier individuals may also have computational aids, such as an accountant on speed-dial, that are unavailable to those of more modest means. . . . [E]ducation seems likely to improve taxpayers capacity to observe and compute taxes, and, again, education is strongly correlated with wealth.”).

77. *Id.* at 100 (“[I]f the likelihood that a consumer will pay sales taxes rather than shift to a consumption decision that is not taxed correlates with lower income, then imposing hidden sales tax will result in a more regressive tax structure.”).

sive.⁷⁸ If they become hidden, sales taxes' regressive effects might seem smaller than they actually are and thus more socially acceptable. While this doesn't affect the actual regressive impact of the taxes, it could cause society to be less willing to address the impact.

Of course, high taxes and regressive taxes do not necessarily represent a social welfare cost. If taxpayers are receiving benefits greater than or equal to the costs of the taxes, then there is not necessarily a social welfare problem, though the taxes may not be optimal. High taxes could lead to a plethora of public benefits, and these benefits could disproportionately accrue to the poor, removing net regressivity concerns. However, in a federal system, horizontal fiscal externalities—the effects of the abilities of citizens to move to different taxing jurisdictions—can cause states to compete away redistributive objectives, thus causing state tax regimes to be more regressive.⁷⁹ Thus, to the extent that hidden taxes are state sales taxes, they are likely to be regressive.⁸⁰

Additionally, there is an idea that the undervaluation of hidden taxes causes consumers to express pre-tax preferences, thereby reducing the substitution effects of the taxes.⁸¹ The substitution effect of a tax is the change in preferences it causes.⁸² For example, Kosha may prefer a \$5.00 magazine in a tax-free world, but once a tax is imposed on magazines, the substitution effect of the tax causes her to prefer a \$4.50 magazine instead. The actual economic effect the market feels from this substitution effect is referred to as “deadweight loss”.⁸³ When pre-tax consumer preferences are expressed, the effect that society feels from the deadweight loss associated with taxes may be lessened because of the absence (or

78. See, e.g., RICHARD A. MUSGRAVE & PEGGY B. MUSGRAVE, PUBLIC FINANCE IN THEORY AND PRACTICE 331–32 (2d ed. 1976).

79. See Robin Boadway, Maurice Marchand & Marianne Vigneault, *The Consequences of Overlapping Tax Bases for Redistribution and Public Spending in a Federation*, 68 J. PUB. ECON. 453, 474 (1998); see also David E. Wildasin, *Income Redistribution in a Common Labor Market*, 81 AM. ECON. REV. 757, 761–65 (1991) (outlining model showing externalities of localized redistribution in federal system).

80. One instance where there would not be net regressivity concerns would be where taxes that have become hidden are solely dedicated to funding social welfare programs.

81. Galle, *supra* note 17, at 77–79, 93; Congdon et al., *supra* note 67, at 10. The idea here is that hidden taxes presumably have no effect on purchasing decisions, thus removing the expected substitution effect of the tax and eliminating the resulting deadweight loss.

82. Bankman & Griffith, *supra* note 64, at 1919–20.

83. See JONATHAN GRUBER, PUBLIC FINANCE AND PUBLIC POLICY 33–35, 49–50, 548–58 (2005); MUSGRAVE, *supra* note 78, at 461–71. For a graphical representation of deadweight loss, see Figure 1 in the attached Appendix.

lessening) of the substitution effect, creating a more robust marketplace.⁸⁴ While this argument has some initial appeal, it rests on tenuous ground because its focus on aggregate markets ignores the income effect on welfare felt by individual consumers.⁸⁵ The income effect of a tax refers to the fact that it costs more to buy the same goods after they have been taxed, thus more income is required to maintain pre-tax preferences, causing the value of income to increase.⁸⁶ This income effect could cause individuals to work more, save less, or take on more debt—any of which could have negative impacts on social welfare. What is important to take from these ideas is that the social welfare costs of the undervaluation of hidden taxes are somewhat unclear, but it is likely that over-taxation and regressive taxation result. The next Section examines the harm to individuals resulting from the undervaluation of taxes. Compared with the social harms, the individual harms are rather straightforward.

B. *Individuals*

Given that consumers undervalue hidden taxes, tax-exclusive pricing injures consumer welfare by causing consumers to over-consume.⁸⁷ Overconsumption causes consumers to suffer due to the unexpected income effects of the hidden taxes.⁸⁸ As a result, consumers must save less money than they expected, work more than they would prefer, take on more debt than they would prefer, or consume less overall than they initially preferred when they thought their tax burden would be smaller. Authors of many studies have noted that such results are expected in the real world, as evidenced by sellers' anticipations that sales would decline with tax-inclusive pricing.⁸⁹

Even though the actual income effect of the undervaluation of each individual hidden tax may be slight, even slight harms such as these spread over such a large number of purchases and a large

84. For a detailed analysis of this argument, see Chetty et. al., *supra* note 3, at 1173.

85. Galle, *supra* note 17, at 79.

86. Bankman & Griffith, *supra* note 64, at 1920.

87. See *infra* app. fig.2.

88. Congdon et al., *supra* note 67, at 8 (“[W]ith a binding budget constraint, spending too much on the good with a hidden tax will leave less income for subsequent purchases—distorting individual consumption and decreasing welfare.”).

89. Chetty et al., *supra* note 3, at 1150 (“The grocery chain’s managers expected that posting tax-inclusive prices would reduce sales.”); see also Bird, *supra* note 14, at 70–71.

number of people can create substantial harm.⁹⁰ Consumers are undervaluing taxes on a number of purchases every day, and one could estimate that each individual consumer's yearly harm is not insignificant, even if each purchase only results in a few pennies worth of harm.⁹¹ However, the effect of losing a few pennies on each purchase does not make the harm salient enough to compel action from any individual consumer, creating a classic collective action problem.⁹²

Further, the harm from hidden taxes is not outweighed by any countervailing benefits. The undervaluation itself does not directly benefit individuals in their capacity as consumers but could indirectly benefit them by increasing social welfare. However, as already discussed, social welfare is not necessarily raised by the undervaluation of hidden taxes. Hidden taxes appear to be regressive in nature, thus causing individual welfare losses for poorer consumers. These welfare losses are relatively more painful to poorer consumers, who should be expected to undervalue hidden taxes with a greater degree of error than wealthier individuals. Also, if the undervaluation of hidden taxes results in over-taxation, then individuals are not receiving an optimal set of benefits from the government. Thus, overtaxed consumers are being harmed because their income could be better spent. Finally, even if the undervaluation of hidden taxes reduces the deadweight loss of the substitution effect of taxes, individual consumers suffer from the income effects.

It is clear therefore that tax-exclusive pricing harms individual consumer welfare by creating hidden taxes that cause consumers to over-consume. This practice does not have any countervailing benefits for individual welfare, and naturally leads to the question of

90. See Chetty et al. Working Paper, *supra* note 32, at 3 ("Even though *individual* welfare may be minimally affected by ignoring taxes, the same taxes can have large impacts on *social* welfare and tax revenue."). The Federal Trade Commission has also recognized that small individual harms can lead to substantial harms when a large number of people are affected. For example, one purpose for enacting the FTC Octane Rule was to prevent millions of consumers from spending a few extra pennies per gallon when buying gasoline. See Posting of Minimum Octane Numbers on Gasoline Dispensing Pumps, 36 Fed. Reg. 23,871, 23,872 (Dec. 16, 1971) (to be codified at 16 C.F.R. pt. 306).

91. For example, if a consumer made five purchases a day and underestimated the cost by an average of \$0.05 on each purchase, then he alone would suffer \$91 of harm yearly.

92. *E.g.* Posting of Minimum Octane Numbers on Gasoline Dispensing Pumps, 36 Fed. Reg. at 23,875 (discussing how consumers were unable to individually obtain information that they needed, creating a collective action problem that the rule sought to remedy).

whether consumers should be protected from tax-exclusive pricing. That question is answered affirmatively in the next Part.

IV. POLICY IMPLICATIONS

Given the welfare implications of undervaluing hidden taxes established in the last Part, consumers should be protected from tax-exclusive pricing on consumer goods. Individual welfare is being harmed by the systematic undervaluing of hidden taxes. Consumers' inability to overcome their biases and accurately consider the cost of hidden taxes prohibits market forces from eliminating the harm. If consumers are unable to correctly value the tax, then sellers who include tax will suffer in open competition because their prices will be perceived as higher than the prices of sellers who do not include the tax. Therefore, a competitive market will lead to tax-exclusive pricing to the ultimate detriment of the consumer.⁹³ Market failures such as this one "open[] the door to the possibility of welfare-enhancing legal intervention."⁹⁴ As the federal consumer protection agency, the Federal Trade Commission has the power and responsibility to intervene and should use its rulemaking authority to establish a minimally intrusive rule mandating tax-inclusive pricing.

A. *Who Should Act, and Why*

1. The Federal Trade Commission

The federal agency charged with protecting consumers is the FTC.⁹⁵ Using its authority granted in the Federal Trade Commission Act (FTCA),⁹⁶ the FTC has tackled a wide range of practices causing consumer injury. Examples of previous actions include mandating disclosure of octane levels in gasoline,⁹⁷ mandating the disclosure of proper care standards for garments,⁹⁸ creating strict rules for the presentation of the prices for funeral parlor services

93. It is admitted that some sellers do present tax-inclusive prices, but this observation about the effects of competition should explain why so few sellers do choose to present tax-inclusive prices.

94. Oren Bar-Gill, *Informing Consumers About Themselves* 43 (N.Y. Univ. Law & Econ. Research Paper Series, Working Paper No. 07-44, 2007), available at <http://ssrn.com/abstract=1056381>.

95. 15 U.S.C. § 45(a)(2) (2006).

96. Federal Trade Commission Act, 15 U.S.C. §§ 41–58 (2006).

97. 16 C.F.R. § 306.10 (2010).

98. 16 C.F.R. § 423.3 (2010).

and goods,⁹⁹ creating strict disclosure requirements for franchisors,¹⁰⁰ and regulating credit practices.¹⁰¹

The FTC's position as the federal consumer protection agency makes it the best actor to tackle the problems surrounding hidden taxes. While it may seem unclear who should act since the problems involve taxes, this murkiness is easily dissipated. The problem arises not from the taxes themselves, but from the fact that the taxes are not disclosed on the price tag. Therefore, the injury arises from the choice of sellers not to include taxes in their prices. Since this is a trade practice that is adversely affecting consumers, the FTC has jurisdiction.

Granted, there is the issue that taxes are not imposed by sellers on consumers. It is therefore unclear that sellers should be responsible for disclosing taxes. Even so, sellers are responsible for the practice of tax-exclusive pricing and benefit from it at the expense of consumers. The FTC has regulated similar practices in the past, as exemplified by rules concerning fees such as shipping and handling fees¹⁰² or per call fees.¹⁰³ Even though such costs are not considered part of the price of the product, the Commission still requires that sellers disclose them to consumers, since they are part of the total cost. It follows from this observation that since taxes are part of the total cost, though not part of the price, they also are prime for similar FTC action. Also, at the most basic level, sellers are the least-cost providers of such information, as it would cost them next to nothing to provide the tax information.¹⁰⁴ On this basis alone, they are ideal candidates to disclose the information.¹⁰⁵

99. 16 C.F.R. § 453.2 (2010).

100. 16 C.F.R. § 436.2 (2010).

101. 16 C.F.R. § 444 (2010).

102. 16 C.F.R. § 310.3(a)(1) (2010) (requiring total cost to consumer to be disclosed before payment in telemarketing sales).

103. 16 C.F.R. § 308.3 (2010) (requiring disclosure of cost of call for pay-per-call service).

104. Bar-Gill, *supra* note 94, at 63.

105. *Cf.* Stephen G. Gilles, *Negligence, Strict Liability, and the Cheapest Cost-Avoider*, 78 VA. L. REV. 1291, 1306 (1992) (discussing general understanding that assigning liability to least-cost avoider in tort actions leads to efficient level of care). The argument for efficient results in tort by assigning liability to the least-cost avoider is easily translatable to the current discussion. The harm from tax-exclusive pricing could be corrected by sellers, the government, or consumers themselves. In any scenario, the cost of fixing the harm is likely to fall on consumers. Sellers will pass any costs along in price, the government will collect more taxes or decrease spending (although neither of these technically have to affect consumers) to account for any costs, and consumers would bear their own costs, which have been shown to be high. Since sellers can disclose taxes most cheaply, as

2. Consumer Protection

It is important to understand how the FTC exercises its consumer protection authority. Consumer protection laws are designed to ensure that consumers have the ability to choose effectively between different options in the marketplace. Such choices require that consumers possess the information reasonably necessary to accurately express their preferences and that they are not misled by certain trade practices.¹⁰⁶ Therefore, the federal law of consumer protection as found in section 5 of the FTCA prohibits “unfair and deceptive trade practices.”¹⁰⁷

Consumer protection law and interpretation of the FTCA are guided by the principle of protecting consumer sovereignty.¹⁰⁸ Consumer sovereignty is an expression of the ideal that consumers should have both the opportunity to define their desires and the opportunity to act to fulfill those desires.¹⁰⁹ Implicit in the definition of consumer sovereignty and in the current analysis of consumer injuries is that consumers make rational decisions.¹¹⁰ Thus, when consumers are given choices and the information needed to choose, they will perfectly express their preferences, and markets will operate efficiently by responding to consumer demand rather than government directives or business preferences.¹¹¹

The FTCA in its current form is a rather vague grant of power. Congress intended for the FTC to develop a framework for identifying unfair trade practices and recognized the Commission’s need for flexibility in this area, since prohibiting specific actions would have simply driven offenders to find loopholes and new methods of

they are already responsible for displaying price to consumers, they are the least-cost provider of the information (or least-cost avoider of the harm). Therefore, they should be required to provide the tax information.

106. Averitt & Lande, *supra* note 6, at 713–14.

107. Federal Trade Commission Act § 5. 15 U.S.C. § 45 (2006).

108. Neil W. Averitt, *The Meaning of “Unfair Acts or Practices” in Section 5 of the Federal Trade Commission Act*, 70 GEO. L.J. 225, 281 (1981).

109. Averitt & Lande, *supra* note 6, at 713.

110. *See Int’l Harvester Co.*, 104 F.T.C. 949, 1061 (1984) (“The Commission does not ordinarily seek to mandate specific conduct or specific social outcomes, but rather seeks to ensure simply that markets operate freely, so that consumers can make their own decisions.”); Thomas B. Leary, *Unfairness and the Internet*, 46 WAYNE L. REV. 1711, 1713 (2000) (noting that current standard “gives primacy to economic factors, and introduces the notion of consumer responsibility”); Joel Waldfoegel, *Does Consumer Irrationality Trump Consumer Sovereignty?*, 87 REV. ECON. & STAT. 691, 691 (2005) (stating that consumer sovereignty and rationality is central to economic theory, although empirical evidence suggests otherwise).

111. Averitt & Lande, *supra* note 6, at 715–16.

exploiting consumers.¹¹² The original framework, the “S&H” test,¹¹³ was used until 1980. It examined whether the act or practice in question 1) violated established public policy, 2) was immoral or unethical, or 3) resulted in substantial consumer injury.¹¹⁴ A 1980 letter from the FTC to Congress outlined a modified version of this test.¹¹⁵ The letter explained that unjustified consumer injury would

112. See S. REP. NO. 63-597, at 13 (1914) (“The committee gave careful consideration to the question as to whether it would attempt to define the many and variable unfair practices which prevail in commerce and to forbid their continuance or whether it would, by a general declaration condemning unfair practices, leave it to the commission to determine what practices were unfair. It concluded that the latter course would be the better, for the reason . . . that there were too many unfair practices to define, and after writing 20 of them into the law it would be quite possible to invent others.”); see also H.R. REP. NO. 63-1142, at 19 (1914) (“It is impossible to frame definitions which embrace all unfair practices. There is no limit to human inventiveness in this field. Even if all known unfair practices were specifically defined and prohibited, it would be at once necessary to begin over again. If Congress were to adopt the method of definition, it would undertake an endless task.”). The Supreme Court has also acknowledged this need for flexibility. See *FTC v. R.F. Keppel & Bro.*, 291 U.S. 304, 310 (1934) (“Neither the language nor the history of the Act suggests that Congress intended to confine the forbidden methods to fixed and unyielding categories.”); *FTC v. Raladam Co.*, 283 U.S. 643, 648 (1931) (“[U]nfair methods of competition. . . belongs to that class of phrases which does not admit of precise definition.”). The Court has also recognized that the Commission is responsible for developing the framework for discovering unfair trade practices. See *id.* at 649 (“[T]he Commission is called upon first to determine, as a necessary prerequisite to the issue of a complaint, whether there is reason to believe that a given person, partnership or corporation has been or is using any unfair method of competition in commerce.”); *Atl. Ref. Co. v. FTC*, 381 U.S. 357, 367 (1965) (“In a broad delegation of power it empowers the Commission, in the first instance, to determine whether a method of competition or the act or practice complained of is unfair.”).

113. The “S&H” test derives its name from the Supreme Court case *FTC v. Sperry & Hutchinson Co.*, 405 U.S. 233 (1972). In that case, the Supreme Court acquiesced in the FTC’s framework for finding unfair trade practices. *Id.* at 244–45 & n.5.

114. *Id.* at 244 n.5 (1972) (recognizing framework as established in *Unfair or Deceptive Advertising and Labeling of Cigarettes in Relation to the Health Hazards of Smoking*, 29 Fed. Reg. 8324, 8355 (July 2, 1964) (to be codified at 16 C.F.R. pt. 408)).

115. 4 Trade Reg. Rep. (CCH) ¶ 13,203 (Dec. 17, 1980) [hereinafter *FTC Unfairness Policy Statement*] (letter from the FTC to the Consumer Subcommittee of the Senate Committee on Commerce, Science, and Transportation). The S&H framework is reproduced in the *FTC Unfairness Policy Statement*, although in slightly different order. *Id.* ¶ 20,908. However, the letter effectively eliminates the “immoral or unethical” prong of the original test and modifies the import of the other two prongs.

thereafter be the main focus of unfairness investigations.¹¹⁶ Public policy was to be used only as evidence that certain types of conduct cause consumer injury,¹¹⁷ and the immorality question was considered duplicative.¹¹⁸ This was all codified as 15 U.S.C. § 45(n),¹¹⁹ consequently superceding the S&H test.

The Commission has stated that for a consumer injury to be legally unfair it “must be substantial; it must not be outweighed by any countervailing benefits to consumers . . . that the practice produces; and it must be an injury that consumers themselves could not reasonably have avoided.”¹²⁰ An injury’s harm can be substantial if it harms either a small group of consumers in a large way¹²¹ or a large group of consumers in a small way.¹²² When considering the net effects of the practice to determine if its benefits outweigh its costs, the Commission may choose to consider the effects on the ultimate consumer only, ignoring the effects the practice has on sellers or even the government.¹²³ The Commission seeks to deter-

116. *Id.* at 20,908 (“Unjustified consumer injury is the primary focus of the FTC Act, and the most important of the three *S & H* criteria.”).

117. *Id.* at 20,909 to -2 (describing how public policy is used in consumer protection cases); see Federal Trade Commission Act § 5(n), 15 U.S.C. § 45(n) (2006) (“In determining whether an act or practice is unfair, the Commission may consider established public policies as evidence to be considered with all other evidence. Such public policy considerations may not serve as a primary basis for such determination.”).

118. FTC Unfairness Policy Statement, *supra* note 115, at 20,909-3.

119. 15 U.S.C. § 45(n).

120. FTC Unfairness Policy Statement, *supra* note 115, at 20,908.

121. For an example of a situation where substantial injury from large harms to a small group of people could occur consider the rationale behind the FTC Franchise Rule, 16 C.F.R. § 436 (2010). See Disclosure Requirements and Prohibitions Concerning Franchising and Business Opportunity Ventures, 43 Fed. Reg. 59,614, 59,621-39 (Dec. 21, 1978) (to be codified at 16 C.F.R. pt. 436) (finding that franchising practices posed significant threats to franchisees and therefore fell under scope of FTCA).

122. See *Orkin Exterminating Co.*, 108 F.T.C. 263, 362 (1986), *aff’d sub nom. Orkin Exterminating Co. v. FTC*, 849 F.2d 1354 (11th Cir. 1988) (“[A]n injury may be substantial if it does a small harm to a large number of people.”); see also *Posting of Minimum Octane Numbers on Gasoline Dispensing Pumps*, 36 Fed. Reg. 23,871, 23,872 (Dec. 16, 1971) (to be codified at 16 C.F.R. pt. 306) (describing potential harm to consumers from overpaying by a few cents for each gallon of gasoline).

123. The Supreme Court has approved this practice in relation to the competition side of the FTCA. See *Atl. Ref. Co. v. FTC*, 381 U.S. 357, 371 (1965) (“Upon considering the destructive effect on commerce that would result from the widespread use of [the practice,] we conclude that the Commission was clearly justified in refusing the participants an opportunity to offset these evils by a showing of economic benefit to themselves.”). This is basically a ban on considering the exter-

mine if terminating the practice would actually make consumers worse off, a consideration that includes the costs of implementing the remedy, as well as the effect of the remedy itself.¹²⁴

The final part of the Commission's consideration, that consumers cannot reasonably avoid the injury on their own,¹²⁵ carries the most weight. The FTC has explained that "whether some consequence is 'reasonably avoidable' depends, not just on whether people know the physical steps to take in order to prevent it, but also whether they understand the necessity of actually taking those steps."¹²⁶ This framework allows the FTC to step in when the market fails to eliminate potential injuries because consumers are unable to accurately express their choices. When consumers can accurately express their choices, the market will self-correct any potential injuries because consumers will simply avoid harmful practices.

This Note is concerned with two of the most common types of consumer injuries: incomplete information and confusing information.¹²⁷ The two are related in that they both obstruct the ability of the consumer to express her preferences accurately. Both drive up the costs of receiving the information reasonably necessary to make a purchasing decision to a point where the consumer is effectively unable to use that information.¹²⁸

Incomplete information is a relative concept. Total disclosure of all information would be costly, if not impossible. Thus, in general, only basic information thought necessary for informed decisionmaking—namely, the price and function of the good—is

nalities of the harm on any parties other than consumers. This is logical; one would not want to allow an exploitative practice because the exploiter gains from it in an equal or greater amount as the exploited loses.

124. FTC Unfairness Policy Statement, *supra* note 115, at 20,909.

125. *Id.*

126. *Int'l Harvester Co.*, 104 F.T.C. 949, 1066 (1984); *accord Orkin Exterminating Co.*, 849 F.2d at 1365 (quoting *Orkin Exterminating Co.*, 108 F.T.C. at 366) ("Consumers may act to avoid injury before it occurs if they have reason to anticipate the impending harm and the means to avoid it, or they may seek to mitigate the damage afterward if they are aware of potential avenues toward that end.").

127. *Averitt & Lande*, *supra* note 6, at 733 ("The most common internal market failures fall into five categories: (1) overt coercion; (2) undue influence; (3) deception; (4) incomplete information; or (5) confusing information."). *Averitt & Lande* give brief descriptions of all five categories of harm. *See id.* at 733–34.

128. *See id.*; *Averitt*, *supra* note 108, at 258–61 (explaining that consumers are unable to make rational decisions without disclosure of material information, even though such information is technically discoverable but only at very high costs to individual).

required to be disclosed.¹²⁹ Specific disclosures of non-basic information may be required on a case-by-case basis.¹³⁰ Only the price at the time of the sale is required to be disclosed because almost every bit of information regarding the future value of a product could be monetized, and to require disclosure of such values would essentially require full disclosure of all tangentially related information.¹³¹ Thus, things such as maintenance costs and scrap value are typically not required to be disclosed. Following the same logic, the information on the function of a product required for disclosure is typically only the primary function of the product, not any secondary functions.¹³²

A final basic category of required disclosures includes those necessary to correct misconceptions about the product or service where the misconceptions are causing consumer injury.¹³³ The seller would have to know that the misconceptions exist and are injuring consumers. The Commission has long required these sorts of disclosures, as seen in various rules and cases, particularly in instances concerning misleading advertising.¹³⁴ These disclosures are not central to this Note, as it is assumed that consumers are aware that taxes exist.¹³⁵

129. See Averitt, *supra* note 108, at 260–65 (noting that “[t]here are obviously a great many pieces of information about a product that consumers might want[,] . . . [y]et the Commission surely would not require such time-consuming data to be compiled” and discussing basic disclosures of price, function, and information needed to correct misconceptions about products).

130. See, e.g., 16 C.F.R. § 436.5 (2010) (describing required disclosures of wealth of non-basic information about franchising opportunities); 16 C.F.R. § 423.3 (2010) (requiring disclosure of instructions for proper care of garments).

131. Averitt, *supra* note 108, at 262.

132. *Id.* at 263–64.

133. *Id.* at 264–65.

134. *Id.*; see, e.g., 16 C.F.R. § 424.1 (2010) (requiring disclosure that advertised goods may be limited in availability, if such is the case); Heavenly Creations, Inc. v. FTC, 339 F.2d 7, 8–9 (2d Cir. 1964) (upholding FTC’s cease and desist decree against petitioner for failing to disclose certain information in advertisements while noting that “if the Commission . . . thinks it best to insist upon a form of advertising clear enough so that, in the words of the prophet Isaiah, ‘wayfaring men, though fools, shall not err therein,’ it is not for the courts to revise their judgment”).

135. To the extent that consumers are actually unaware that taxes exist, there is a case that their disclosure should be mandated under this theory of curing misconceptions.

Also, information should not be disclosed in a confusing way.¹³⁶ Disclosed information would serve little purpose if consumers were unable to understand what the information meant. Confusing information can be seen as a subset of incomplete information because it is effectively taken out of the picture, or as a deceptive trade practice, if the confusion misleads consumers.¹³⁷ Examples of attempts to prevent confusing information are seen in the FTC's guides concerning environmental marketing claims¹³⁸ and the use of endorsements and testimonials in advertisements.¹³⁹

3. Tax-Exclusive Pricing is an Unfair Trade Practice

Given its welfare implications, tax-exclusive pricing is an unfair trade practice under the FTCA because it causes consumer injury. Tax-exclusive pricing is a clear case of incomplete information in that it impedes consumers' ability to form accurate preferences in a way that consumers cannot overcome on their own because of their cognitive biases discussed above.¹⁴⁰ Part II of this Note showed that tax-exclusive pricing leads to taxes being hidden and to consumers not correctly valuing their costs. This undervaluation leads to substantial aggregate consumer harm, as consumers suffer from every purchase that has a hidden tax. Because the individual harm on each purchase is so small, many consumers may not realize that they are being harmed, thus not realizing the necessity of correcting their actions even if they had the ability to do so. Finally, as developed in Part III, individual consumers' welfare is being harmed, and there are no countervailing benefits to tax-exclusive pricing that would remove it from the realm of consumer injury. However, in addition to finding consumer injury, the FTC must also be assured that the potential remedy's costs do not outweigh its benefits before it takes action.

136. *See Bakers Franchise Corp. v. FTC*, 302 F.2d 258, 262 (3d Cir. 1962) (upholding Commission's order to bread company to stop using language likely to confuse consumers).

137. *Averitt & Lande*, *supra* note 6, at 733–34; *Gulf Oil Corp. v. FTC*, 150 F.2d 106, 109 (5th Cir. 1945) (quoting *Florence Mfg. Co. v. J.C. Dowd & Co.*, 178 F. 73, 75 (2d. Cir. 1910) (“The law is not made for the protection of experts, but for the public—that vast multitude which includes the ignorant, the unthinking and the credulous, who, in making purchases, do not stop to analyze, but are governed by appearances and general impressions.”)).

138. 16 C.F.R. § 260.7 (2010).

139. *Id.* § 255 (2010).

140. *See Trout*, *supra* note 56, at 412 (noting that “serious doubts” exist concerning “the ability of normal individuals to make decisions that are” as accurate as “tested institutional decisions” because of cognitive biases).

B. How Should the FTC Fix It?

1. The Ideal Disclosure Rule

Given that the FTC has the jurisdiction to tackle the injuries created by tax-exclusive pricing, what action should it take? The FTC could best address this issue through a disclosure rule. Mandatory disclosure of hidden taxes would create substantial benefits for individuals by eliminating the potential for consumer undervaluation of hidden taxes while imposing only a minimal cost on sellers.¹⁴¹

Disclosure requirements are frequently simple and relatively cost-efficient means of regulation,¹⁴² and are often used in regard to consumer products.¹⁴³ Disclosure does not regulate market forces, but merely disseminates information with the hope that individuals will use the information to better define their preferences. Disclosure facilitates efficient markets by removing barriers to information, thus reducing the cost of information. In this case, disclosure would prevent undervaluation by removing the costs associated with calculation. However, disclosure is not without costs. The direct costs of disclosure include the costs to sellers of collecting, compiling, and distributing information.¹⁴⁴ In addition, too much disclosure could flood consumers, and consumers may not be able to process complex or confusing disclosures.¹⁴⁵ As will be seen, however, there is a simple way to disclose the costs of hidden taxes with minimal costs to sellers and little danger of flooding or confusing consumers.

When considering the disclosure scheme it should be observed that the taxes that should be disclosed are those that a consumer without any special tax exemptions would have to pay.¹⁴⁶ While it would be ideal for each individual to know the total tax-inclusive price before purchasing, such a system would be impractical.

141. See Stern, *supra* note 15, at 87 (“Providing disclosure at the outset of a transaction ameliorates the information-processing deficits elicited by latecoming disclosure . . . [and] does not appear to [be] burdensome for sellers and their agents.”).

142. Bar-Gill, *supra* note 94, at 44, 46, 63.

143. *E.g.*, 16 C.F.R. § 306.10 (2010) (requiring disclosure of automotive fuel ratings); 16 C.F.R. § 423.3 (2010) (requiring disclosure of care requirements on apparel labels).

144. Bar-Gill, *supra* note 94, at 63.

145. *Id.* at 64; Colin Camerer et al., *Regulation for Conservatives: Behavioral Economics and the Case for “Asymmetric Paternalism,”* 151 U. PA. L. REV. 1211, 1235 (2003).

146. That different goods may not be taxed the same is tax policy that this Note is not concerned with.

Therefore, it must be assumed that consumers with special tax exemptions will be sophisticated enough to realize that they have those exemptions and that they will be able to calculate their actual burden. While such an assumption would be contrary to the argument of this Note, it is the safest assumption to make when creating the rule. It will cause the least amount of harm until the day when consumers can find their individual total prices before making purchasing decisions. While the exemptions differ from state to state, non-profit organizations are one of the main classes of exempt consumers. Non-profits may be more sophisticated than individual consumers due to institutional knowledge and protections. Indirect taxes on goods, such as corporate taxes that are already worked into price to some degree, will not need to be disclosed.¹⁴⁷

One can conceive of four general schemes for the disclosure of the price and taxes for consumer goods: 1) fully tax-exclusive prices, 2) tax and price both stated separately but not added together, 3) tax and price both stated separately and added together, and 4) fully tax-inclusive prices with no breakdown. To optimize consumer sovereignty, price and taxes should be both stated separately and added together. To arrive at this conclusion, consider the effects of each scheme.

The first two means of presenting prices—fully tax-exclusive and separate disclosure of price and tax—are tax-exclusive pricing schemes and cannot be allowed because they will cause consumer injury. Although the second scheme would present the cost of taxes, the cognitive biases discussed in Part II would still lead to undervaluation of the total cost of goods even with that disclosure. Since undervaluation of cost is the target of the disclosure requirement, it would not make sense to implement such a scheme if it would not prevent consumer injury.

A fully tax-inclusive scheme is appealing due to its simplicity but would create its own problems by hiding taxes in another way—making their existence less salient.¹⁴⁸ Consumers would then have trouble discovering what their tax burden is in comparison to the price for a product. This would make it difficult for consumers to

147. Fees are analogous to taxes. Some fees are already required to be disclosed by the FTC, *see, e.g.*, 16 C.F.R. § 453.2 (2010) (requiring disclosure of fees for funeral parlor goods and services), but many are not. This Note does not address fees, but one could surmise that the analysis here would easily apply to fee-exclusive pricing.

148. *See* Bird, *supra* note 14, at 66–68 (describing possible effects of keeping existence of taxes salient by excluding them from posted prices and exposing them separately at the register).

challenge taxes and might lead to over-taxation by the government.¹⁴⁹ Sellers might also inflate prices and blame the total cost rise on taxes, although this would most likely constitute a deceptive act and be illegal as such under the FTCA. While a fully tax-inclusive scheme could viably solve the undervaluation problem, it is not the best option given its costs.

The tax-inclusive pricing scheme with a breakdown of price and tax represents the best solution for consumers. This tax-inclusive scheme would ensure that the total price of the product would be presented to the consumer before the purchasing decision is made while also giving her the opportunity to understand her actual tax burden, alleviating over-taxation concerns.¹⁵⁰ This solution may also cause the taxes that have become hidden to become less regressive by making them more salient if the public takes notice of them and wishes to make a change.

There are two possible ways to disclose prices and taxes under this scheme, and sellers should be given the choice between the two. The first option would be to state the price, the taxes, and the total cost before the consumer makes the purchase, presumably by putting all the information on the price tag. Sellers wanting to distinguish between their prices and taxes would prefer this method. This type of disclosure could lead to consumer confusion, but any confusion could easily be avoided by mandating that the total cost be in a more noticeable font. The second way to disclose under the scheme is simple, easy, and practically cost-free. It would be to merely change the posted price to the tax-inclusive price and maintain the current tax and price breakdown on receipts.¹⁵¹ Since disclosure requirements should impose as small a burden as possible

149. See Finkelstein, *supra* note 3, at 1002–05 (discussing reduced consumer response to increases in tolls, which allowed for increases in the tolls, even during election years, without negative political impact).

150. The concern arising from the disaggregation bias that consumers will undervalue their total tax burden will still exist, but the cost of providing consumers with constant reminders of what they have spent would far outweigh the benefits under this new system.

151. It is admitted that the current price and tax breakdown on receipts has not effectively counteracted the behavioral processes that lead to undervaluation of hidden taxes, and valid concerns exist that failing to change this structure will still encourage over-taxation because individuals will not be aware of their tax burden. However, by presenting the tax-inclusive price to consumers on the price tag rather than the tax-exclusive price, consumers will technically become aware of the cost of the taxes, and can determine the exact cost by looking at their receipts. Over-taxation will be less of a concern because consumers will realize the total costs of products before deciding to purchase them, and if they feel that those costs are too high, they can lobby the government to lower the tax rate.

on sellers,¹⁵² implementing a scheme that requires only a change in posted prices—while also affording the option to present price breakdowns on the price tag—is ideal. Additionally, as the costs of implementing this rule would be small and the benefits huge, it would pass the FTC’s cost-benefit analysis.

2. Implementation Challenges

Though a tax-inclusive scheme with a breakdown of price and tax appears to be the best way to prevent consumer injuries from tax-exclusive pricing, it will undoubtedly face some challenges to its implementation. Such regulation might be attacked as overly paternalistic for failing to give consumers the chance to de-bias themselves. However, consumers have proven unable to de-bias themselves, and this disclosure rule would constitute only “weak paternalism,” where regulators provide individuals with the information they need to decide what they want.¹⁵³ Also, given the negligible costs of disclosure to sophisticated consumers who are relatively unharmed by undervaluation and the substantial benefits it would bring to the unsophisticated consumers suffering the most from undervaluation, the regulation would fit the definition of “asymmetrical paternalism.”¹⁵⁴ Asymmetrically paternalistic regulation has been used approvingly in the past, especially in the consumer protection context.¹⁵⁵

The regulation might also be attacked for its reliance on behavioral economics to define the consumer injury. However, the FTC’s goal in all cases is to accurately understand how trade prac-

152. See Bar-Gill, *supra* note 94, at 63 (noting that if disclosure requirements impose too large a burden on seller, cost will be passed to consumer).

153. See Cass R. Sunstein, *Boundedly Rational Borrowing*, 73 U. CHI. L. REV. 249, 258 (2006) (describing de-biasing efforts through law as the “weakest form of weak paternalism [because] the relevant steering operates directly on bounded rationality and allows people to act as they see fit”); Klick & Mitchell, *supra* note 59, at 1621 (noting that under “softer forms of paternalism . . . the government regulates the form in which information and options are presented to citizens and restricts the role of laypersons in the market, legal, and political systems without completely controlling choices”).

154. Camerer et al., *supra* note 145, at 1212 (“A regulation is asymmetrically paternalistic if it creates large benefits for those who make errors, while imposing little or no harm on those who are fully rational.”); see Sunstein, *supra* note 153, at 256.

155. See Camerer et al., *supra* note 145, at 1212 (describing asymmetrically paternalistic regulation in consumer protection context and using Truth in Lending Act, 15 U.S.C. §§ 1601–1667f (2006), as example). Indeed most, if not all, of the FTC disclosure regulations cited in this Note, *e.g.*, *supra* notes 97–103, 138–39, fit the definition of asymmetrically paternalistic regulation.

tices affect consumers. In this instance, behavioral economics is used to supplement the rational response to hidden taxes in a way that gives a more accurate picture of consumers' reactions. Such analysis is a valuable tool when used methodically and should not be ignored by the FTC. Behavioral economics provides valuable insight into the way that trade practices will affect consumer decision-making,¹⁵⁶ and the FTC can use behavioral analysis to help effectively de-bias consumers, which will lead to more robust and accurate consumer protection.¹⁵⁷ However, behavioral economic analyses must be carefully used. There are dangers inherent in crafting regulation from such tools, and the Commission should be cautious when so doing.¹⁵⁸ A cautious approach would mean crafting regulations based on behavioral economic analysis only where the analysis leads to clear conclusions about consumers, as has been done here regarding the effects of tax-exclusive pricing.

A further problem is that the FTC has been reluctant to engage in substantive unfairness rulemaking, although it has not expressed a similar reluctance towards disclosure rules.¹⁵⁹ While the current trend of the FTC's rulemaking has involved disclosure requirements, as advocated here, most of those rules have followed congressional mandates.¹⁶⁰ Therefore, it may be necessary for Congress to mandate that prices be tax-inclusive before the Commission will act. Congressional action may be difficult, given the challenges of the United States political structure.¹⁶¹ However, given that all voters are presumably consumers, such a provision could be expected to have a reasonable prospect of success if it made it to votes in the Senate and House of Representatives.

156. See, e.g., Congdon et al., *supra* note 67, at 2 ("Imperfectly rational people will respond to taxes in a way that is mediated by psychology.")

157. See Trout, *supra* note 56, at 433–34 (explaining that de-biasing through government regulation will enhance both autonomy and welfare).

158. See Klick & Mitchell, *supra* note 59, at 1626 (bringing attention to the potential "moral and cognitive hazards" of paternalistic interventions); Camerer, *supra* note 145, at 1253–54 (discussing need for carefully analysis of cognitive biases in order to determine if a seemingly irrational behavior is actually rational).

159. See Alan M. White & Cathy Lesser Mansfield, *Literacy and Contract*, 13 STAN. L. & POL'Y REV. 233, 259 (2002) ("The FTC has not issued a substantive consumer-contract regulation since the 1984 Credit Practices Rule. Recent rules tend to emphasize disclosures or the avoidance of misrepresentations and to regulate abusive conduct, rather than regulating substantive terms.")

160. MARY DEE PRIDGEN, CONSUMER PROTECTION AND THE LAW § 12:14 (2009).

161. See, e.g., Daniel A. Farber & Philip P. Frickey, *Legislative Intent and Public Choice*, 74 VA. L. REV. 423, 425–37 (1988) (detailing and responding to commonly accepted problems attributed to legislative action by public choice theory).

As a practical matter, the regulation will need to have a preemption clause since taxes on consumer products are mostly state imposed rather than federally imposed.¹⁶² Serious problems would result if the tax-inclusive scheme were not uniform. Individuals would be drawn to tax-exclusive prices and the whole scheme would be undermined. This would raise consumer protection concerns in itself as individuals could be misled by the tax-exclusive prices if they expect that all prices are tax-inclusive. Therefore, the rule would need a preemption clause to ensure that sellers in all states would fall under its mandate. Agency preemption actions can be problematic though, and may present a hurdle to adoption of the rule.¹⁶³

Monitoring sellers also presents a challenge. There could be a large incentive to cheat the system by not including taxes because of the advantages just described. However, doing so would certainly fall under the jurisdiction of the FTCA, and sellers that violate the rule could be subjected to fines and injunctive relief.¹⁶⁴ This could be a strict liability offense in order to remedy the situations quickly and cheaply. Individuals would presumably be able to act as watch-dogs. Price tags and receipts would allow for easy proof of violations.

Even given the challenges, the FTC should still take action. FTC Commissioners have noted that one of the strengths of the Commission is that it is charged with developing innovative regulatory solutions to potentially harmful market practices.¹⁶⁵ A rule mandating disclosure of taxes on consumer goods, as proposed, would play to that strength and reaffirm the Commission's commitments to innovative thinking and consumer protection.

V. CONCLUSION

This Note has proposed that the FTC create a rule mandating tax-inclusive pricing of consumer goods. Such a dramatic shift in the way prices of consumer goods are presented may appear hope-

162. For an example of a preemption clause in a FTC rule, see 16 C.F.R. § 306.4 (2010).

163. See Ernest A. Young, *Executive Preemption*, 102 NW. U. L. REV. 869, 871–81 (2008) (discussing executive preemption and federalism doctrine).

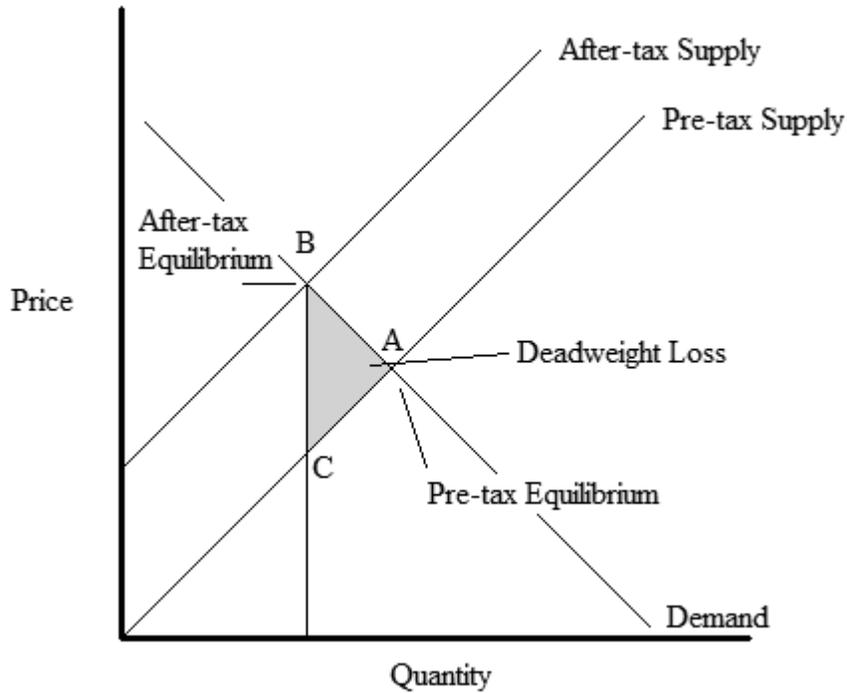
164. Federal Trade Commission Act § 5(m), 15 U.S.C. § 45(m) (2006).

165. See generally *More Than Law Enforcement: The FTC's Many Tools—A Conversation With Tim Muris and Bob Pitofsky*, 72 ANTITRUST L.J. 773 (2005) (edited version of the original conversation at the FTC 90th Anniversary Symposium (Sept. 22, 2004)).

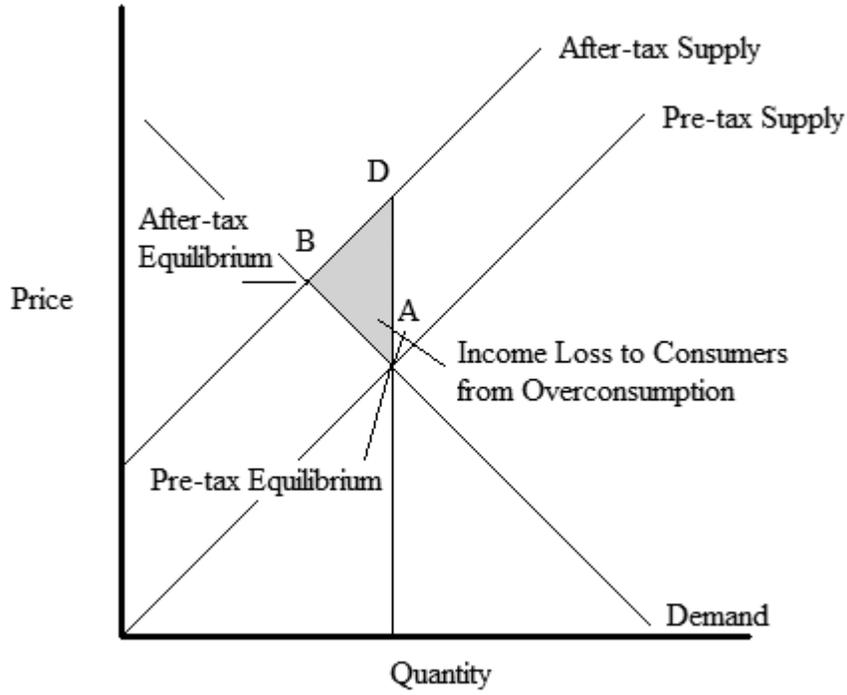
less at first glance, but the experience of the European Community and other countries provides evidence that it can work. While the United States is certainly different from its neighbors worldwide, it should not ignore the examples those neighbors provide. This Note by no means advocates that the FTC follow other countries' laws over U.S. laws, but when deciding how to draft regulation, the Commission would be wise to consider the difficulties and successes of those countries that have mandated tax-inclusive pricing.

This Note has applied the FTC's determination that incomplete information can create serious injuries to consumer sovereignty—the ability of consumers to freely choose from a variety of options on the market—to the practice of tax-exclusive pricing. Tax-exclusive pricing creates incomplete pricing information by lowering levels of tax salience such that taxes on consumer goods become hidden from and undervalued by consumers. Undervaluation leads to overconsumption and significant harm to individuals' welfare. The FTC must step in and provide consumers with the price information they are unable to discover on their own. Requiring sellers to present tax-inclusive prices is the cheapest and most efficient way of correcting this problem, and thus the Commission should mandate this practice.

APPENDIX



In this traditional model, the tax is fully salient to the consumer and its cost is considered in purchasing decisions, creating a deadweight loss to society. The deadweight loss from the tax is represented by area ABC, which is created because consumers who would be willing to purchase pre-tax are not willing to purchase after-tax.



In this hidden tax model, consumers ignore the cost of the tax and continue to purchase at pre-tax levels, causing overconsumption given the total price. The income loss to consumers from overconsumption is represented by area ABD, which is created because consumers pay more for the pre-tax equilibrium quantities than they anticipate.

REVISING JUDICIAL APPLICATION OF THE SINGLE SUBJECT RULE TO INITIATIVE PETITIONS

FLORIN V. IVAN*

The idea that an impartial branch of government protects the rule of law is appealing to many—so much so that the requirement of judicial impartiality is enshrined in our federal Constitution.¹ While both federal and state courts may be vulnerable to accusations of judicial activism tainted by partisan politics,² state courts are seen as more susceptible to political pressure.³ Few situations are as sensitive as cases in which state courts invalidate action by political branches.⁴ It stands to reason, then, that when state courts invalidate citizen-initiated ballot petitions in states that recognize the people's right to legislate directly, the judiciary would be highly susceptible to charges of partisanism. The charges might intensify when amorphous doctrines such as the single subject rule and equally amorphous jurisprudence combine to strike down popular initiatives before they appear on the ballot.

This Note answers critics and supporters of the single subject rule by exploring in depth the single subject jurisprudence in one state: Colorado. The Note examines the case law through the lens of four different hypotheses and tries to determine whether initia-

* J.D., New York University School of Law. Law clerk to the Hon. David G. Campbell, United States District Court for the District of Arizona. This Article does not reflect the Author's work as a law clerk nor the opinion of any court. The Author wishes to thank Professor Richard H. Pildes for providing invaluable advice on this project, and Professor Burt Neuborne for his impassioned accounts of the history of democracy. The Author is deeply grateful to the talented editorial staff of the Annual Survey of American Law, without whose support this Note would not have been possible.

1. Pamela S. Karlan, *Electing Judges, Judging Elections, and the Lessons of Caperton*, 123 HARV. L. REV. 80, 86 (2009).

2. See Erwin Chemerinsky, *The Meaning of Bush v. Gore: Thoughts on Professor Amar's Analysis*, 61 FLA. L. REV. 969 (2009) (discussing political pressure in federal court); Helen Hershkoff, *State Courts and the "Passive Virtues": Rethinking the Judicial Function*, 114 HARV. L. REV. 1833, 1886 (2001) (discussing political pressure in state courts).

3. Hershkoff, *supra* note 2, at 1886.

4. See, e.g., Devera B. Scott et al., *The Assault on Judicial Independence and the Uniquely Delaware Response*, 114 PENN. ST. L. REV. 217, 227–34 (2009) (listing high-profile examples of controversies following state court invalidation of political branch action).

tive petitions challenged on single subject grounds are struck down or upheld based on a consistent, intelligible interpretive framework—or whether other considerations account for case outcomes. The Note also suggests methodological enhancements—applicable to all states with similar rules—that would serve to dispel the perception that courts are subject-matter vetogates, while still enabling them to perform the important function of checking direct democracy.⁵

I. INTRODUCTION

Rousseau once wrote, “If there were a people of gods, it would govern itself democratically. A government so perfect is not suited to men.”⁶ As if to disprove his thesis and demonstrate that a nation of brave men and women *can* have political heaven on Earth, the great experiment that is our nation blended republicanism and democracy to reach a delicate balance.⁷ Maintaining the balance has not been easy, however: waves of populist fervor often sought to place more power in the hands of the people, giving rise to waves of counter-populism and staunch resistance from status quo power players.⁸ For example, when some states began granting their citizens the right to petition for changes to state laws and constitutions via ballot initiatives, many prominent scholars, politicians, and practitioners of law winced noticeably.⁹ They saw direct democracy as a threat to stable governance in the States; as one attorney framed it, the problem was that, henceforth, “[a]ny malcontent could initiate an amendment to the [state] constitution.”¹⁰ In response to the backlash, the same state governments that had given soon began to

5. Julian N. Eule, *Judicial Review of Direct Democracy*, 99 YALE L.J. 1503, 1549 (1990) (citing processual concerns that suggest judiciary must play some role in reviewing direct democracy measures).

6. JEAN-JACQUES ROUSSEAU, *THE SOCIAL CONTRACT* 114 (Maurice Cranston trans., Penguin Books 1963) (1762).

7. See Gary Thompson & Paul Wilkinson, *Set the Default to Open: Plessy’s Meaning in the Twenty-First Century and How Technology Puts the Individual Back at the Center of Life, Liberty and Government*, 14 TEX. REV. L. & POL. 48, 73 (2009) (examining balance between citizen and government); Eule, *supra* note 5, at 1548–49.

8. See Kenneth P. Miller, *Constraining Populism: The Real Challenge of Initiative Reform*, 41 SANTA CLARA L. REV. 1037, 1039–45 (2001) (discussing resistance to Progressives and Populists in attempting to initiate direct democracy in California). See generally Henry M. Campbell, *The Initiative and Referendum*, 10 MICH. L. REV. 427 (1912).

9. See, e.g., Campbell, *supra* note 8, at 427 (exemplifying one strand of academic reaction to direct democracy).

10. *Id.* at 430.

2011] JUDICIAL APPLICATION OF THE SINGLE SUBJECT RULE 827

take away political power from the masses—or, at the least, to make it more difficult for individual malcontents to undermine the established political order. State legislators and executives devised many restrictions: annual or biennial ballots,¹¹ signature-gathering requirements,¹² restrictions on signature gatherers,¹³ short timeframes for obtaining signatures,¹⁴ executive review of petition qualification,¹⁵ executive preparation of impact statements whose data may be difficult to refute,¹⁶ and the like. But the more critical development—and the aspect that forms this Note’s focus—is the fact that politically agnostic courts began entering the “political thicket,”¹⁷ not merely to enforce the other branches’ rules but also, arguably, to fashion rules of their own.

A. *The emergence of courts as referees of direct democracy*

Before the crackdown on popular-democracy power across the states intensified, the state judiciary typically performed its review of ballot measures at procedural margins: it determined whether petition processes were followed appropriately, whether executive officials abused discretion in invalidating petitions or in allowing them to move forward, and whether petition titles and explanations prepared by the state were accurate.¹⁸ The majority of petitions appeared to survive such deferential scrutiny. But such deference was, in the case of many states, short-lived.

As if sensing that courts were on their side, and seizing popular dissatisfaction with state legislatures, the malcontents about whom we were warned eventually built enough popular momentum through patchworked deals to wage an attack on legislative agendas across several states. From the “ham and eggs”¹⁹ California Bill of Rights, a logrolled initiative that comprised subjects as diverse as “pensions, taxes, right to vote for Indians, gambling, oleomarga-

11. *See, e.g.*, Miller, *supra* note 8, at 1068.

12. *Id.* at 1061–63.

13. *Buckley v. Am. Constitutional Law Found., Inc.*, 525 U.S. 182, 187–89 (1999).

14. *Id.* at 188.

15. Miller, *supra* note 8, at 1071–72.

16. *See Developments in State Constitutional Law: 2001*, 33 RUTGERS L.J. 1593, 1602 (2002).

17. “Courts ought not to enter [the] political thicket.” *Colegrove v. Green*, 328 U.S. 549, 556 (1946) (Frankfurter, J., plurality opinion).

18. Without a single subject rule to prompt review of an initiative’s substance, courts were likely limited to reviewing other less substantial aspects of a petition.

19. FRANK A. PINNER ET AL., *OLD AGE AND POLITICAL BEHAVIOR: A CASE STUDY 4* (Brian Quinn ed., Arno Press 1980) (1959).

rine, the health professions, reapportionment of the State Senate, fish and game, and surface mining,"²⁰ to the Taxpayer Bill of Rights (TABOR) in Colorado,²¹ citizens across initiative states made their voices heard, often preserving them in state constitutions. In one sense, this popular empowerment was a sign of a vibrant democracy. But in a deeper sense, the movement was disruptive to the stable structure of state government; not as disruptive as the Dorr Rebellion was for Rhode Island,²² but an unsettling rattling nonetheless. One by one, legislatures realized that the power of popular initiative needed stronger checks, and state constitutions were amended to require petitions to conform to a single subject.

The single subject rule may be a descendant of the Roman prohibition on omnibus legislation,²³ and is used in many states to formalistically limit the scope and scale of legislative bills.²⁴ Expansion of the rule to initiative petitions forced—or, perhaps, enabled—courts to play a more active role in the battle between popular will as expressed through initiatives on the one hand and popular will as expressed through the people's elected representatives on the other. In applying the new rule, courts began to examine the content of ballot initiatives.²⁵ And thus the brief era of near-plenary popular power came to a close, and the age of channeled policy began.

The troubling part was that the channeling was being performed by the judiciary, an institution outwardly dedicated to fairness and partisan agnosticism.²⁶ Moreover, the rule's application was often clothed in objectivity, despite appearances that subjective judgments were at work, causing legal scholars to question judicial

20. Daniel H. Lowenstein, *California Initiatives and the Single Subject Rule*, 30 UCLA L. REV. 936, 949–50 (1983).

21. See COLO. CONST. art. X, § 20.

22. For a detailed history of the Dorr Rebellion, see generally MARVIN E. GETTLEMAN, *THE DORR REBELLION: A STUDY IN AMERICAN RADICALISM, 1833–1849* (1973).

23. ALLAN CHESTER JOHNSON ET AL., *ANCIENT ROMAN STATUTES* 45–46 n.37 (Clyde Pharr ed., 2003).

24. See Michael D. Gilbert, *Single Subject Rules and Public Choice Theory* 1–4 (bepress Legal Series Paper 816, 2005), available at <http://law.bepress.com/expresso/eps/816/> (listing cases in which single subject challenges have resulted in limiting potential ballot initiatives).

25. E.g., *infra* Part II.

26. Jonathan H. Steinberg, *Congressional Redistricting, Served Two Ways*, 6 ELECTION L.J. 322, 328 n.44 (2007) (book review).

2011] JUDICIAL APPLICATION OF THE SINGLE SUBJECT RULE 829

transparency if not judicial integrity.²⁷ The single subject rule is, like most legal rules, seemingly neutral on its face:²⁸ an initiative that contains multiple unrelated themes cannot stand.²⁹ However, critics allege that the term “single subject” is ambiguous and that courts have full reign to strike down measures they dislike.³⁰ Some frame the problem as being only “aggressive” application by courts, along with the resulting trample upon popular sovereignty.³¹ But aggressiveness can hardly be a problem for the courts if, in fact, the single subject rule was meant to be a check on popular power. As long as the aggressive rule is applied in a transparent, consistent fashion by the judiciary, its consistent outcomes will be visible to the people for what they are, and the rule can always be changed by the people through initiative if it fails to meet the majority’s need. If left to stand, an aggressive rule can create a predictable, albeit narrow, mold within which petition sponsors can channel their measures with confidence. In other words, to those whose main concern is judicial transparency and integrity, aggressiveness is not the problem: inconsistency or political subjectivity is.

B. *Charting the course of in-depth research*

And so begins the first task of this Note, accomplished in Part II: to analyze whether the single subject rule has led to a clear, consistent, and predictable jurisprudence, or if it has turned courts into judicial vetogates³² for initiatives with certain subject matter. The Note thus analyzes the entire fifteen-year history of single subject initiative jurisprudence in one state: Colorado.³³ The re-

27. John G. Matsusaka & Richard L. Hasen, *Aggressive Enforcement of the Single Subject Rule* 2–3 (Loyola–LA Legal Studies Paper No. 2010–4), available at <http://ssrn.com/abstract=1549824>.

28. See Lowenstein, *supra* note 20, at 938–42 (laying out schema to determine what constitutes a “subject”).

29. *Id.*

30. Gilbert, *supra* note 24, at 4.

31. Daniel H. Lowenstein, *Initiatives and the New Single Subject Rule*, 1 ELECTION L.J. 35, 36–39 (2002).

32. The use of the term “judicial vetogate” is an application of William Eskridge’s concept of “legislative vetogates” to the judicial sphere. William N. Eskridge, Jr., *Vetogates, Chevron, Preemption*, 83 NOTRE DAME L. REV. 1441 (2008).

33. Colorado was selected on account of several factors. The state adopted a single subject rule for petitions in 1994, COLO. REV. STAT. § 1-40-106.5(1)(b) (2010), relatively recently compared to other states. This makes for a body of jurisprudence that is both sufficiently meaningful and reasonably sized for a thorough analysis. Moreover, the measure that extended single subject to initiatives received broad popular support—a 65% approval rate. *Ballot History, Year 1994, Ballot Number: A*, COLO. GEN ASSEMBLY, <http://www.leg.state.co.us/lcs/ballothistory.nsf/> (last

search³⁴ concludes that, at least in Colorado, the single subject rule has become a judicial vetogate for initiative petitions.³⁵ The outcome in a single subject challenge appears to be more a function of the petition's substantive content than an impartial application of a robust jurisprudential framework. While outcomes may have served the common good, the Note argues that the absence of a well-defined interpretive framework opens courts to charges of discretionary power plays or, worse, political partisanship—charges that any court whose claim to legitimacy is its objectivity must dispel.

Part III accomplishes the second task: offering suggestions for reshaping the single subject interpretive framework so that it is more transparent and more closely tailored to the practical interests that animate the rule. The Note asserts that the single subject rule is a countermajoritarian check aimed to limit the scale of the burden imposed on a state and its citizens via initiatives, not a drafting device meant to prevent logrolling or fraud on the voters. Accordingly, the key proposal is adoption of a tiered-scrutiny model that begins with a deferential stance toward initiatives and increases the level of scrutiny proportional to the level of burden that an initiative aims to foist upon the state's traditional powers as sovereign or upon the liberties of its citizens. It should be noted that Part III's suggestions are applicable to any state that allows direct democracy, not just to Colorado.

II.

SINGLE SUBJECT APPLICATION IN COLORADO

A. *Brief primer on the initiative petition process in Colorado*

Different devices exist for giving citizens a voice in the passage of statutes and constitutional amendments: initiative, citizen-initi-

visited Dec. 30, 2010). Third, Colorado is one of only four single subject states that does not carve out specific subjects for special treatment, nor are any subjects exempt from the single subject rule; therefore, the analysis is not tainted by unnecessary complications. COLO. CONST. art. V § 1(5.5); NAT'L CONFERENCE OF STATE LEGISLATURES, INITIATIVE AND REFERENDUM IN THE 21ST CENTURY: FINAL REPORT AND RECOMMENDATIONS OF THE NCSL I&R TASK FORCE 17–19 (2002), http://www.ncsl.org/Documents/legismgt/irtaskfc/IandR_Report.pdf. Finally, Colorado's single subject jurisprudence appears to be relatively unstructured, thereby allowing for the multiple-hypothesis testing that is a highlight of this Note.

34. This primary research has been performed by the Author, and involved reviewing every case featuring single subject challenges to initiative petitions. The research statistics are summarized in the Appendix, and key cases are discussed in Part II.

35. *See infra* Part II.

2011] JUDICIAL APPLICATION OF THE SINGLE SUBJECT RULE 831

ated referendum, and legislative-initiated referendum.³⁶ The initiative is a means for citizens to pass statutes or changes to their state constitutions by putting proposals on the ballot. If a majority of voters accept the initiative, the proposed measure becomes law. Initiatives are of two types: direct and indirect.³⁷ Since Colorado law only permits the former, the term initiative in this Note refers only to direct initiative. Moreover, Colorado's single subject requirement extends only to initiatives, thereby obviating the need to examine referenda.³⁸

Citizens must follow a set of time-sensitive steps in order put an initiative petition³⁹ on the state ballot.⁴⁰ Some steps will be summarized herein because they are instrumental to subsequent discussion, and other steps are omitted for clarity. As a starting point, the proponent (i.e., petition sponsor) submits the petition's text to two government departments: the Legislative Council and the Office of Legislative Legal Services.⁴¹ These departments consult with any other state executive agency they deem appropriate, and prepare comments regarding the substantive content, drafting style, and formatting of the initiative.⁴² The heads of the above-mentioned departments deliver their comments to the petition proponent at a public meeting.⁴³ Upon revising the petition to conform to said comments, the proponent submits the revised text to the Secretary of State.⁴⁴ The Secretary then forwards the petition draft to the Title Board (Board), comprising the Secretary, the Attorney General, and a representative from the Office of Legislative Legal Services.⁴⁵

36. The political recall, whereby citizens can remove public officials from office, is also a direct-democracy device; however, it does not involve the passage of laws and therefore is not germane to this Note's discussion.

37. A direct initiative is put before the people, whereas an indirect initiative is submitted to legislators, who choose whether to pass the law themselves or to put the measure on the popular ballot.

38. As an informational note, the citizen-initiated referendum is a means for citizens to require that the legislature put a proposed bill up for vote to the public before the bill can become law. By contrast, the legislative-initiated referendum involves a legislature voluntarily submitting a bill to a public vote; some constitutions go so far as requiring the legislature to do so for certain types of legislation.

39. The terms "petition" and "measure" will be used interchangeably with "initiative" in this Note, though they can be used to describe referenda as well, in direct-democracy parlance.

40. Colorado law allows initiatives at the local level as well, but the focus of this Note is on statewide petitions.

41. COLO. REV. STAT. § 1-40-105(1) (2010).

42. *Id.*

43. *Id.*

44. *Id.* § 1-40-105(4).

45. *Id.* § 1-40-106(1).

The Board takes one of two actions: (1) it approves the petition and creates—or sets—a title and submission clause⁴⁶ for the petition; or (2) it refuses to set the title and submission clause because the petition is legally deficient, such as would occur if the Board deems the petition to contain multiple subjects, misleading or unclear language, or other deficiencies.⁴⁷ The Board's decision is communicated in a public meeting, where opponents of the petition are likely in attendance.⁴⁸ Once the Board approves the petition by setting the title and submission clause, the proponent must gather sufficient signatures to qualify the petition for placement on the ballot.⁴⁹ The Colorado Constitution calls for signatures from a minimum of 5% of the number of citizens who voted in the last election for Secretary of State.⁵⁰ Once the petition qualifies, it is placed on the ballot and must be proclaimed by the governor as law if more than 50% of the votes cast approve it; the governor has no veto power over initiatives.⁵¹

As noted above, the Board announces its decision in a public meeting. This allows those who disagree with the Board's decision—be they citizens who oppose the initiative (in case of approval), or the petition proponent (in case of refusal)—to object before the Board.⁵² If the Board does not change its decision, the objectors may request rehearing by the Board and later file for judicial review before the Supreme Court of Colorado.⁵³ If an objector seeks judicial review, work on the petition—including the gathering of signatures—is suspended until the court makes a ruling as to the validity of the Board's action.⁵⁴ This essentially sounds the death knell for most initiatives because, even when the court finds in favor of the Board, the time it takes to hear the case and render a decision diminishes the proponent's remaining time to qualify the petition.⁵⁵ The inference is that few challenged petitions qualify for the ballot in the year in which they are challenged.

46. The submission clause is the prefatory text that introduces the initiative to voters.

47. COLO. REV. STAT. § 1-40-106(1), (3)(b); § 1-40-106.5(3).

48. *Id.* § 1-40-106(1), (3)(b); § 1-40-107.

49. *Id.* § 1-40-109.

50. COLO. CONST. art. V, § 1(2).

51. *Id.* art. V, § 1(4).

52. COLO. REV. STAT. § 1-40-106(1), (3)(b); § 1-40-107.

53. *Id.* § 1-40-107.

54. *Id.* § 1-40-107(4).

55. Answer Brief of Respondents at 3–4, *In re* Proposed Initiative for 1997–98 No. 30, 959 P.2d 822 (Colo. 1998) (No. 97SA319).

2011] JUDICIAL APPLICATION OF THE SINGLE SUBJECT RULE 833

B. Introducing the single subject challenge

Between the introduction of the single subject rule in 1994⁵⁶ and December 31, 2009, 104 petitions were brought before the Colorado Supreme Court under challenges related to petition language, title language, single subject violations, and related issues.⁵⁷ Eighty-five percent of the petitions were challenged on single subject grounds.⁵⁸ Of the 88 single subject challenges before it, the court found a rule violation 56% of the time.⁵⁹ These statistics raise two significant questions: (1) “Why are so many petitions being challenged in court?”; and (2) “Why does the court invalidate a significant percentage of the petitions?”

This part of the Note evaluates four hypotheses to determine an answer to both questions.⁶⁰ The first hypothesis postulates that a significant number of initiative proponents and opponents misunderstand the law with regard to what the single subject rule requires. The second alternative is that the law is clear, but a significant number of initiative proponents attempt to pull the wool over voters’ eyes through cunning drafting and are caught in time by overeager watchdogs. The third explanation is that the single subject rule has no real legal meaning: it is simply a delay tactic used by opponents to knock out a petition from a given year’s ballot, and the court uses a randomizing strategy to decide which initiatives win the lottery. Finally, it is also possible that, rather than using a process of randomization, the courts have become subject-matter vetogates, rejecting initiatives based on their substantive content.

A fifth distinct hypothesis may also exist, namely that initiative sponsors know the law, yet draft non-compliant initiatives, fully expecting that the measures will not withstand review. Such a possibility can be dealt with swiftly, on account of the presence or absence of two key telltale signs: invalidation of the measure by the Title Board (the first line of defense for invalidation), and an upholding of the Board’s decision by the court with little or no commentary or analysis. Out of all single subject challenges, none of them fit this

56. COLO. REV. STAT. § 1-40-106.5(1)(b).

57. Some cases consolidated complaints related to separate petitions. Research reveals a total of 72 separate Colorado Supreme Court opinions.

58. Many of the other challenges, 68%, also raised other issues in addition to single subject violations.

59. *See infra* Appendix.

60. While hybrid explanations may exist, this Note will analyze each hypothesis independently. Readers may draw their own conclusions about whether a combination of factors provide a better explanation for the statistics.

pattern. While there are three petitions where the court summarily upheld the decision of the Board without giving a formal legal analysis (i.e., without opinion), the Board had approved—rather than invalidated—the measure in every case.⁶¹ As such, the fifth hypothesis appears to be a mirage. We are left, then, with evaluating the first four.

C. *Hypothesis #1:*

Many initiative proponents and opponents do not understand the law

Misunderstanding the law can result from two causes: (1) the law itself is not clear; or (2) proponents and opponents are not able to grasp the law as defined.⁶² The second cause can be eliminated as a strong contender, at least in Colorado. Even if initiative proponents are not sufficiently skilled in the arcane practices of legal interpretation, there is another actor who has the adequate skills to grasp the law: the Title Board.⁶³ As mentioned earlier, the Board consists of the Secretary of State, the Attorney General, and a representative from the Office of Legislative Legal Services—which means that at least two, if not all three, Board members have routine engagement with election law and state law in general. Moreover, even if initiative opponents are not sufficiently skilled and simply challenge petitions in court just in case, the court would defer to the Board a significant percentage of the time. Given that the court disagrees with the Board's legal decision, this can only mean—at least in the context of this hypothesis—that the law itself is unclear. I test this assertion by analyzing individual single subject cases and trying to tease out clear rules of decision.

In reviewing Colorado's case law, certain tests appear as a running thread throughout judicial opinions: "single purpose," "central theme," "connection among provisions," "necessary connection between provisions and theme," and the like.⁶⁴ However, many of these tests appear to be only restatements of the phrase "subject," and do not provide a clear, prescriptive model for reaching a legal conclusion. Moreover, the tests appear to be used selectively: one

61. *In re* Proposed Initiative No. 97, 962 P.2d 973 (Colo. 1998) (en banc) (per curiam); *In re* Proposed Initiative No. 112, 962 P.2d 255 (Colo. 1998) (en banc) (per curiam); *In re* Proposed Initiative No. 80, 961 P.2d 1120 (Colo. 1998) (en banc) (per curiam).

62. Without delving into the complex debate about what "law" is, this Note adopts a positivist definition of law: "law," here, is the text of constitutions or statutes, along with the judiciary's explicit rules derived from interpreting said texts.

63. See *supra* Part II.A.

64. E.g., *In re* Proposed Initiative for "Public Rights in Waters II" (*Public Rights in Waters II*), 898 P.2d 1076 (Colo. 1995) (en banc).

2011] JUDICIAL APPLICATION OF THE SINGLE SUBJECT RULE 835

set of tests is applied to one fact pattern but is omitted from another where it would be equally applicable.⁶⁵ This Section aims to rationalize the case law by identifying the different factors or elements that comprise the jurisprudential framework, their application, and their boundaries. Since this framework is not explicitly stated in case law, it is only one means of rationalization; other alternatives may exist.

1. A plain meaning baseline

As with most matters of legal interpretation, we begin with the text of the Colorado Constitution.⁶⁶ Article V sets forth the single subject requirement as follows: “No measure shall be proposed by petition containing more than one subject.”⁶⁷ The dictionary defines “subject” as “material or essential substance.”⁶⁸ Colorado’s statutes, passed at the time of the referendum, further clarify the dual purpose of the rule: (1) to prevent misleading or defrauding voters; and (2) to prevent patchwork legislation, whereby “incongruous subjects” are sown together into the same measure.⁶⁹ The latter practice has two well-known variants: logrolling, whereby measures that each enjoy less-than-majority support are combined into a proposal that garners majority vote; and riding, whereby unattractive features are piled on top of proposals that already enjoy majority support, thereby gaining otherwise unattainable electoral victory.⁷⁰

Accordingly, an initial, plain language definition might read thus: “A measure contains a single subject if the essential substance of the measure represents one theme, and the measure does not join incongruous provisions for purposes of logrolling, riding, or otherwise misleading or defrauding voters.”

The first in a trio of single subject cases, *Suits Against Nongovernmental Employers*,⁷¹ seems in line with the above definition, though the case does not go to great lengths to develop a formalis-

65. *E.g.*, *infra* Part II.D.3.

66. COLO. CONST. art. V § 1(5.5); *see* Akhil Reed Amar, *The Second Amendment: A Case Study in Constitutional Interpretation*, 2001 UTAH L. REV. 889, 901 (2001) (noting, in another context, that methodologically “[p]roper constitutional interpretation begins with the text of the Constitution itself”).

67. COLO. CONST. art. V § 1(5.5).

68. WEBSTER’S THIRD NEW INTERNATIONAL DICTIONARY 2275 (1993).

69. COLO. REV. STAT. § 1-40-106.5(1)(e) (2010).

70. Robert D. Cooter & Michael D. Gilbert, *A Theory of Direct Democracy and the Single Subject Rule*, 110 COLUM. L. REV. 687, 706–09 (2010).

71. *In re* Proposed Initiated Constitutional Amendment “Concerning Suits Against Nongovernmental Employers Who Knowingly and Recklessly Maintain An

tic rule and deals with the single subject test only in an indirect manner. The initiative at issue in the case allowed employees to sue employers who “knowingly or recklessly maintain[ed] an unsafe work environment.”⁷² The measure also had a remittitur provision, whereby awards in such lawsuits were reduced by workers’ compensation benefits received by the plaintiff.⁷³ The challenge against the initiative alleged that the title of the measure did not comply with single subject requirements.⁷⁴ After a verbatim restatement of constitutional and statutory provisions governing the single subject rule, the court concluded that the measure’s title met single subject requirements because, inter alia, it expressed the “central features” and “true intent” of the measure.⁷⁵ No discussion was entertained about why the “central features” were deemed to comprise only a single subject, or whether the “true intent” was in any way drafted in a misleading fashion.⁷⁶ Such succinct yet opaque rationale was, perhaps, a foreshadowing of things to come. This said, at least one useful jurisprudential outcome did result from this decision: the court adjured the Title Board to apply to initiatives the single subject jurisprudence that the court had developed for legislative bills.⁷⁷ The outcome in the case implies that the Board had done so successfully here.⁷⁸

2. The necessary connection test: disaggregating potentially interdependent provisions

(a) Public Rights in Waters II

The second single subject challenge, decided on the same day as the foregoing case, clarified matters to some extent but also raised important analytical questions. The case was called *In re Proposed Initiative for “Public Rights in Waters II” (Public Rights in Waters II)*.⁷⁹ The Colorado Supreme Court dealt with the single subject

Unsafe Work Environment” (*Suits Against Nongovernmental Employers*), 898 P.2d 1071 (Colo. 1995) (en banc).

72. *Id.* at 1075. At the time of the petition, suits were prevented by the state’s Workers’ Compensation laws.

73. *Id.* at 1074.

74. *Id.*

75. *Id.*

76. *Id.* at 1074–75.

77. *Id.* at 1074.

78. Colorado’s single subject jurisprudence with respect to legislative bills may merit a separate, independent Note. This Note eschews discussion of the legislative dimension. A fruitful discussion occurs in Gilbert, *supra* note 24.

79. *In re Proposed Initiative for “Public Rights in Waters II” (Public Rights in Waters II)*, 898 P.2d 1076 (Colo. 1995) (en banc).

2011] JUDICIAL APPLICATION OF THE SINGLE SUBJECT RULE 837

rule head on, finding multiple subjects in what appeared, at least to the Title Board, to be related provisions.⁸⁰

The case involved a petition that established a public trust for natural streams (a provision sympathetic to the public interest) and revised the election scheme for water-conservation and water-conservancy districts at the local level to make them more accountable to the people (and thereby more likely to implement a public trust doctrine that favored voters).⁸¹ The court, borrowing from its single subject jurisprudence on legislative bills, declared that an initiative contains multiple subjects when it has “at least two distinct and separate purposes which are not dependent upon or connected with each other.”⁸² The court held that the present petition consisted of two purposes: (1) the establishment of a statewide public trust and (2) “election and boundary rules for water conservancy and conservation districts.”⁸³ The fact that there were two purposes did not appear to be dispositive, however; the fatal flaw was that the two purposes were disconnected.⁸⁴ Since the local districts had “little or no power over the . . . development of a *statewide* public trust doctrine,” and since public trust powers would have vested in the state (not the local districts), the two provisions were not “*necessarily* connected”; therefore, they were deemed to be separate subjects.⁸⁵

To better understand the extent to which the connection between local districts and the state was not necessarily present, a brief summary of the responsibilities of each is in order. A public trust would require the state to ensure that non-navigable waters that belong to the trust are used for the benefit of the public rather than solely for private interests.⁸⁶ Public interest concerns could well include conservationist considerations, such as preserving water levels for generations, protecting fish and wildlife species inhabiting the water, and other similar concerns.⁸⁷ They could also involve recreational use considerations, such as allowing citizens to fish or swim in non-navigable rivers and streams, even if these streams flow across private lands—as long as citizens’ entry point onto the stream does not involve trespass.⁸⁸ The Colorado Water Conservation Board is

80. *Id.*

81. *Id.* at 1077 n.1.

82. *Id.* at 1078–79.

83. *Id.* at 1080.

84. *See id.*

85. *Id.* (emphasis added).

86. *See* JOHN W. JOHNSON, UNITED STATES WATER LAW: AN INTRODUCTION, 14–15 (2009).

87. *See id.*

88. *See id.*

the state's main agency for managing water, water uses, and water-related development.⁸⁹ Conservancy and conservation districts also have a part to play in managing water levels, though not much involvement—other than, perhaps, educational—in regulating recreational use.⁹⁰ Water conservancy districts are subunits of the state government tasked with providing a water supply for their respective service areas, including buying/selling water and building “water resource projects.”⁹¹ In contrast, a water conservation district is part policy organization, part planning group, and part special projects entity that helps plan apportionment of natural water streams in the state.⁹²

To the layperson, it may appear that once the Water Conservation Board establishes policies declaring a public trust for a set of streams, water conservation districts would plan apportionment such that more water is conserved, and conservancy districts would limit the amount of irrigation water they supply to agricultural entities. The extent to which these local districts would go to implement trust policy depends on the directors who manage the districts—directors who are nominated by local judges.⁹³ If the directors were instead elected by the people—which is the change sought by the present initiative—the districts might well adopt more public-interest-leaning policies. There is always the possibility, of course, that business interests could capture a majority of elected directors, and that election would not necessarily result in more public-friendly policies. In fact, elections may well have the opposite effect, whereby local districts could fight the public-trust state policy through their captured elected directors. Accordingly, changing the electoral scheme does not necessarily implement the state's public trust policy.

Though logical, the above explanation is not only too subtle, but also puts the court in the position of evaluating social dynamics and predicting what is more or less likely to happen. Moreover, the fact that the court seemingly places the disconnectedness of this measure on par with that of a legislative statute which contained 46 separate provisions—provisions as disparate as creating a “commis-

89. *Mission & Strategic Plan*, COLO. WATER CONSERVATION BD., <http://cwcb.state.co.us/Home/Mission/> (last visited Dec. 31, 2010).

90. See *Background*, SAN JUAN WATER CONSERVANCY DIST., <http://www.sjwcd.org/Background.html> (last visited Dec. 31, 2009); *About Us*, COLO. RIVER DIST., http://www.crwcd.org/page_1 (last visited Dec. 31, 2010).

91. *E.g.*, SAN JUAN WATER CONSERVANCY DIST., *supra* note 90.

92. See COLO. RIVER DIST., *supra* note 90.

93. Ed Quillen, Editorial, *Public Vote a 'Trend' for Water Districts*, DENVER POST, Apr. 1, 2001, <http://extras.denverpost.com/opinion/persp0401b.htm>.

2011] JUDICIAL APPLICATION OF THE SINGLE SUBJECT RULE 839

sion on information management in the department of administration,” charging inmates for medical visits, reducing contributions by public employers to retirement funds, and the like⁹⁴—reinforces the idea that subtle or potential disconnectedness is not what the court had in mind.

An alternative explanation is that changes to the director-election mechanism can stand independently of a public trust doctrine: neither provision is required in order to implement the other, even though one may enhance the other’s effectiveness. This interpretation could well address logrolling concerns—concerns that were highlighted by the court as animating the passage of the single subject rule.⁹⁵ However, for logrolling to be a possibility, there generally must exist two or more distinct camps of voters, neither of which would vote for the other camp’s proposal if submitted standalone.⁹⁶ In this petition, one camp might be election reformers whose sole interest is to revise the selection scheme for district directors from judge-appointed to voter-elected. The other camp could be environmentalists who care about aspirational ideas such as public resource trusts, but are not moved by the political means of implementing them.⁹⁷ The court did not speculate about the existence of multiple interest groups, nor did it make a finding that two or more distinct voting blocs exist—though such a finding is not required since a single subject is an issue of law, not of fact. Perhaps we must presume that the existence of multiple logical subjects necessarily implies multiple distinct voting blocs. Or perhaps the single subject rule is a prophylactic device that strikes down anything that raises even the remotest possibility of logrolling.

Another possibility is that the petition was not really about a strong public trust after all: public trust provisions could have simply been a rider on top of a more popular call for reforming district elections. (The reverse could also have been the case.) Unlike the rather lengthy logrolling discussion, however, there was not even one mention of riding. As such, we may be safe in assuming that riding was not a material consideration for the court.

(b) Petition Procedures

The necessary connection test introduced in *Public Rights in Waters II* left several important questions open. How independent

94. *In re Proposed Initiative for “Public Rights in Waters II (Public Rights in Waters II)*, 898 P.2d 1076, 1079 (Colo. 1995) (en banc).

95. *Id.* at 1079.

96. See *infra* Part III.A.2 (discussing aspects of logrolling in more detail).

97. *Id.*

must the provisions be in order for a court to determine that they are not necessarily connected? And how would the court treat independent provisions that are nonetheless guaranteed to enhance each other (there was no such guarantee for the elections provision in *Public Rights in Waters II*)? Some answers are provided in a companion case, *In re Proposed Initiative on Petition Procedures (Petition Procedures)*,⁹⁸ decided on the same day.

In *Petition Procedures*, the court struck down a measure that changed the processes involved in initiatives, citizen-initiated referenda, and recall petitions.⁹⁹ The fact that the provisions were all aimed at strengthening the staying power of direct democracy (i.e., citizen-initiated petition measures) was not sufficient to satisfy single subject requirements. The court held that the measure comprised multiple subjects, including the “retroactive creation of substantive fundamental rights” in the matters contained within constitutional petitions,¹⁰⁰ the requirement that courts treat the term “shall” as a “mandatory command” in petition cases,¹⁰¹ the requirement that challenges to initiatives on single subject grounds be entertained only if a unanimous court deems “beyond a reasonable doubt” that the petition contains multiple subjects,¹⁰² and others.¹⁰³ The court also noted that the three mechanisms of direct democracy—initiative, referendum, and recall—are addressed by separate sections of the state constitution, implying that “placement” may have influenced its determination.¹⁰⁴

One could well interpret the court as saying that two or more purposes are not necessarily connected if at least one purpose does not necessarily require one or more of the others. For example, unlike *Suits Against Nongovernmental Employers*—where the remittur provision necessarily required a right to sue, a right provided only by the petition’s suit provision—each of the rights and commands in this case can stand on their own, without requiring the existence of the others. Besides clarifying the “necessary connection” test, this interpretation also answers the two questions left open in *Public Rights in Waters II*. Unlike *Public Rights in Waters II*, where multiple voting blocs could theoretically be imagined, it is highly unlikely

98. *In re Proposed Initiative on Petition Procedures (Petition Procedures)*, 900 P.2d 104 (Colo. 1995) (en banc).

99. *Id.* at 109–11.

100. *Id.* at 109.

101. *Id.*

102. *Id.*

103. *Id.* at 109–11.

104. *Id.*

2011] JUDICIAL APPLICATION OF THE SINGLE SUBJECT RULE 841

that separate voter blocs exist for each provision of the present petition; the same people who want to see one provision pass would want the others to pass as well.¹⁰⁵ Moreover, if one of the purposes of the present petition is to strengthen direct democracy, each provision is guaranteed to do so, even if implemented as part of another purpose—unlike the risk from *Public Rights in Waters II*, whereby electing directors could backfire on implementing an aggressive public trust. These observations suggest refining the definition of “necessarily connected” as follows: (1) exclude any consideration as to whether logrolling could have occurred, and focus strictly on textual analysis; and (2) do not evaluate whether fulfilling one purpose is guaranteed to further the goals of another: as long as two or more purposes’ provisions can stand independently, they represent different subjects.

Notwithstanding the above, *Petition Procedures* is also amenable to another interpretation: namely that two or more purposes are not necessarily connected if their provisions relate to multiple objects. Two variants of this exist, the first of which is an object-dependent subject determination. Since the present petition affects the procedures (including rights, processes, and judicial interpretations) for initiatives, referendum, and recall, we may say that each set of procedural changes—even if syntactically similar—represent three distinct logical processes: procedures for initiative, procedures for referendum, and procedures for recall. Each set of processes can stand by itself, and there are rational reasons for allowing initiative processes to be different from referendum or from recall processes. In referendum cases, for example, the bill to be voted on has been drafted by expert lawmakers, and underwent lengthy deliberations (unlike initiatives). In recall petitions, as another example, there are no textual subtleties or other drafting concerns that can mislead voters.

The second variant of the multiple-object relationship is the single object test.¹⁰⁶ To summarize, such a test would deem any measure that acts on multiple objects as violating the single subject rule when independent purposes are at play. Applying this test here, we see three purposes: changing initiative procedures, changing referendum procedures, and changing recall procedures. Since each purpose can be fulfilled without requiring the other, three

105. Statistical evidence about voter preferences in this scenario does not exist. Nonetheless, the Author believes that the conclusion is reasonable based on the interrelatedness of the provisions and general knowledge about voter policy preferences with respect to petition procedures.

106. See *infra* Part III.

separate subjects exist. There may well be separate voting blocs who only care about initiatives, or referenda, or recalls—which means that logrolling may well have occurred. Which interpretation is more likely correct is answered months later in a case with similar facts: *In re Proposed Initiative on Petitions (Petitions)*.¹⁰⁷

(c) Petitions

Petitions contained a host of procedures, rights, requirements, and commands with regard to initiative and referendum (but not recall) petitions.¹⁰⁸ The court appeared to apply the necessary connection test used to strike the *Petition Procedures* measure, but reached a result opposite to the latter despite similar facts.¹⁰⁹

In reaching its result, the court distinguished between major and minor provisions, declaring that “*minor* provisions necessary to effectuate the purpose of the measure” do not represent different subjects.¹¹⁰ Though *Petitions* included most of the same provisions as *Petition Procedures*, two were missing: the establishment of retroactive fundamental rights, and the requirement that single subject violations be found only if a unanimous court so decides.¹¹¹ Accordingly, one or both of these provisions might well have been major, and therefore subject to the stricter necessarily connected test. We can only surmise.

The above appears to foreclose the object-dependent subject determination theory discussed above with regard to *Petition Procedures*. In other words, since initiative and referendum were involved in *Petitions* as well, the idea that procedures for initiative are logically independent of procedures for referendum or recall may have no merit. Moreover, since objectors did not raise the object argument, and the court hinted that “initiative and referendum . . . are commonly associated with each other and reflect a common interest,”¹¹² we have only adumbrations with regard to the result under a *single-object* test, if such a test were even recognized. For example, would *Petitions* have come out differently had recall also been included? The court’s hint that initiative and referendum “are com-

107. *In re Proposed Initiative on Petitions (Petitions)*, 907 P.2d 586 (Colo. 1995) (en banc).

108. *Id.* at 593–94.

109. *Id.* at 590.

110. *Id.*

111. *Compare Petitions*, 907 P.2d at 593–94, with *In re Proposed Initiative on Petition Procedures (Petition Procedures)*, 900 P.2d 104, 109–11 (Colo. 1995) (en banc).

112. *Petitions*, 907 P.2d at 591 n.3.

2011] JUDICIAL APPLICATION OF THE SINGLE SUBJECT RULE 843

monly associated”¹¹³ implies an affirmative answer. If so, however, the court would have to recognize at least one form of the single object test.

2. Declining, and later adopting, a single object test

A single object test was raised as a possible interpretation in *Petition Procedures*. While *Petitions* may have later foreclosed one variant of such a test (i.e., object-dependent subject), the second was left open by the fact that the court may well have treated initiative and referendum as a composite/single object in said case.¹¹⁴ The ambiguity appeared to be resolved in *In re Amend Tabor No. 32 (Amend Tabor 32)*,¹¹⁵ where the court declined to adopt a single-object test.¹¹⁶

On December 18, 1995, the Colorado Supreme Court decided, in *Amend Tabor 32*, that an initiative which provided a tax credit did not violate the single subject rule, even though the credit pertained to six separate, unrelated state and local taxes.¹¹⁷ Since the initiative involved only one \$60 tax credit, the fact that it could be applied to six disparate and unconnected taxes—such as telecommunications taxes, personal property taxes for business, state income taxes, vehicle taxes, levies spent on courts or social services, and abatements levies—was not enough to cause conceptual issues.¹¹⁸ Unlike the objects that the *Petitions* court noted were “commonly associated with each other” (i.e., initiative and referendum),¹¹⁹ one cannot reasonably argue that the six disparate tax objects were commonly associated in people’s mind—unless the broad subject of “taxes” is a sufficient basis for common association. But if “water” was not a sufficient basis of association in *Public Rights in Waters II*,¹²⁰ one must question why “taxes” would be. Moreover, the taxes here were assessed at either state or local levels, so one could not even reframe the connection as “state taxes” or “local taxes,” even if reframing would have resulted in a more cohesive

113. *Id.*

114. *Id.*

115. *See* 908 P.2d 125 (Colo. 1995) (en banc).

116. *See id.*

117. *Id.* at 129. The initiative also called for the state to transfer funds to local governments in the amount of the revenue shortfall. *Id.* Said provision is not relevant in the analysis of the present petition, but will be mentioned for comparison purposes in subsequent comments.

118. *Id.*

119. *Petitions*, 907 P.2d at 591 n.3.

120. *In re Proposed Initiative for “Public Rights in Waters II” (Public Rights in Waters II)*, 898 P.2d 1076, 1080 (Colo. 1995) (en banc).

connection. Therefore, *Amend Tabor 32* must stand for the propositions that subject and object are not conflated, and that the court declined the invitation to adopt a single object test. Instead, the court applied its necessary connection jurisprudence as refined in *Petitions*: the subject here was “tax credit,” and all of the provisions were necessary, or plainly adapted, to effectuate the purpose of providing said credit.

Notwithstanding the above, one is left to wonder whether the single subject rule thus framed wards off the evils it was intended to dispel, with logrolling¹²¹ being chief among them. Though *Petition Procedures* is clear about the absence of logrolling not being a defense to multiple subjects, it did not address whether a measure can be deemed to contain a single subject even if had been (or could have reasonably been) logrolled. In *Amend Tabor 32*, the possibility exists that proponents logrolled different types of taxes into one measure in order to accumulate votes. For example, suppose there are six groups of voters, each paying \$10 in different types of taxes. They all know that an initiative to reduce their respective tax categories would not win a popular vote. So they band together and enact a tax credit that comprises all of their situations. To sweeten the deal even more and attract voters for whom \$10 would have been too little, they raise the credit from \$10 to \$60. This hypothetical has all of the signs of traditional logrolling, yet the court does not engage in such speculation.

In addition to practical considerations, the court’s decision also raises a logical inconsistency: when two measures call for nearly identical terms but are drafted differently, should they result in different single subject outcomes? For example, what if instead of calling for a \$60 credit for six types of taxes, the measure had called for six \$10 credits—one for each type of tax? Would that constitute a multiple subject situation? It most likely would, under the rule thus developed: each \$10 credit, likely a major provision, can stand on its own and is not necessarily connected with the others. Though there may be one distinction between the two (i.e., the credit for any given tax is higher in the \$60 scenario), is it a distinction without a difference?

The foregoing raises the question as to whether the original intent behind the single subject rule has given way to wooden applications of a linguistic principle. Some of the Justices on the court

121. See *infra* Part III.A.2.

2011] JUDICIAL APPLICATION OF THE SINGLE SUBJECT RULE 845

raised this issue in dissents,¹²² and later the idea received a warm welcome from the majority of the court, as the sections on subject specificity and hidden subjects discuss.¹²³ It also appears that the court realized the difficulty posed by declining to adopt—or at least allow for some semblance of—a single object test. And nearly ten years later, it began to change course.

The court appeared to erase, or at least fuzz, the distinction between subject and object in *In re Proposed Initiative for 2005–2006 No. 74 (Initiative No. 74)*.¹²⁴ The initiative put a ten-year expiration date on new fiscal measures voted on by the people, measures such as higher taxes, new public debt that exceeded a certain threshold, and suspension of spending limits.¹²⁵ (The Taxpayer Bill of Rights requires such fiscal measures to be submitted to popular referendum.) The court held that each type of program to which the sunset provision applied represented a separate subject.¹²⁶ The decision appears to be a reversal of the course charted in *Amend Tabor 32*. Now, the objects on which the subject of the restriction acted (the subject being “sunsetting fiscal measures”) would, if many, serve to deem the initiative as covering multiple subjects—provided that the multiple objects could stand alone.

3. Subject specificity

One issue raised by foregoing cases is the issue of subject specificity: how atomic must a subject be in order to be considered unitary? The issue received a healthy dose of attention from Justice Howard Kirshbaum in a dissent to the opinion *In re Proposed Initiative on Parental Rights (Parental Rights)*¹²⁷ in 1996. The initiative called for recognizing parents’ inalienable rights to “direct and control the upbringing, education, values, and discipline of their children.”¹²⁸ The majority held that the initiative satisfied the single subject requirement because the unifying theme was the right of parents to “direct and control the upbringing, education, values, and discipline of their children.”¹²⁹ Since *Amend Tabor 32* had rejected an object test, the result is in line with precedent at that time.

122. *In re Proposed Initiative on Parental Rights (Parental Rights)*, 913 P.2d 1127, 1137–40 (Colo. 1996) (en banc) (Kirshbaum, J., dissenting).

123. See *infra* Part II.C.4, 6.

124. *In re Proposed Initiative for 2005–2006 No. 74 (Initiative No. 74)*, 136 P.3d 237 (Colo. 2006).

125. *Id.* at 238, 242.

126. *Id.* at 242.

127. *Parental Rights*, 913 P.2d at 1137–40 (Kirshbaum, J., dissenting).

128. *Id.* at 1132–33 (majority opinion).

129. *Id.* at 1131.

However, the apparent grouping of disparate aspects of childrearing into one subject prompts one to inquire as to the granularity that a proper subject should have.

After reframing the subject as the “general relationship of parent and child,”¹³⁰ Justice Kirshbaum contends in his dissent that “parental rights” is too broad of a subject to withstand single subject scrutiny; he would rather disaggregate and treat education, discipline, values, and upbringing as separate subjects.¹³¹ Justice Kirshbaum distinguishes *Petitions* on the ground that “petitions,” albeit aggregating initiated and referred petitions, represent a “clearly defined” subject, whereas broad concepts such as “parental rights” are too vague and confusing for voters.¹³² He argues for a test to address specificity, though a “clearly defined” test hardly offers substantial clarity because of its subjectivity. The majority declined to adopt such a test, but future jurisprudence nonetheless inched in this direction.

Less than two months after *Parental Rights*, the court declared, in *In re Proposed Initiative 1996–4 (Repeal Tabor)*,¹³³ that “[g]rouping the provisions of a proposed initiative under a *broad* concept that potentially misleads voters will not satisfy the single subject requirement.”¹³⁴ The concept in question was the Taxpayer Bill of Rights (TABOR), which encompassed diverse topics such as spending limits; revenue limits; and elections for new taxes, revenue limit adjustments, and spending limit adjustments.¹³⁵ The *Repeal Tabor* holding does not appear to be a great departure from *Public Rights in Waters II*, where a broad concept such as “water” was too broad to be a unifying theme.¹³⁶ But it does represent a new issue for the court to address in light of *Parental Rights*: namely, where to draw the *specificity* line. Justice Kirshbaum’s invitation to adopt a “clearly defined” specificity test (which, incidentally, may have served to uphold the measure since TABOR is clearly defined in article X, section 20 of the Colorado Constitution)¹³⁷ was implicitly rejected in favor of an encompassing principle test:¹³⁸ that an initiative consists of multiple

130. *Id.* at 1137–40 (Kirshbaum, J., dissenting).

131. *Id.*

132. *Id.*

133. *In re Proposed Initiative 1996–4 (Repeal Tabor)*, 916 P.2d 528, 532 (Colo. 1996) (en banc) (emphasis added).

134. *Id.*

135. COLO. CONST. art. X, § 20.

136. *In re Proposed Initiative for “Public Rights in Waters II” (Public Rights in Waters II)*, 898 P.2d 1076, 1080 (Colo. 1995) (en banc).

137. COLO. CONST. art. X, § 20.

138. *Repeal Tabor*, 916 P.2d at 533.

2011] JUDICIAL APPLICATION OF THE SINGLE SUBJECT RULE 847

subjects if its provisions are “disconnected from any *encompassing principle*.”¹³⁹ While “encompassing principle” might simply be a synonym for “subject” rather than a descriptive phrase, it was sufficient for the court to find that the principle, or subject, advanced here was “limiting government spending,” and that it was too general—akin to “water” in *Public Rights in Waters II*—to qualify as a single subject.¹⁴⁰

The weakness of the encompassing principle test is very visible in the controversial case *In re Proposed Initiative for 2005–2006 No. 55 (Initiative 55)*.¹⁴¹ *Initiative 55* restricted the provision of non-emergency services by the state of Colorado to only “citizens of and aliens lawfully present in the United States,” unless federal law required otherwise in specific situations.¹⁴² It appears that the encompassing principle of “government services to unlawful aliens” was unsatisfactory; the court ruled that the category “government services” is too broad of a subject area.¹⁴³ Instead, the category should be disaggregated until one arrives at a “cohesive purpose” for which services, or groups of services, are provided.¹⁴⁴ The court wasted no time disaggregating services into two sub-categories: (1) those “benefit[ing] individual welfare,” achieved by “medical and social services”; and (2) those “facilitat[ing] economic transactions,” achieved by “administrative services.”¹⁴⁵

We are left to wonder why the court stopped at this level of specificity and did not drill further. One possibility is that the court simply chose to stop as soon as it found multiple subjects. Another is that all of Colorado’s government services comprise only two “cohesive purposes” with regard to individuals: medical/social and administrative. If the latter were the case, then a future petition targeted at eliminating “government services that benefit individual welfare for illegal aliens” might pass a single subject challenge, whereas if the former applies then the court might well drill down to the next level of specificity should such a petition come before it. Alas, a clearer test for specificity has not made its way into the jurisprudence.

139. *Id.* (emphasis added).

140. *Id.*

141. *In re Proposed Initiative for 2005–2006 No. 55 (Initiative 55)*, 138 P.3d 273 (Colo. 2006) (en banc).

142. *Id.* at 282.

143. *Id.*

144. *Id.* at 280.

145. *Id.* at 277.

4. The effects test and hidden purposes

The idea that crafty initiative proponents would surreptitiously insert subtle provisions in petitions to implement hidden agendas was identified by the legislature as a motivating factor for the single subject rule.¹⁴⁶ Though the issue had been lurking in the background since *Public Rights in Waters II*, early single subject tests were not able to root out the subtlest of the subtle. For example, the court in *Petitions* allowed a measure that introduced changes to processes for initiative and referendum petitions.¹⁴⁷ The measure also made initiative and referendum mandatory for all districts as a constitutional matter.¹⁴⁸ Prior to *Petitions*, citizens had initiative and referendum constitutional rights only at the level of the state, city, town, and municipality.¹⁴⁹ The hunch that municipality would cover counties is incorrect, as the Colorado Supreme Court confirmed in a separate case unrelated to *Petitions*.¹⁵⁰ Therefore, prior to *Petitions*, there was no constitutional right to initiative at the county level—and yet the generic term “district” (used in the *Petitions* measure) encompassed all types of local government.¹⁵¹ Though the *Petitions* court did not highlight the inconsistency, the measure could well have been aimed at introducing mandatory initiative and referendum to counties in addition to changing existing processes for state, city, town, and municipality petitions. As discussed earlier, the *Petitions* court upheld the measure. Consequently, the problem of hidden purposes remained in the shadows until it emerged two years later in two key cases: *In re Proposed Initiative for 1997–98 No. 30 (Initiative No. 30)*¹⁵² and *In re Proposed Initiative for 1997–98 Nos. 84 & 85 (Initiatives 84 & 85)*.¹⁵³ The device for locating hidden purposes was the effects test, a variant of the object test rejected in *Amend Tabor 32*.

146. See COLO. REV. STAT. § 1-40-106.5(1)(e)(II) (2009) (noting that single subject is intended to, *inter alia*, “prevent surprise . . . from being practiced upon voters”).

147. *In re Proposed Initiative on Petitions (Petitions)*, 907 P.2d 586, 593–94 (Colo. 1995) (en banc).

148. *Id.*

149. COLO. CONST. art. V, § 1(9).

150. See *Board of County Comm’rs v. County Road Users Ass’n*, 11 P.3d 432, 436 (Colo. 2000) (declaring that article V, section 1(9) of the Colorado Constitution “does not include counties, and this court has not recognized any constitutional initiative powers reserved to the people over countywide legislation”).

151. *Petitions*, 907 P.2d at 593.

152. *In re Proposed Initiative for 1997–98 No. 30 (Initiative No. 30)*, 959 P.2d 822 (Colo. 1998) (en banc).

153. *In re Proposed Initiative for 1997–98 Nos. 84 & 85 (Initiatives 84 & 85)*, 961 P.2d 456 (Colo. 1998) (en banc).

2011] JUDICIAL APPLICATION OF THE SINGLE SUBJECT RULE 849

(a) Hidden effects: Initiative 30 and Initiatives 84 & 85

The facts in *Initiative 30* are almost indistinguishable from *Amend Tabor 32*. Just like the latter, *Initiative 30* applied a tax cut to a group of unrelated taxes, and called for the state to transfer funds to local governments in the amount of the revenue shortfall (i.e., to replace revenue).¹⁵⁴ The problem, however, was that a couple of the taxes to which the cut was applied would not benefit from full state replacement of revenue, on account of a subtlety in the measure's language as interpreted in concert with existing state law.¹⁵⁵ The context is as follows. The Taxpayer Bill of Rights required year-over-year tax-rate increases that exceeded a certain adjustment index to be submitted to voters for approval via referendum measures.¹⁵⁶ The measures were not required to state fixed maximum tax rates or fixed dollar ceilings.¹⁵⁷ When *Initiative 30* came along, it applied certain cuts to certain types of taxes, except taxes that had been increased via measures that listed fixed tax rates and fixed dollar ceilings.¹⁵⁸ Moreover, it provided for state replacement of lost local revenue from said cuts, but only to the extent of the revenue collectible by localities without the need for special referendum measures.¹⁵⁹ Together, the petition's measures served to annul, *inter alia*, many of the local non-standard tax increases from the past several years¹⁶⁰—almost as subtle as *Petitions*, but a subtlety to which the court did not look kindly.

Rather than retaining the measure's subject as "tax cuts"—with the subtle provision representing a bigger-than-expected tax cut—the court concluded that two subjects were present: (1) tax cuts; and (2) imposing new criteria for tax-increase measures that had already been passed by voters.¹⁶¹ This reframing raises the question whether the outcome would have been different if the new criteria had been merely prospective. One argument for the affirmative would point out that there are no prospective limitations per se: local governments can always choose to draft their special tax measures such that they are not affected by this provision. A counterargument would be that strong-arming the government into limiting its future revenue-raising measures through implicit con-

154. *Initiative No. 30*, 959 P.2d at 823.

155. *Id.*

156. COLO. CONST. art. 10 § 20(4), (8).

157. *Id.*

158. *Initiative 30*, 959 P.2d at 826–27.

159. *Id.*

160. *Id.*

161. *See id.* at 826.

straints represents a subterfuge that should be deemed a separate, independent purpose from tax cuts.

The hidden purpose test took a more expanded meaning in a case decided just a couple of months later: *Initiatives 84 & 85*.¹⁶² This pair of measures proposed a set of cuts in state and local taxes on utilities, vehicles, real property, and income;¹⁶³ they also had a standard revenue-replacement provision, just like *Initiative 30* (and others).¹⁶⁴ Unlike *Initiative 30*, there was no hidden effect on account of subtle criteria because the state was required by the initiative's text to replace revenue lost by the locality on account of the tax cuts. There was, however, another secondary effect: the state would likely have to cut spending for state programs in order to make up for localities' shortfall.¹⁶⁵ Per existing TABOR provisions, the state could not simply raise new taxes without obtaining voter approval, a move that was unlikely to succeed given citizens' feelings about taxes at the time.¹⁶⁶ Without raising taxes, the state could not meet its revenue replacement mandate under *Initiatives 84 & 85*, which meant one thing: the state would have to cut funding for state programs.¹⁶⁷ Given this parallel effect, the initiatives were struck down for comprising two subjects: (1) cutting state and local taxes; and (2) cutting state programs to make up for localities' revenue shortfall.¹⁶⁸ Apparently, if left on the ballot, voters would have been confused, all the while thinking that lowering taxes meant that the state could continue to spend as before on all state programs.¹⁶⁹

(b) Direct and logical effects are permitted: Initiative 258(A)

The effects test was somewhat narrowed in 2000, in a case called *In re Proposed Initiative No. 258(A) (Initiative 258(A))*.¹⁷⁰ The measure called for English-only instruction for students in public

162. *In re Proposed Initiative for 1997–98 Nos. 84 & 85*, 961 P.2d 456, 461–63 (Colo. 1998) (en banc).

163. *Id.*

164. *Id.*

165. *Id.* at 460. To raise taxes beyond a standard index that adjusts for inflation and population growth, state law requires voter approval. *Id.*

166. *Id.*

167. *Id.*

168. *Id.*

169. *See id.* at 460–61 (indicating that voters would have been “surprised” to learn that tax cuts actually meant reduction in amount of money spent on state programs).

170. *In re Proposed Initiative for 1999–2000 No. 258(A) (Initiative 258(A))*, 4 P.3d 1094 (Colo. 2000) (en banc).

2011] JUDICIAL APPLICATION OF THE SINGLE SUBJECT RULE 851

schools, with exceptions for bilingual education programs (subject to parental waiver) and foreign-language classes.¹⁷¹ The measure also explicitly stated that schools are not required to provide bilingual programs, thereby leaving open the possibility that some schools would phase out such initiatives.¹⁷² This possibility might seem similar to the effect in *Initiatives 84 & 85*, whereby the state might have to cut some of its spending as part of the measure.¹⁷³ There is a difference, however, between automatic effects and permissive effects, even if both are hidden. In *Initiatives 84 & 85*, the state would have to cut spending if it could not obtain new tax increases, whereas schools in *Initiative 258(A)* did not have to phase out their bilingual programs—they simply had the option to do so. Rather than adopt a permissive effects test as the rationale for upholding the measure, the court formulated a different rule: effects that “follow[] *directly and logically* from the central focus of [the initiative]” are “logical[ly] incident” to the measure, and do not constitute a separate subject.¹⁷⁴ One might well ask whether this rule would have resulted in a different outcome in *Initiatives 84 & 85*.

(c) Uncertain effects: Initiatives 25–27

One scenario that had not been squarely addressed so far is the uncertain effects case, where the court cannot determine the likely effects of a measure. Prior to the outright adoption of an effects test, the court did not appear troubled by measures whose effects were uncertain. For example, in the case *In re Proposed Initiative 1996–6 (Public Rights In Waters III)*,¹⁷⁵ the court upheld the measure calling for a “strong public trust doctrine” despite the fact that the “precise meaning” of such a doctrine—and, implicitly, the precise effects of the doctrine—were not defined by the initiative.¹⁷⁶

Approximately three years later, the court adopted a more rigorous approach vis-à-vis uncertain effects, beginning with *In re Proposed Initiative for 1999–2000 Nos. 25, 26, & 27 (Initiatives 25–27)*.¹⁷⁷

171. *Id.* at 1097.

172. *Id.*

173. *In re Proposed Initiative for 1997–98 No. 30 (Initiative 30)*, 959 P.2d 822, 823, 826–27 (Colo. 1998) (en banc).

174. *Initiative 258(A)*, 4 P.3d at 1097 (emphasis added).

175. *In re Proposed Initiative 1996–6 (Public Rights in Waters III)*, 917 P.2d 1277 (Colo. 1996) (en banc).

176. *Public Rights in Waters III*, 917 P.2d, at 1281.

177. *In re Proposed Initiative 1999–2000 Nos. 25, 26, & 27 (Initiatives 25–27)*, 974 P.2d 458 (Colo. 1999) (en banc).

Here, the court was faced with proposed tax cuts.¹⁷⁸ Like other prior tax cut measures that were deemed to comprise multiple subjects, these initiatives called for state replacement of local revenue.¹⁷⁹ But unlike those measures, the replacement here was conditional: it would only occur after the state revenue exceeded a certain threshold figure, thereby making the actual effect on the state indeterminable *ex ante*.¹⁸⁰ The court blocked the measures and ordered the Title Board to determine the precise effect on state spending.¹⁸¹

If *Initiatives 25–27* does not represent an evolution of the jurisprudence, but rather should be read in concert with *Public Rights in Waters III*, the uncertainty itself is not dispositive: the critical issue is the government's control of the outcome. In *Public Rights in Waters III*, the legislature would have to define—and thereby control—the effects of the public trust doctrine. By contrast, the legislature's hands would be bound in *Initiatives 25–27*; it could not avoid unpleasant effects on state expenditures if revenue exceeded the threshold. Accordingly, we cannot be certain whether uncertain effects necessarily invalidate a measure.

5. Procedure-vs.-substance test

Another subtle distinction that the court has drawn is between provisions that pertain to procedure versus those that relate to substance. Even if both types of provisions relate to a single logical theme that might otherwise be deemed a single subject (e.g., petitions in *Petitions*), the measure is nonetheless deemed to relate to separate subjects if procedure and substance are mixed. The case in point is *In re Proposed Initiative 2001–02 Nos. 43 & 45 (Initiatives 43 & 45)*.¹⁸² *Initiatives 43 & 45* involved changes to initiated and referred petitions.¹⁸³ In addition to the types of procedural changes found acceptable in *Petitions*,¹⁸⁴ *Initiatives 43 & 45* also added two provisions: a presumption that measures which change one section of the constitution are a single subject unless the court had previously ruled otherwise; and a prohibition on referred petitions for

178. *Id.* at 459–60.

179. *Id.*

180. *Id.*

181. *Id.* at 466.

182. *In re Proposed Initiative 2001–02 Nos. 43 & 45 (Initiatives 43 & 45)*, 46 P.3d 438 (Colo. 2002) (en banc).

183. *Id.* at 444–45.

184. *In re Proposed Initiative on Petitions (Petitions)*, 907 P.2d 586, 593–94 (Colo. 1995) (en banc).

2011] JUDICIAL APPLICATION OF THE SINGLE SUBJECT RULE 853

zoning issues.¹⁸⁵ One might expect the measure to sail through, since the subject formulated by the court in *Petitions* was “petitions,”¹⁸⁶ not “petition procedures.” However, the court drew a subject distinction between petition procedures and substantive content.¹⁸⁷ The lesson: since the measure affected the procedure of the initiative process and the substantial law, there were two subjects.

6. Exhortative statements

Initiatives, or parts of initiatives, do not necessarily create binding law: they can also request the legislature to do something—a formal means of petitioning the government to redress grievances. How should measures containing these types of provisions be handled? The court had the opportunity to rule on this issue in *In re Proposed Initiative on Parental Choice in Education (Parental Choice in Education)*.¹⁸⁸ The measure’s goal was to enact a school voucher program, and the majority of provisions were aimed to that end.¹⁸⁹ One provision, however, called on the legislature to improve the quality of public schools, and read thus: “The General Assembly is encouraged to repeal those laws and/or regulations that are determined to be impediments to the ability of public schools to provide a quality of education equal to or greater than that provided in non-public schools.”¹⁹⁰ One could well argue that this provision can stand by itself, and is not necessarily related to implementing a voucher program. However, the provision, even if enacted, does not represent binding law: it is merely an exhortative statement. The General Assembly can well decline the invitation without legal repercussions. Such a statement should not be allowed to invalidate an otherwise-valid petition, especially since it is related to the very reason why some citizens would opt for private schools. Accordingly, the court declined the invitation to strike down the measure.¹⁹¹ Though one might question whether such exhortative statements are a means of soft logrolling, the court does not appear concerned.

185. *Initiatives 43 & 45*, 46 P.3d at 444–45.

186. *Petitions*, 907 P.2d at 590.

187. *Initiatives 43 & 45*, 46 P.3d at 444–45.

188. *In re Proposed Initiative on Parental Choice in Education (Parental Choice in Education)*, 917 P.2d 292 (Colo. 1996) (en banc).

189. *Id.* at 294–95.

190. *Id.* at 295.

191. *Id.*

7. Conclusion

Even if the interpretive layer offered by the foregoing comments appears to create a predictable framework, two further observations may lessen the strength of such a conclusion. First, it appears that the tests are applied selectively rather than consistently. Second, an analysis of the single subject cases reveals that in at least 50% of the petitions, the court announced refinements of the rules applied previously. Accordingly, one may be unsurprised if petition proponents claim difficulty in determining whether their measures will stand up to judicial scrutiny. But even if proponents are not well versed in the law, there is a gatekeeper who is: the Title Board. The fact that even the experienced Title Board has difficulty getting it right is unsettling, even though we acknowledge that reasonable people can disagree over legal interpretation.

The interpretive layer provided in this section could well be off the mark with regard to the court's intended methods for interpreting single subjects. As such, there is a strong probability that "single subject" is a term of art rather than a descriptive phrase. Under the case law, it appears that anything can be reframed as comprising multiple subjects, depending on how fine of a specificity line we draw¹⁹² or on the level of scrutiny with which we examine the effects of a measure.¹⁹³ Since there is no clearly articulable test for specificity or effect, we are left with the solution of recognizing single subjects when we see them. This might work for a limited time while jurisprudence stabilizes, but becomes somewhat disingenuous if the rules keep evolving ad infinitum; after all, the single subject doctrine had been evolving for over one-hundred years with respect to legislative bills before it was merged into the interpretive space for initiatives.¹⁹⁴

Since the purely formalistic avenue for rationalizing single subject jurisprudence is unfulfilling, other hypotheses should be evaluated.

192. See *supra* Part II.C.4.

193. See *supra* Part II.C.5.

194. One of the early cases involving single subject challenges to legislative statutes was *People ex rel. Thomas v. Goddard*, 7 P. 301 (Colo. 1885), a full 104 years before Coloradoans extended single subject to initiatives. The Colorado Supreme Court had merged its legislative single subject jurisprudence into that of the newly-established initiative single subject jurisprudence. See *In re Proposed Initiative "Public Rights In Waters II"* (*Public Rights in Waters II*), 898 P.2d 1076, 1078-79 (Colo. 1995) (en banc).

*D. Hypothesis #2:**The single subject rule as a trap for cunning initiative sponsors*

Rather than see the single subject rule as strictly analytical, perhaps it is a means of ferreting out clever initiatives that try to pull the wool over voters' eyes. In reviewing some of the cases above, one might agree that the initiatives contained subtleties that were rightly caught. There are several difficulties with this hypothesis as an exclusive explanation, however. The most important one is that subtleties in drafting that result in broader-than-expected effects have not consistently caused the court to find that an initiative contained multiple subjects, even after the effects test was adopted in 1998.

Take, for example, *In re Proposed Initiative for 1999–2000 No. 255 (Initiative 255)*, a measure that required background checks at gun shows.¹⁹⁵ The provisions seemed straightforward, yet few voters would realize—as the challengers did—that unlicensed gun sellers would likely stop selling their wares at such events because federal law would work to prevent them from obtaining background checks via the means prescribed by the measure.¹⁹⁶ Even if such an outcome were favored on policy grounds, this would not render the measure's effect any less subtle. The measure was upheld notwithstanding the aforementioned subtlety due to the court's decision to “not interpret [a measure's] language or predict its application if it is adopted.”¹⁹⁷

Another example is *In re Proposed Initiative for 2005–2006 No. 73 (Initiative 73)*, a measure that aimed to stop pay-to-play contributions to tax and debt elections.¹⁹⁸ The measure stated that if a district provides any benefit—including employment, contracts, or other transfer payments in return for services—to an individual who contributed more than \$500 to support a tax and debt election, the tax or debt authorized by the election would be canceled.¹⁹⁹ While the idea of preventing corruption is very appealing, this measure's provisions could, inter alia, prevent existing government employees from making substantive contributions to tax or debt elections. For example, teachers who believe that additional government spending on schools serves the community might not

195. *In re Proposed Initiative for 1999–2000 No. 255 (Initiative 255)*, 4 P.3d 485, 495 (Colo. 2000) (en banc).

196. *Id.*

197. *Id.*

198. *In re Proposed Initiative for 2005–2006 No. 73 (Initiative 73)*, 135 P.3d 736, 739 (Colo. 2006) (en banc).

199. *Id.* at 741–42.

be able to substantially support a tax or debt election to that effect. Moreover, the measure makes the benefit prohibition effective for the entire duration of the tax or debt program passed by the election. For multi-year programs, this might force districts to not hire individuals who contributed to the program's election years earlier. Or, as Justice Gregory Hobbs stated in his dissent, the measure would also "prevent districts from awarding a public contract to the lowest bidder because that individual or entity has contributed more than five hundred dollars to an issue committee."²⁰⁰ These effects are subtle, and yet the court deemed the measure to contain a single subject.²⁰¹

Several other examples exist of uncaught subtlety, but they need not be discussed in detail; this Section is not meant to contain an exhaustive account. Instead, the cases cited above raise sufficient doubts about the notion that the single subject rule is a sieve that catches all cunning initiative drafters. The very fact that some appear to slip through the cracks prompts us to examine other hypotheses.

*E. Hypothesis #3:
Randomization perspective of the single subject rule*

A third hypothesis is that in hard cases, where it is difficult to determine whether a petition contains multiple subjects and precedent is not squarely on point, the court may choose an outcome by using tie-breakers that are not directly related to the merits, thereby resulting in controlled randomization. Such suggestions have been made in other contexts.²⁰² This Note casts doubt on a randomization hypothesis as applied to Colorado's single subject cases by presenting two sets of statistics.²⁰³

200. *Id.* at 742 (Hobbs, J., dissenting).

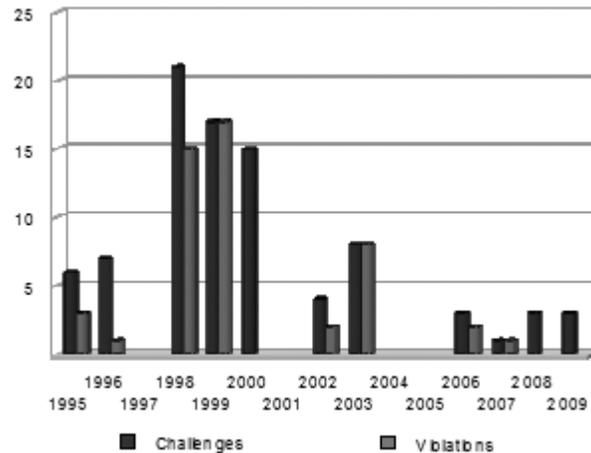
201. *Id.* at 739.

202. Adam M. Samaha, *Originalism's Expiration Date*, 30 *CARDOZO L. REV.* 1295, 1354–63 (raising the possibility, admittedly unorthodox, that originalism as an interpretive basis serves a randomizing function).

203. The statistics were compiled as part of this Note's research.

2011] JUDICIAL APPLICATION OF THE SINGLE SUBJECT RULE 857

Figure 1. Single subject challenges vs violations, 1995-2009



The first is a count of the number of single subject challenges brought since the rule was introduced, compared with the number of violations found; this relationship is depicted in Figure 1. I analyze the graph in light of two assumptions. First, it is assumed that the Title Board²⁰⁴ weeded out initiatives that clearly contained multiple subjects. Therefore, if a petition was left standing, was later challenged in court, and the court did not summarily dismiss the challenge without comment, then the case may fairly be characterized as a hard case that required in-depth analysis from the court.²⁰⁵ Second, if all challenges were hard cases, then a purely random decision theory may be expected to result in a 50% violation rate, whereby 50% of petitions are deemed to violate the single subject rule. Fifty percent is not necessarily an arbitrary figure. Binomial probability distribution theory suggests that, in a random toss of a coin (a two-state variable), heads will result in half of the tosses after a certain number of trials.²⁰⁶ One may assume the principle to result in violations being found in half of the cases during any given year if decisions were reached randomly. That said, the number of single subject challenges brought in a given year do not meet the requirements of a normal statistical sample,²⁰⁷ and there-

204. See *supra* Part II.A.

205. Of all cases challenged on single subject grounds, only four were dismissed summarily without comment; this occurred in 1998. See *infra* Appendix.

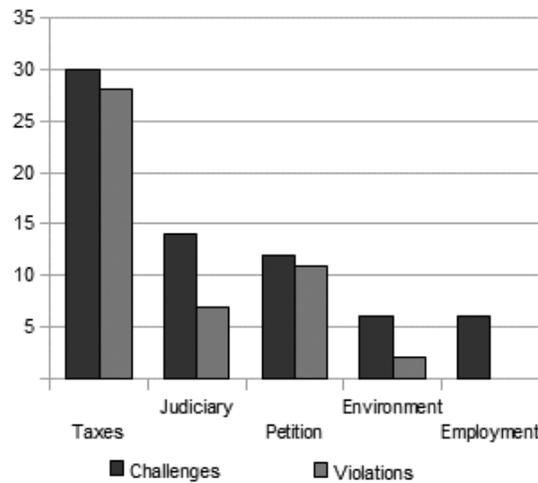
206. See JAY L. DEVORE, *PROBABILITY AND STATISTICS FOR ENGINEERING AND THE SCIENCES* 108–13 (7th ed. 2008).

207. *Id.*

fore reference to binomial distributions may be misplaced. Nonetheless, the 50% figure provides a convenient theoretical pivot point for what one might expect if judicial decisions in these cases were purely random.

When Figure 1 is viewed with the foregoing assumptions in mind, the hypothesis that cases are decided randomly shows signs of weakness. Out of eleven years that featured single subject challenges, the 50% ratio was met on only two occasions: in 1995 and in 2002. Moreover, the figures show a significant skewing toward either ends of the violation/no-violation spectrum. While one may not be able to draw firm conclusions from this analysis, Figure 1 at least serves to raise doubts about the viability of a pure-randomization hypothesis.

Figure 2. Single subject challenges vs violations in the top 5 theme categories, 1995-2009



Since the first diagram is inconclusive, we now turn to a different cross-section of the data. The second set of statistics, depicted in Figure 2, categorizes every petition by theme and presents the challenges-vs.-violations data with regard to the top five themes, which account for 78% of the cases. A couple of assumptions merit clarification. First, if the subject matter of petitions were irrelevant—as this Note assumes would be the case in a pure-randomization hypothesis—we might expect to see roughly the same percentage of

2011] JUDICIAL APPLICATION OF THE SINGLE SUBJECT RULE 859

violations across the different themes. Second, we may expect to also see no correlative policy trend among themes.

A brief review of Figure 2 suggests that the data do not conform to the foregoing assumptions. Other than petitions involving the judiciary (e.g., measures that would either enhance or restrict the judicial power through term limits, judicial qualifications, and the like), all other categories appear to be heavily skewed toward either a finding of violation (e.g., taxation- and petition-related initiatives) or finding of no-violation (e.g., environment- and employment-related initiatives). Moreover, the data appear to show that, with some exceptions, the petitions that were deemed to not violate the single subject rule represented policy positions that favored the environment, labor, and taxation—and disfavored drastic changes to the process of direct democracy.²⁰⁸

Similar to the Section prior, this Section is not meant to contain an exhaustive analysis of the pure-randomization theory. The author believes that the data do not support a randomization hypothesis in this case, and the explanation above—albeit brief—directly addresses the point. Because the above observations raise sufficient doubt about the viability of a pure-randomization hypothesis, we must turn to another explanation.

F. Hypothesis #4:

The single subject rule as a judicial vetogate for certain subject matter

A final alternative this Note explores is that, rather than the single subject rule representing an analytical device aimed to stop logrolling or fraud on the voters, it either is or has become a judicial subject-matter vetogate. Unlike traditional legislative politics, where vetogates are institutional devices that typically involve intentional conduct by political actors,²⁰⁹ the judicial vetogate need not involve judicial activism. A perfectly valid alternative to activism is the possibility that some subjects are intrinsically complex so as to not be amenable to meaningful initiative petitions. In other words, these types of subjects could never be narrow enough to pass judicial scrutiny—not in any form that would attract meaningful interest from the voting public. This Note does not attempt to assign a reason for why court decisions come out more favorably for some subjects as opposed to others; instead, identifying the existence of a vetogate is sufficient for suggesting a solution.

208. See *supra* Figure 2 in conjunction with the Appendix and the cases cited therein.

209. Eskridge, *supra* note 32, at 1442–43.

Figure 2 above serves as a starting point for substantiating the judicial vetogate claim. Since subjects appear to be difficult to define, this Note has instead organized single subject challenges by theme. The top five themes, in descending order by initiative count, are: taxes (30), judiciary (14), direct democracy (represented by the term “petition”) (12), environment (6), and employment (6). As the chart shows, measures related to taxes (predominantly tax cuts) and direct democracy (predominantly enhancing direct democracy) are likely to be invalidated, whereas those related to the environment and employment are upheld, at least on single subject grounds. Measures related to the judiciary are evenly split, however.

Having established an apparent pattern, the next issue to explore is why petitions on certain subjects tend to fare more (or less) favorably than others. Given that several possible explanations had been ruled out earlier in the Note, we may be left to explain the statistics by inferring that the rejected petitions were disliked by the court on substantive grounds. Recent research by Professors John Matsusaka and Richard Hasen goes even further, suggesting that judges’ partisan policy preferences affect single subject outcomes, especially in states that apply the single subject rule “aggressively,” such as Colorado.²¹⁰ The research concludes that “in aggressive states, judges upheld initiatives they *disagreed* with only 42.1 percent of the time,” while they “uph[eld] initiatives they *agreed* with 83.2 percent of the time.”²¹¹ “Disagreement” in this context means a mismatch between the ideological label attached to an initiative (e.g., conservative or liberal/progressive) and the political affiliation of judges (e.g., Republican or Democrat).²¹² Matsusaka and Hasen do not appear to argue that judges’ partisan motivations are driven by political ambition; in fact, they maintain that a re-election motivation may not satisfactorily explain the data.²¹³ Instead, they contend that judges’ underlying ideological beliefs on social policy are doing the work.²¹⁴

There are reasons to agree with such a conclusion. As this Note’s research suggests with regard to Colorado, initiatives in cer-

210. Matsusaka & Hasen, *supra* note 27, at 2. Matsusaka and Hasen characterize Colorado as a state with “aggressive” application of the single subject rule. *Id.* at 26.

211. *Id.* at 27 (emphasis added).

212. *Id.* at 23–24, 36.

213. “The voting behavior of judges is not reliably different when an election is close than when it is distant.” *Id.* at 25.

214. *Id.* at 22.

2011] JUDICIAL APPLICATION OF THE SINGLE SUBJECT RULE 861

tain categories have a higher likelihood of being upheld than others: pro-labor and pro-environment initiatives were upheld more frequently than initiatives seeking to decrease taxes, for example.²¹⁵ Moreover, if a subject-matter judicial vetogate exists—as this Note argues is the case—it might be intuitively obvious to conclude that the vetogate opens or closes based on judges' policy preferences. Finally, in other legal subject areas with flexible standards or wide judicial discretion, judges' legal pronouncements have long been suspected of being influenced by such extra-legal factors as justice, morality, and good public policy.²¹⁶ Political ideology may be yet another extra-legal factor that is added to the mix.

However, there are also reasons to stop short of completely embracing the conclusion of the Matsusaka and Hasen study—at least on the facts presented by the study thus far—on both theoretical and practical grounds. Methodologically, the categorization of certain initiatives as “conservative,” “liberal” or “other” appears too stereotypical for comfort—and sometimes also counterintuitive. For example, the subject “initiative procedures” is listed by the study in the “other” category, along with “medical insurance,” “term limits,” “smoking prevention,” “tobacco education,” and “campaign finance, disclosure.”²¹⁷ To some, most or all of the aforementioned causes might be liberal or progressive in nature, and not labeling them as such may have affected statistical outcomes. Moreover, the authors themselves highlight the fact that the classification of some of the subjects tagged as “conservative” or “liberal” may be disputed.²¹⁸ Second, we are not told, as a descriptive matter, why political ideology has been assumed to take precedence over other extra-legal considerations such as justice, morality, or the common good in a judge's hierarchy of internal decisionmaking.²¹⁹ In other words, the study's statistics may well be explained by extra-legal factors other than social policy views. Finally, we may wish to stop short of asserting that, in a battle between social-policy views of the people and the judiciary, the latter typically wins. After all, the same judges who pass on single subject matters also have regular cases on their dockets. Accordingly, the shadow cast over the former cases may well extend over the latter, thereby challenging judicial integrity wholesale. When alternative explanations may exist, and when

215. See *supra* Part II.C; *infra* Appendix.

216. See, e.g., RICHARD A. POSNER, *HOW JUDGES THINK* 19–57 (2008).

217. Matsusaka and Hasen, *supra* note 27, at 36.

218. *Id.*

219. See, e.g., RICHARD A. POSNER, *supra* note 216.

the statistics themselves may be questioned, we may not necessarily want to assume the worst.

In addition to the theoretical reasons cited above, practical considerations may also counsel stopping short of designating political ideology as the exclusive motivation behind the judicial veto-gate effect. For starters, many initiatives are multi-faceted insofar as extra-legal considerations are concerned. It suffices to review only a couple of cases in Colorado in order to illustrate the point. For example, we can look at *Suits Against Nongovernmental Employers*, an initiative that allowed employees to sue employers despite workers' compensation laws.²²⁰ It was proposed by a litigator who specialized in workers' compensation and personal injury cases,²²¹ and was challenged before the Colorado Supreme Court by a former Republican state representative turned "reform[er] [of] Colorado's workers' compensation laws," who became head of the Colorado Department of Labor and Employment soon after the above-mentioned case.²²² We might be prompted to make several inquiries, none of which may be solely dispositive of the outcome: whether the court was composed predominantly of Justices nominated by Democratic governors whose policy was pro labor;²²³ whether allowing such lawsuits enhances the power of the legal profession and the courts; whether Colorado's workers' compensation laws at the time had a disparate impact on minorities; whether workers' com-

220. *In re Proposed Initiated Constitutional Amendment "Concerning Suits Against Nongovernmental Employers Who Knowingly and Recklessly Maintain An Unsafe Work Environment"* (*Suits Against Nongovernmental Employers*), 898 P.2d 1071 (Colo. 1995) (en banc).

221. Initiative proponents were Neil D. O'Toole—of Dallas, Holland & O'Toole, P.C.—and Sharyn E. O'Toole. *Suits Against Nongovernmental Employers*, 898 P.2d at 1071. Neil O'Toole's experience includes "Workers' Compensation" and "Personal Injury." *Curriculum Vitae of Neil D. O'Toole*, LAW OFFICE OF O'TOOLE & SBARBARO, PC, http://www.otoole-sbarbaro.com/ndo_cv.php (last visited Dec. 31, 2010).

222. Vickie Armstrong is listed as the objector in *Suits Against Nongovernmental Employers*. *Suits Against Nongovernmental Employers*, 898 P.2d at 1071. In 1998, Armstrong was appointed by Colorado's Republican governor-elect as Executive Director of the Colorado Department of Labor and Employment. Press Release, Governor-elect Owens Transition Office, Owens Names Vickie Armstrong to Lead Labor Department (Dec. 23, 1998), http://www.state.co.us/owenspress/transition_releases.htm#12/23. The press release hails Armstrong as leading "the fight to reform Colorado's workers' compensation laws." *Id.*

223. A list of justices present on the Colorado Supreme Court is available at <http://www.state.co.us/courts/sctlib/1995.htm>. Justices Anthony Vollack, Howard Kirshbaum, Luis Rovira, were appointed by Democrat governor Richard Lamm. Justices Mary Mullarkey, were appointed by Democrat governor Roy Romer. Justice William Erickson was appointed by Republican governor John Arthur Love.

2011] JUDICIAL APPLICATION OF THE SINGLE SUBJECT RULE 863

pensation laws were just when they deprived workers of plenary damages as determined by a court of law when they suffered injury as a result of negligent employer behavior; and the like.

We can also evaluate a more cited case, such as *Public Rights in Waters II*, where the measure would have prospectively changed property rights in non-navigable water streams, a disruption to Colorado's long-held property rights scheme.²²⁴ Even before joining the Union, Colorado had adopted the prior appropriations doctrine for managing rights with regard to use of non-navigable waterways,²²⁵ and the Colorado Supreme Court repeatedly declined to permit public interest concerns to interfere with private appropriations of water.²²⁶ The court also declined to allow public access to non-navigable waters, unlike courts of other prior-appropriation states who built recreational use exceptions into their appropriations doctrines.²²⁷ Legislative policies evidenced by state statutes support a similar view of water rights, weighed toward private appropriation.²²⁸ A public trust would mark a significant shift in policy: it would not only allow public use (e.g., fishing, hunting, bathing, swimming, and recreation) of non-navigable waters, but also may require the state to ensure that waters are used for the benefit of the public rather than solely for private interests.²²⁹ Such public-interest concerns could well involve conservationist considerations, such as preserving water levels and protecting wildlife species inhabiting the waters. One may also surmise that conflicts between conservationist and private interests, especially with regard to irrigation, could have negatively affected the state's agricultural economy. Perhaps the court believed that such wide-ranging policy changes should not be made via citizen initiatives or that such policies were detrimental to the interests of the people as a whole. Or perhaps such considerations had no bearing on the outcome. The latter thought might be supported by the court's ruling in a subsequent case that involved strictly the subject of a public trust.²³⁰ In said case, the court upheld the petition, though the issuance of the decision in June²³¹ posed difficulties for qualifying the initiative on

224. *In re Proposed Initiative for "Public Rights in Waters II"* (*Public Rights in Waters II*), 898 P.2d 1076 (Colo. 1995) (en banc).

225. *See generally* Wyoming v. Colorado, 259 U.S. 419 (1922).

226. *See* People v. Emmert, 597 P.2d 1025, 1028–30 (Colo. 1979) (en banc).

227. *Id.*

228. *Id.*

229. *See* JOHNSON, *supra* note 86.

230. *In re Proposed Initiative 1996–6* (*Public Rights in Waters III*), 917 P.2d 1277 (Colo. 1996) (en banc).

231. *Id.*

the ballot; consequently, the measure was not put before the voters that year.

But *Public Rights in Waters II* was not just about a public trust: the measure would also have changed how directors for water conservation and conservancy boards are selected.²³² At the time of the measure—and even now—the directors for water districts were not elected, unlike directors of other special utility districts. Instead, they were appointed by local judges.²³³ Local citizens can request elections, but they must amass signatures equal to ten percent of the registered voters in the district.²³⁴ To better understand the policy implications of switching to an election scheme permanently, we might ask several questions: whether local judges' power to appoint directors of water conservation and conservancy districts is an important status symbol for the judiciary; whether judges see appointment as a chore or as a privilege; whether the elected directors share the policy preferences, if any, of their appointing judges; whether an election model would result in directors who are more friendly to private appropriation interests rather than less; and the like. Again, one, many, or none of these considerations may have affected the outcome here.

While the aspects discussed above are somewhat teleological (i.e., highlighting conflicts with the purposes/ends of an initiative), there are also deontological considerations that may well influence single subject decisions. For example, instead of concluding that judges voted down a tax-reducing initiative because they disagreed with low taxes on ideological grounds, we may as well conclude that they rejected the initiative because of the means through which the petition sought to reduce taxes. For example, decreasing local taxes while forcing the state to replace revenue in a way that automatically cuts state spending is an improper means of cutting taxes because it burdens a state's sovereign discretion in spending for the general welfare. Equally invalid are petitions that force judges to automatically reach a specific result in single subject cases based on a formula. We have seen other courts, namely federal courts, react negatively to such incursions into their inherent powers of adopting appropriate rules of decision.²³⁵

232. *In re Proposed Initiative for "Public Rights in Waters II"* (*Public Rights in Waters II*), 898 P.2d 1076, 1079 (Colo. 1995) (en banc).

233. Quillen, *supra* note 93.

234. *Id.*

235. See, e.g., *United States v. Klein*, 80 U.S. 128 (1871) (resisting Congressional attempts to prescribe rule of decision for federal courts in case involving legal significance of Presidential pardons).

2011] JUDICIAL APPLICATION OF THE SINGLE SUBJECT RULE 865

In sum, given the multi-faceted nature of every initiative, explaining single subject outcome statistics solely in terms of policy differences may be little more than polarized speculation. Over the past two decades, the United States has seen increased political polarization, at least in the political branches of our government—polarization that may have had deleterious effects for public governance.²³⁶ We may well ask whether extending the polarizing shadow over courts—whose legitimacy may be rooted in fairness and political impartiality—is in society’s interest. It may suffice to raise the inference that a subject-matter vetogate exists and seek to redress it through various solutions rather than speculate further as to judges’ possible animating rationales.

G. *Final thoughts on Colorado’s case law*

In retrospect, the court may well have gotten it right in every case; had the court not stepped in, voters could have been fooled into adopting laws they did not wish. But just like “I know it when I see it” was not a workable jurisprudential test for the United States Supreme Court in obscenity cases,²³⁷ it does not appear sufficient as a legal test in cases of direct democracy—a domain that the Colorado Supreme Court deemed a fundamental right²³⁸ for Colorado citizens. Not only is the lack of a coherent test fuel for the fire of realists who allege that judges simply do what they wish,²³⁹ it places the court in the undesirable position of being perceived as stifling the sovereign voice of the people. Moreover, it would also be ironic that voters, in an attempt to avoid confusing initiatives, had adopted a single subject provision that has turned out to be so confusing to implement.

236. Richard H. Pildes, *Why the Center Does Not Hold in American Democracy: Persons, History, Institutions* (NYU Sch. of Law Pub. Law Research Paper No. 10–47, 2010), available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1646989 (noting that level of political polarization has increased in past two decades, at least in political branches of federal government).

237. Justice Potter Stewart’s concurring opinion in *Jacobellis v. Ohio* famously stated, with regard to the definition of obscenity, “I shall not today attempt further to define the kinds of material I understand to be embraced within that shorthand description [of obscenity] . . . [b]ut I know it when I see it.” *Jacobellis v. Ohio*, 378 U.S. 184, 197 (1964) (Stewart, J., concurring).

238. *Loonan v. Woodley*, 882 P.2d 1380, 1383 (Colo. 1994) (en banc).

239. E.g., Felix S. Cohen, *Transcendental Nonsense and the Functional Approach*, 35 COLUM. L. REV. 809 (1935).

III. SOLUTIONS

Though this Note's research is focused on Colorado, the solutions suggested in this part may well apply to many or all states that use single subject rules to check citizen initiatives. Before delving into solutions, the Note examines the traditional animating causes behind the single subject rule—i.e., fraud upon voters and patchwork legislation—in a critical light, and concludes that the traditional reasons may not fully justify the single subject rule as applied. The Note also posits that the rule is actually—or should be, normatively—a countermajoritarian check that regulates the level of burden that is foisted upon the state and its citizens via initiative. Based on this reframing of the rule, the Note evaluates two interpretive reforms: a well-defined, consistently applied prophylactic model, and a partially deferential tiered-scrutiny model. The latter model is more desirable because it balances republican and democratic ideals. The Note also mentions briefly two process solutions: an evidentiary model and a revision of the ballot. Because the latter ideas would likely require legislative changes at the state level, they are outside the scope of this Note's focus on interpretative theory, and are therefore discussed only summarily.

A. *The evils of fraud upon voters and patchwork legislation*

The two concepts of fraud upon voters and patchwork legislation are related, in that both involve the majority voting for something they do not want, but are analytically distinct: there can exist patchworked measures that are not fraudulent, fraudulent measures that have not been logrolled, and measures that contain hidden provisions so as not to detract significant numbers of voters. This Section will briefly evaluate what each term means, as well as identify the aspects that most need protection by a court, as a normative matter.

1. Fraud upon voters

Voters can be misled through affirmative statements that turn out to be false or through omissions. With regard to initiatives, however, affirmative misstatements pose less of a problem, for reasons discussed herein.

It might be said that fraud is perpetrated upon voters through affirmative misstatements when a measure's clearly stated provisions have an effect that the voters do not want—e.g., the tax cut that reduces taxes and also has the effect of reducing state spend-

2011] JUDICIAL APPLICATION OF THE SINGLE SUBJECT RULE 867

ing, with the latter being undesired. Such an assertion is not analytically satisfying, because voters who do not want something need not vote for it;²⁴⁰ in the end, this is a choice for voters, not courts. Moreover, the fact that voters do not like an outcome *ex post*, once a measure has been implemented, does not necessarily mean that they did not want the measure *ex ante*, at the time it was enacted. Invalidating a proposed measure simply because voters might not want the results (and do not currently realize their future discontent) turns the court into a paternalistic seer—a role that it might not be well equipped to fulfill. But even if it were, it is difficult to see how limiting a measure to a single subject would increase the likelihood that voters will like—and, therefore, want—the effects of the measure.

The more practical interpretation of fraud upon voters, then, is where a measure turns out to have an effect the voters did not expect. But even this description is too broad, because *ex ante* and *ex post* expectations are different, and the latter cannot normatively serve as a basis for single subject invalidation. Voters are not omniscient, and neither are judges; therefore, one cannot judge a measure by the possibility that the *ex post* effect could be different from the stated effect. Some measures could conceivably cause unexpected results not because the provisions themselves were unclear but because they were not implemented appropriately, or because another supervening cause changed the game. If a single subject rule were to protect against suboptimal implementation or supervening events, courts would again be in the position of paternalistic seers. Therefore, the only unexpected effect that is left to be protected in cases of affirmative misstatements is where the measure asserts that the law works in a certain way and, as a consequence of such operation, it will deliver a certain set of outcomes. But such situations are less problematic than they seem. Since initiatives create or change law, the supposed misstatement about the operation of a law should, in fact, serve to change the law so that it works the way in which it has been framed. This is, after all, what the voters understood the future to be. Courts can always give effect to voters' wishes and interpret enacted provisions in concert with the promised effects, achieving as much of the latter as possible. For the reasons mentioned above, affirmative misstatements do not appear to provide a compelling case for judicial protection.

240. See Arthur Lupia & John G. Matsusaka, *Direct Democracy: New Approaches to Old Questions*, 7 ANN. REV. POL. SCI. 463, 469 (2004) (noting that some voters reject initiatives "when information is lacking or when worries about general state conditions are greatest").

The second means of misleading voters, omissions, present a much stronger need for *ex ante* intervention. In reviewing a certain provision that is less than forthcoming about details, a voter might well make incorrect assumptions about substantive and procedural aspects, leading to unreasonable conclusions about the resulting effects of the measure. Since no affirmative statements are made, the law does not change to match any given voter's assumptions. Guarding against omissions may protect voters who are not knowledgeable about the law or who know the text of the law but misinterpret it. But this realization is only the beginning of the inquiry.

The next question that arises is, "When is an omission, in this context, a material omission?" For example, is there fraud upon voters if, despite the omission and incorrect assumptions by voters, the measure still has the effect that a reasonable voter wants it to have? In an omission context, it is difficult to determine what a reasonable voter wants, because different voters have diverse reasons for approving a measure. For example, some citizens who want higher taxes may desire an increase in social programs, defense programs, or infrastructure-building programs. Others might desire higher taxes so that government may employ more of its citizens, regardless of what programs are enacted. Still others might intend higher taxes to narrow the gap between rich and poor. The fact that a measure results in higher taxes does not necessarily mean that the effect voters wanted has been achieved. The same holds true for a mirror-opposite provision: cutting taxes. Some might want lower taxes so that they may keep more of their income. Others might simply want to direct some of their tax cut dollars to different activities than those funded by the government. Still others may want to cut certain taxes because they believe the government should not be in a certain business, but might be fine with raising taxes, as a compensating measure, for other areas they support. Therefore, a reduction in certain taxes does not necessarily effectuate the policy causes voters may have thought they were supporting.

According to the observations above, asserting that an omission is material if it can reasonably lead to a different outcome may seem to not account for the diverse outcome expectations of voters. However, outcome is relevant. For example, assume a scenario where every outcome desired by every approving voter materializes. In such a case, omissions are irrelevant—at least from a substantive perspective—because every voter realized the result he wanted. Since single subject analyses are conducted prior to any measure taking effect, however, we must develop a definition of outcome-oriented materiality that can be applied *ex ante*. Such a definition

2011] JUDICIAL APPLICATION OF THE SINGLE SUBJECT RULE 869

might read thus: “An omission is material if it leads a voter to *reasonably expect* a larger or smaller *set* of *probable* outcomes from the measure as compared to the set expected if the omission had not existed.”

The above definition does not treat as material those omissions that do not change the probable set of outcomes, because even laws drafted with the utmost precision may conceivably be interpreted differently by different, equally competent legal scholars. In fact, we see this variability take place in state and federal appellate courts across the nation whenever judgments are not unanimous—especially in close cases where one judge’s vote determines the outcome. Therefore, even the most competent voter need not (and, if the above assumption is correct, cannot) understand the exact effect at which the law aims: only the range of probable effects can reasonably be understood. It is up to the voter, then, to determine whether to take the risk and approve a measure whose effects would fall in said range. As long as the omission did not hide the set of reasonably probable effects, it should not be material.

We might ask whether a measure violates the materiality rule above if it implies, by omission, a larger set of outcomes than will reasonably occur. One may argue that since voters approved a broader set of outcomes and no voter can be certain about which outcomes will be effected during implementation (as long as all of them are reasonably probable), voters cannot complain if the actual outcome set is narrower. For this argument to hold, further refinement of the term “outcome set” is needed. An outcome set is a group of one or more outcomes, and an outcome is a combination of a bundle of benefits and their attending costs. For example, a \$25 tax cut (benefit) and a resulting reduction in educational programs (cost) represent a simple outcome. A complex measure has many outcome combinations, some of them complex; it is likely that voters do not analyze all combinations systematically. However, as long as voters understand that any of those outcomes are fair game, where is the complaint?

Voters, like all decisionmakers, base their decisions not only on the knowledge that a certain outcome might win, but also on the probability that a certain outcome will win. This is one reason why the materiality definition proposed by this Note includes only reasonably probable outcomes, not all possible outcomes. In a reasonably probable scenario, every outcome has roughly the same chance of taking effect. When voters approve a larger outcome set, some of them could well be rooting for the outcomes that turned out to be falsely implied, even though they realize that other outcomes have

an equal probability of selection. Due to the omission, however, some outcomes actually had a 0% probability of selection while others had a higher probability on account of the smaller element set. As a result, some voters were misled into approving an outcome set that was rigged: some voters' choices were never going to materialize as a result of the omission. Given such false expectations, the omission should be deemed material in this case.

An inverse question also arises: "Do omissions that cause a larger-than-expected outcome set violate the materiality rule?" This analysis is much more straightforward. Since some of the outcomes present in the larger set could have been considered undesirable by voters who voted for the smaller set, the omission deprived such voters of the opportunity to vote against the measure. Consequently, the omission is material.

Finally, it is important to distinguish the concepts "voter confusion" and "fraud upon voters." Though omission of material information can be confusing, voters can just as easily be confused by correct affirmative statements that appear to conflict with other correct statements. Voters can also be confused by correct, non-conflicting statements that are nonetheless complex and cognitively difficult to process, such as complex processes, formulas, and the like.²⁴¹ That said, one should question whether invalidation of a measure by courts is the best remedy for voter confusion that arises from conflicting or complex statements. Voters can always reject a measure if they believe that the language does not allow them to form a reasonable set of expectations about the measure's outcomes.²⁴² Therefore, initiative drafters are (or should be) intrinsically motivated to write provisions so that they are neither conflicting nor overly complex. Moreover, unlike in the case of omissions, in which voters may not realize that information has been omitted, voters are aware of the potential conflict or the cognitive difficulty they experience while reading confusing affirmative provisions. As such, voters should decide whether to take the risk of approving the measure, or instead reject it and maintain the status quo. By taking away voters' freedom to choose, a court is not only treading the separation-of-powers line, but is also discounting the ability of voters to make decisions. As several studies have observed, voters may be more competent than critics of direct democracy contend; for example, voters sometimes use extra-textual information, such as knowledge about a measure's supporters, to decide

241. *See generally id.* at 467–469 (noting that voters can be confused by complex propositions, such as proposed initiatives that regulate insurance industry).

242. *Id.* at 469.

2011] JUDICIAL APPLICATION OF THE SINGLE SUBJECT RULE 871

whether to vote for a complex measure.²⁴³ Accordingly, as long as voter confusion is not due to a material omission, a court need not interfere.

The counterargument with regard to conflicting and complex provisions is that voters may well take the risk, misinterpret the provisions, and enact a law that creates outcomes they did not expect. It naturally follows that someone should protect voters from such a possibility. Perhaps this is so, but then voters or the legislature should explicitly empower the courts to do this. When empowerment has been extended only to single subject regulation—not to clear language regulation²⁴⁴—voters and their representatives may well have wanted to reserve to the people the right to deal with situations such as the one mentioned above. When courts reframe single subject *sua sponte* and extend its reach beyond language clarity, they might be seen by some to engage in self-empowerment.

The preceding answer is not fulfilling in cases where voter risk-taking on interpretation could result in significant unexpected burdens on the state and its people: there are some risks that we may not trust the demos to take, even if it wanted to—risks that lead to discrimination, for example. But legislators arguably misinterpret provisions all the time—and yet no one suggests using the single subject rule for bills to strike down proposals that result in significant unexpected burdens. Why should the single subject rule operate differently where the legislators are average voters instead of professional politicians? Since one can never be certain about how many politicians base their votes for a bill based on a personal review of its text, we need not compare the cognitive aptitude of the average voter and the politician; in the eyes of the law, they are equally capable legislators.²⁴⁵ The difference instead lies in the institutional practices surrounding legislative bills versus petitions. When legislators realize that they have misunderstood certain provisions, they can attempt to amend the law and clarify it with much greater ease than initiative petitioners.²⁴⁶ As discussed at the beginning of this Note, many roadblocks stand in the way of petitioners making amendments. Given this process disparity, courts may be

243. *Id.* at 467–70.

244. As an aside, Colorado voters had enacted a form of the clear language rule, but the discussion in this part of the Note is state-agnostic and contained to the single subject theme.

245. See Lupia & Matsusaka, *supra* note 238, at 467–70 (noting that voters for ballot initiatives often successfully use extra-textual cues—such as political signals and interest-group preferences—to decide how to cast their vote).

246. Whether legislators can obtain majority support for such changes post-enactment is another story.

justified in imposing—*sua sponte* if need be—a clear language rule for initiative measures whose effects are substantial. The Note adopts this view in the solution proposed below in Part III.B.2.

2. Patchwork legislation

Another evil that the single subject rule professedly combats is patchwork legislation. The most prevalent form of this is logrolling, a practice whereby multiple measures that are each popular with a separate minority voter group—minority here meaning popular minority—are combined into one petition that gains majority support through aggregation.²⁴⁷ A second, analytically distinct practice is riding, whereby an unpopular measure is combined with an overwhelmingly popular measure in order to pass.²⁴⁸ Protecting voters from both of these bad practices is said to be virtuous because it allows passage only of measures that gain majority support independently.²⁴⁹ Any claim to virtuosity deserves a closer look. What exactly are we trying to avoid and why?

This Note argues that the goal of preventing patchwork legislation should be solely to avoid costs imposed by one group of citizens over another. The terms shall be used as follows in this context. Costs refers to both fiscal burdens as well as psychic burdens imposed by regulating behavior or otherwise limiting freedom. A popular minority is any group whose preferences with regard to a unitary, single subject measure are the opposite of the popular majority.

Logrolling and riding will be examined in turn. Two variants of logrolling exist. The first involves two (or more) completely unrelated measures that are merged into one.²⁵⁰ A hypothetical scenario is illustrative. Assume two measures, X and Y. Each measure is supported by a separate popular-minority group; for the sake of simplicity, let us assume that each group represents 26% of the population. Let us also assume the worst-case scenario: if each group brought its measure independently, the measure would receive 26% yes votes (representing the sponsor group) and 74% no votes (representing everyone else). The no votes may be dominated by a wide range of factors, including political animosity and ignorance. But in a world where voters are rational, self-interested actors, the strongest motivator might well be self-interest: the nays receive no benefit from the measure while having to bear some or

247. Cooter & Gilbert, *supra* note 70, at 706.

248. *Id.* at 707 & n.99.

249. *Id.* at 714–718.

250. *Id.* at 706, 713.

2011] JUDICIAL APPLICATION OF THE SINGLE SUBJECT RULE 873

all of the costs. Let us also assume that the two groups decide to compromise—which often happens in a pluralistic society—by merging measures. Together, the two groups represent an ad hoc majority of 52%, and by voting for the measure can obtain the benefits of their respective provisions while shouldering only part of the burden (if any). The remaining 48% of the population is left to shoulder much of the burden without receiving any benefit (worst-case scenario). Does this hypothetical present a problem in which courts should interfere?

One answer is that the hypothetical cannot be meaningfully distinguished from the scenario where a unitary measure that conforms to the single subject rule passes by a 52% yes vote. The latter measure can also result in 48% of the population shouldering much of the burden without receiving any benefit.²⁵¹ Is there a significant difference? If not, then both single subject and multi-subject measures can suffer from the same substantive evil: namely, unbalanced apportionment of benefits and costs among the population.

One may imagine that in the second hypothetical, the cost may well be less than if the two disparate measures had been combined. However, the costs are measure-specific, and there is no necessary connection between the number of subjects and the resulting costs. Moreover, the second hypothetical benefits the same percentage of people as the first, albeit through one set of provisions instead of two. In a situation where the per capita benefit, regardless of measure, is fueled by one unit of cost, the cost may well be similar in both situations since the magnitude of the benefit (population-wise) is the same. Granted, these are hypothetical situations with theoretical measurements, which means that real life might well be different. Nonetheless, they serve to point out that logrolling in the abstract is not necessarily the evil that it may be painted as being.

Another type of logrolling is micro-logrolling, whereby initiative proponents add implementation details to a single subject measure in order to accumulate voter support.²⁵² Unlike macro-logrolling, which involves disparate measures, this one simply piles enhancements onto a single subject proposal to make it more attractive. Take, for example, the \$60 tax credit found by the Colorado Supreme Court in *Amend Tabor 32* to have satisfied the single

251. Granted, there are often trickle-down benefits from any measure in real life, but this does not distinguish the two scenarios: both can have trickle-down benefits to the nays.

252. See Cooter & Gilbert, *supra* note 70, at 712 (“Logrolling can take place within a measure that embraces one logical subject.”).

subject requirement.²⁵³ The credit was applied to six different taxes.²⁵⁴ One could argue, as the Note did earlier, that a ten-dollar (or even a 60-dollar) credit applied to one type of tax may not have gathered sufficient support, since the individuals who would have qualified for one type of tax may not have been numerous. Hence, initiative proponents may well have simply ratcheted up the number of taxes (and, perhaps, even the credit amount) until they reached figures they believed were attractive to a majority. There appears to be no analytical difference between this type of logrolling and the kind mentioned earlier in this section provided that both turn minority measures into majority ones.

Riding is a different type of problem. A rider provision is never needed to compromise, because the provision on which it rides has enough majority support by itself.²⁵⁵ However, some riders may, in fact, be a form of implementing the common good; not all riders are detestable. Assume, for example, the hypothetical where two proposals, X and Y, are supported by 70% and 30% of the population, respectively. In the combined measure XY, Y is the rider; X would pass independent of Y. Assuming that the combined measure passes with close to 100% vote, it represents a desire of the majority to ensure that everyone gets a benefit, not just the supporters of X. In this sense, riders have the potential of bringing society together and, in a way, compensate Y supporters for carrying part of the burden for X (which they would have carried anyway even without receiving any benefits). Automatic exclusion of riders, therefore, can lead to a polarized society: a negative outcome if one's goal is to avoid the tyranny of the majority.

There is also at least one scenario that, while it might appear to represent logrolling or riding, it is neither: a multiple subject measure in which most voters have inseparable preferences for all subjects.²⁵⁶ Assume, for example, a measure containing two disparate provisions: R and S. (We will call the measure RS, for simplicity). Let us also assume that the same voting bloc who prefers R also prefers S, and that bloc represents 60% of the population. We may also assume that a separate voting bloc, accounting for 20%, prefers only measure S, but would nonetheless vote for RS. Even if the votes of RS would rise to 80%, RS would pass regardless of whether the separate 20% bloc supports it or not; as such, combining the mea-

253. *In re Amend Tabor No. 32 (Amend Tabor 32)*, 908 P.2d 125 (Colo. 1995) (en banc).

254. *Id.* at 131.

255. Cooter & Gilbert, *supra* note 70, at 707 n.99.

256. *Id.* at 717.

2011] JUDICIAL APPLICATION OF THE SINGLE SUBJECT RULE 875

asures was done for purposes of efficiency, not passage. Therefore, RS is not a logrolled measure. Moreover, neither R nor S is a rider, because each has the support of 60% of the population. A purely analytical single subject test may treat RS as presumptively logrolled, whereas a test that permits evidence to be brought with regard to support levels for any given provision might be better tailored to separate such situations. This Note recommends such a test below in Part III.C.1.

As a final thought, perhaps the problem with logrolling and riding is not the majoritarian aspect, or the costs upon those who do not receive benefits, but rather the enormity of the burden on the population as a whole from logrolled measures. In Colorado, for example, the single subject rule was passed after TABOR was enacted under a process unrestricted by said rule. TABOR changed procedural and substantive aspects of both state revenue and state spending activities. Given the significant change foisted upon state government through such a comprehensive measure, the operative lesson that gave birth to the single subject rule²⁵⁷ appeared to be “Never let another measure effect this scale of change upon the polity as a whole,” not “Never let an ad-hoc majority burden everyone else.” Therefore, we may reasonably conclude that what the people are trying to do when enacting single subject rules is to prevent large-scale burdens from being foisted upon the state and its people via initiative. The tiered scrutiny model proposed in Part III.B.2 satisfies this goal more directly than other models.

Having discussed the concerns that a state should address through judicial review of initiatives as a normative matter, we turn now to evaluating several solutions.

B. Interpretive frameworks

1. The prophylactic model

One’s instinct might well suggest that narrowly drawn measures will typically impose limited burdens on the state and its denizens; therefore, the ideal model might be one that aggressively and consistently screens out all but the narrowest measures. To attain its prophylactic effect, the model will necessarily need to be over-inclusive, and cannot draw the fine-grained distinctions discussed in above. Borrowing from the interpretative layer applied to Colorado’s case law, as well as some of the earlier discussion on judicial protection, the model calls for several elements/prongs to

257. Ironically, the single subject rule locked in TABOR and made it un-repealable by a single initiative, but that is another story.

be met in order for a measure to be deemed as comprising a single subject:

1. *fine-grained subject specificity*—the level of acceptable specificity depends on what a reasonable voter deems to be the lowest-level subject that logically results from the measure.²⁵⁸
2. *necessary connection*—each of the provisions of the measure must have a necessary connection with the central theme of the measure and with each other.²⁵⁹
3. *single object*—the measure must affect only one object; an object represents the logical target which the provisions of a measure aim to affect.²⁶⁰
4. *related effects*—provisions that cause multiple unrelated effects on a single object are presumptively dealing with multiple subjects.²⁶¹
5. *procedure vs. substance*—provisions that deal with the procedure of a central theme are presumptively dealing with a separate subject than those dealing with substantive aspects of the same theme.²⁶²
6. *material omissions*—any measure that omits information which leads a voter to reasonably expect a larger or smaller set of probable outcomes is presumptively dealing with multiple subjects.²⁶³
7. *clear statement test*—a measure must be stated in clear language that can be understood by a lay person; otherwise, the measure will be invalidated as a prudential matter.²⁶⁴
8. *exhortative statements*—must be only loosely related to the central theme.²⁶⁵

The difficulty with such a rule is manifold. First, if single subject was meant to be a scalpel that excised only initiatives that were logrolled or those that defrauded or misled voters, then the rule as devised above would be more of a chainsaw. It fails to differentiate between simple measures that cause clearly understood changes to a variety of government functions, and multi-faceted measures that are difficult for average voters to grasp but nonetheless change only one narrow aspect of government operation. Second, the rule thus

258. *See supra* Part II.C.4.

259. *See supra* Part II.C.2.

260. *See supra* Part II.C.3.

261. *See supra* Part II.C.5.

262. *See supra* Part II.C.6.

263. *See supra* Part II.C.5.a.

264. *See supra* Part II.C.5.

265. *See supra* Part II.C.7.

2011] JUDICIAL APPLICATION OF THE SINGLE SUBJECT RULE 877

defined would prevent comprehensive measures with broad effects from being passed by initiative, regardless of what band(s) in the political spectrum proponents may occupy. Third, such a rule would also place courts in the sensitive position of stifling the voice of the people (rather than channeling said voice), thereby raising separation-of-powers and democratic concerns. In states with direct democracy, the people are essentially a fourth branch of government with legislative power; stepping too far over the line to limit popular legislation may have significant repercussions even for apolitical courts.

In its defense, the model is clear and, if applied consistently, would not suffer from the perception that courts act as judicial vetogates for policy. If the people wish to change it, they may do so through their legislature, a branch toward which a court may offer greater deference. A transparent and consistent model would be, after all, easy to understand and easy to change if desired.

2. Tiered scrutiny model

The uniform model described above can be much improved by a tiered scrutiny framework. The latter would vary the level of judicial scrutiny of a measure proportionally with the burdens the measure seeks to foist upon the state and its people. Unlike the prophylactic model, tiered scrutiny recognizes—as the Colorado Supreme Court recognized—that the people’s right of initiative is a “fundamental right.”²⁶⁶ As with other fundamental rights—most notably those present in the federal constitution—the right of initiative should be neither absolute nor erased by overly deferential or overly oppressive judicial tests. Instead, the burdens imposed on the right must be weighed in light of the burdens the right itself imposes on other fundamental rights of the people.

(a) Rational Scrutiny

The court should give the greatest deference to measures that are entirely hortatory; impose minimal burdens on citizens’ non-fundamental rights; impose minimal burdens on functions of state or local government that are not traditional functions of government; or create new non-fundamental rights for citizens. Under this level of deference, measures are presumed to contain a single subject unless their provisions relate to multiple themes that are not logically related. The themes can be construed broadly (i.e., no

266. *Loonan v. Woodley*, 882 P.2d 1380, 1383 (Colo. 1994) (en banc).

specificity requirement), which means that most measures would pass this level of scrutiny.

The rationale for applying this level of scrutiny in the situations above is that the costs of a measure that meets the above criteria are typically minimal. If the goal of the single subject rule is to prevent substantial cost from being foisted upon the people and their government via initiative, as postulated by Part III.A, then the goal is not defeated by this approach. Instead, the societal benefit of allowing the people to shape the environment in which they live without roadblocks from courts would outweigh the minimal burdens that may be imposed if the measure passes.

(b) Intermediate Scrutiny

Intermediate scrutiny should be applied to measures that: impose minimal burdens on fundamental rights of citizens; impose substantial burdens on non-fundamental rights of citizens; impose minimal burdens on traditional functions of state or local government; or introduce new fundamental rights for the people.

Given that such measures begin to ratchet up costs for the people as a whole and for their government, a variety of checks—borrowing from the arsenal highlighted in the prophylactic measure—can ensure that the costs are not out of line with the common good. The costs here are burdens on people's non-fundamental rights and, to a lesser extent, the minimal burdens on fundamental rights and the state. Clear expectations are critical to ensuring that voters understand what they would be taking on. Therefore, a clear statement test should be adopted as a prudential matter. An effects test may also be employed to further the same purpose: multiple unrelated effects on the same object—where object is defined by the right or state function being burdened—may be deemed to constitute different subjects. Furthermore, the necessary connection test should also be applied, whereby multiple provisions must be necessarily related to the main theme of the measure and to each other. This test would ensure that the burden is reasonably narrow, localized, and clear. If a measure meets these tests, its burdens will likely be limited, and voters should be permitted to organize for or against it at the polls.

(c) Strict Scrutiny

The strictest scrutiny should apply to measures that: impose substantial burdens on fundamental rights of citizens; impose substantial burdens on a traditional function of state or local government; or represent an overhaul of the state constitution. The full

2011] JUDICIAL APPLICATION OF THE SINGLE SUBJECT RULE 879

array of tests described in the prophylactic model should be applied here²⁶⁷ in order to ensure that only a narrow set of fairly localized majoritarian measures are put before the voters. Otherwise, the state risks a significant disruption in operation and stability—a disruption that might have much greater negative effect on the majority than the benefit gained by passing such a measure. Given its power, the strict scrutiny test should be reserved for the largest of burdens; otherwise, the same criticisms may be lodged against it as apply to the prophylactic model.

(d) The value of an interpretive model

However satisfying an interpretive model might seem from a formalistic perspective, why should we assume that it will eliminate the judicial vetogate effect? Perhaps if courts simply adopted a generally deferential stance toward petition initiatives, the “judicial vetogate” effect would disappear, or at least be less pronounced. Such an approach is proposed in a recent study by Professors Matsusaka and Hasen.²⁶⁸ The study suggests that judges in states with restrained enforcement of the single subject rule (i.e., deferential toward initiatives) are less likely to strike down initiatives due to partisan leanings.²⁶⁹ Accordingly, the study suggests that “the most effective way to promote objectivity may be to adopt a restrained approach rather than to seek additional interpretive ‘tests’ that operationalize the concept of a single subject.”²⁷⁰ The suggestion of deference appears appealing, especially if we assume that judges are unconstrained by restrictive interpretive frameworks.²⁷¹ The tiered scrutiny model is compatible with the Matsusaka/Hasen pure deference theory at the first tier (i.e., rational scrutiny). However, the two models may be in tension as the weight of an initiative’s burden on a state’s functions or its citizens’ liberties increases. Hence, the question is whether the tiered scrutiny model offers any advantages at this heightened level of burden, as opposed to the pure deference theory.

In addressing the question, this Note makes two assumptions about the pure deference theory. First, the theory would uphold an

267. See *supra* Part III.B.1.

268. Matsusaka & Hasen, *supra* note 27, at 4; see also *supra* Part II.D.

269. Matsusaka & Hasen, *supra* note 27, at 4.

270. *Id.* at 5.

271. See Daryl Levinson, *Parchment and Politics: The Positive Puzzle of Constitutional Commitment* (2010) (unpublished working paper), available at http://ssrn.com/abstract_id=1577749 (addressing question of why constitutional rules would have staying power in light of fact that they can be “reinterpreted” if different substantive results were desired in certain situations).

initiative that significantly burdens the state or its citizens, as long as the initiative does not make the mistake of blatantly aggregating clearly disparate subjects.²⁷² Second, we cannot rely on popular majorities to reject burdensome initiatives, even though many initiatives that are allowed on the ballot are not eventually approved by voters. Therefore, if the court is deferential, a burdensome initiative may well become law.

Before continuing, it is worthwhile to acknowledge that distinguishing between minimal, reasonable, and heavy burdens is no easy feat. The Supreme Court of the United States has itself wrestled with the issue in election contexts.²⁷³ Though there is no bright-line rule for evaluating burdens, courts can establish a variety of *per se*, evidentiary, and presumptive rules that simplify the process somewhat.

Assuming that burden can be reasonably defined, the pure deference theory's failure to distinguish between heavy and reasonable burdens is problematic. The first challenge is that a deferential approach may drain the single subject requirement of independent vitality. Inasmuch as a state's citizens passed such a provision to tie themselves and their followers to the mast in the future and limit the scale of change brought about in a single initiative, deference may well serve to loosen the ropes. Second, a deferential approach may relegate courts to the role of rubber stamping popular petitions rather than engaging in meaningful judicial review of popular legislation. Our Founding Fathers realized the risks inherent in unchecked majorities,²⁷⁴ and aimed to protect against such threats by establishing certain countermajoritarian devices, including judicial review.²⁷⁵ Inasmuch as we extend their wisdom from the federal to

272. California and Washington, cited by Matsusaka and Hasen as having a "deferential" approach to single subject challenges, uphold initiatives 94% and 91% of the time, respectively. Matsusaka & Hasen, *supra* note 27, at 19. Since the study does not discuss the substantive content of the initiatives that were struck down under the deferential approach—and does not suggest using "burden" as a differentiating factor—we may be justified in assuming that deference may serve to uphold a burdensome initiative.

273. *Crawford v. Marion Cnty. Election Bd.*, 553 U.S. 181 (2008); *Burdick v. Takushi*, 504 U.S. 428 (1992).

274. "It is of great importance in a republic not only to guard the society against the oppression of its rulers, but to guard one part of the society against the injustice of the other part. Different interests necessarily exist in different classes of citizens. If a majority be united by a common interest, the rights of the minority will be insecure." *THE FEDERALIST* NO. 51, at 351 (James Madison) (Jacob E. Cooke ed., 1961).

275. *See id.* ("There are but two methods of providing against this evil[, one of which involves] . . . comprehending in the society so many separate descriptions of

2011] JUDICIAL APPLICATION OF THE SINGLE SUBJECT RULE 881

the state context—a move that may not be unreasonable in light of the Guarantee Clause²⁷⁶—we can understand the single subject rule to be a countermajoritarian check rather than simply a drafting device. A purely deferential application of this rule may serve to annul such a check, thereby potentially leaving the door open to the same risks at the state level that Madison envisioned federally, resulting in significant harm to popular minorities.²⁷⁷

Aside from protecting minority rights, we may also ask whether courts should protect the majority from its own folly in cases where governors and legislatures are legally prohibited from interfering with popular will. For example, the chief justice of the California Supreme Court, Ronald M. George, criticized the state's direct-democracy process in 2009, blaming it for "rendering [California's] state government dysfunctional."²⁷⁸ Justice George criticized the effects of the initiative and referendum process on civil rights as well as on the state's ability to function in the face of economic challenges.²⁷⁹ Perhaps suggesting that the courts should protect the majority from itself gives the court much more authority than it should have, especially since the majority can easily reverse itself during the next ballot cycle. However, we may ask whether a stable legal order, which is within the province of the courts, deserves protection from the whims of popular legislative will. The pure deference theory's silence on the foregoing issues does not necessarily make the model undesirable; it simply causes us to hesitate before embracing it uncritically.

In contrast, a tiered interpretive model that increases judicial scrutiny proportionally to the societal burden of an initiative enables courts to play a prudential countermajoritarian function while enhancing legitimacy through transparency and notice.²⁸⁰ While an interpretive model can, in theory, be molded by partisan judges to fit a desired policy outcome in every case, we cannot say with cer-

citizens as will render an unjust combination of a majority of the whole very improbable, if not impracticable.”).

276. U.S. CONST. art. IV, § 4 (“The United States shall guarantee to every State in this Union a Republican Form of Government.”).

277. See Cooter & Gilbert, *supra* note 70, at 700 (noting that in initiative context, “no matter how much harm a proposition inflicts on minority interests, that minority cannot bargain with members of the majority and convince them not to pass it”).

278. Jennifer Steinhauer, *Lead Judge Denounces State's Glut of Measures*, N.Y. TIMES, Oct. 10, 2009, at A23.

279. *Id.*

280. See Claire R. Kelly, *Institutional Alliances and Derivative Legitimacy*, 29 MICH. J. INT'L L. 605, 609 (2008) (discussing legitimacy in non-judicial context).

tainty that such a model is toothless or useless. At a minimum, it can increase the political and legitimacy costs of jurisprudential sleight of hand due to the more transparent nature of a structured interpretive model,²⁸¹ thereby limiting deviations from prior applications of the model to cases that are worth the cost. Moreover, since many state judges are either elected or subject to a retention vote, a track record of manipulating a legal rule illegitimately—a record that is more clearly perceived when the legal rule has a well-defined formalistic structure—may have some moderating effect. Furthermore, a clear rule provides notice to petition sponsors, conceivably increasing the likelihood that initiatives will be crafted so that they have a predictable chance of at least getting on the ballot. Finally, if rejection of an initiative is unavoidable, a clearly defined interpretive rule will shift this undesirable task to political branches (such as the Title Board in the case of Colorado) and away from politically impartial courts.²⁸² This will, at a minimum, protect the legitimacy of courts against charges of political partisanship. In sum, while an interpretive solution is by no means perfect, it may be more desirable than a purely deferential approach.

C. *Process solutions*

An interpretive model is not the exclusive method for fulfilling the precepts of the single subject rule. In fact, solutions that incorporate statistical evidence or that allow for alternative ballot structures might enhance the interpretive exercise or replace a purely analytical model wholesale. Two such solutions are briefly mentioned below.

1. Evidentiary model

Unlike an analytical model that is based on textual analysis, an evidentiary approach uses actual voter behavior to inform courts whether a given petition violates the single subject rule. Under this model, objectors to a petition must establish by a preponderance of the evidence that the measure is guilty of the evils the objector

281. *See* *Planned Parenthood v. Casey*, 505 U.S. 833, 869 (1991) (plurality opinion) (“A decision to overrule [precedent would be done] at the cost of both profound and unnecessary damage to the Court’s legitimacy, and to the Nation’s commitment to the rule of law.”).

282. In Colorado, the Title Board is supposed to not place a measure on the ballot if the measure violates the single subject rule. Other states likely have a similar procedure. If the single subject jurisprudence featured a clearer test, the Board may be forced to perform the “dirty work of invalidating violative initiatives, rather than such work being handed off to the courts.

2011] JUDICIAL APPLICATION OF THE SINGLE SUBJECT RULE 883

claims—evils such as logrolling, riding, voter confusion, and fraud upon voters. Evidence can consist of polling data, historical statistics from other states, and past ballot results in the state. Petition proponents can bring their own evidence to persuade the court otherwise. Given that different types of scenarios exist that might only appear to be impermissible—but upon closer inspection are not, as discussed in Part III.A above—evidence would help the court make a decision that better comports with reality.

An extension of this model could well involve putting questions of fact to a jury, especially for claims involving fraud on the voters. Since the issue in such claims is whether voters are misled as to the effects of a measure, what better mechanism to determine the views of the average voter than a jury.

2. Revising the ballot

Another way to resolve challenging measures is to shift the last word from the court to voters. A court may identify the different subjects in a convoluted measure, for example, or prompt a petition supporter to clarify the text in a certain fashion. Once that happens, voters can be allowed to vote for each separate subject or provision of the measure rather than require voting at the petition level. Whichever provisions or subjects obtain a majority vote are enacted, and the rest are not. This approach is the functional equivalent of separating one petition into mini-petitions that meet the single subject test. A methodology is still needed to define how to parcel out the provisions; this Note suggests that either the tiered scrutiny model or the evidentiary model could serve this purpose.

One might observe that revising the ballot will not necessarily address the issue of burdens imposed by petitions: voters may well vote the costliest mini-petitions into law. Like the evidentiary model, revising the ballot aims to faithfully implement the underlying formalistic evils—such as logrolling and fraud upon voters, but not burdensome initiatives—that the single subject rule facially aims to eliminate.²⁸³ When presented with mini-petitions, voters will at least realize that they can choose how much change to effect.

283. This Note does not presume to decree with certainty that only one policy motivates the single subject rule in every jurisdiction. Accordingly, the Note presents different solutions that are more closely hewn to meet different policy interests.

IV. CONCLUSION

Unlike other issues where people take to the streets in protest of judicial restrictions on citizen rights, there have not been too many riots, tea parties, or marches against judicial application of nebulous single subject rules. Moreover, petitions are making it to the ballot in nearly every state that permits direct democracy; although many are rejected, some are approved. Even after the single subject rule was adopted in Colorado, for example, measures have been approved by voters on such controversial topics as campaign finance, the minimum wage, marriage, corruption, medical marijuana, school funding, abortion, term limits, and the environment.²⁸⁴ What's more, citizens in little more than half of U.S. states do not even have the means to pass laws or constitutional amendments through initiative, and yet they still survive and flourish.²⁸⁵ Accordingly, one may counsel leaving well enough alone: citizens are passing some measures, the judiciary and statutory restrictions keep things under control, and all is well. Why should anything change?

Many may have said the same about other aspects of jurisprudence. No one was marching asking for expanded rights for criminal defendants, even though everyone knew that many states' practices had a disparate oppressive impact on racial and ethnic minorities. There were no tea parties demanding "One person, one vote," though it may have been apparent that districting practices were keeping incumbents in power despite their falling out of touch with voters' needs.²⁸⁶ There were no riots before the passage of the right of initiative in the 24 states that have it. Yet each and every improvement to the legal and political system mentioned above became, over time, ensconced in fundamental jurisprudence as a mechanism to empower citizens.

Revising the application of the single subject rule, though less momentous than the aforementioned examples, serves important societal functions. In a nation where partisan politics has resulted in increased polarization,²⁸⁷ political frustration can arise in states where both the executive and legislative branches are captured by

284. *Colorado Ballot Issue History (By Year)*, COLO. GEN ASSEMBLY, <http://www.leg.state.co.us/lcs/ballothistory.nsf/> (last visited Jan 1, 2010).

285. Twenty-six states do not have initiative processes. See INITIATIVE AND REFERENDUM INST., INITIATIVE USE (2009), <http://www.iandrinstitute.org/IRI%20Initiative%20Use%20%281904-2008%29.pdf>.

286. See *Thornburg v. Gingles*, 478 U.S. 30 (1986).

287. Pildes, *supra* note 234.

2011] JUDICIAL APPLICATION OF THE SINGLE SUBJECT RULE 885

one political party. Allowing losing party members to leverage the initiative process without undue interference from courts restores some of the political balance and gives concerned citizens the ability to pursue policy goals refused by the legislature. Moreover, clearly articulated rules allow citizens to plan and coordinate their grassroots movements with reasonable certainty. Unclear rules, on the other hand, may demoralize citizens and diminish political participation. They also position the judiciary as yet another government roadblock that stifles the voice of the people. Furthermore, balanced and consistently applied rules also increase judicial legitimacy,²⁸⁸ promoting the courts as fair, apolitical adjudicators of issues that are, by their nature, political. They also allow people to fully realize the political rights their states have recognized, especially people who do not have sufficient economic resources to influence state legislatures—the same insular minorities that *Carolene Products*²⁸⁹ brought to our attention. Finally, inasmuch as initiatives aim to unclog legislative roadblocks and lessen political corruption, limiting initiatives through obfuscated or oppressive rules fails to succeed in either respect, while at the same time furthering the notion that courts are simply kowtowing to the interests of the legislature.

In the end, this Note does not advocate leaving laws, constitutions, and the good of the people as a whole to the unfettered will of voters who can change the dynamics of a state through an undeliberated push of the button—or press of a stylus—every year or two. As Jean-Jacques Rousseau noted about popular majorities, “*The people is never corrupted, but it is often misled, and only then does it seem to will what is bad.*”²⁹⁰ Limiting substantial burdens on people’s rights or on the government’s ability to fulfill its traditional responsibilities is part of America’s venerable political tradition of trying to avoid the tyranny of the majority. In a nation that believes in separation of powers as a means to a stable society, popular legislative power—when appropriately channeled—may serve to further the common good and correct the machinery of representative government. In the end, however, it is representative government—with all of its institutional machinery, compensating measures, and expertise—that is the most resilient means of coordinating across large masses: the means guaranteed by our Constitution to the states. Direct democracy, however valuable, can only be a pressure

288. See *Planned Parenthood v. Casey*, 505 U.S. 833, 869 (1991) (plurality opinion).

289. *United States v. Carolene Prods. Co.*, 304 U.S. 144, 152–53 n.4 (1938).

290. ROUSSEAU, *supra* note 6, at 72.

886 NYU ANNUAL SURVEY OF AMERICAN LAW [Vol. 66:825

valve. This Note thus ends as it began, with Rousseau's admonition that "if there were a nation of gods, it would govern itself democratically. A government so perfect is not suited to men."²⁹¹

291. *Id.* at 114.

2011] JUDICIAL APPLICATION OF THE SINGLE SUBJECT RULE 887

APPENDIX A:
 COLORADO'S CASE LAW ON INITIATIVES, 1995-2009

BALLOTED	INIT.#	CASE	REPORTER	YR	DATE	THEME	Struck down Initiative	No explanation	Single subject	Misleading initiative	Misleading title/summary	Catchphrase	Conflicting titles	Board jurisdiction	Procedural	Inadequate impact stmt
1995		In re Title, Ballot Title and Submission Clause, and Summary for Proposed Initiated Constitutional Amendment Concerning Suits Against Nongovernmental Employers Who Knowingly and Recklessly Maintain an Unsafe Work Environment. (en banc)	898 P.2d 1071	1995	06/30/1995	Employment	0									
1995		In re Title, Ballot Title, Submission Clause, and Summary Adopted April 5, 1995, by Title Board Pertaining to a Proposed Initiative Public Rights in Waters II (en banc)	898 P.2d 1075	1995	06/30/1995	Environment	1		1							
1995		In re Title, Ballot Title and Submission Clause, and Summary With Regard to a Proposed Petition for an Amendment to Constitution of State Adding Section 2 to Article VII (Petition Procedures) (en banc)	900 P.2d 104	1995	06/30/1995	Petition	1		1	1						
1995		In re Title, Ballot Title and Submission Clause, and Summary with Regard to a Proposed Petition for an Amendment to Constitution of State of Colo. Adding Subsection (10) to Sec. 20 of Art. X (Amend Labor 25). (en banc)	900 P.2d 121	1995	07/19/1995	Taxes	1		1	0						
1995		In re Title, Ballot Title and Submission Clause, and Summary With Regard to a Proposed Petition for an Amendment to the Constitution of State of Colo. Adding Section 2 to Article VII (Petitions) (en banc)	907 P.2d 586	1995	12/04/1995	Petition	0		1	0						

C = Challenges; V = Violations

1995	In re Title, Ballot Title and Submission Clause, and Summary with Regard to a Proposed Petition for an Amendment to the Constitution of the State of Colo. Adding Paragraph (d) Subsection (8) of Section 20 of Article X (Amend Labor No.32) (en banc).	908 P.2d 125	1995 12/18/1995	Taxes	0	1	0	1	0	1	0	1	0
1996	In re Title, Ballot Title and Submission Clause, and Summary Adopted November 1, 1995, By Title Bd. Pertaining to a Proposed Initiative on Trespass-Streams with Flowing Water (en banc)	910 P.2d 21	1996 01/22/1996	<not single-subject issue>	0	1	0	1	0	1	0	1	0
1996	In re Proposed Ballot Initiative on Parental Rights (en banc)	913 P.2d 1127	1996 04/01/1996	Children	0	1	0	1	0	1	0	1	0
1996	I In re Proposed Initiative Bingo-Raffle Licenses (I) (en banc)	915 P.2d 1320	1996 04/29/1996	Gambling	0	1	0	1	0	1	0	1	0
	II		1996 04/29/1996	Gambling	0	1	0	1	0	1	0	1	0
1996	In re Proposed Initiative 1996-4 (en banc)	916 P.2d 528	1996 05/13/1996	Taxes	1	1	1	1	1	1	1	1	1
1996	In re Title, Ballot, Title and Submission Clause, and Summary Pertaining to Proposed Initiative on Parental Choice in Educ. (en banc)	917 P.2d 292	1996 05/28/1996	Children	0	1	0	1	0	1	0	1	0
1996	In re Title, Ballot Title and Submission Clause, and Summary for Proposed Initiated Constitutional Amendment 1996-3 Adopted on April 3, 1996, and Motion for Rehearing Denied on April 17, 1996 (en banc) (per curiam)	917 P.2d 1274	1996 06/10/1996	<not single-subject issue>	0	TIE	0	1	0	1	0	1	0
1996	In re Title, Ballot Title, Submission Clause, and Summary Adopted March 20, 1996, By the Title Bd. Pertaining to Proposed Initiative 1996-6 (en banc)	917 P.2d 1277	1996 06/10/1996	Environment	0	1	0	1	0	1	0	1	0
1996	In re Title, Ballot Title, Submission Clause, and Summary for Proposed Constitutional Amendment Concerning 1996-15 (en banc) (per curiam)	918 P.2d 225	1996 06/17/1996	<not single-subject issue>	0	0	0	1	0	1	0	1	0
1996	In re Title, Ballot Title, Submission Clause, and Summary Adopted Apr. 17, 1996 by Title Setting Bd. Pertaining to a Proposed Initiative Statute Proposed by Apple and Meeker (1996-17) (en banc)	920 P.2d 798	1996 06/24/1996	Environment	1	1	0	1	0	1	0	1	0

THE POSTER'S PLIGHT: BRINGING THE PUBLIC DISCLOSURE TORT ONLINE

JAIME A. MADELL*

INTRODUCTION

On December 26, 2009, “Lee” posted the following to a Facebook¹ discussion board:

Recently, someone I had denied acces [sic] as a friend wound up friending a friend of mine and gained access to my pictures and videos. She then posted them on her wall and now everyone [sic] can see them. These are pictures and videos of my kids. I contacted facebook and I see the account is no longer valid. My suspicion is that she will create a new profile with different information and repost the videos/pictures as they are surley [sic] saved to her hard drive. Is this illegal and what can I do about it. The person in question is my mother and me and my siblings were taken away as a result of neglect. This was close to 16 yrs [sic] ago and I do not want any contact with her.²

Lee’s question is remarkably nuanced. We might rephrase it like so: Can I hold somebody legally liable for (a) downloading a picture I have posted to an online social network (OSN) with the intent that it be viewed only by a specified group of people and (b) re-posting it so that it can be seen by people to whom I have not provided access?

There are two answers here, one more obvious than the other. The simple answer can be found on the discussion board itself. A sympathetic “Mathew” responded that: “The reality is you should not post stuff to the Internet you are worried about people seeing.

* J.D. Candidate 2011, New York University School of Law; Notes Editor, New York University Annual Survey of American Law; M.M., Honors, Northwestern University School of Music; B.A., *summa cum laude*, Columbia University. The Author wishes to thank Katherine Strandburg for her dedicated advising, the staff of the *New York University Annual Survey of American Law* for their assiduous editing, and Helen Nissenbaum, Ira Rubinstein, and the Privacy Research Group for their helpful comments.

1. As most readers know, Facebook is a popular online social network. For a discussion of online social networks, see *infra* Part I.

2. Posting of Lee, to *Stealing Pics from Someone and Reposting Them on Your Wall*, FACEBOOK (Dec. 26, 2009, 4:07 PM) (on file with author).

Once it's online, it's almost impossible to keep it contained."³ There's a lot right about this straightforward answer. Broadly speaking, the law does not protect information that people freely disclose, even if the extent of that disclosure is not commensurate with the eventual scope of dissemination. This is because privacy law in the United States is defined in large part by what many privacy scholars have termed the public-private dichotomy.⁴ This dichotomy stems from a notion, embedded in Fourth Amendment jurisprudence, that what one discloses to third parties is no longer private in the eyes of the law. According to this line of thinking, the answer to Lee's question is simply "no."

This is not to say that Lee does not have a cause of action to pursue. Hence the second answer: Lee might look to the public disclosure tort for help. A relatively recent addition to American tort law, this tort aims to police those who give publicity to private facts.⁵ In practice, however, the tort is both weak and doctrinally unstable, due in large part to varying approaches to its "legitimate concern" and "reasonableness" prongs.⁶ Even before the days of the Internet, these requirements—protected information must be (a) non-newsworthy and (b) reasonably expected to remain private—were vague. This lack of clarity has been exacerbated by recent technological developments, many of which pose serious challenges to norms of information flow. For example, whereas it was once possible to assume that the newsworthiness of a given piece of information could be determined by using the media's decision to publish it as a benchmark, the proliferation of blogs and micro-journalists has trounced the media-as-gatekeeper norm and rendered this proxy unreliable. The interpretative morass that has followed such developments significantly hampers the ability of harmed individuals to seek and receive legal redress.⁷

The failure of the public disclosure tort might not be a devastating loss in the online context, were information flows on online utilities to more thoroughly protect and preserve users' privacy ex-

3. Posting of Mathew, to *Stealing Pics from Someone and Reposting Them on Your Wall*, FACEBOOK (Jan. 9, 2010, 2:58 PM) (on file with author).

4. See, e.g., HELEN NISSENBAUM, *PRIVACY IN CONTEXT* 141 (2010).

5. The tort is thusly defined in the Second Restatement:

"One who gives publicity to a matter concerning the private life of another is subject to liability to the other for invasion of his privacy, if the matter published is of a kind that (a) would be highly offensive to a reasonable person, and (b) is not of legitimate concern to the public."

RESTATEMENT (SECOND) OF TORTS (1977). See *infra* Part II.

6. See *infra* Part II.A.

7. NISSENBAUM, *supra* note 4; see *infra* Part II.

pectations. A “perfect” Facebook would have preserved Lee’s expectation that his posting decision would not result in sharing with his mother. This hypothetical Facebook would implement principles of information flow that satisfy a user’s expectation that his decision to share only with specific parties will be respected.

If principles of information flow (expressed through OSN source code) preserved such expectations perfectly, there would perhaps be no need for law—technology would suffice. Unfortunately, prevailing code falls far short of such a lofty goal. Instead of bridging the gap between online information handling and individuals’ expectations, code widens the divide. As Helen Nissenbaum explains in her book *Privacy in Context*, technologies such as OSNs handle information in ways that conflict with prevailing norms of information flow.⁸ Unlike OSN policies,⁹ these highly specific norms develop within a particular context (e.g., friendship) to match the expectations of those contexts (e.g., a photo shared with a close friend will not be disseminated).

Broadly speaking, this Note explores how law and code can help each other protect privacy online. As Lee’s story demonstrates, technological change can encroach upon our entrenched notions of privacy. But by updating law and code to better match the behavioral realities of online culture, we can ensure that people do not find themselves caught in a web of unfamiliar norms. More specifically, this Note aims to demonstrate how the law-and-code collaboration can reinvigorate the public disclosure tort and help plaintiffs in Lee’s situation.

Thus, though the problem of privacy online is far-reaching, this Note deals with only the issue of nonconsensual re-posting described in Lee’s thread, a problem I call the Poster’s Plight. I focus on the Poster’s Plight for two reasons. First, the lack of legal protection for victims of nonconsensual re-posting clearly demonstrates the inadequacy of the current public disclosure tort. This inadequacy is troubling, as the public disclosure tort is one of the few legal tools available to plaintiffs in situations such as Lee’s. Second, the Poster’s Plight sheds light on the intersection of four important issues in the analysis of online privacy: (1) technological developments and the privacy risks they pose, (2) legal doctrines and their

8. *Id.*

9. For the purposes of this Note, a “policy” can be defined as a technologically determined principle of information flow. An example familiar to most readers would be the policy—built into email protocol—that senders cannot tell if their emails have been received.

ability to mitigate those risks, (3) code-based structures that mediate such risks, and (4) prevalent notions of privacy.

In Part I of this Note, I examine the psychology of privacy on OSNs and the divide between the contextually rich public perceptions of privacy and the dichotomous view built into OSN code. In particular, I demonstrate how OSN privacy controls fail to provide users with protections that fit their complex expectations of privacy. I also argue that because code alone cannot solve OSN privacy problems, law must intervene. In Part II, I examine the use of the public disclosure tort as a privacy tool, explain how its unstable doctrine fails to capture our basic intuitions of privacy, and suggest a more context-friendly approach to the reasonableness analysis. In Part III, I develop a hybrid legal–technical solution to the problem of nonconsensual picture re-posting on OSNs; in short, I propose a new OSN functionality that allows users to explicitly express disclosure preferences for the pictures they post.¹⁰ I argue that, in conjunction with a modified reasonableness analysis, this functionality can go a long way to resuscitating the public disclosure tort in the digital age and helping restore context to interactions on online social networks.

I.

CONTEXTUALIZING THE PROBLEM

A. OSNs, Picture Posting, and Related Risks

This Section provides a brief overview of OSNs and some basic functionalities central to the present discussion. We start with a basic definition: “[OSNs are] web-based services that allow individuals to (1) construct a public or semi-public profile within a bounded system, (2) articulate a list of other users with whom they share a connection, and (3) view and traverse their list of connections and

10. It is important to stress that this Note deals only with the re-posting of pictures. One might reasonably ask why this would be so, as the re-posting of text is obviously possible and problematic. My answer stems from an intuitive sense that pictures convey information with more immediacy than text. This intuition is at least partly reflected in the psychological literature. See, e.g., Kevin S. Douglas, David R. Lyon & James R. P. Ogloff, *The Impact of Graphic Photographic Evidence on Mock Jurors' Decisions in a Murder Trial: Probative or Prejudicial?*, 21 LAW & HUM. BEHAV. 485, 492 (1997) (showing that mock jurors arrive at more guilty verdicts when shown explicit photographs of mutilated victims than when simply given textual descriptions of a victim's physical condition); Adina Shmidman & Linnea Ehri, *Embedded Picture Mnemonics to Learn Letters*, 14 SCIEN. STUD. READING 159 (2010) (finding that children learn letter-sound combinations more efficiently when letter orthography is mapped onto familiar pictures).

those made by others within the system.”¹¹ Depending on the OSN, other functionalities can enrich this basic design.¹² Most OSNs allow users to decide who can access the material they share. The typical default setting ensures that “friends” have access to information, and strangers do not.

Although Facebook and MySpace tend to garner the most media attention, there are hundreds of OSNs in operation.¹³ OSN use represents a significant, and continually growing, portion of the leisure time pie. According to The Nielsen Company, as of December 2009, Internet users in the United States, the United Kingdom, Australia, Brazil, Japan, Switzerland, Germany, France, Spain, and Italy spend more than 5.5 hours per month on OSNs. This represents an 82% increase from 2008.¹⁴ Facebook, the top-ranked social networking destination, received approximately 109 million unique visitors in the month of December alone.¹⁵ According to Facebook, its members spend more than fifty-five minutes per day on the site.¹⁶

Nearly all OSNs provide an opportunity for users to submit small personal pictures to adorn their profiles. Facebook boasts particularly extensive photo features,¹⁷ allowing users to upload hundreds of high-resolution digital photos to their accounts. These photos are organized into albums, each of which features its own

11. danah boyd & Nicole B. Ellison, *Social Network Sites: Definition, History, and Scholarship*, 13 J. COMPUTER-MEDIATED COMM. (2007), <http://jcmc.indiana.edu/vol13/issue1/boyd.ellison.html>.

12. Facebook, in particular, offers users a multitude of communicative tools. Basic user activities include posting status updates (e.g., “I’m cooking dinner”), commenting on users’ home pages, posting pictures, playing games, and searching the full user database. FACEBOOK, <http://www.facebook.com> (last visited Oct. 7, 2010).

13. *List of Social Networking Websites*, WIKIPEDIA, http://en.wikipedia.org/wiki/List_of_social_networking_websites (last visited Jan. 18, 2011).

14. *Led by Facebook, Twitter, Global Time Spent on Social Media Sites up 82% Year over Year*, NIELSENWIRE (Jan. 22, 2009), <http://blog.nielsen.com/nielsenwire/global/led-by-facebook-twitter-global-time-spent-on-social-media-sites-up-82-year-over-year>.

15. *Top U.S. Web Brands and Site Usage: December 2009*, NIELSENWIRE (Jan. 14, 2010), http://blog.nielsen.com/nielsenwire/online_mobile/top-u-s-web-brands-and-site-usage-december-2009/. Facebook’s in-house statistics are equally impressive. Facebook claims to have 500 million active users, 250 million of whom update their status posts every day. *Statistics*, FACEBOOK, <http://www.facebook.com/press/info.php?statistics> (last visited Oct. 7, 2010).

16. *Statistics*, FACEBOOK, <http://www.facebook.com/press/info.php?statistics> (last visited Oct. 7, 2010).

17. See, e.g., *Help Center*, FACEBOOK, <http://www.facebook.com/help/?page=830> (last visited Oct. 7, 2010).

customizable privacy profile; that is, users can dictate exactly who can view each specific album. After the pictures are posted, those with access can comment on the pictures or “tag” people they recognize in the pictures.¹⁸ Tagging has the effect of making pictures searchable by the name(s) of the depicted person(s). Facebook allows tagging not only of Facebook users, but also of members of the general public. All users are able to “untag” themselves from photos (though non-members are required to join before they can untag themselves). For the purposes of clarity, I will refer to this sort of tagging as “name-tagging.”

A more advanced feature of Facebook allows a user to “share” a photo belonging to someone else. This means that the photo is displayed on the sharer’s page, though it cannot be viewed by those not granted access by the original poster. This form of re-posting respects the poster’s original privacy preferences. This is not to say, however, that Facebook has solved the re-posting problem. A user may circumvent privacy preferences by executing a print-screen command while viewing a photo to which he has access, saving it to his hard drive, and then uploading it to his own account over which he retains control of privacy settings. Alternatively, a user could simply click the “download” link that appears underneath photos in the Facebook viewing console.

Even if we ignore overt circumventions of expressed privacy preferences, each sharing decision on an OSN such as Facebook is loaded with risk. This is because OSN privacy settings are often not user-friendly. On Facebook, for example, a fair degree of technological competence is required to effectively calibrate privacy settings. Given the ease of information flow on Facebook, even a small technical oversight could lead to oversharing of private information. Such missteps are easy to make. Facebook itself encourages oversharing. First, default privacy settings upon joining Facebook are minimal, at best.¹⁹ Second, Facebook’s “recommended” settings provide very weak barriers to information flow. This perhaps reflects Facebook’s vested interest in keeping information flow on the site fluid, so as to encourage repeat visits to stay abreast of social “news” and content.

18. Recently, Facebook revised its photo viewing console. The newest version enables users to download photos posted to Facebook. This is particularly troubling from the perspective of the Poster’s Plight.

19. The All Facebook site chronicles Facebook’s default settings in graphic form. Nick O’Neil, *INFOGRAPHIC: The History Of Facebook’s Default Privacy Settings*, ALL FACEBOOK, <http://www.allfacebook.com/infographic-the-history-of-facebooks-default-privacy-settings-2010-05> (last visited Feb. 21, 2011).

One need only skim the headlines for tales of inadvertent over-sharing. Take the case of high school teacher AP, for example, who was forced to resign in November 2009 after school officials received an anonymous email complaining that she posted inappropriate language and photographs on her Facebook page.²⁰ The allegedly inappropriate photographs (only 10 out of 700 posted) merely depicted AP, apparently sober, visiting beer gardens on a trip abroad.²¹ According to news reports, AP's profile was visible only to a select group of friends that did not include students or parents.²² Furthermore, her status post simply remarked that she was on her way to "Bitch Bingo" at a local bar.

It is hard to find the justly punishable transgression here. AP seemingly took a great deal of care to maintain her Facebook privacy. If indeed she used the full panoply of technical protections offered by Facebook to block students from accessing her information, holding her responsible for a leak that could well have been the result of malicious intermeddling is patently unfair. Countering this intuition is an equally strong belief that what is online is not private. This belief fits neatly within the United States' model of at-will employment—a simple Google search for "fired because of Facebook" will reveal that surveillance of online postings has become standard practice.

Employment disputes are only a subset of the privacy mayhem fostered by picture posting on OSNs. It is beyond the scope of this Note to address the full range of privacy concerns stemming from OSN picture posting. It suffices to say that OSN picture posting is a widespread activity with significant risks to privacy. Given the popularity of OSNs and the expectations of their users, these are risks that both the code underlying OSNs and privacy law should mitigate.

In the following pages, I focus on code's role in mediating OSN privacy. In Part B, I argue first that code should be recognized as a malleable design feature and not an immutable characteristic. I also introduce the theory of contextual integrity to frame the potentially transformative effects of new technologies such as OSNs on prevailing norms of information flow. In Part C, I explain how current OSN code does not map adequately onto users' deeply contextual notions of privacy and thus violates contextual integrity.

20. Maureen Downey, *Facebook Flap in Barrow Raises Troubling Fairness Issues*, GET SCHOOLED (Nov. 13, 2009, 4:44 PM), <http://blogs.ajc.com/get-schooled-blog/2009/11/13/facebook-flap-in-barrow-raises-troubling-fairness-issues/>.

21. *Id.*

22. *Id.*

B. *The Digital Difference*

In 1996, when Amazon.com still looked a little bit like a Geocities homepage,²³ Judge Frank Easterbrook sparked a debate. Speaking at a conference on cyberlaw, Easterbrook suggested that rather than develop new modes of legal analysis specific to cyberlaw, legal scholars should simply apply existing doctrine to the novel problems posed by the Internet.²⁴ Easterbrook advanced a Coasean solution to the problems at the intersection of technology and copyright law. Instead of attempting to anticipate the effects of technology on marketplace outcomes and tailor policy to those predictions, Easterbrook suggested that we simply provide parties with a stable set of rules and an opportunity to bargain.²⁵

Easterbrook reasoned that “if we are so far behind in matching law to a well-understood technology such as photocopiers . . . what chance do we have for a technology such as computers that is mutating faster than the virus in *The Andromeda Strain*?”²⁶ This statement depends on a flawed assumption about the relationship between law and technology, namely, that technology exists as a complex creature beyond our control and elusive of law. As Lawrence Lessig pointed out in rebuttal to Easterbrook’s speech, many people assume either that “the nature of cyberspace is fixed—that its architecture, and the control it enables, cannot be changed—or that government cannot take steps to change this architecture.”²⁷ This is not the case. Human beings create code. And as a result, “code can change.”²⁸ In assessing the frequent collisions between privacy and technology, we should keep in mind that the systems we

23. *How 20 Popular Websites Looked When They Launched*, TELEGRAPH (Sep. 2, 2009, 5:04 PM), <http://www.telegraph.co.uk/technology/6125914/How-20-popular-websites-looked-when-they-launched.html>.

24. See Frank H. Easterbrook, *Cyberspace and the Law of the Horse*, 1996 U. CHI. LEGAL F. 207 (1996).

25. *Id.* at 210.

26. *Id.*

27. Lawrence Lessig, Commentary, *The Law of the Horse: What Cyberlaw Might Teach*, 113 HARV. L. REV. 501, 505 (1999). For a more thorough discussion of the relationship between code and law, see LAWRENCE LESSIG, CODE: AND OTHER LAWS OF CYBERSPACE, VERSION 2.0 (2006). For a similar discussion of code, architecture, and OSNs, see Gordon Hull, Heather Richter Lipford & Celine Latulipe, *Contextual Gaps: Privacy Issues on Facebook*, ETHICS & INF. TECH. 6–7 (April 2010), available at <http://www.springerlink.com/content/072730305020wm26/>. Because “the architecture of an online environment is a function of the code that creates it, and because that code can be changed, the coding and interface of a site can make an enormous difference both in how much privacy users have, and how they experience their privacy.” *Id.* at 7.

28. Lessig, *supra* note 27, at 506.

confront are not autonomous, natural, or immutable. They are subject to our control and influence.

Easterbrook's speech is problematic in another important regard. Implicit in Easterbrook's argument that there should not be a distinct law of cyberspace is the suggestion that principles of law can be developed outside the technological context in which the law applies. With regard to copyright law, Easterbrook suggested that in reducing the effective cost of copying, contemporary computing technology simply "continues a trend that began when Gutenberg invented movable type."²⁹ Turning to privacy, a similar point could be made about posting pictures online; in reducing the effective cost of sharing pictures, picture posting continues a trend that began when people started sharing pictures.

These assertions are weak because they overlook the fact that the internet has done far more than lower transaction costs. Rather, by lowering transaction costs, it has also effectuated a degree of information transfer and processing that alters the nature of contemporary data-handling protocol. Consider the practice of "aggregation," defined as "the gathering together of information about a person."³⁰ While data aggregation is not a new practice, today "the data gathered about people is significantly more extensive, the process of combining it is much easier, and the computer technologies to analyze it are more sophisticated and powerful."³¹ Even if all the aggregated information is in the public domain, the aggregation still presents a challenge to individual privacy insofar as the very nature of that information is altered. It is a difference of kind *and* degree.

Aggregation, readily accessible database technologies, easily portable data, and statistical-analysis and data-mining tools represent technological transformations that "shape the many different ways computerized record-keeping systems and practices impinge on privacy and affect experiences."³² By changing the way information can be accessed and used, these transformations "have affected the state and practice of electronic engagement with personal information, which, in turn, are experienced as threats to privacy."³³ In other words, as technological transformation continues to accelerate, people find themselves racing to rein technology

29. Easterbrook, *supra* note 24, at 208.

30. Daniel J. Solove, *A Taxonomy of Privacy*, 154 U. PA. L. REV. 477, 506 (2006).

31. *Id.*

32. NISSENBAUM, *supra* note 4, at 38–42.

33. *Id.* at 44–45.

back within boundaries that fit prevailing privacy norms.³⁴ If and when they are unable to do so, they might be forced to accept new, less optimal privacy norms. Often, new norms are “forced” to develop because people simply don’t realize that their offline expectations are violated by online policies of information flow.

Consider the example of courtroom videotaping. Unlike many countries, the United States guarantees its citizens a front-row seat to the criminal adjudication process. As recently emphasized in *Presley v. Georgia*,³⁵ both the First and Sixth Amendments require trial courts “to take every reasonable measure to accommodate public attendance at criminal trials.”³⁶ In January 2010, Judge Vaughn Walker suggested videotaping the now-famous Proposition 8 case, *Perry v. Schwarzenegger*.³⁷ Although this idea might seem like a logical step forward in light of the public’s right to view a trial, when considered with an eye to technology’s ability to transform norms, the notion of videotaping becomes much more troublesome. Imagine the effect of a piece of testimony taken out of context. Were a courtroom visitor to transmit a videotape to friends and colleagues, the information would likely enjoy limited viewership and do little damage to the witness. In contrast, were a fifteen-second clip of the same testimony distributed on YouTube, the scope of dissemination and its resultant effects could be astounding. What is in one case harmless gossip can, via technological magnification, become a potentially damning news story at a national level.

We have arrived at two important principles. First, the drawbacks, benefits, and features of web-based technologies (including OSNs) are mediated by code. This code is within our control. Second, new technologies can violate and transform established pri-

34. In attempting to understand the implications of such rapid transformation, Viktor Mayer-Schönberger’s work on digital memory is helpful. He contextualizes the history of technology in the human species’ collective effort to remember. He writes: “Since the early days of humankind, we have tried to remember, to preserve our knowledge, to hold on to our memories, and we have devised numerous devices and mechanisms to aid us.” VIKTOR MAYER-SCHÖNBERGER, *DELETE: THE VIRTUE OF FORGETTING IN THE DIGITAL AGE* 48–49 (2009). Language, books, and computers all facilitate memory, to a different degree. Whereas “through millennia, forgetting has remained just a bit easier and cheaper than remembering,” we are now, by virtue of the switch from analog to digital technology, required to confront the question “of whether we would like to remember everything forever if we could.” *Id.*

35. 130 S. Ct. 721 (2010).

36. *Id.* at 725.

37. Lisa Leff, *Judge: Gay Marriage Trial Can Be Shown on YouTube*, ABCNews (Jan. 7, 2010), <http://abcnews.go.com/Business/wireStory?id=9501129> (discussing *Perry v. Schwarzenegger*, 704 F. Supp. 2d 921 (N.D. Cal. 2010)).

vacy norms that people might reasonably assume still prevail despite technological transformation. How can we methodically apply these principles to specific cases, such as picture posting on OSNs? Here is where contextual integrity comes into play.

Contextual integrity provides a heuristic for both predicting when a given transformation tramples upon existing norms and concluding normatively whether the underlying technology's benefits outweigh its detrimental impact on norms. The normative analysis centers on the relationship between new technologies and "context-relative information norms" (information norms).³⁸ Information norms are made up of four components: contexts, actors, attributes, and transmission principles. Contexts are defined as "structured social settings" that give rise to particular roles and relationships.³⁹ Actors fall into one of three categories: senders, receivers, or subjects.⁴⁰ Attributes are types of information transferred (e.g., "medical information" or "contact information").⁴¹ Transmission principles are constraints on information flow; the notion of confidentiality is a clear example of such a principle.⁴² When a technology impinges on any of these information-norm components as they existed prior to that technology, the technology "is flagged as violating entrenched informational norms and constitutes a prima facie violation of contextual integrity."⁴³

The normative analysis compares "entrenched normative practices [and] novel alternatives or competing practices on the basis of how effective each is in supporting, achieving, or promoting relevant contextual values."⁴⁴ According to Nissenbaum, "if the practices prescribed by entrenched informational norms are found to be less effective . . . than challengers . . . [or] novel practices," existing norms and practices can be justifiably replaced.⁴⁵ Procedurally, this analytic method should sit comfortably with legal scholars, given its use of a balancing test. Substantively, however, the normative component of contextual integrity will likely run up against objections.

Skeptics might complain that contextual integrity merely creates problems to solve. In other words, some shifts in norm struc-

38. NISSENBAUM, *supra* note 4, at 140.

39. *Id.* at 132.

40. *Id.* at 141.

41. *Id.* at 143.

42. *Id.* at 145.

43. *Id.* at 150.

44. *Id.* at 166.

45. *Id.*

tures are entirely natural. Take instant messaging, for example. As any user of Google Chat or America Online Instant Messenger knows, the rules of etiquette online are quite different from those in physical space. Consider the common acronym “brb.” Although this expands to “be right back,” people use “brb” for a wide variety of purposes. While they sometimes use it to pause a conversation while they do something quickly away from the computer, they might actually use it to terminate a conversation indirectly. Such callous behavior would never be tolerated in physical space—imagine calling someone and then leaving them on hold for two hours.

To the extent that people have grown accustomed to these new rules of conduct, they have arguably been forced into a new set of norms by the underlying technology. Yet not all shifts in norm structure are unnatural. The semantics of instant messaging—the creation of “lol,” “brb,” and their cousins—developed from within the instant messaging user community itself. And while this new semantics was born in response to new technology, it was not dictated to the masses by that technology; in other words, people didn’t develop “lol” because they were prohibited from typing “laughing out loud.”

Contextual integrity does not police such user-motivated behaviors. On the contrary, contextual integrity is about challenges to norm structure that are made through top-down technology-to-user mandates. Whereas instant messaging provides a good example of the former, OSN privacy represents the latter. As explained below, OSN users are not collectively developing notions of online privacy that are blended into OSN code structure. Users’ notions are being excluded from the code structure, and users are not always aware of the forced norm shift.

To use an example discussed in more detail below, when users post pictures on an OSN, they might assume that the norms that govern offline sharing apply equally on the OSN, despite the fact that the OSN code does not let them express such norms. Because users cannot express these norms, over time, the community might assume they simply do not apply. Eventually, a new set of norms will develop in response to the technical limitations of the OSN code. Because this new set of norms will be born from technical limitation, not user preference, it will be both artificial and suboptimal from a privacy perspective. This is the dangerous process that contextual integrity attempts to prevent.

Even accepting this, one can still mount what Solove calls an “I’ve got nothing to hide” argument.⁴⁶ Solove notes that in the context of government surveillance, “many people believe that there is no threat to privacy unless the government uncovers unlawful activity, in which case a person has no legitimate justification to claim that the [activity] remain private.”⁴⁷ The argument is that if you have not posted pictures or statements that could tarnish your image or land you in trouble if revealed, you have no reason to be worried that the information will reach a broader audience than you originally anticipated.

This argument fails because it assumes “that privacy is about hiding bad things.”⁴⁸ As Solove notes, the harm here is not a question of “dead bodies.”⁴⁹ Rather, the harm is a matter of dignity and autonomy as enshrined in contextually relevant norms of information transfer. And, to the extent that users might actually share less and refrain from maximizing the full communicative and creative potential of OSNs for fear of privacy violations, it is a question of social utility.⁵⁰ If, as contextual integrity argues, people rely on contextually relevant and community-prevalent norms of information flow to mitigate harms to autonomy and dignity, then understanding when technological systems violate these norms is an important social goal. When, as is the case with OSNs, such a violation takes place, a remedy is necessary. If no remedy is provided, the technological transformations might end up forcing users to adopt new, less optimal information norms.

To review, applying the contextual integrity heuristic to OSNs requires a comparison of information norms as they exist in users’ collective conscious and the norm-impinging features of OSN code. Insofar as they exist before the technological transformation, such previously entrenched norms can be understood to cognitively pre-date the transformation. In the next section, I will attempt to high-

46. Daniel J. Solove, *“I’ve Got Nothing To Hide” and Other Misunderstandings of Privacy*, 44 SAN DIEGO L. REV. 745 (2007).

47. *Id.* at 746.

48. *Id.* at 764.

49. *Id.* at 768.

50. Readers will likely note the apparent contradiction between this assertion and the argument below that OSN users are myopic. *See infra* Part I.C. A point of clarification is useful here. In arguing that users are myopic, I am not suggesting that they all are actively aware of the privacy risks of OSN participation or the technologically driven shift in norms. Rather, I am arguing that, even if users understand the privacy risks, the free market will be unable to select for privacy-friendly firms because users will irrationally discount all long-term privacy costs (including the ones they thoroughly understand).

light and identify information norms that cognitively predate OSNs, explain how OSNs violate contextual integrity by forcing users to unwittingly abandon these information norms, and explain the particular role code plays in effecting this violation.

C. *What Web Am I Surfing?*

Recall Mathew's response to Lee's post: if you want something kept private, do not put it on the web. The implication here is clear. Assuming that the user has perfect information about the prevailing norms and technical details of an OSN, the user can be counted on to make rational decisions protecting her privacy. However convenient, this "user beware" attitude only holds if the web the user *thinks* he is surfing is the one that he *is* surfing. In other words, Mathew's logic is only appropriate if Lee's intuitions with regard to the relationship between privacy norms on Facebook and those that cognitively predate Facebook's code are respected by the prevailing code structure. As it turns out, this essential condition is not satisfied by current Facebook code; to understand why, we need to look at both empirical evidence on information norms that predate Facebook and the transformations effected by current code.

Revealing information norms is difficult detective work. It is no easy task to understand exactly what people expect from OSNs in terms of privacy, even when we ask them. For example, despite expressing significant privacy concerns, people seem to love to disclose information. The popular press likes to call this apparent contradiction the "privacy paradox."⁵¹ Examples abound in the academic literature. For example, Acquisti and Gross report that nearly 16% of surveyed members of an American university who expressed concern for a hypothetical scenario in which a stranger discovered his or her schedule of classes and home address on an OSN nevertheless listed *both* pieces of information on their OSN profiles.⁵² This same study revealed that over 89% of undergradu-

51. See, e.g., Andy Greenberg, *The Privacy Paradox*, FORBES.COM (Feb. 15, 2008, 6:00 AM), http://www.forbes.com/2008/02/15/search-privacy-ask-tech-security-cx_ag_0215search.html; Brad Stone, *Our Paradoxical Attitudes Toward Privacy*, BITS BLOG (July 2, 2008, 3:56 PM), <http://bits.blogs.nytimes.com/2008/07/02/our-paradoxical-attitudes-towards-privacy/>.

52. Alessandro Acquisti & Ralph Gross, *Imagined Communities: Awareness, Information Sharing, and Privacy on the Facebook*, in PRIVACY ENHANCING TECHNOLOGIES: SIXTH INTERNATIONAL WORKSHOP 36, 51 (George Danezis & Philippe Golle eds., 2006), available at <http://privacy.cs.cmu.edu/dataprivacy/projects/facebook/facebook2.pdf>.

ate respondents who expressed the most concern for threats to personal privacy joined Facebook.⁵³

Similarly, Stutzman found that while undergraduate and graduate students expressed concern about the consequences of sharing identity information on OSNs, more than 50% of these respondents listed their name, academic classification, gender, email, picture, major, birthday, home town, high school, relationship status, address information, interests, and political views on their Facebook profiles.⁵⁴ While Stutzman does not directly address privacy settings, Gross and Acquisti do: they were able to publicly view all but three of over 4000 profiles studied on their school network.⁵⁵ Govani and Pashley report similar results. While 84% of undergraduate respondents were aware that they could restrict who could view their Facebook profiles, fewer than 48% actually used Facebook's privacy settings.⁵⁶

Despite these 2005 findings, there is a fair amount of evidence that the use of privacy settings to mediate the openness of OSNs is on the rise. A 2007 report by the Pew Internet & American Life Project notes that 59% of teens with active profiles on OSNs claim that these profiles are visible only to friends.⁵⁷ Similarly, Young and Quan-Haase found that while large percentages of undergraduate survey participants included school name, email address, birth date, personal photos and photos of friends on their Facebook profile, 64% limited profile visibility to "only friends."⁵⁸ This trend seems equally prevalent on European OSNs.⁵⁹

53. *Id.* at 46.

54. Frederic Stutzman, *An Evaluation of Identity-Sharing Behavior in Social Network Communities*, 3 J. INT'L DIGITAL MEDIA & ARTS ASS'N 10, 15–16 (2006).

55. Ralph Gross & Alessandro Acquisti, *Information Revelation and Privacy in Online Social Networks*, in PROCEEDINGS OF THE WORKSHOP ON PRIVACY IN THE ELECTRONIC SOCIETY 71, 77 (Sabrina De Capitani di Vimercati & Roger Dingledine, eds., 2005).

56. Tabreez Govani & Harriet Pashley, *Student Awareness of the Privacy Implications When Using Facebook* 8 (2005) (unpublished manuscript), available at <http://lorrie.cranor.org/courses/fa05/tubzhlp.pdf>.

57. Amanda Lenhart & Mary Madden, *Teens, Privacy, and Online Social Networks: How Teens Manage Their Online Identities and Personal Information in the Age of MySpace*, PEW INTERNET AND AMERICAN LIFE PROJECT 26 (2006).

58. Alyson L. Young & Anabel Quan-Haase, *Information Revelation and Internet Privacy Concerns on Social Network Sites: A Case Study of Facebook*, in C&T '09: PROCEEDINGS OF THE FOURTH INTERNATIONAL CONFERENCE ON COMMUNITIES & TECHNOLOGIES 268 (2009).

59. Sonja Utz & Nicole Krämer, *The Privacy Paradox on Social Network Sites Revisited: The Role of Individual Characteristics and Group Norms*, CYBERPSYCHOLOGY: J. PSYCHOSOCIAL RES. ON CYBERSPACE (2009), <http://cyberpsychology.eu/view.php?cisloclanku=2009111001&article=2>.

It is certainly good news that people are becoming more active in protecting their privacy on OSNs. Yet this trend, however fortunate, does not offer much insight into how OSN users psychologically relate to the often-fragile balance between public and private information online—it might just show that, by means of using privacy settings, OSN users are explicitly recognizing that the degree of publicity or privacy on OSNs is, to some degree, within their control. Put differently, unless the privacy settings examined are co-extensive with users' actual notions of privacy, evidence of their use or neglect tells us little about what web users think they are surfing.

Filling this data void, numerous cutting-edge studies demonstrate that people expect congruence between the information norms that cognitively predate OSNs and the transmission principles that OSN code allows.⁶⁰ Reporting results from interviews with teenage OSN users, Livingstone found that users' notion of "friends" online tracks similar offline notions.⁶¹ For example, one student drew a distinction between "best friends," "friends I'm good friends with," "friends that I see every so often," and "people that I do not really talk to."⁶² Similarly, Lampe, Ellison, and Steinfield found that "despite changes in the technical ability of non-university people to join Facebook," student users feel that Facebook is a "student-only" site.⁶³ Thus, students seem to expect student-centric offline information norms to persist in the online context.

As several commentators have pointed out, OSN code has failed to respect these expectations.⁶⁴ One of the most powerful examples springs from the very nature of online social engagement. In describing OSNs, danah boyd has used the term "networked publics."⁶⁵ A networked public is at least in part defined by its mediated nature—"the network mediates the interactions between members of the public."⁶⁶ In unmediated environments, "the boundaries and audiences of a given public are structurally defined" by the real-world physics. Thus, as boyd points out, the audi-

60. See *supra* notes 55–57 and accompanying text.

61. Sonia Livingstone, *Taking Risky Opportunities in Youthful Content Creation: Teenagers' Use of Social Networking Sites for Intimacy, Privacy, and Self-Expression*, 10 *NEW MEDIA & SOC.* 393, 405 (2008).

62. *Id.*

63. Cliff Lampe, Nicole B. Ellison & Charles Steinfield, *Changes in Use and Perception of Facebook*, ASS'N FOR COMPUTING MACHINERY CONF. 721, 729 (2008).

64. See generally NISSENBAUM, *supra* note 4, at 231–56.

65. danah boyd, *Social Network Sites: Public, Private, or What?*, 13 *KNOWLEDGE TREE* 6 (2007), http://kt.flexiblelearning.net.au/tkt2007/wp-content/uploads/2007/05/edition_13.pdf.

66. *Id.* at 8.

ence that watches you trip on the curb is “restricted to those present in a limited geographical radius at a given moment in time.” Mediating technologies such as OSNs, in contrast, “change everything.” In short, the “scale of the public” becomes magnified—now you have to worry about “all the people who might witness a reproduction” of your fall.⁶⁷

Note the symmetry between boyd’s point that technology changes everything and the notion of transformation central to contextual integrity. How information is actually handled is changing, not necessarily our expectations of how that information will be handled. Boyd’s falling example demonstrates how the visibility of OSNs broadens the scope of actors, increasing the number of information receivers; alters transmission principles, encroaching on the physical-space notion of reciprocity, where you see the people who see you fall; and conflates contexts, blurring the limited public sphere of the curb location with the unlimited public sphere of the online forum.

These transformations are not obvious. Rather, “the abstraction involved in asynchronous, online social networking encourages a gap between a user’s perceived audience and actual audience.”⁶⁸ A more specific example of violation can be seen in OSNs’ tendency to flatten the “nuances of face-to-face interactions” by forcing users to classify fellow OSN members as friends and non-friends.⁶⁹ Nissenbaum notes that this binary approach represents “a failure to grasp some of the subtle ways people share and withhold certain types of information in the complex web of their relationships.”⁷⁰ As Hull, Lipford, and Latulipe point out, in requiring users to stuff people into two categories, OSNs such as Facebook force users to “determine a set of *ex ante* rules for determining how information should flow in and between contexts . . . before knowing any of the details about the information itself.”⁷¹ Offline, we are able to adjust these transmission principles on a more flexible, case-by-case basis in accordance with our expectations of relevant information norms.

One important takeaway from all this is that the problem with privacy on OSNs—specifically, Lee’s problem—is not so much that there is a cognitive disconnect between what OSN users want (privacy) and what they do (disclose information) in binary terms as it is that OSN users are applying a multicolor privacy approach to a

67. *Id.*

68. Hull, Lipford & Latulipe, *supra* note 27, at 6.

69. *Id.* at 12–13.

70. NISSENBAUM, *supra* note 6, at 226.

71. Hull, Lipford & Latulipe, *supra* note 27, at 6.

functionally dichromatic (private vs. public) system that doesn't accommodate it.⁷² As the contextual integrity heuristic reminds us, we should support only those transformations that preserve the spirit of information norms that cognitively predate such transformations, *unless* new norms offer a benefit that outweighs the privacy costs. To return to boyd's falling example, we would have to balance the benefit to society of sharing a humorous photo with the violence done to principles of dignity enshrined in the original information norms. As explained above, OSNs clearly do not meet this standard; as a result, contextual integrity guides us to seek a remedial strategy.

One such strategy would be to educate consumers and allow them to select the most privacy-conscious firms on the open market. Assuming that consumers will act rationally (in the economic sense), we would predict that they will "pay" an amount of privacy commensurate with the utility they receive from the OSN they "purchase."⁷³ Assuming a competitive market, the privacy cost to join an OSN would drop to the marginal cost of maintaining an OSN that provides the demanded utility. Those OSNs that "charge" a higher privacy price (or a privacy price not aligned with users' expectations) will perish.⁷⁴

As researchers in behavioral economics have recognized, however, people do not always behave rationally.⁷⁵ For example, scholarly research suggests that consumers tend to be myopic—that is to say, they focus more on the here-and-now than the apparently dis-

72. *Id.* at 6–7, 12. Nissenbaum points out that the public–private dichotomy “can be understood as a cruder version of contextual integrity, postulating only two contexts with distinct sets of informational norms for each—privacy constraints in the private, anything goes in the public.” NISSENBAUM, *supra* note 4, at 141.

73. As any Facebook or MySpace member knows, most OSNs are nominally free. The quotes above are included to underscore the fact that the illusion of a free lunch obscures high privacy costs. Viewed in this light, it is appropriate to consider the OSN signup process as tantamount to a point of sale transaction.

74. *See, e.g.*, Oren Bar-Gill, *Seduction By Plastic*, 98 NW. U. L. REV. 1373 (2004). “When price equals marginal cost, a buyer will buy if and only if she values the good or service more than its cost. Marginal-cost pricing aligns private incentives with the social objective of welfare maximization. Goods and services are produced only when the benefit exceeds the cost, and an optimal allocation of resources is achieved.” *Id.* at 1377.

75. *See, e.g.*, Daniel Kahneman, Jack L. Knetsch & Richard H. Thaler, *Experimental Tests of the Endowment Effect and the Coase Theorem*, in BEHAV. L. & ECON. 211 (Cass R. Sunstein ed. 2000).

tant future.⁷⁶ If offered a product with a benefit of x at T1 and a total cost of zero at T1 and $x + 1$ at a later time T2, consumers will likely choose to purchase the product at T1 despite the fact that its total cost outweighs its benefit. For example, when deciding whether to obtain a credit card, consumers are likely to focus more on the high up-front benefits (e.g., rewards points) than on the high back-end costs (e.g., high interest rates).⁷⁷

In the OSN context, this means that consumers are likely to focus more on the features being offered (the T1 benefit) and the up-front price (the T1 cost) than on the privacy cost (the T2 cost), which is by nature a long-term cost. Reacting to this, OSNs will compete on features and up-front prices—not on privacy costs, which matter less to consumers. To compensate for the low up-front price and extensive features, which have high production costs, OSNs will backload the production costs to the price dimension that consumers heavily discount (i.e., the privacy price). In light of this distortion, the market will fail to select adequately for privacy-friendly firms. From an OSN's perspective, this strategy works best when users do not understand privacy risks at all; but the discounting effect holds even when consumers understand the nature of the risk.

Even if consumers behave rationally, they might not end up choosing the most privacy-conscious OSNs. Consider, for example, two OSNs—one with abysmal privacy protections and a million users and another with ideal privacy protections but only fifty users. If offered a choice between these two networks, most users would choose the former. The explanation for this decision is fairly straightforward. OSNs are fun and socially valuable because they allow users to benefit from shared information and experiences; thus, the more people that use the network, the more valuable the network becomes.⁷⁸ Such effects are particularly significant given the fact that the US OSN market is dominated by a few key players with hundreds of millions of users.⁷⁹

76. Oren Bar-Gill, *The Law, Economics and Psychology of Subprime Mortgage Contracts*, 94 CORNELL L. REV. 1073, 1120 (2009).

77. See Bar-Gill, *supra* note 74.

78. In the terminology of economics, OSNs are subject to network effects. OSNs become harder to leave as the number of users increases because "each user's payoff from the adoption of [the OSN], and his incentive to adopt it, increase[s] as . . . others adopt [the OSN]." Paul Klemperer, *Network Effects and Switching Costs*, (The New Palgrave Dictionary of Economics, Working Paper, 2005), available at www.paulklemperer.org.

79. See, e.g., *Statistics*, FACEBOOK, <http://www.facebook.com/press/info.php?statistics> (last visited Oct. 7, 2010).

Alternatively, the administrative state could force more contextual privacy controls. Of course, barring a perfect match between privacy settings and privacy expectations, conflict will still arise. Consider the Poster's Plight: The re-poster might be distributing because the OSN did not allow the poster to convey a context-specific transmission principle—"please do not spread this around." But the re-poster might also just be ignoring all the signals the poster provided. Thus a full consideration of the Poster's Plight requires both a technical and a legal analysis. So where does the law fit in?

As mentioned earlier, one weapon that the law provides for plaintiffs suffering from the Poster's Plight is the public disclosure tort. While hypothetically available to plaintiffs such as Lee, the public disclosure tort is of limited use in the OSN context. This is largely the result of vague elements that were built for a pre-internet age and that have not been adequately updated to accommodate contemporary technology. Understanding how these elements work is a prerequisite of our analysis of the tort's interaction with the Poster's Plight.

In the next Section, I outline the public disclosure tort, focusing predominantly on the reasonableness and legitimate concern prongs, the diversity of interpretative approaches to these prongs, and the transformative effects of OSN technology on the underlying analyses. Chief amongst these effects is the fundamental alteration of information norms that have accreted in response to existing doctrine.

II. CONTEXTUALIZING THE LAW

It is quite difficult to saddle OSNs with liability. This difficulty stems from section 230 of the Communications Decency Act, which provides an extremely powerful shield for providers of interactive computer services (i.e., OSNs).⁸⁰ Our discussion of the law should therefore start with this important provision.

Because OSN providers are responsible for the code that makes our online experiences possible, it might seem reasonable to hold them accountable for the shortcomings. Section 230 of the Communications Decency Act and recent section 230 jurisprudence, however, make it clear that targeting the OSN provider is a wasted effort. Motivated by an intent to "promote the continued

80. 47 U.S.C. § 230 (2006).

development of the Internet,”⁸¹ section 230 holds that “no provider or user of an interactive computer service shall be treated as the publisher or speaker of any information provided by another information content provider.”⁸²

In short, what this means is that platforms such as Facebook are largely immune to liability for the torts of clients and users. For example, in *Doe v. MySpace, Inc.*,⁸³ the Fifth Circuit refused to hold MySpace liable for negligent failure to “implement basic safety measures to prevent sexual predators from communicating with minors on its web site.”⁸⁴ Similarly, in *Chicago Lawyers’ Committee for Civil Rights Under Law, Inc. v. Craigslist, Inc.*,⁸⁵ the Seventh Circuit refused to hold Craigslist liable for hosting housing advertisements in violation of The Fair Housing Act. Although by no means bulletproof,⁸⁶ section 230 is undoubtedly robust. For example, in *Finkel v. Facebook, Inc.*,⁸⁷ the New York Supreme Court recently held that section 230 protects Facebook from liability for a third party’s defamatory statements despite the fact that Facebook retains an ownership interest in information posted on its site.

Our search for liability therefore should focus on the re-poster, not the OSN. The privacy torts defined by Prosser in the Second Restatement of Torts⁸⁸ and inspired by Samuel Warren and Louis Brandeis’s classic article *The Right to Privacy*,⁸⁹ are the chief weapons in this fight. The tort most relevant to the Poster’s Plight is the pub-

81. *Id.* § 230(b)(1).

82. *Id.* § 230(c)(1).

83. 528 F.3d 413 (5th Cir. 2008).

84. *Id.* at 416.

85. *Chicago Lawyers’ Comm. for Civil Rights Under Law, Inc. v. Craigslist, Inc.*, 519 F.3d 666 (7th Cir. 2008).

86. For a detailed overview of section 230 jurisprudence in 2009, see Eric Goldman, *47 USC 230 Year-in-Review for 2009*, TECH. & MARKETING L. BLOG (Jan. 5, 2010), http://blog.ericgoldman.org/archives/2010/01/47_usc_230_year_2.htm. As the cases listed demonstrate, section 230 remains a strong but not impermeable bar to liability. Of particular interest are cases such as *Almeida v. Amazon*, in which the potential application of section 230’s intellectual property exception to privacy torts is given somewhat detailed theoretical treatment. 456 F.3d 1316 (11th Cir. 2006).

87. *Finkel v. Facebook, Inc.*, No. 102578/09, 2009 WL 3240365 (N.Y. Sup. Ct. Sept. 15, 2009).

88. RESTATEMENT (SECOND) OF TORTS § 652A-E (1977).

89. Samuel D. Warren & Louis D. Brandeis, *The Right to Privacy*, 4 HARV. L. REV. 193 (1890). Solove and Richards stress that “Warren and Brandeis did not invent the right to privacy from a negligible body of precedent but instead charted a new path for American privacy law.” See Neil M. Richards & Daniel J. Solove, *Privacy’s Other Path: Recovering the Law of Confidentiality*, 96 GEO. L.J. 123, 125 (2007).

lic disclosure tort. The Second Restatement provides the following explanation:

One who gives publicity to a matter concerning the private life of another is subject to liability to the other for invasion of his privacy, if the matter published is of a kind that (a) would be highly offensive to a reasonable person, and (b) is not of legitimate concern to the public.⁹⁰

As this description makes clear, the public disclosure analysis is tricky business. “Private life,” “reasonable,” “highly offensive,” and “legitimate concern” are concepts that beget uncertainty and awaken judicial imagination. Imagination, which thus far has been used in many jurisdictions to shape these terms to restrict the tort’s overall ambit.⁹¹ As a result, public disclosure suits are all but impossible to win in more restrictive jurisdictions. What motivates this interpretative trend? As Lior Strahilevitz explains, while “[o]n one hand, the law seeks to encourage the expressive and psychological benefits that people derive from disclosing sensitive information about themselves to others,” on the other hand, “the law seeks to regulate the further dissemination of this information.”⁹²

The doctrinal ambiguities concerning the elements themselves, and the jurisdictional inconsistencies in interpretation, result in a tort that Solove aptly calls “one of the most fascinating puzzles of tort law.”⁹³ The puzzle originates in the fundamental tension between our dedication to free speech on one hand and the inevitable feeling that we have a “right to be let alone,” on the other.⁹⁴ This tension is most explicitly felt in the public disclosure tort’s “legitimate concern” element, which the Supreme Court has broadened to include many pieces of information that we wouldn’t expect to see on the eleven o’clock news. It also comes to bear, though less immediately, on the necessary determination of what, precisely, “private life” means.

I discuss each element separately. In Part A, I describe the current state of the legitimate concern test and explain the transforma-

90. RESTATEMENT (SECOND) OF TORTS § 652D (1977).

91. See, e.g., Patrick J. McNulty, *The Public Disclosure of Private Facts: There is Life After Florida Star*, 50 *DRAKE L. REV.* 93, 100 (2001).

92. Lior J. Strahilevitz, *A Social Networks Theory of Privacy* (Univ. of Chi. Law Sch. Pub. Law & Legal Theory, Working Paper No. 79, 2004), available at http://ssrn.com/abstract_id=629283.

93. Daniel J. Solove, *The Virtues of Knowing Less: Justifying Privacy Protections Against Disclosure*, 53 *DUKE L.J.* 967, 971 (2003).

94. See Warren & Brandeis, *supra* note 89 (citing THOMAS M. COOLEY, *A TREATISE ON THE LAW OF TORTS: OR THE WRONGS WHICH ARISE INDEPENDENT OF CONTRACT*, 29, 2d ed. 1888).

tive effects of technology on the analysis. In Part B, I do the same for the reasonableness prong.

A. *All the News that's Fit (or Not) to Print*

In light of the First Amendment, it is both constitutionally and ethically difficult to determine when a given piece of information is so newsworthy that the value of disclosure trumps the privacy interests of those who wish to keep the information secret. To the extent that courts have settled on a few approaches to this challenge, information norms that stem from the legal tests have developed. For example, where a community standard is applied, that community standard dictates the expectations of the population with respect to disclosures. In disrupting the technological-media framework that gave rise to these legal rules, OSNs threaten information norms that developed in response to the old framework and thereby violate contextual integrity. The present Section will chronicle this violation. Before breaking the doctrine apart, however, let us first take a look at how it plays out in real cases.

On September 22, 1975, Oliver Sipple saved President Gerald Ford's life. Just before Sara Jane Moore pulled the trigger on her .38 caliber revolver, Sipple, a decorated Vietnam War veteran living on 100% disability, somehow managed to push Moore's arm down just enough to reroute her shot and save the President from a potentially fatal wound.⁹⁵ After the dust settled, Sipple was questioned by the Secret Service and the FBI for three hours; he was interviewed by reporters sporadically for days. When asked what he wanted in return for his heroic deed, Sipple replied: "I just want a little peace and quiet."⁹⁶ What he got was the full scrutiny of the press with regard to his sexuality.

Despite Sipple's refusal to answer questions regarding his sexuality, the news media continued to press the issue and eventually "outed" Sipple. In retaliation, Sipple filed a lawsuit against several newspapers and other involved parties for publication of private facts.⁹⁷ Sipple lost the suit, in part because the court determined that the facts running to Sipple's sexuality were of legitimate public concern. In reaching this holding, the court noted that "the record shows that the publications . . . were prompted by legitimate politi-

95. Richard West, *President Escapes Assassin's Bullet*, L.A. TIMES, Sept. 23, 1975, at 1.

96. Daryl Lembke, *Hero in Ford Shooting Active Among S.F. Gays*, L.A. TIMES, Sept. 25, 1975, at A3.

97. *Sipple v. Chronicle Publ'g Co.*, 201 Cal. Rptr. 665, 666-67 (Cal. Ct. App. 1984) (upholding trial court's grant of summary judgment for defendants).

cal considerations”⁹⁸ such as the need “to dispel the false public opinion that gays were timid, weak and unheroic figures and to raise the equally important political question whether the President of the United States entertained a discriminatory attitude or bias against a minority group such as homosexuals.”⁹⁹

The *Sipple* court applied the legitimate concern, or newsworthiness, test described in the Second Restatement.¹⁰⁰ The Second Restatement states that “when the publicity ceases to be the giving of information to which the public is entitled, and becomes a morbid and sensational prying into private lives for its own sake, with which a reasonable member of the public, with decent standards, would say that he had no concern,” information ceases to be newsworthy.¹⁰¹ This is only one way to answer the newsworthiness question. Commentators have identified no fewer than five prominent approaches to the question,¹⁰² including simply rejecting the public disclosure tort outright.¹⁰³

An alternative approach is marked by deference to the media—the “leave it to the press approach.” As Diane Zimmerman describes, this approach is rooted in the belief that “social norms that govern acceptable behavior in the exchange of information are better communicated through the marketplace than through the courtroom.”¹⁰⁴ Because “the economic survival of publishers and broadcasters depends upon their ability to provide a product that the public will buy,” the press “must develop a responsiveness to what substantial segments of the population want (and perhaps

98. *Id.* at 670.

99. *Id.*; see DANIEL J. SOLOVE & PAUL M. SCHWARTZ, *INFORMATION PRIVACY LAW* 122 (3d ed. 2009) (noting that the President hadn’t taken time to call Sipple and thank him).

100. *Id.* at 669–70.

101. RESTATEMENT (SECOND) OF TORTS § 652D cmt. h (1977).

102. See Geoff Dendy, Note, *The Newsworthiness Defense to the Public Disclosure Tort*, 85 KY. L.J. 147, 157–64 (1997) (describing five approaches to newsworthiness question). Daniel Solove offers a different categorization in his recent article on the public disclosure tort and the First Amendment. See Solove, *supra* note 93, at 1001. Solove presents three approaches: (1) deferring to the media; (2) focusing on the status of the individual; and (3) examining the nature of the information. *Id.*

103. Although most states recognize the tort, Nebraska, New York, North Carolina, North Dakota, Rhode Island, Utah, and Virginia do not. See SOLOVE & SCHWARTZ, *supra* note 99, at 106.

104. Diane L. Zimmerman, *Requiem for a Heavyweight: A Farewell to Warren and Brandeis’s Privacy Tort*, 68 CORNELL L. REV. 291, 354 (1983). See, e.g., *Berg v. Minneapolis Star & Tribune Co.*, 79 F. Supp. 957, 960–61 (D. Minn. 1948) (asserting that under law of supply and demand, what press publishes can be seen as proxy for what people want reported).

even need) to know.”¹⁰⁵ In short, if it is in the news, it is there because people want it there.

Yet another approach joins a “nexus” component to the Restatement test. As articulated by the Fifth Circuit in *Campbell v. Seabury Press*,¹⁰⁶ privacy can be protected “by requiring that a logical nexus exist between the complaining individual and the matter of legitimate public interest.”¹⁰⁷

Although these tests stifle at least some intrusive reporting at the fringes of reasonableness, they all provide news media with a fair degree of leeway. The strength of this leeway is obvious in the Supreme Court’s privacy jurisprudence, which holds that “where a newspaper publishes truthful information which it has lawfully obtained, punishment may lawfully be imposed, if at all, only when narrowly tailored to a state interest of the highest order”¹⁰⁸ Yet, in marking the boundaries with respect to community norms, the predominant Second Restatement test at least recognizes the sorts of entrenched values central to contextual integrity. This is a tenuous shield for prospective plaintiffs, to be sure. But in an era in which violations were limited to major institutional actors (e.g., newspapers, magazines, and TV stations), people enjoyed the benefit of reinforcing community norms via decisions that reward good reporting (by purchasing a paper or watching a broadcast) and punish bad reporting (by ignoring it).

All of this changes in the digital age, given the democratizing effect of blogging and microblogging.¹⁰⁹ As Lauren Gelman has eloquently argued, Internet technologies have provided a “technological megaphone” that individuals can use to “broadcast their story [and those of others] to the world.” Prior to widespread use of the Internet, “content was filtered through news or other publishing

105. Zimmerman, *supra* note 104, at 353–54.

106. 614 F.2d 395 (5th Cir. 1980).

107. *Id.* at 397. The Tenth Circuit treads a similar path. As the court explained in *Gilbert v. Medical Economics Co.*, a “newsworthy publication must have some substantial relevance to a matter of legitimate public interest.” 665 F.2d 305, 308 (10th Cir. 1981). Applying this test in *Shulman v. Group W Productions*, the California Supreme Court held that video and audio footage depicting a car crash victim’s “injured physical state” and “disorientation and despair” were “substantially relevant” to an emergency response documentary’s inherently “newsworthy subject matter,” despite the fact that the victim never consented to the broadcast. 955 P.2d 469, 488 (Cal. 1998).

108. *Florida Star v. B.J.F.*, 491 U.S. 524, 541 (1989).

109. Microblogging is a relatively recent phenomenon, epitomized by services such as Twitter, which allow users to broadcast short messages. See *Microblogging*, WIKIPEDIA, <http://en.wikipedia.org/wiki/Microblogging> (last visited Oct. 13, 2010); TWITTER, <http://twitter.com> (last visited Feb. 22, 2010).

intermediaries.” These intermediaries “played an important social role in balancing the newsworthiness of information against the privacy interests of third parties who were indented.”¹¹⁰ Now that anyone can effortlessly broadcast to the world, a more precise notion of newsworthiness is required.¹¹¹ It is necessary both because people have lost the protection of an institutional news media and because the very essence of a community-norms approach is obliterated by the global context of the Internet.

Prior to the internet era, newsworthiness determinations relied on a particular set of information norms tailored to the prevailing news media business model and technological infrastructure. Blogs and OSNs tore open transmission principles at the core of these old-fashioned information norms. For example, whereas in the old model citizens were receivers, in the new model they can be both receivers and senders. Traditionally, when a newspaper made a decision based on community norms, it could apply a local context tailored to its circulation area. Now, a conceivably massive number of diverse communities can be served information.¹¹² The result is a clear prima facie violation of contextual integrity.

The problem is not limited to the newsworthiness analysis. Technological transformation has also affected the publicity requirement, a close cousin of newsworthiness. According to the Second Restatement, publicity “means that the matter is made public, by communicating it to the public at large, or to so many persons that the matter must be regarded as substantially certain to become one of public knowledge.”¹¹³ Courts tend to take two approaches to this, one quantitative and one qualitative. The quantitative test hews to the line drawn in the Restatement and “contemplates that a large number of persons must be aware of the intimate and embarrassing information before an actionable claim of invasion of privacy exists.”¹¹⁴ The qualitative test ignores the magnitude of the disclosure

110. Lauren Gelman, *Privacy, Free Speech, and “Blurry Edged” Social Networks*, 50 B.C. L. REV. 1315, 1333 (2009).

111. It should be noted that Internet-era developments also have implications for the disclosure tort’s publicity requirement. See *infra* Part II.A.

112. Note the clear parallel to similar problems with community-based obscenity standards. As Justice Breyer explained in *Ashcroft v. American Civil Liberties Union*, “adopting the community standards of every locality in the United States would provide the most puritan of communities with a heckler’s Internet veto affecting the rest of the Nation.” 535 U.S. 564, 590 (2002) (Breyer, J., concurring).

113. RESTATEMENT (SECOND) OF TORTS § 652D cmt. a (1977).

114. McNulty, *supra* note 91, at 100. Under the first test, courts have not held the publicity requirement satisfied when a defendant discloses plaintiff’s debts to an employer, *Yoder v. Smith*, 112 N.W.2d 862, 864–65 (Iowa 1962); when a defen-

and holds that the publicity requirement “may be satisfied by proof that the plaintiff has a special relationship with the public to whom the information is disclosed.”¹¹⁵

As Gelman’s “megaphone” metaphor makes clear, the publicity requirement changes dramatically in the Internet age. Consider these numbers: USA Today has a daily circulation of about two million, and the New York Times has a daily circulation of about one million.¹¹⁶ Facebook, in contrast, claims that more than 50% of its 500 million users sign in on any given day. These users collectively upload more than three billion photos each month, post sixty million status updates each day, and share five billion pieces of content (including photos) each week.¹¹⁷ As these statistics demonstrate, the information posted to popular OSNs can, depending on the poster’s privacy settings, reach an audience larger than that of the print versions of major domestic and international newspapers.¹¹⁸ Given the danger that a re-poster can circumvent privacy settings to broadcast information beyond the scope originally intended by the poster, this disclosure risk reaches beyond the limits of current OSN security measures.

Under a quantitative test, satisfaction of the publicity requirement in an OSN case will likely hinge on factors such as the privacy settings of the original poster’s account. For example, if the origi-

dant contacts plaintiff’s employer in an attempt to collect a debt, *Vogel v. W.T. Grant Co.*, 327 A.2d 133 (Pa. 1974); or when a defendant discloses to plaintiff’s superiors information relevant to a decision to terminate for cause, *Rogers v. Int’l Bus. Machs. Corp.*, 500 F. Supp. 867, 870 (W.D. Pa. 1980). In contrast, courts have upheld a suit when a creditor posted a large sign on his shop window publicizing plaintiff’s debt, *Brents v. Morgan*, 299 S.W. 967 (Ky. 1927), and when a store employee loudly and obviously interrogated and accused a patron of shoplifting in front of the store, *Bennet v. Norban*, 151 A.2d 476, 479 (Pa. 1959).

115. McNulty, *supra* note 91, at 100. *Miller v. Motorola* provides a good example of how the second test works. Joy Miller underwent a mastectomy and reconstructive surgeries and had to take three leaves of absence from her job as a result. Although Motorola’s resident nurse assured Miller that her medical information wouldn’t be disseminated, a co-employee informed Miller that she knew about the mastectomy. Miller’s subsequent belief that other employees surely must have known as well caused her to prematurely retire from her twenty-three-year employment with Motorola. The court upheld Miller’s privacy claim against a motion to dismiss. *Miller v. Motorola, Inc.*, 560 N.E.2d 900, 904 (Ill. App. Ct. 1990).

116. Joseph Plambeck, *Newspaper Circulation Falls Nearly 9%*, N.Y. TIMES, April 26, 2010, at B2, available at <http://www.nytimes.com/2010/04/27/business/media/27audit.html>.

117. *Statistics*, FACEBOOK, <http://www.facebook.com/press/info.php?statistics> (last visited Oct. 7, 2010).

118. Whether this information will actually spread depends on a variety of factors, such as the nature of the information and the popularity of the poster.

nal poster allows friends *and* friends of friends to view the material in question, far more people will likely view the material than if the original poster had allowed viewing access to her friends alone. Under the qualitative test, in contrast, the question of “special relationships” as they exist in the OSN context will dominate the issue. The archetypal qualitative case involves a defendant whose disclosure was to a party “whose knowledge of [the] facts would be embarrassing to the plaintiff.”¹¹⁹ How this will play out in OSN cases will likely depend on the degree of knowledge required of the defendant.

While *Miller v. Motorola*, the seminal case on the issue, does not clearly establish the degree of knowledge required by a defendant in a privacy suit, certain jurisdictions seem to consider whether or not the defendant knew about the special relationship.¹²⁰ Under a strict liability framework, the issue is resolved easily, as the outcome will not depend on whether the re-poster knew that the people to whom he was disclosing information had a relationship with the plaintiff. Yet if the re-poster discloses the information to a contact under a negligence framework, how much due diligence must he complete before he will escape liability? If the threshold is low (e.g., merely checking if the contact and the plaintiff list similar schools or communities in their profiles), the publicity test will provide minimal protection. If the threshold is high, the test will provide more protection—however, it should be noted that a too-burdensome “checking” requirement risks overly chilling the sharing that is central to OSNs’ psychological and social-economic benefits.

How all this comes together in Lee’s case is difficult to predict. In light of the standards established above, it would seem that the identity of the people to whom Lee’s re-poster disseminated the photographs and the nature of their relationship with Lee (and the subjects of the photograph) would critically affect the publicity analysis. Unfortunately, Lee’s post gives us no details on who the re-poster’s target audience was. If we apply a quantitative analysis, publicity will likely be satisfied regardless. Even if the re-poster’s initial dissemination provided access only to a few people, the re-posting actions of those subsequent actors would lead to an exponential increase in viewers. A qualitative approach, in contrast, would need

119. See *Miller*, 560 N.E.2d at 903 (holding both that “an invasion of a plaintiff’s right to privacy is important if it exposes private facts to a public whose knowledge of those facts would be embarrassing to the plaintiff,” and that said public might be “fellow employees, club members, church members, family, or neighbors, if the person [is] not a public figure”).

120. See, e.g., *Pachowitz v. Ledoux*, 666 N.W.2d 88 (Wis. Ct. App. 2003).

to consider the relationship between Lee and those given viewing access by the re-poster. We do not have this information either, though one could surmise that because Lee's mother has acted as re-poster, there is some substantial relationship between Lee and those people to whom his mother permitted access.

Our hands are similarly tied with respect to the newsworthiness question. This is largely because most public disclosure cases center on news media. Unlike a typical re-poster, newspaper publishers, bloggers, and television stations do not just broadcast pictures—there's usually a story attached. Applying any of the tests to the Poster's Plight is thus incredibly difficult. It is hard to contemplate, however, what interest the public would have in photos of Lee's children; unless there is a cloud hanging over the situation not revealed in Lee's post (such as parental abuse), it strains credulity to believe that Lee's photos are newsworthy.

I have outlined the degree to which OSN technology exerts a transformative effect both on the newsworthiness and the publicity requirement of the public disclosure tort. In the next section, I consider another critical element of the tort—the privacy requirement itself. Like the newsworthiness test, the privacy inquiry is fraught with difficulty (even outside the Internet context). This difficulty stems from the necessary analytic step of determining whether an individual who discloses information retains a reasonable expectation of privacy in that information.

Faced with this difficult decision, some courts prefer hard line rules that treat the precipitating context (and any information norms that go along with it) as irrelevant. Other courts appear to attempt a more contextual analysis but do so in a haphazard manner. Whether these courts consciously are applying a contextual decision heuristic is unclear. Nevertheless, they offer evidence that a contextual reasonableness test is not wholly impracticable. In the next Section, I outline both approaches to demonstrate this point.

B. Beyond Binary Notions of Privacy

Earlier in this Note, I concluded that a binary notion of privacy lacks the fidelity of a more robust, context-dependent approach. The same story plays out in courts' attempts to decide why a given fact is private rather than public. While it is generally acknowledged that American courts are all over the map when it comes to "determining whether an individual has a reasonable expectation of privacy in a particular fact that has been shared with one or more

persons,”¹²¹ what Strahilevitz calls a “hard-line”¹²² approach seems to have developed a significant following. This hard line approach is isomorphic with the public-private dichotomy discussed above; it takes as its general assumption that disclosure—even to a few people—precludes any future privacy interest in the disclosed information.

In the Fourth Amendment context, the hard line approach is referred to as the third-party doctrine.¹²³ The basic idea is that if you disclose private information to one person, you bear the risk that that person will further broadcast the information. As Justice White explained in his concurrence in *Katz v. United States*, “[W]hen one man speaks to another he takes all the risks ordinarily inherent in so doing, including the risk that the man to whom he speaks will make public what he has heard.”¹²⁴ Thus, the Court has held that parties do not have reasonable expectations of privacy in dialed phone numbers because the numbers have already been dis-

121. Strahilevitz, *supra* note 92, at 3.

122. *Id.* at 22.

123. See, e.g., SOLOVE & SCHWARTZ, *supra* note 99.

124. 389 U.S. 347, 363 n.** (1967) (White, J., concurring). *Katz* established the framework under which modern Fourth Amendment privacy analysis has evolved. In *Katz*, the Court determined that the “Fourth Amendment protects people, not places.” *Id.* at 351. Thus, the government invaded Katz’s privacy when it surreptitiously placed (without a warrant) a recording device outside a phone booth that Katz used to transmit wagering information in violation of a federal statute, even though there was no physical intrusion into the booth during the call. *Id.* at 348–51. In his famous concurrence, Justice Harlan explained that for Fourth Amendment purposes, privacy is established only when a person shows “an actual (subjective) expectation of privacy. . .that society is prepared to recognize as ‘reasonable.’” *Id.* at 361 (Harlan, J., concurring).

The third-party doctrine has come under considerable fire for the problems it presents in the digital age. See, e.g., David Couillard, Note, *Defogging the Cloud: Applying Fourth Amendment Principles to Evolving Privacy Expectations in Cloud Computing*, 93 MINN. L. REV. 2205, 2215 (2009) (noting that the doctrine “is particularly relevant in the cloud-computing world, where information is turned over to cloud service providers for remote storage and other quasi-transactional purposes with increasing frequency”). See also Matthew D. Lawless, Note, *The Third Party Doctrine Redux: Internet Search Records and the Case for a “Crazy Quilt” of Fourth Amendment Protection*, 11 UCLA J.L. & TECH. 1, 15 (2007) (arguing that a pragmatic “operational realities” test is more equitable than the third-party doctrine in determining a person’s reasonable expectation of privacy).

Note, however, that the doctrine is not without its supporters. Orin Kerr has written an article defending the doctrine both for its value in maintaining the “technological neutrality of Fourth Amendment rules” and providing “ex ante clarity” of Fourth Amendment doctrine. See Orin S. Kerr, *The Case for the Third-Party Doctrine*, 107 MICH. L. REV. 561, 564–65 (2009).

closed to the phone company;¹²⁵ nor do they have reasonable expectations of privacy in bank records, for the same reason.¹²⁶

One of the most famous hard-line cases is *Nader v. General Motors*.¹²⁷ In 1965, Ralph Nader published *Unsafe at Any Speed*, a book detailing the dangers of the Chevrolet Corvair. In retaliation, General Motors embarked on a smear campaign to undermine Nader's credibility.¹²⁸ According to the complaint, General Motors conducted interviews with Nader's associates and friends, questioning them about his political, social, racial, and religious views, his integrity, his sexual inclinations, and his personal habits. General Motors also allegedly spied on Nader in public places, made threatening phone calls to him, and tapped his phone.¹²⁹ Nader sued for, among other things, invasion of privacy.

Predictably, the court had no problem analyzing the wiretapping claim.¹³⁰ The interviews with Nader's associates, however, was a more bitter pill to swallow. The court found it "difficult to see how [the interviews] may be said to have invaded [Nader's] privacy." Because Nader "had previously revealed the information to such other persons," he "necessarily assume[d] the risk that a friend or acquaintance in whom he had confided might breach the confi-

125. See *Smith v. Maryland*, 442 U.S. 735, 743-45 (1979).

126. *United States v. Miller*, 425 U.S. 435, 442-45 (1976). The Court notes that "the Fourth Amendment does not prohibit the obtaining of information revealed to a third party and conveyed by him to Government authorities, even if the information is revealed on the assumption that it will be used only for a limited purpose and the confidence placed in the third party will not be betrayed." *Id.* at 443.

Note that the holdings in *Miller* and *Smith* have been limited by statutes. Two years after *Miller* was decided, Congress enacted the Right to Financial Privacy Act of 1978 (RFPA). 12 U.S.C. §§ 3401-22 (2006). The RFPA prohibits a search of "the financial records of any customer [held by] a financial institution" unless the customer has authorized disclosure, the Government has obtained a subpoena or court order, or the Government acts by formal written request in limited circumstances. 12 U.S.C. §§ 3401-22. Following the decision in *Smith*, Congress passed the Pen Register Act, which holds that "no person may install or use a pen register or a trap and trace device without first obtaining a court order." 18 U.S.C. § 3121 (2006).

127. *Nader v. Gen. Motors Corp.*, 255 N.E.2d 765 (N.Y. 1970).

128. See Ian McDonald, *Nader's Raiders*, *TIMES* (Lon.), Feb. 6, 1971, at 15, available at http://archive.timesonline.co.uk/tol/viewArticle.arc?articleId=ARCHIVE-The_Times-1971-02-06-15-001&pageId=ARCHIVE-The_Times-1971-02-06-15.

129. See *Nader*, 255 N.E.2d at 767.

130. The court upheld Nader's wiretapping claim, noting that the claim "most clearly meets" the requirements of an actionable invasion of privacy claim. *Nader*, 255 N.E.2d at 570.

dence.”¹³¹ And here’s the kicker: “[I]nformation about the plaintiff which was already known to others could hardly be regarded as private to the plaintiff.”¹³²

The court directly endorsed the third-party doctrine and indirectly endorsed the practice of unmitigated aggregation. As discussed earlier, aggregation is only weakly justified by the fact that aggregated information, whether in the form of court records, addresses, phone numbers, college degrees, or email addresses, is technically public information; far from simply taping things together, aggregation makes the end product far more informative than the sum of its component parts.

Judge Breitel, in a concurring opinion, located this weakness in the court’s conclusion that “the mere observation of the plaintiff in a public place does not amount to an invasion of his privacy.”¹³³

Although acts performed in “public,” especially if taken singly or in small numbers, may not be confidential, at least arguably a right to privacy may nevertheless be invaded through extensive or exhaustive monitoring and cataloguing of acts normally disconnected and anonymous.¹³⁴

History has proven Breitel’s point. It is hardly news that by statistically analyzing impersonal user profile data such as friend lists on OSNs such as Facebook, computers can draw inferences about very personal characteristics (such as political party preference or sexual orientation).¹³⁵ Yet the binary distinction—public versus private—that the *Nader* court applies cannot accommodate such nuanced analysis; to be more specific, the *Nader* court disregards the possibility that particular information norms existed between Nader and his friends that mediated Nader’s privacy expectations about the likelihood of disclosure.

In adopting the third-party doctrine and forcing Nader to shoulder the risk of confiding in friends, the court neatly sidesteps a more complex, context-dependent analysis. Not all courts, however, draw such sharp lines. Consider *Sanders v. American Broadcasting Companies*.¹³⁶ Mark Sanders was a telepsychic employed by a Los Angeles company. He worked in a large room with approximately

131. *Id.*

132. *Id.*

133. *Id.* at 771.

134. *Id.* at 772 (Breitel, J., concurring).

135. Carolyn Y. Johnson, *Project ‘Gaydar,’* BOSTON GLOBE, Sept. 20, 2009, http://www.boston.com/bostonglobe/ideas/articles/2009/09/20/project_gaydar_an_mit_experiment_raises_new_questions_about_online_privacy/.

136. *Sanders v. ABC*, 978 P.2d 67 (Cal. 1999).

ninety-nine other psychics, each of whom took calls in a personal cubicle. As part of an exposé for *PrimeTime Live*, ABC reporter Stacey Lescht applied for a job at the same outfit. Without telling Sanders, Lescht recorded her conversations with Sanders using a hidden camera and subsequently included portions of the video in the *PrimeTime Live* broadcast. Sanders sued ABC for, among other things, intrusion upon seclusion.¹³⁷

A close cousin of the tort of public disclosure, the tort of intrusion upon seclusion requires a plaintiff to show that the defendant intruded, physically or otherwise, upon the solitude or seclusion of the plaintiff or his private affairs or concerns, and that such intrusion would be highly offensive to a reasonable person.¹³⁸ Although this tort features elements different from those required by the public disclosure tort, it relies on a similar determination of whether the plaintiff had a reasonable expectation of seclusion or solitude in the place, conversation, or data source in question.¹³⁹

The key factual issue in *Sanders* was that, given the office environment, it was quite likely that someone other than Sanders or Lescht would have overheard the videotaped conversation. Under a third-party doctrine approach, Sanders would have had no claim—the conversation would be deemed public because it had been already revealed to the office at large. But the California Supreme Court thought otherwise:

[P]rivacy, for purposes of the intrusion tort, is not a binary, all-or-nothing characteristic. There are degrees and nuances to societal recognition of our expectations of privacy: The fact that the privacy one expects in a given setting is not complete or absolute does not render the expectation unreasonable as a matter of law.¹⁴⁰

In other words, the California court was reluctant to adopt a purely binary attitude. As the court asserted, “[T]he possibility of being overheard by coworkers does not, as a matter of law, render unreasonable an employee’s expectation that his or her interactions within a nonpublic workplace will not be videotaped in secret by a journalist.”¹⁴¹

There is a hint of contextual integrity in the *Sanders* opinion. As the *Sanders* court suggests, our disclosure decisions are not decisions to open or close an information faucet to the world. They are

137. *See id.* at 69–71.

138. *See* RESTATEMENT (SECOND) OF TORTS § 652B (1977).

139. *See Sanders*, 978 P.2d at 71–72.

140. *Id.* at 916.

141. *Id.* at 923.

decisions to share particular pieces of information with others pursuant to the belief that the people with whom the information is shared operate under similar transmission principles, roles, and other contextual values. Solove frames the issue of transmission principles nicely, stressing that “not all people and entities have the same obligations in maintaining the confidentiality of information.”¹⁴² One would expect more of her parent or best friend, regardless of the lack of a legal fiduciary relationship.

Although confidentiality is clearly at play in the public disclosure tort, it is not the only active transmission principle. Disclosure decisions are also linked to culture-mediated judgments about how well the recipient can handle the information disclosed and to what degree the recipient will want that information. Thus, as Katherine Strandburg highlights, “[T]he interplay between self-control, temptation, human cognitive limitations, and the theory of social norms” fosters the development of a natural hesitation to accept “too much information” in a given context.¹⁴³ This natural hesitation leads to the creation of another important transmission principle—the willpower norm.¹⁴⁴

Strandburg’s theory of willpower norms brings us to a separate analysis of *Sanders*. Because the conversation between Sanders and Lescht involved “ordinary workplace chat,” it was thus “a socially acceptable discussion within the workplace community of matters that might be expected not to be shared with the world at large.”¹⁴⁵ Because “the social norms that might ordinarily restrict disclosure of workplace discussions to outsiders were ineffective against the defendant journalist,” the court needed to step in to fill the norm gap. At least insofar as the *Sanders* court was concerned, the “law protects the social norms of workplace discourse from an intruder who is not reachable by those norms.”¹⁴⁶

Obviously, the scope of information norms is vast. The point in this Section is not to pick and choose. Rather, the goal is to demonstrate how essential the preservation of context is to a realistic assessment of privacy. Consider Lee’s situation: The *Nader* court would throw the case out immediately because, in posting the photos online, Lee crossed the third-party doctrine’s threshold of

142. Solove, *supra* note 93, at 1014.

143. Katherine J. Strandburg, *Privacy, Rationality, and Temptation: A Theory of Willpower Norms*, 57 RUTGERS L. REV. 1235, 1238 (2005).

144. *See id.*

145. *Id.* at 1300.

146. *Id.* at 1301.

privacy. In binary speak, he flipped from private to public.¹⁴⁷ Such an approach would deny Lee the opportunity to prove that his mother's use of a common "friend" connection to view the photos violated an offline confidentiality norm.¹⁴⁸

Under a *Sanders*-style analysis, Lee would at least get a chance to argue that his partial disclosure did not vitiate a privacy claim. Instead of asking only whether Lee disclosed the pictures to anyone, a court would likely investigate both the nature of the relationship between Lee and those to whom he provided viewing access and the disclosure norms (i.e., transmission principles) undergirding those relationships. The court in *Multimedia Wmaz v. Kubach*¹⁴⁹ provides a stellar example of such an analysis in its determination that an HIV patient's disclosure of his condition to friends, family, medical personnel, and members of a support group did not preclude him from suing a TV station when it failed to adequately blur his face in an AIDS documentary. Relying on the implicit understanding of confidentiality between Kubach and those to whom he disclosed his condition, the court stressed that the two disclosures (close contacts versus TV audience) "were similar in neither degree nor context."¹⁵⁰

However encouraging, the approach taken in cases such as *Sanders* and *Kubach* falls short of ensuring adequate protection, especially in light of the transformative effects of OSNs. This is primarily because the analysis is, at its core, unprincipled. Although judges seem to be exploring the contextual values at stake, they are doing so on the basis of intuition, not through consistent application of a clearly expressed and thoroughly considered decision heuristic. Thus, it is conceivable that a court that treats one case contextually will, by virtue of a lack of familiarity with underlying technology, apply a less contextually appropriate analysis to another case.

147. This prediction is not a flight of fancy. In a recent case involving a college student who briefly posted a derogatory poem about her hometown on an unprotected MySpace page, the court determined that the student had no reasonable expectation of privacy in the disclosure. *See* *Moreno v. Hanford Sentinel, Inc.*, 91 Cal. Rptr. 3d 858, 862–63 (Cal. Ct. App. 2009). The determination was made solely on the grounds that the student "made her article available to any person with a computer and thus opened it to the public eye." *Id.*

148. Recall that Lee was not linked to his mother as a "friend." Lee's mother was linked to one of Lee's friends as a friend. Because Lee presumably set his privacy settings such that "friends of friends" could see posted pictures, Lee's mother was able to view the pictures without becoming Lee's "friend."

149. 443 S.E.2d 491, 493 (Ga. 1994).

150. *Id.* at 494.

An alternative account of judicial thought focuses on probabilities of disclosure. Judges might, as Strahilevitz suggests, actually be asking themselves, “Had the defendant not become involved, would I have expected this information to remain private were I in the parties’ shoes?”¹⁵¹ The appeal of such an approach lies in its attachment to a basic and empirically verified premise of social networks theory—namely, that information flows through human networks in predictable ways. Of course, contextual integrity would hold that the principles dictating these flows stem from entrenched information norms. Therefore, insofar as judges applying this approach fail to consciously, transparently, and methodically confront contextual values, decisions will hew to a court’s personal understanding of social dynamics, not the organic reality of information norms as they exist in the relevant social space.

In the final analysis, it is possible that courts are doing none of these things—they could simply be going with what “feels right.” It is clear that even the most open-minded courts have not thoroughly embraced contextual integrity as a decision heuristic. Formal adoption of a reasonableness analysis that centralizes questions of contextually rooted values would grant public disclosure cases a much needed degree of predictability and, more importantly, congruence with peoples’ actual privacy values. One possible approach would require the deciding court to ask if, in light of prevailing information norms, the plaintiff was justified in expecting the disclosed information to remain private absent the defendant’s actions. Such an approach would join the intuition of the probabilistic method with the guidance of contextual integrity.

Most importantly, the reasonableness analysis should measure behavior against prevailing information norms. Thus, in Lee’s case, the court should not simply ask: What was the probability that Lee’s picture would reach the public absent the re-poster’s actions? Rather, the court should ask: Given the information norms that obtained between Lee and those to whom he provided access, what was the probability that Lee’s picture would reach the public absent the re-poster’s actions?

The approach just outlined requires some judicial access to contextual information. Unfortunately, as I highlighted in Part I, new technologies such as OSNs frequently impose new principles of information flow on users who either have not yet adjusted or do not fully appreciate that the rules of the game have changed. Therefore, a court that looks to evidence of a plaintiff’s privacy set-

151. Strahilevitz, *supra* note 92, at 14–15.

tings for contextual information, for example, will end up with a grossly inadequate and unrealistic determination of operative information norms. As a result, any reasonableness analysis predicated on these false information norms will be deeply flawed.

If we want courts to embrace an analytic approach that recognizes realistic information norms, we should encourage parties to create objective evidence of such information norms at the time of original disclosure, not merely after the fact at the evidence stage, when the pressure of litigation might cause parties to be less than candid about their norms, values, and expectations. With respect to the Poster's Plight, we should encourage OSNs such as Facebook to allow users to express more than just a binary "post/do not post" or "friend/not friend" preference when deciding whether to upload a picture. This objective is motivated by two important points. First, contextually accurate privacy settings are needed to preserve fundamental dignity and autonomy values and to ensure a safer and more consumer-friendly social environment that maximizes social utility through sharing. Second, a revamped public disclosure analysis that relies on contextual values needs *ex ante* evidence of those values. By recording preferences at the time of information disclosure, OSNs can help build a valuable record of evidence.

Our goal should be to add context to the OSN privacy regime. In the next section, I propose a preference expression tool that I hope will throw at least a few hues on the canvas. In Part A, I outline a tool that will help picture posters more thoroughly express their privacy preferences with regard to posted pictures. In Part B, I discuss how this tool can be expected to meet both objectives discussed above.

III. PUTTING THE BLURRY EDGES IN FOCUS

Let us start by examining how the law interacts with posting decisions on OSNs. Taking our lead from law and economics literature, we assume that all OSN members are perfectly rational, utility-maximizing, and self-interested actors.¹⁵² We begin our analysis with a three-person hypothetical OSN, of which X, Y, and Z are members (fig. 1). X is the picture poster. X and Y are offline friends. Y and Z are offline friends. X does not know Z. Assume that the OSN, like Facebook, gives users the ability to determine who

152. See Robert H. Frank, *Departures from Rational Choice: With and Without Regret*, in *THE LAW AND ECONOMICS OF IRRATIONAL BEHAVIOR* 13 (Francesco Parisi & Vernon L. Smith eds., 2005) (summarizing classical theory of law and economics).

can view their pictures. Assume also that the OSN, like Facebook, does not allow users to convey any privacy preferences relevant to the posting decision other than access. Finally, assume that all photos posted on this OSN feature only the poster.

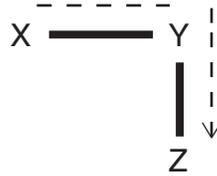


Figure 1: OSN schematic. Solid lines indicate offline friendships; dotted lines indicate (1) X’s willing online disclosure to Y and (2) Y’s deceptive online dissemination to Z.

Imagine further that X posts a picture and gives only Y viewing access. At this stage, Y must rely exclusively on contextual cues (i.e., transmission principles) that obtain in X and Y’s offline relationship in deciding whether or not to re-post the picture. Recall that Y’s re-posting decision is not limited by OSN technology. If Y so desires, she can download the photo to her hard drive and re-post it via her own account. In this case, the picture will be subject to only Y’s privacy preferences. If Y decides to allow Z to view the picture, X’s preference that only Y should have access will be vitiated. Furthermore, if a friend of Z’s joins the network and Z follows the same re-posting procedure as Y, Z’s contacts will have access to the picture.

X thus runs the risk that Y will misinterpret the disclosure norms implicit in their offline relationship or that Y will maliciously breach these norms. Because the OSN does not allow expression of any privacy preference outside of “share this with Y” or “do not share this with Y,” Y’s knowledge of operative norms is only as strong as her offline relationship with X. If X and Y are best friends, the likelihood of norm misinterpretation is fairly low. But if X and Y are not particularly close, the probability of misinterpretation markedly increases. Thus, assuming she is perfectly rational, X will base her decision whether to share on a calculation of the expected probability that Y will non-consensually re-post, the expected probability of prevailing in litigation precipitated by the re-posting, the expected privacy cost of a non-consensual re-posting incident, and the expected benefits obtained by sharing with Y.

This calculation works out differently depending on the jurisdiction’s treatment of the privacy torts. Under a binary regime, there is no chance that X will prevail in litigation because X has no

legally cognizable expectation of privacy (disclosure to one friend is viewed as tantamount to disclosure to the general public). Under such a regime, X will share if, and only if, the expected probability of Y's re-posting multiplied by the expected cost of Y's nonconsensual re-posting is less than the expected benefit of sharing with Y. Given the significant costs of privacy breaches (think about the school teacher who was fired), we would expect members of this OSN to be fairly timid sharers of information. Such an OSN would likely have no party pictures, no lively political commentary, and no flirting—it would be a communicative graveyard.

Like X's OSN, today's prominent OSNs (Facebook included) exist in a binary privacy space. Yet they are hardly barren and anti-social landscapes. We're left, then, with a pressing question: If OSNs do not provide a context-friendly privacy regime, and the law does not reliably take into account contextual information, why are people willing to share so much on OSNs? In other words, why are OSNs the bustling, vivid, and powerful sources of digital life that they are? Assuming that participation in OSNs is voluntary and not necessary, the most plausible explanation is that people simply aren't perfectly rational. In other words, whereas traditional law and economics assumes that individuals are "possessed of sufficient cognitive capacity to solve relatively simple optimization problems,"¹⁵³ individuals might actually be imperfect in their ability to consistently choose the efficient solution.

As discussed earlier, one irrational tendency is myopia. The benefits of joining an OSN (fun applications, increased social capital, romantic flings, etc.) dwarf the up-front costs of joining (\$0 + time spent registering and linking to OSN friends). Because the larger costs of joining (privacy risks down the line, commercial trading of user information by the OSN, etc.) are not usually felt until well after the time of registration, people have a tendency to discount them. To return to the example above, prospective OSN users may irrationally deflate the expected cost of a non-consensual re-posting incident. Thus, even if the cost to X is quite high (again, remember the school teacher who lost her job), it will appear small to X in comparison to the up-front benefit.

Another irrational tendency is optimism. Researchers in behavioral economics have identified an optimism bias in a variety of contexts, including criminal justice, litigation, and credit card

153. *Id.*

borrowing.¹⁵⁴ Put simply, “people exhibit a strong tendency to underestimate the probability that negative events will happen to them as opposed to others.”¹⁵⁵ The difference between myopia and optimism is subtle; whereas myopia acts on the total expected privacy cost (the probability of a nonconsensual re-posting incident multiplied by the expected cost of such an incident), optimism acts on the expected probability of a nonconsensual re-posting incident.¹⁵⁶ In the example above, an optimistic X will underestimate the probability of Y’s nonconsensual re-posting. This judgment error will distort X’s decision whether to participate in an OSN and share information.

At least on Facebook, these psychological tendencies are likely exacerbated by the OSN provider’s dauntingly complex approach to privacy. Over the course of Facebook’s ascent, the site’s privacy policy has expanded from 1004 words to 5830 words. As the New York Times pointed out in a recent article, this makes it longer than the United States Constitution, sans amendments.¹⁵⁷ The document is dense, technical, and by no means easy to navigate. As Oren Bar-Gill has argued in the context of subprime mortgage contracts, while “the rational [actor] is unfazed by complexity, the imperfectly rational [actor] might be misled by complexity.”¹⁵⁸ The New York Times reported in May 2010 that there were no fewer than fifty privacy settings on Facebook’s privacy page, with more than 170 options total.¹⁵⁹ While the rational OSN user will have no problem

154. See Oren Bar-Gill, *The Evolution and Persistence of Optimism in Litigation*, 22 J.L. ECON. & ORG. 490 (2006); Bar-Gill, *supra* note 74; Christine Jolls, *On Law Enforcement with Boundedly Rational Actors*, in *THE LAW AND ECONOMICS OF IRRATIONAL BEHAVIOR* 268 (Francesco Parisi & Vernon L. Smith eds., 2005).

155. Jolls, *supra* note 154, at 270.

156. *Id.* at 271 (“At least in some contexts, the empirical evidence makes clear that optimism bias reflects not only underestimation of the probability of a negative event relative to the average person’s probability of that event, but also underestimation of the probability of a negative event relative to the actual probability of that event.”).

157. Nick Bilton, *Price of Facebook Privacy? Start Clicking*, N.Y. TIMES, May 12, 2010, <http://www.nytimes.com/2010/05/12/technology/personaltech/13basics/html>.

158. Bar-Gill, *supra* note 76, at 1122.

159. Bilton, *supra* note 157. Some readers might find confusing the dual assertions that OSNs sponsor a “dichotomous” approach to privacy and also feature complex and numerous privacy controls. The argument is not that OSNs feature only two options—“privacy off” versus “privacy on.” The argument is that OSNs feature numerous options that are expressed in binary terms. In physical space, we make disclosure decisions by relying on norms of information flow (e.g., when we tell our best friends secrets, we rely on the relationship of trust to prevent further disclosure). Ideally, our online decisions could rely on the continued operation of

navigating these privacy settings and arriving at her personal optimal expected privacy result, the imperfectly rational user, intimidated by complexity, will ignore or weakly apply these privacy settings (and will suffer down the line as a result).¹⁶⁰

Of course, before we can confidently apply such an analysis, we need more empirical data linking OSN use patterns with the predictions and expectations of behavioral economics. Nevertheless, myopia, optimism, and irrational handling of complexity do offer a compelling set of explanations for why OSN users brave the privacy storm despite the immense privacy costs; in short, they just do not accurately perceive these costs. Yet the cost-benefit analysis proposed above is relevant only if OSN participation is viewed as a voluntary action and not something required of life in contemporary society.

As evidenced by the usage statistics, OSNs are central to modern life. [D]anah boyd makes this point by comparing Facebook to a utility company. She argues that “[p]eople’s language suggests that people are depending on Facebook just like they depended on the Internet a decade ago.”¹⁶¹ While boyd acknowledges that “Facebook may not be at the scale of the Internet (or the Internet at the scale of electricity),” she stresses that this size differential “doesn’t mean that [Facebook is] not angling to be a utility or quickly becoming one.”¹⁶² In making this point, boyd doesn’t merely rely on “people’s language.” Rather, she highlights the fact that despite Facebook’s multiple privacy faux pas,¹⁶³ people still use the service. According to boyd, there is no longer any reason to waste time wondering whether or not there will ever be enough user revolt to make Facebook turn back from its privacy-threatening policies— “there won’t be.”¹⁶⁴

such norms. At present, however, we must make a multitude of binary decisions (e.g., “share with best friend X,” “let X share with Y,” “do not let Y share with Z”) that are based on OSN codes’ dichotomous (i.e., “off versus on”) approach to privacy.

160. *Id.* In response to such reports and user discontent, Facebook recently implemented a more streamlined privacy setting interface. While marginally better, it still is far from simple.

161. danah boyd, *Facebook is a utility; utilities get regulated*, ZEPHORIA.ORG (May 15, 2010), <http://www.zephoria.org/thoughts/archives/2010/05/15/facebook-is-a-utility-utilities-get-regulated.html>.

162. *Id.*

163. A full history of Facebook privacy disputes is well beyond the scope of this Note. Many websites have chronicled Facebook’s struggle with privacy issues. *See, e.g.*, Caroline McCarthy, *Facebook’s follies: A brief history*, THE SOCIAL (May 13, 2010, 4:00 AM), http://news.cnet.com/8301-13577_3-20004853-36.html.

164. boyd, *supra* note 161.

If we assume that “buying” from Facebook is like buying from the local energy company, risky OSN participation becomes much easier to understand. As boyd explains, “when it comes to utilities like water, power, sewage, Internet, etc., I am constantly told that I have a choice. But like hell I’d choose Comcast if I had a choice. Still, I subscribe to Comcast. Begrudgingly. Because the ‘choice’ I have is Internet or no Internet.”¹⁶⁵

Necessity aside, it is at least plausible that the benefits of using OSNs such as Facebook really are significant, even in light of the hefty risks. Thus a third explanation might simply be that OSNs provide users high utility. In a recent post, technology blogger Nancy Baym explains why she hasn’t yet withdrawn from the Facebook community. In her eyes, Facebook provides a platform through which she “gain[s] real value.”¹⁶⁶ She writes:

I actually like the people I went to school with. I know that even if I write down all their email addresses, we are not going to stay in touch and recapture the recreated community we’ve built on Facebook. I like my colleagues who work elsewhere, and I know that we have mailing lists and Twitter, but I also know that without Facebook I won’t be in touch with their daily lives as I’ve been these last few years. I like the people I’ve met briefly or hope I’ll meet soon, and I know that Facebook remains our best way to keep in touch without the effort we would probably not take of engaging in sustained one-to-one communication.¹⁶⁷

As Baym explains, the value of Facebook is immense. For her, at least, this value exceeds the expected costs of membership and participation. Yet despite her tech savvy, Baym confesses to some irrationality, pointing out that “the rewards of Facebook are concrete and immediate,” and the “costs are abstract and ideological.”¹⁶⁸

It is likely that each of these explanations tell a piece of the overall story. What is important for our present purposes is not to identify a precise answer, but rather to appreciate that all these explanations are predicated on the assumption that there is at least *some* benefit to OSN use, and that this benefit entails *some* privacy tradeoff. This becomes clear upon revisiting our simple OSN. Just

165. *Id.*

166. Nancy Baym, *Why, despite myself, I am not leaving Facebook. Yet.*, ONLINE FANDOM (May 13, 2010, 12:40 PM), <http://www.onlinefandom.com/archives/why-despite-myself-i-am-not-leaving-facebook-yet>.

167. *Id.*

168. *Id.*

as there is information that X might not want re-broadcasted or re-posted, there is also information that X might not care to protect. Existence-on-the-network is a good example of this. If X is interested in the music of Arnold Schoenberg, for example, she might welcome the ability to sift through her OSN's member pages to find someone else (the proverbial needle in a haystack) who shares her atypical interest (atonal music is not exactly easy listening). Taking another step, she might also want other Schoenberg fans to be able to find her. A privacy solution that locks down all information would destroy the permeability that makes a good OSN work.

Lauren Gelman makes a similar argument in her piece on "blurry edged" social networks.¹⁶⁹ Addressing the question of why people "post content on a medium available to the whole world when that content is not intended for the whole world," Gelman argues that "Internet users are calculating that they are unlikely to identify *a priori* all the people they intend to reach with their posts because their social network is undefined." Thus, because X is unlikely to be able to identify all the Arnold Schoenberg fans on her OSN, she might be willing to unprotect the information on her profile running to her Schoenberg interest. In other words, X recognizes that the boundaries of the set "people I know who are interested in Schoenberg" are best left blurry (or permeable) to leave open the possibility that others with a shared interest can join the group.

So how do we simultaneously encourage socially optimal sharing and discourage socially troublesome over-sharing? One way would be through contract law. Woodrow Hartzog has suggested a "privacy box" application that would allow users to "enter information they wish to share with other connected 'friends' but request a promise of confidentiality before the information is divulged."¹⁷⁰ These promises of confidentiality could then be used to invoke reliance via the equitable remedy of promissory estoppel.¹⁷¹ This is a clever solution, if a slight miss. I would argue that by requiring users to entangle their interactions in a web of contracts, such a solution would not only take the fun out of networking online, but would also itself exert a transformative effect on norms. Confidentiality agreements are the stuff of business transactions; widespread appli-

169. See Gelman, *supra* note 110, at 1317. In fact, the Schoenberg example here is a riff off Gelman's own explanatory device.

170. Woodrow Hartzog, *The Privacy Box: A Software Proposal*, FIRST MONDAY (Nov. 2, 2009), <http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/2682/2361>.

171. *Id.*

cation of these agreements to a social context would ultimately conflate two separate spheres of interaction that do not belong together.

A more context-friendly solution would preserve the social nature of OSNs, while still giving users an opportunity to put the contextual stuffing back into transactions contextually eviscerated by OSN technology. Gelman's suggestion is spot on:

One option would be a tool for users to express and exercise privacy preferences over uploaded content. It would permit users to express their intentions by tagging any uploaded content with an icon that immediately conveys privacy preferences to third parties.

As Gelman points out, this tool "would provide immediate visual feedback to third parties about the content owner's preferences." Thus, just as search engines skip over certain webpages that contain privacy-requesting metatags, individuals can be generally expected to be "hesitant to abuse user privacy preferences when such preferences appear clearly alongside the relevant content."¹⁷² Returning to our basic example, if X can tag her content with extra guidance for Y, the risk that Y will make a disclosure decision based on a misinterpretation of prevailing norms will be significantly reduced. This matters because the less worried X is about unwanted disclosure, the more X will engage in socially optimal posting behavior.

The devil, of course, is in the details. The success or failure of such a tool will depend greatly on its narrow tailoring to the privacy problem it aims to solve. Our chief concern here is the Poster's Plight. To refer back to our basic model, we want to reduce the risk that Y will misperceive the norms on which X is relying when she posts her picture. We also want to generate a body of evidence that will chronicle X's expectations at the time of posting the picture. We have been spending a fair amount of time in the air; let us now move into the weeds.

A. *Color-Coding Privacy*

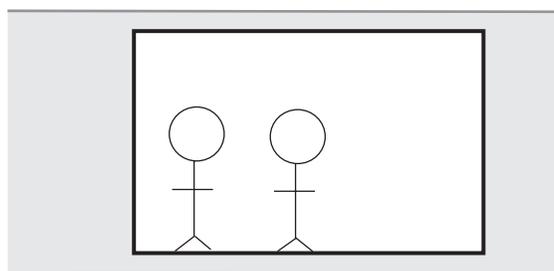
Currently, OSNs protect photo privacy by allowing users to limit viewing access either to OSN friends generally (MySpace) or to particular, handpicked users (Facebook). I propose that OSNs take another step and provide picture posters with an opportunity to disclose highly detailed preferences prior to (and after) posting pictures. The system would work as follows: Upon uploading a

172. Gelman, *supra* note 110, at 1343.

2011] BRINGING THE PUBLIC DISCLOSURE TORT ONLINE 939

photo, users would receive a stock privacy warning, reminding them of the potential downsides of disclosure (lost job, information leakage, etc.). Such a warning would have a link to the OSN's privacy policy page. After checking a checkbox labeled "I understand the potential risks of disclosure," the user would be presented with five descriptive choices, each placed next to a checkbox and a corresponding color. In the Facebook layout style, the chosen color would appear either next to the poster's name below the photo (if he or she is name-tagged in the photo) or in the place where the name would otherwise be (if he or she is not name-tagged) (fig. 2).

Photo 5 of 9 [Back to Album](#) - User's Photos [Previous](#) [Next](#)



In this photo: User A ● Untagged User ●

Figure 2: Mock-up of picture viewing screen in Facebook layout style. User A is name-tagged, so his color-coded preference is visible next to his name. The other user is not name-tagged, but his color-coded preference is nevertheless still visible. Note that on Facebook, when the mouse pointer scrolls over "User A" or "Untagged User," a box appears framing the person to whom the name-tag (or anonymous tag) refers.

While the descriptions themselves should be evaluated empirically for accuracy and contextual integrity, a preliminary group might look like this:

- **Green:** This picture can be disseminated throughout this OSN to any user.
- **Yellow:** This picture can be disseminated to anybody to whom I am linked on this OSN.
- **Orange:** This picture can be disseminated only to those to whom I explicitly provide access via this OSN's privacy settings.
- **Red:** This picture can be disseminated only by those who have asked me explicitly for re-posting permission.

- **Black:** This picture cannot, under any circumstances, be disseminated.¹⁷³

Skeptics will likely have at least two pressing questions. First, how can we depend on OSNs to do this? In short, we cannot. But luckily, law provides a useful incentivizing force. There are many ways in which United States law can compel OSNs to adopt such a system. Perhaps most straightforwardly, the FTC can incentivize adoption via its adjudicatory authority. Subchapter I of the Federal Trade Commission Act provides that the FTC is “empowered and directed to prevent persons . . . [and] corporations . . . from using unfair methods of competition in or affecting commerce and unfair or deceptive acts or practices in or affecting commerce.”¹⁷⁴ If the FTC were to determine that failure to provide a preference expression system is an “unfair” act or practice, the Commission could issue a complaint to the offending OSN.¹⁷⁵ Even if the OSN were to successfully challenge the complaint, by flagging the system as something the FTC will look for when evaluating an OSN’s privacy practices, such a tactical move would send a strong signal to other OSNs that they can minimize the risk of conflict with the FTC (and the bad press this sort of conflict entails) by providing the system.¹⁷⁶

173. In light of the well-documented fact that users tend not to thoroughly read privacy policies, a color code might facilitate knowledge through experience for those who do not read the descriptions. In other words, over time, the colors will attain contextual meaning through continued use and recognition. Also, the colors are mapped (in part) onto the traffic light scheme with which most users will presumably be familiar.

174. 15 U.S.C. § 45 (2006).

175. *Id.* § 45(b).

176. In fact, the FTC has been relatively active in using its adjudicatory authority to improve industry privacy practice. See *Legal Resources*, BUREAU OF CONSUMER PROTECTION BUSINESS CENTER, http://www.ftc.gov/privacy/privacy_initiatives/promises_enf.html (last visited Mar. 20, 2011) (listing FTC actions charging parties with deceptive or unfair practices under Case Highlights). It should be stressed, however, that the FTC is not unrestrained in its ability to condemn practices it deems as “unfair.” According to the FTC Act, the Commission can find unlawful only those acts or practices that are “likely to cause substantial injury to consumers which is not reasonably avoidable by consumers themselves and not outweighed by countervailing benefits to consumers or to competition.” 15 U.S.C. § 45(n) (2006). One might reasonably doubt, however, whether the FTC would train its enforcement weaponry on such a narrow target. One might still expect the agency to exert some influence on industry practice via “nonenforcement regulatory tools.” As Kenneth Bamberger and Deidre Mulligan demonstrate, such tools (e.g., publicity, best-practice guidance, the encouragement of certification regimes) can be powerful in shaping privacy policy. Kenneth A. Bamberger & Deidre K. Mulligan, *Privacy on the Books and on the Ground*, 63 STAN. L. REV. (forthcoming 2011).

A second approach would be to alter section 230 of the Communications Decency Act to offer a particularized safe harbor for OSNs providing tools enabling users to express privacy preferences for all materials they post. Currently, section 230(c)(2)(A) exempts OSNs from liability for subpar or ineffective efforts taken in “good faith to restrict access to or availability of material that the provider or user considers to be obscene, lewd, lascivious, filthy, excessively violent, harassing, or otherwise objectionable, whether or not such material is constitutionally protected.”¹⁷⁷ Similarly, section 230(c)(2)(B) exempts OSNs from liability for “any action taken to enable or make available to information content providers or others the technical means to restrict access to material described [above.]”¹⁷⁸ A slight alteration of this safety net would allow plaintiffs to include OSNs in their lawsuits against tortious re-posters. Consider the following addition to section 230(c)(2):

- (C) No provider of an Online Social Network shall be held liable on account of any reasonable action taken to enable or make available to users the technical means to express disclosure preferences with respect to all information posted to the Online Social Network.
- (D) Nothing in (C) shall be interpreted to provide exemption from liability for providers of Online Social Networks whose actions to provide users the technical means to express disclosure preferences are either unreasonable or willfully incomprehensive.

Ideally, courts would read this modification to mean that when X sues Y for nonconsensual reposting, X can also sue the OSN provided that the privacy expression tools provided by the OSN are unreasonable or willfully incomprehensive. Ideally, the court would determine reasonableness by applying the sort of cost-benefit test familiar to all first-year torts students.¹⁷⁹ Thus, over time, OSNs would be incentivized to add only those privacy-protecting tools that are beneficial at the margin (i.e., efficient in light of the costs of implementing them).

The differences between the adjudicative regulatory approach and the safe-harbor approach are subtle. While the adjudicative regulatory approach avoids the slow gears of legislative decision-making, it also decouples the oversight process from the tort sys-

177. See 47 U.S.C. § 230(c)(2)(A) (2006).

178. 47 U.S.C. § 230(c)(2)(B) (2006).

179. *United States v. Carrol Towing Co.*, 159 F.2d 169, 173 (2d. Cir. 1947) (the algebraic tort liability analysis known as the Hand Formula, articulated by Judge Learned Hand).

tem. By doing so, it takes significant wind out of plaintiffs' sails; whereas under the safe harbor regime, angry plaintiffs could simply join the negligent OSN to their lawsuit, under the adjudicative regulatory approach, they would have to complain to the FTC or wait for the FTC to act on its own. On the other hand, the adjudicative regulatory approach has the benefit of centralized and unambiguous decisionmaking, whereas the safe harbor approach subjects the law to the diffuse interpretation of the various courts.

Yet another related approach would borrow concepts from products liability law. As James Grimmelmann has recently suggested, defective design jurisprudence offers a useful model for compelling OSNs to behave in socially responsible ways.¹⁸⁰ For example, the Third Restatement of Torts provides that a product "is defective in design when the foreseeable risks of harm posed by the product could have been reduced or avoided by the adoption of a reasonable alternative design by the seller . . . , and the omission of the alternative design renders the product not reasonably safe."¹⁸¹ These principles could be used to find that the cost to an OSN of providing a preference expression system would be so minimal relative to the benefits to users that failure to institute such a system necessitates liability.

Even if proponents of this tool find it difficult to mobilize the legislative and administrative machinery to compel OSNs to provide better tools, they can likely provide some benefits via independent action. For example, Facebook's celebrated application programming interface makes it relatively easy for developers to create applications that operate within the Facebook environment. Drawing on Hartzog's privacy box concept, third-party developers could promote the color-coding system themselves via a free application that users could download.¹⁸² Unlike the core solution, which ideally would be engineered by Facebook and embedded in the posting process, this approach would require users to seek out the tool and remind themselves and fellow users to use it. Thus, while a potential expedient, this independent option is a distant second best to an OSN-implemented approach.

Up to this point, we have been assuming that the photos posted on OSNs feature only the poster. This simplifies our task

180. James Grimmelmann, *Privacy as Product Safety*, 19 WIDENER L. REV. 793, 813 (2010).

181. RESTATEMENT (THIRD) OF TORTS: PRODS. LIAB. § 2(b) (1998).

182. I am currently in the process of developing such an application with the gracious support of the New York University Privacy Research Group. Details on the project can be found at www.postpref.com.

markedly, as the posting decision is made by the person who bears the risk of that decision. What happens when multiple people are in the posted picture? How would the proposed privacy tool account for the privacy externalities of a poster's decision?

One possibility is for the tool to take a cue from Facebook's tagging procedure. As mentioned earlier, Facebook provides the opportunity for users to name-tag photos (i.e., draw an invisible box around a face and link a name to the framed face). The proposed tool could piggyback on this tagging functionality. For example, all parties depicted in a photo could be given the opportunity to add disclosure preferences to the photo. The major difference between this sort of tagging and simple name-tagging would be that while anyone can name-tag anyone in a picture, only those people depicted in the picture would be able to add a disclosure preference to a picture in which they are depicted.

In the Facebook layout, when depicted parties express preferences and remain name-tagged, their color-coded preferences could be depicted at the bottom of the picture next to their names (fig. 2). In the event that these parties *do* untag themselves, the color-coded preferences could still be depicted at the bottom, however the parties' names could be deleted (fig. 2). To encourage active preference expression, the standard email sent to tagged users when someone name-tags them could explicitly remind these users to add their disclosure preferences to the picture. An email could also be sent to all persons depicted in a photo whenever the original poster updates his or her disclosure preferences for that photo.

A significant weakness of this system is that it would rely on the tagging behavior of OSN users; absent facial recognition technology, the system would be able to send notification emails only to users already name-tagged in the photo. Another weakness is that the system would have no way of contacting individuals who are depicted in posted photographs but who are not members of the OSN.

Weaknesses aside, this preference tool goes a long way toward reintroducing some context to context-naked OSN privacy settings. Understanding just how far this tool will take us, however, requires a consideration of how it can be expected to interact with privacy law. This is because the full potential of the preference tool I propose requires a one-two punch of code plus law. In the next section, I explain the necessity of this cooperative relationship and argue that the tool will play well with a reformed reasonableness analysis.

B. *Reintroducing Context*

To be successful, the preference tool needs to accomplish two objectives. First, it needs to reduce the probability that users will misunderstand the operative norms governing posters' posting decisions. Second, it needs to clarify the reasonableness analysis that will take place should a re-posting conflict make its way to court. On its own, the preference tool can address the first objective; the law will need to address the rest.

As Gelman suggests, “[S]imple neighborliness requires that we honor each other’s privacy preferences.”¹⁸³ In other words, Gelman predicts that “Internet users will respect the social force of a plea for privacy if they are faced with such a request at the time they access online content.” Absent any enforcement mechanism, “simple social signals” can be relied on to “exert their own force across forums.”¹⁸⁴ This optimism is tempting, but we should be wary of designing our policies to fit ideals instead of pragmatics. On what theoretical ground can we expect users to take the poster’s preferences into account in deciding whether to re-post?

Again, contextual integrity comes to the rescue. Recall danah boyd’s “tripping on the curb” example. Recall also that I argued that OSN technology caused a fundamental transformation of the transmission principles governing the visual interaction between the person who trips and those who see him fall. In physical space (that is to say, offline), a principle of reciprocity dominates—the tripping party is likely to take note of those who see him fall. When a photo of the fall is distributed online, this sense of reciprocity fades away because the tripper has no way of knowing who is viewing his embarrassing moment. This matters because reciprocity has important accountability effects. If Y knows that the tripper saw him witness the fall, Y will be much less likely to feel that she can disseminate the fact, or other representation, of the fall with impunity. We might call this contextual deterrence.

In a sense, the preference tool I propose will help bring the offline principle of reciprocity back into the online privacy fold. Experience with human nature encourages us to expect that a user faced with a photo bearing a red mark will think twice before re-posting, *even if* there is no system of legal liability to add bite to the bark. Empirical evidence and economic analysis lead to similar conclusions. For example, Fehr, Fischbacher, and Gächter demonstrate that people “have a tendency to voluntarily cooperate, if treated

183. Gelman, *supra* note 110, at 1343.

184. *Id.*

fairly, and to punish noncooperators.”¹⁸⁵ This “strong reciprocity” might have its roots in the struggles of ancient human groups to survive; groups with a disproportionately large number of strong reciprocators were probably “much better able to survive” the many threats that marked early human life.¹⁸⁶

Even if we assume that the average OSN member will respect the norms expressed by posters, we have to confront the possibility that some will not. While many people respect privacy, some do not. Thus, while code can do much to make OSNs more context-friendly, law is necessary to complete the privacy circle. As Gelman suggests, “[I]f individuals were able to tag content with their preferences . . . one could envision the privacy torts evolving to take account of individual privacy expectations.”¹⁸⁷ How exactly would the analysis play out?

To return to our simplified example, imagine that X has posted a picture to a three-person OSN consisting of X, Y, and Z. Imagine further that X has given Y viewing access to her photo and has used the proposed preference tool to assign a ranking of “red” to her photo. Ignoring this signal, Y re-posts the picture, effectively providing viewing access to Z—Y’s OSN friend and someone with an ability to harm X—and anyone else Z permits. Assume that the picture includes some information that is useful to Z in his effort to harm X. Finally, Assume that Z takes this action and harms X. Should Y be legally accountable to X under the public disclosure tort?

Earlier I proposed that the presiding court should assess, given the information norms that obtained between Lee and those to whom he provided access, the probability that Lee’s picture would have reached the public absent his mother’s actions. I also argued that, absent some injection of contextual information at the time of the original posting, courts would be at a loss in determining the substance of the information norms that obtained. In providing evidence of the information norms supporting X’s disclosure decision, the proposed tool fills the knowledge gap for the courts, providing them with the contextual information they need.

To illustrate, let us add a little more complexity to our hypothetical. Imagine now that A, B, C, and D join the OSN. Imagine also that X has shared her photo and her preferences with A, B, C, and D. In the language of contextual integrity, X has set herself as

185. Ernst Fehr, Urs Fischbacher & Simon Gächter, *Strong Reciprocity, Human Cooperation, and the Enforcement of Social Norms*, 13 HUMAN NATURE 1, 1 (2002).

186. *Id.* at 5.

187. Gelman, *supra* note 110, at 1344.

sender and Y, A, B, C, and D as receivers. Trampling on X's preferences, Y re-posts and thus alters the norms, setting herself as sender and the public—Z—as receiver. Absent Y's actions, the probability that X's picture would have reached Z is likely minimal; because Y, A, B, C, and D have access to X's clearly expressed preferences, we can assume that they would have respected these preferences and refrained from re-posting. Thus, Y should be held liable.

To illustrate further, assume that X's photo depicts X, G, and H, all of whom have expressed privacy preferences with respect to the posted photo. Let us also assume that X assigns a rating of "red," G assigns a rating of "red," and H assigns a rating of "green." As before, Y breaches the trust and re-posts the photo. As before, X (now, along with G) sues Y for public disclosure of private facts. What privacy rating should the court use? Assuming H doesn't care, the court can safely apply the "red" level. But what if H truly wants the picture disseminated to all the world? Is it not unjust (and, in fact, unconstitutional) to dampen H's freedom of expression?

Recall, however, that the remedy here is damages, not an injunction. With public disclosure suits, the oil has already spilled. In short, nobody is preventing H's image from spreading, because it already did. But what about chilling the sharing behavior of future posters? As Gelman has noted, X's "ability to protect [her] privacy may interfere with [Y's] ability to speak [her] life story."¹⁸⁸ In fact, the problem extends beyond the people in the picture. Even if Y is not in the picture, the picture might have some relevance or importance to Y that justifies her interest in taking part in the disclosure decision.

This is a valid concern. Yet I would argue that it is not so much a problem with the proposed approach to the reasonableness test as it is with the tort itself. Recall that the reasonableness analysis is focused only on the question of whether the plaintiff has a reasonable expectation of privacy in the disclosed material. Other elements of the public disclosure analysis—the legitimate concern test, for example—can be expected to help courts judge when the plaintiff's privacy interest is outweighed by the public's need to know. For example, imagine that X, G, and H are all social workers who help people with gambling addictions. Imagine further that the picture depicts X, G, and H at a party of a prominent casino owner. This potential conflict of interest might well justify Y's disclosure decision, X and G's privacy expectations notwithstanding.

188. *Id.* at 1332.

A final concern, though unrelated to free speech, runs to the possibility of a “race to the top.” Insofar as users recognize the utility of the preference tool in deterring unwanted disclosure, they might just always choose the most restrictive setting. This would, of course, derail the tool’s recontextualizing effect and hamper the legal analysis. It seems unlikely, however, that OSN users would take such an approach. OSN interchanges are all about socialization; users might reasonably worry that taking an extreme privacy attitude—one unrelated to specific contexts of disclosure—would mark them as an impediment to social interchange and detrimentally affect their social capital online. This logic also helps address the concern discussed above with regard to photos depicting multiple people. People will be hesitant to post pictures depicting someone who always chooses an unreasonably high rating. If this person truly cares about participating in the posting exchange, he or she will modify his or her posting behavior or risk exclusion from the socially valuable posting activity.

IV. CONCLUSION

When Warren and Brandeis penned their classic article on privacy, the Kodak camera and the telephone were state-of-the-art technologies. Early in their discussion, Warren and Brandeis stress that “instantaneous photographs and newspaper enterprise have invaded the sacred precincts of private and domestic life; and numerous mechanical devices threaten to make good the prediction that ‘what is whispered in the closet shall be proclaimed from the house-tops.’”¹⁸⁹ The privacy torts can be seen as emerging from a struggle between extant social norms and rapidly developing technologies that threatened them.

In our zeal to adopt fantastic (and useful) new devices, we often forget that these marvels of invention are not autonomous but rather are at our beck and call. In advocating against blind reliance on technological progress, Neil Postman stresses that “once a technology is admitted, it plays out its hand; it does what it is designed to do.”¹⁹⁰ He argues that in designing and using technology, “[o]ur task is to understand what that design is.” In other words, “when we admit a new technology to the culture, we must do so

189. Warren & Brandeis, *supra* note 89, at 195.

190. NEIL POSTMAN, *TECHNOPOLY: THE SURRENDER OF CULTURE TO TECHNOLOGY* 7 (1992).

with our eyes wide open.”¹⁹¹ Often, however, we cannot (or do not) foresee the full developmental trajectory of a given technology at the time we invent and implement it. Therefore, we are frequently engaged in a game of catch-up.

Yet it seems that many of the most influential digital hawkers have forsaken our eminently human duty to catch up. No less than Eric Schmidt, the CEO of Google, a company with access to enough information to write my biography, has matter-of-factly stated: “If you have something that you don’t want anyone to know, maybe you shouldn’t be doing it in the first place.”¹⁹² Similarly, Scott McNealy, the CEO of Sun Microsystems is on record as saying that “you have zero privacy.”¹⁹³

Such abdications are disappointing. But they are certainly not demoralizing. As the recent proliferation of blogs, academic papers, and books devoted to digital privacy demonstrates, digital privacy is not so private anymore. Consider Facebook’s most recent privacy bungle. In April 2010, Facebook dramatically changed the way users list information on their profiles.¹⁹⁴ While historically, users were able to add information about their personal lives in plain text and limit access to that information to a select group of people, the update forced users to either broadcast that information to the entire Facebook network or refrain from posting it. This is because Facebook began to treat each bit of personal information as a “connection.”¹⁹⁵ Under the connections model, user information (e.g., “I like football”) shows up as a hyperlink on the user profile. At the time Facebook rolled out this new feature, anybody viewing the home page of the connection (be it “I like football” or “Northern New Jersey Violin Enthusiasts Club”) was able to see the full list of users who list that connection in their profiles.

It doesn’t take too much imagination to realize the potential chilling effects of this policy (just think about the gay teenager who wants to support gay rights but is not ready to come out publicly). This message wasn’t lost on the world. Indeed, the privacy community (and major media institutions such as the New York Times)

191. *Id.*

192. *Google CEO on Privacy*, HUFFINGTON POST (Dec. 7, 2009, 3:43 PM), http://www.huffingtonpost.com/2009/12/07/google-ceo-on-privacy-if_n_383105.html.

193. Polly Sprenger, *Sun on Privacy: Get Over It*, WIRED (Jan. 26, 1999), <http://www.wired.com/politics/law/news/1999/01/17538>.

194. Kurt Opsahi, *Six Things You Need To Now About Facebook Connections*, ELECTRONIC FRONTIER FOUNDATION (May 4, 2010), <http://www EFF.org/deeplinks/2010/05/things-you-need-know-about-facebook>.

195. *Id.*

stepped up to bat in a big way to challenge Facebook's arguably irresponsible behavior. While the blogs ranted and the New York Times reported, the Electronic Privacy Information Center, along with fourteen privacy and consumer protection organizations, almost immediately filed a complaint with the FTC.¹⁹⁶ To fan the flames, Senators Charles E. Schumer, Michael F. Bennet, Mark Begich, and Al Franken wrote a letter to Mark Zuckerberg (Co-founder, CEO, and President of Facebook), in which they "express[ed] . . . concern regarding recent changes to the Facebook privacy policy . . ." ¹⁹⁷ By late May, Facebook had reined in the connections model; users can now employ privacy settings to control who can learn of their membership by viewing connection "home pages."¹⁹⁸

As this slice of current events demonstrates, the privacy landscape changes fast and furiously. Perhaps our best defense against this changing landscape is conscientious innovation. Certainly, the time is ripe for developing creative and effective solutions to the novel privacy problems that today's phenomenally useful technologies produce. As this Note hopefully has demonstrated, it is by no means impossible to bring privacy law up to speed with technological reality. By combining our sophisticated judicial system with the intuition of technologists, incisive creativity of technological philosophers, and nearly limitless possibilities of code, we can help bring the color of context back to an increasingly dichromatic online canvas.

196. *New Facebook Privacy Complaint Filed with Trade Commission*, EPIC (May 5, 2010), <http://epic.org/2010/05/new-facebook-privacy-complaint.html>.

197. Letter from Charles E. Schumer, Michael F. Bennet, Mark Begich, and Al Franken, United States Senators, to Mark Zuckerberg, Co-founder, CEO, and President, Facebook (Apr. 27, 2010), *available at* <http://www.politico.com/news/stories/0410/36406.html>.

198. Scott M. Fulton III, *Facebook CEO: 'We are removing the connections privacy model'*, BETANEWS (May 26, 2010), <http://www.betanews.com/article/Facebook-CEO-We-are-removing-the-connections-privacy-model/1274906695>.

950 NYU ANNUAL SURVEY OF AMERICAN LAW [Vol. 66:895

NEW YORK CITY'S LANDMARKS LAW AND THE RESCISSION PROCESS

JOACHIM BENO STEINBERG*

“No one is questioning the need for the university’s law school to expand, but surely it can be worked out in a way that does not destroy yet another piece of this fast-vanishing area. It is hard for me to believe that a great institution like N.Y.U., which had the foresight and good taste to expel me many years ago, would be insensitive to this situation.”

—Woody Allen¹

INTRODUCTION

Decisions over whether to accord a New York City building or site landmark status are frequently controversial.² Such decisions implicate divergent and often diametrically opposed views of the role the city’s government should play in economic development and in protecting cultural sites, as well as the role the democratic process should play in these decisions. Different views on the aesthetic or historical value of proposed sites often lead to protracted and often highly visible disputes waged in the press, in the political arena, and in the courts.

* J.D., New York University School of Law, 2010; A.B., University of Chicago, 2004. I would like to thank Professor Vicki Been and Professor Clayton Gillette for their thoughtful advice and guidance, the students of the Law of New York City Seminar for their comments and insights, and the staff of the New York University Annual Survey of American Law for all of their hard work and encouragement.

1. Woody Allen, Letter to the Editor, N.Y. TIMES, July 27, 2000, <http://www.nytimes.com/2000/07/27/opinion/1-a-city-so-lovely-through-woody-allen-s-lens-389765.html> (regarding proposed construction of Furman Hall at New York University School of Law).

2. See, e.g., Nicolai Ouroussoff, *New York City, Tear Down These Walls*, N.Y. TIMES, Sept. 28, 2008, <http://www.nytimes.com/2008/09/28/arts/design/28ouro.html> (describing controversy over 2 Columbus Circle); Nicolai Ouroussoff, *In Village, a Proposal That Erases History*, N.Y. TIMES, Apr. 1, 2008, <http://www.nytimes.com/2008/04/01/arts/design/01pres.html> (discussing proposal to demolish O’Toole building); Witold Rybczynski, *Goodbye, 2 Columbus Circle*, SLATE, Jan. 14, 2009, <http://www.slate.com/id/2208529> (criticizing architectural changes to 2 Columbus Circle); Julia Vitullo-Martin, Op-Ed, *Has Landmarking in New York Gone Too Far?*, N.Y. POST, Apr. 25, 2010, http://www.nypost.com/p/news/opinion/opedcolumnists/has_landmarking_in_new_york_gone_SzbdqcXFtBsoZx2F7IHSLN/3 (discussing downside of landmarking).

New York City's Landmarks Preservation and Historic Districts Law (the Landmarks Law) was enacted in 1965.³ As of 2008, New York City's Landmark Preservation Commission (LPC) had designated more than 1400 landmarks, including 1194 exterior landmarks⁴ and 106 historic districts,⁵ which include around 25,000 buildings.⁶ In 2009, the LPC designated 40 landmarks and historic districts, twice the LPC's targeted number.⁷ The sheer number of landmarks and the staggering amount of money at stake⁸ in landmark disputes raise the question of how best to revisit landmark designations.

The Landmarks Law requires that designated sites be at least thirty years old.⁹ This was meant to ensure that there had been time for a consensus to emerge that designated sites had sufficient historical or aesthetic merit.¹⁰ Today, the Landmarks Law is over forty

3. N.Y.C. ADMIN. CODE § 25-301 (1992); see also Cindy Moy, Note, *Reformulating the New York City Landmarks Preservation Law's Financial Hardship Provision: Preserving the Big Apple*, 14 CARDOZO ARTS & ENT. L.J. 447, 479 (1996); James Baron, *Celebrating 45 Years of Preserving New York*, N.Y. TIMES CITY ROOM BLOG (Apr. 19, 2010, 6:15 PM), <http://cityroom.blogs.nytimes.com/2010/04/19/celebrating-45-years-of-preserving-new-york/>.

4. An exterior landmark is a building or object designated as a landmark. An interior landmark is an interior space, which must be commonly accessible to the public. A scenic landmark is a landscape or group of features, such as Central Park, while a historic district is an area representative of a style of architecture prevalent in a particular historical period. N.Y.C. ADMIN. § 25-302.

5. FURMAN CENTER FOR REAL ESTATE AND PUBLIC POLICY, STATE OF NEW YORK CITY'S HOUSING AND NEIGHBORHOODS 26 (2008), available at http://furmancenter.org/files/sotc/State_of_the_City_2008.pdf [hereinafter FURMAN CENTER REPORT]. The exact breakdown was 1194 exterior landmarks, 112 interior landmarks, 10 scenic landmarks, and 106 historic districts. *Id.*

6. Edward L. Glaeser, *Preservation Follies: Excessive Landmarking Threatens to Make Manhattan a Refuge for the Rich*, CITY JOURNAL, Spring 2010, available at <http://www.city-journal.org/printable.php?id=6091>.

7. CITY OF NEW YORK, MAYOR'S MANAGEMENT REPORT 108 (Sept. 2009), available at http://www.nyc.gov/html/ops/downloads/pdf/2009_mmr/0909_mmr.pdf [hereinafter 2009 MAYOR'S REPORT].

8. See, e.g., Elias Wolfberg, *Ninth Avenue Noir*, N.Y. TIMES, Jan. 20, 2002, <http://www.nytimes.com/2002/01/20/nyregion/ninth-avenue-noir.html> (stating that buyer willing to spend between forty-five and fifty-five million dollars would be needed to refurbish The Windermere); see also Janet Lorin, *New York University Challenged on 40% Expansion Plan*, BLOOMBERG BUSINESSWEEK, Apr. 14, 2010, <http://www.businessweek.com/news/2010-04-14/new-york-university-challenged-on-four-fronts-of-40-expansion.html>.

9. ADMIN. § 25-302(m)-(n).

10. See Thomas W. Ennis, *Landmarks Bill Signed by Mayor: Wagner Approves It Despite Protests of Realty Men*, N.Y. TIMES, Apr. 20, 1965, at 28. The thirty-year provision was not in the original law but was added during the City Council debates on

five years old.¹¹ Decisions made in 1967 may no longer reflect the current consensus among historians or cultural critics, or circumstances surrounding the landmark may have changed so dramatically as to make the landmark designation deleterious to the City at large. Yet short of near-total takings under the regulatory takings doctrine, it is virtually impossible to de-designate a building's landmark status. This Note will argue that it should be easier for owners to obtain de-designation of their property when it becomes an "albatross landmark,"¹² a landmark whose value to the City is significantly outweighed by the potential uses of the site.

While there is a process for rescission of a landmark designation,¹³ it is seldom used.¹⁴ Owners of landmarked property can also apply for certificates of appropriateness,¹⁵ certificates of no-effect,¹⁶ and certificates of insufficient return.¹⁷ But these provisions either offer only limited relief for property owners or require the LPC to analyze claims based on concerns well outside of their administrative competence.¹⁸ In response to this, two judicially created doctrines have evolved: the "hardship exception" for nonprofit organizations and the "regulatory takings doctrine" for commercial operations.¹⁹ But these doctrines create their own problems. Both provide incentives for owners of landmarked properties to take sub-optimal care of the properties or their businesses.²⁰ Because of the lack of any real possibility of rescission under the current regime, disputes center on the regulatory takings doctrine, which is not well suited to address whether maintaining a landmark designation is

the legislation. Thomas W. Ennis, *Landmarks Bill Goes to Council: Protective Zone Is Cut but Architectural Rules Stay*, N.Y. TIMES, Mar. 24, 1965, at 50.

11. Barron, *supra* note 3.

12. This phrase is used here to refer to any landmark that creates a significant encumbrance or burden to the city. See OXFORD ENGLISH DICTIONARY (2d ed. 1989) (defining "albatross" as a "source or mark of misfortune, guilt, etc., from which one cannot (easily) be free; a burden or encumbrance."). The point of this Note is not to make substantive judgments about any existing landmarks, but instead to articulate standards by which the LPC could judge landmarks in the future.

13. ADMIN. § 25-303(h)(1). This provision applies to all decisions of the LPC and is more commonly used to rescind orders permitting alterations to landmarks.

14. *Infra* Part I.E.

15. ADMIN. § 25-307.

16. *Id.* § 25-306.

17. *Id.* § 25-309.

18. *Infra* Part II.D.

19. *Infra* Parts II.B–C. The regulatory takings doctrine does not, strictly speaking, apply only to commercial owners, but this is the relevant distinction for the purposes of this Note.

20. *Infra* Part II.D.

desirable. The better question is whether the landmark status of a building should be maintained based on a balancing of historical worth against costs associated with the landmark, and the current process for rescission does not substantively engage this question.

Decisions to rescind are rare. In addition, few claims are made under the reasonable returns standard, the regulatory takings doctrine, and either the statutory or judicial hardship exceptions. Given the history of New York City's preservationist efforts,²¹ this may be desirable. But in cases where landmark status no longer serves the needs or interests of New York City, the status quo creates significant problems: it allows demolition by neglect, creates incentive structures that may cause needless loss, and leads to heavily adversarial disputes between developers and preservationists.

This Note will argue that the Landmarks Law's existing rescission provision should be modified and used more frequently. The current statutory provision is inadequate because it fails to provide clear standards to be used by the LPC in evaluating whether rescission is merited. In addition, the mechanisms used in lieu of rescission do not address the proper issues. The Landmarks Law should be changed to allow for rescission when the party seeking rescission can demonstrate that there has been a change in the circumstances relevant to the landmark, the particular landmark itself no longer adds significant value to the general landmark scheme, or the costs of maintaining the landmark status are unjustifiably high compared to the benefits that the City receives from the landmark. The burden of demonstrating both that there has been a relevant change and that the circumstances justify rescission should be placed on property owners. To balance against concerns of re-litigation and abuse by developers, procedural constraints such as timing provisions should also be added.

Part I addresses the nature of landmarks as public goods, the impact of historic preservation on development, and the institution of the LPC. It then argues that landmark status should not be perpetual. Part II examines the current mechanisms within the Landmarks Law for removal of landmark designations and argues that they are inadequate and create distortions. Part III proposes that the Landmarks Law should be changed to include provisions allowing for rescission under certain circumstances.

21. *Infra* Part I.C.

I.
LANDMARKING IN NEW YORK CITY

A. *Landmarks and Local Government Law*

This Note will proceed from two assumptions regarding landmarks. The first is that landmarks have some value, both intrinsic²² and extrinsic,²³ to New York City.²⁴ The second assumption is that landmarks are a form of public good,²⁵ albeit of a somewhat atypical variety. It is necessary, therefore, to define “public good” and demonstrate how landmarks fit into the category. A public (or collective consumption) good is classically defined as a resource that is both non-rival and non-excludable.²⁶ A good is non-rival if one person’s use of the good does not conflict with another’s use of it,²⁷ and it is non-excludable if no one can be prevented from using

22. See generally, e.g., Elizabeth C. Gutman, Note, *Landmarks as Cultural Property: An Appreciation of New York City*, 44 RUTGERS L. REV. 427 (1992). The claim that cultural preservation has intrinsic value is not uncontroversial. See, e.g., James W. Nickel, *Intrinsic Value and Cultural Preservation*, 31 ARIZ. ST. L.J. 355 (1999) (responding to Sarah Harding, *Value, Obligation and Cultural Heritage*, 31 ARIZ. ST. L.J. 291 (1999)). That particular debate is beyond the scope of this Note. Additionally, the question of what constitutes the correct rationale for historic preservation in general has never been entirely settled. David F. Tipson, *Putting the History Back in Historic Preservation*, 36 URB. LAW. 289, 289 (2004). This Note does not attempt to reconcile the different rationales, beyond assuming that the widespread use of historic preservation and the ferocity with which these controversies are normally contested is sufficient to demonstrate some intrinsic value.

23. These benefits could include anything from increased tourism to a larger tax base because of increased property values. *Infra* Part I.B. It also may be possible to speak of these benefits as being divided into “direct” benefits, such as the preservation of architecture itself, and “indirect” benefits, such as increases in tourism or property values. Tipson, *supra* note 22, at 294. For these purposes, the precise nomenclature is unimportant.

24. The two categories of value are not mutually exclusive. Some of the “intrinsic” value of a landmark may be that it promotes some non-quantifiable social good, which in turn produces “extrinsic” benefits to the city. See generally Carol M. Rose, *Preservation and Community: New Directions in the Law of Historic Preservation*, 33 STAN. L. REV. 473 (1981).

25. Randall Mason, *Economics and Historic Preservation: A Guide and Review of the Literature*, THE BROOKINGS INSTITUTION 3 (2005), http://www.brookings.edu/~media/Files/rc/reports/2005/09metropolitanpolicy_mason/20050926_preservation.pdf (proceeding from the assumption that “historic preservation is a legitimate public good”).

26. Paul A. Samuelson, *The Pure Theory of Public Expenditure*, 36 REV. ECON. & STAT. 387, 387 (1954).

27. N. GREGORY MANKIW, PRINCIPLES OF MICROECONOMICS 224 (4th ed. 2007).

it.²⁸ Examples of public goods include roads, national defense, and tornado sirens.²⁹

Landmarks, or at the very least the historic or aesthetic features that landmark status is designed to protect, fit this definition.³⁰ Landmarks are goods in the sense that they provide cultural or aesthetic benefits to the city at large,³¹ and their presence is both non-rival and non-excludable. They are non-rival in the sense that one person's enjoyment of a landmark does not affect another's ability to enjoy it, and non-excludable insofar as the aesthetic value of preserving a piece of architecture, or the historical value of maintaining the exteriors of a neighborhood, cannot be easily taken away from any one person.³²

Because landmarks, as public goods,³³ benefit everyone, no one person has an incentive to provide the optimal level of care for

28. *Id.*

29. *Id.* The literature surrounding public goods often makes a distinction between "pure" and "impure" public goods. *See, e.g.*, James M. Buchanan, *An Economic Theory of Clubs*, 32 *ECONOMICA* 1 (1965). Each of the examples given here could be considered an impure public good, at least insofar as it is potentially possible for each of them to become rival in the event of scarcity. This distinction is somewhat irrelevant for the purposes of this Note, as there are very few public goods for which the possibility of rivalry is completely non-existent. *See also* Mason, *supra* note 25, at 11 (discussing how historic preservation is both public and private).

30. *See, e.g.*, Per-Olof Bjuggren & Henrik af Donner, *Ownership of a Cultural Landmark: The Case of Gotha Canal*, 21 *INT'L REV. L. & ECON.* 499, 504-05 (terming the canal an "impure public good").

31. *See* Mason, *supra* note 25, at 3. Local government theorists and economists often make a distinction between "public goods" and "local public goods." *See, e.g.*, Charles M. Tiebout, *A Pure Theory of Local Expenditures*, 64 *J. POL. & ECON.* 416, 418 (1956). Since this Note deals entirely with local government landmarks, the distinction is largely irrelevant for these purposes, except insofar as some of the criticisms of national-level distribution of public goods are not as salient when directed at local government expenditures, and insofar as the benefits of landmarking may have differing consequences for residents of different parts of the city.

32. There are some elements of landmark protections that are excludable. If a neighborhood gentrifies as a result of a historic district designation, then some may not receive the full benefits of the designation. Interior landmarks are also potentially excludable and could be subject to entrance fees; however, the Landmarks Law itself defines interior landmarks as places commonly available to the public, N.Y.C. ADMIN. CODE § 25-302(m) (1992), and interior landmarks are a relatively small portion of the total number of landmarks in the city. *See* FURMAN CENTER REPORT, *supra* note 5, at 26.

33. One other possible definitional problem is that it is possible to consider parts of the current landmark scheme, especially the historic districts, as part of an exclusionary zoning system. That discussion, while occasionally relevant, is largely beyond the scope of this Note. For a discussion of exclusionary zoning in the context of cities and neighborhoods as public goods, see generally Lee Anne Fennell,

them.³⁴ Further, local residents may be unwilling to express their actual level of preference for having landmarks, hoping others will subsidize the costs.³⁵ While some landmarks may be amenable to exclusion mechanisms,³⁶ most are not.³⁷ As such, without some form of government intervention, one would expect there to be fewer landmarks.³⁸ But government intervention is also problematic: as landmarks create externalities, governments may not be able to accurately measure the public preference for landmarks; and unlike with other public goods, control over a landmark is typically in private hands.

The first problem with government intervention to preserve landmarks is that landmark decisions create significant externalities, both positive and negative.³⁹ In general, the property owner will bear a disproportionate share of the negative externalities and fail to recoup his or her entire investment. If landowners were able to capture the value of the positive externalities, then perhaps they would have an incentive to preserve buildings themselves. Instead, the burdens fall onto owners of landmarked property, while the benefits, such as increased land values in the surrounding area⁴⁰ and the historic value of the site itself, largely benefit the public.

Exclusion's Attraction: Land Use Controls in the Tieboutian Perspective, in THE TIEBOUT MODEL AT FIFTY: ESSAYS IN PUBLIC ECONOMICS IN HONOR OF WALLACE OATES 163, 175 (William A. Fischel ed., 2006).

34. See, e.g., Jide O. Nzeli, *Interest Groups, Power Politics, and the Risks of WTO Mission Creep*, 28 HARV. J.L. & PUB. POL'Y 89, 94 n.19 (2004) (citing MANCUR OLSON, *THE LOGIC OF COLLECTIVE ACTION: PUBLIC GOODS AND THE THEORY OF GROUPS* (1965)).

35. Ariel Porat, *Private Production of Public Goods: Liability for Unrequested Benefits*, 108 MICH. L. REV. 189, 205 (2009) (proposing system for restitution in certain other contexts); see also LYNN A. BAKER & CLAYTON P. GILLETTE, *LOCAL GOVERNMENT LAW: CASES AND MATERIALS* 33–35 (3d ed. 2004).

36. See, e.g., Bjuggren & Donner, *supra* note 30, at 505.

37. As noted above, of the four categories of Landmarks in New York City, interior landmarks are the only form for which it is plausible that this could work.

38. Porat, *supra* note 35, at 205–06; cf. John P. Conley & Christopher S. Yoo, *Nonrivalry and Price Discrimination in Copyright Economics*, 157 U. PA. L. REV. 1801, 1802 (2009) (explaining that “a core policy implication of public goods theory is that markets tend to produce too few public goods and underutilize those that are produced,” and arguing for different conception of public goods in context of copyright law). This appears to be borne out by the history of landmarks in New York City as well. See *infra* notes 71–73.

39. See *infra* Part I.B.

40. This is somewhat different in the context of historic districts, rather than exterior landmarks, where the owners of the property are far more able to see the benefits of the designation themselves. These owners also tend to be families holding their own homes rather than large commercial or nonprofit entities, which may also alter the theoretical framework.

Thus, property owners have a significant incentive to fight landmark designations in the first instance.

The second problem is that the government may not be able to accurately account for the public's preferences for landmarks. The costs and benefits may be hard to quantify,⁴¹ and voting patterns may reveal very little about preferences for landmarking.⁴² Worse yet, there is reason to believe that landmarking in particular may cause politicians to be particularly responsive to "repeat players" such as preservationists and developers.⁴³

The third problem is that, unlike more archetypal public goods such as roads or national defense, property being considered for landmark status is typically in private hands.⁴⁴ In this sense, the designation of a building as a landmark is an expropriation: the public benefits of protecting a landmark come at the expense of the private property owner's rights to exploit the property.⁴⁵ While courts have addressed the constitutionality of landmark regimes as regulatory takings,⁴⁶ the tests that they have created are by no means clear, reflecting significant tension between the existing law and the redistributive consequences of landmark decisions. In addition, while courts have referred to "average reciprocity of advantage,"⁴⁷ or the idea that restrictions that might otherwise be a taking are permitted if their application generally benefits all property owners,⁴⁸ there does not appear to be data to support the claim that this phenomenon actually occurs. Furthermore, the ferocity

41. Mason, *supra* note 25, at 11–18 (comparing different econometric techniques for valuation of benefits from landmarks); *see also* Paolo Rosato et al., *Redeveloping Derelict and Underused Historic City Areas: Evidence from a Survey of Real Estate Developers*, 53 J. ENVTL. PLAN. & MGMT. 257, 261–63 (2010).

42. *Infra* Part I.D.

43. *Infra* Part I.D.

44. Some of the burdens of other public goods may occasionally fall on private parties. For example, cities may mandate that property owners shovel snow from in front of their properties in order to maintain the safety of roads. N.Y.C. ADMIN. CODE § 16-123 (1992). But these examples tend to be either *de minimis*, as in the snow-shoveling example, or related to exceptional circumstances, such as the need to quarter troops in times of war. *See* U.S. CONST. amend. III. Examples in which the burden of consistently maintaining an expensive public good is permissibly placed almost entirely on a private party are extremely rare.

45. Rose, *supra* note 24, at 497.

46. *See* Penn Cent. Transp. Co. v. New York City, 438 U.S. 104 (1978) (discussed *infra* Part II.C.1.).

47. *See, e.g.*, Tahoe-Sierra Pres. Council, Inc. v. Tahoe Reg'l Planning Agency, 535 U.S. 302, 324 n.19 (2002) (discussing Lucas v. South Carolina Coastal Council, 505 U.S. 1003, 1017–18 (1992)).

48. Richard A. Epstein, *The Disintegration of Intellectual Property? A Classical Liberal Response to a Premature Obituary*, 62 STAN. L. REV. 455, 478–79 (2009).

with which owners sometimes fight these designations provides at least indirect evidence that owners do not agree with the claim that they receive sufficient benefits from the landmarking scheme.⁴⁹

B. Landmarks and Development

Landmark decisions also have distributional consequences and impact the general economic development of the city.⁵⁰ The aggregate impact is hard to determine, and the question of who benefits and who loses in that equation is similarly murky. On one hand, a landmark district can accelerate the process of gentrification.⁵¹ Similarly, property values surrounding historic landmarks tend to increase,⁵² and advocates of preservation argue that economic indicators like retail sales in stores in historic districts tend to in-

49. For reasons that are largely political, the LPC prefers to landmark buildings by consent. For a discussion of an example that includes the political factors involved in landmarking, the consent policy, and a successful fight against landmarking, see Eliot Brown, *After Push by Extell, Landmarks Backs Down Over West 57th Street Building*, N.Y. OBSERVER, Nov. 10, 2009, <http://www.observer.com/2009/real-estate/after-push-extell-landmarks-backs-down-57th-st>. The tendency towards consent seems somewhat illusory, as there is very little the owner of a property that the LPC wants to designate can offer in a negotiation. For another example of the intensity of the negative reaction to a proposed designation, see Eliot Brown, *Oh, God, No*, N.Y. OBSERVER, Feb. 16, 2010, <http://www.observer.com/2010/real-estate/oh-god-no>.

50. Robin M. Leichenko, N. Edward Coulson & David Listokin, *Historic Preservation and Residential Property Values: An Analysis of Texas Cities*, 38 URB. STUD. 1973, 1984 (2001); David M. Stewart, Note, *Constitutional Standards for Hardship Relief Eligibility for Nonprofit Landowners Under New York City's Historic Preservation Act*, 21 COLUM. J.L. & SOC. PROBS. 163, 167–68 (1988); see also Glaeser, *supra* note 6 (asserting that large historic districts “are associated with a reduction in housing supply, higher prices, and increasingly elite residents”).

51. Loretta Lees, *Super Gentrification: The Case of Brooklyn Heights, New York City*, 40 URB. STUD. 2487, 2494 (2003). While the question of whether gentrification itself is net positive or net negative for the city as a whole is thankfully beyond the scope of this Note, it is fair to say that gentrification generally increases the city's tax base. Jacob L. Vigdor, *Does Gentrification Harm the Poor?*, BROOKINGS-WHARTON PAPERS ON URBAN AFFAIRS 133, 133 (2002).

52. Leichenko et al., *supra* note 50, at 1976. The authors compared the previous empirical studies of the impact of designation on property values. Of the fourteen prior studies, seven suggested a positive impact, while four found the impact to be neutral. Two showed a negative impact, and one showed a mixed impact. Leichenko, Coulson, and Listokin concluded that landmark designations tend to have a positive impact on property value but “may displace less affluent residents of historic areas.” *Id.* at 1984; see also N.Y.C. INDEP. BUDGET OFFICE, *THE IMPACT OF HISTORIC DISTRICTS ON RESIDENTIAL PROPERTY VALUES* 8 (Sept. 2003), available at http://www.dlnhs.org/IBO_HistoricDistricts03.pdf (concluding that while property values tended to be higher within historic districts, there was insufficient causal evidence).

crease.⁵³ On the other hand, landmark status can slow the development of an area⁵⁴ by maintaining potentially sub-optimal⁵⁵ use of the land.⁵⁶ Because potential investors cannot realistically expect further development of the site itself, landmark status also discourages increased investment in the surrounding neighborhood. While the area around the landmark should benefit from increased tourism both from visitors and city residents,⁵⁷ it is not clear that these benefits outweigh the drop-off in investment.⁵⁸ While precise numbers are often hard to find and separate from other factors, “heritage tourism” is a large industry.⁵⁹ Advocates of preservation have produced studies suggesting that heritage tourism can bring in millions of dollars in a given year.⁶⁰ It is important to note that landmarking is often used in ways that are not among the intended purposes of the Landmarks Law. For example, landmark status can be used as an explicitly ideological, anti-development tool⁶¹ or as a political tool for purposes other than historic preservation.⁶²

53. See, e.g., NAT’L HISTORIC TRUST FOR PRES., ECONOMIC IMPACT OF HISTORIC DISTRICT DESIGNATION: LOWER DOWNTOWN DENVER, COLORADO 7–8 (1990).

54. Glaeser, *supra* note 6 (arguing that historic districts face slowed rates of new construction). While this should be unsurprising (after all, part of the reason to designate something as a historic district is to slow new construction), Glaeser argues that this drives up prices artificially, creating a significant obstacle to affordable housing in historic districts. See *id.*

55. Here “sub-optimal” is used purely to describe economic efficiency. MANKIW, *supra* note 27, at 148.

56. Stewart, *supra* note 50, at 166–67.

57. DONOVAN D. RYPKEMA, VIRGINIA’S ECONOMY AND HISTORIC PRESERVATION 4–5 (1995).

58. Mason, *supra* note 25, at 21 (“Historic preservation . . . produces certain economic benefits for both private actors and the public at large [P]reservation policies do make sound fiscal sense. However, the economic impacts and measures of historic preservation activities are too situational to be able to extrapolate widely.”).

59. ADVISORY COUNCIL ON HISTORIC PRESERVATION, HERITAGE TOURISM AND THE FEDERAL GOVERNMENT 6–7 (2002), available at http://www.achp.gov/heritage_tourismsummit.pdf.

60. RYPKEMA, *supra* note 57, at 4.

61. See, e.g., Norman Oder, *Census of Places that Matter, art opening, and the (upcoming) “vanished site” of Freddy’s*, ATLANTIC YARDS REPORT (Mar. 8, 2010, 2:21 PM), <http://atlanticyardsreport.blogspot.com/2010/03/census-of-places-that-matter-and-art.html> (rejecting suggestion that landmark status would be obtainable for Freddie’s Back Room, in blog post on site generally opposed to Brooklyn Atlantic Yards Project).

62. David B. Fein, Note, *Historic Districts: Preserving City Neighborhoods for the Privileged*, 60 N.Y.U. L. REV. 64, 90 (1985); see also Editorial, *Ground Zero Tolerance: Future New York Mosque a fitting symbol of American Values*, HOUSTON CHRONICLE, Aug. 12, 2010, <http://www.chron.com/disp/story.mpl/editorial/7151485.html> (discussing LPC’s refusal to landmark site proposed for “Ground Zero Mosque”);

There are also other benefits to historic preservation, both quantifiable and qualitative.⁶³ Historic preservation can help inculcate a sense of community,⁶⁴ which can produce other benefits as a result, including strengthened social ties⁶⁵ and greater stability in neighborhood populations.⁶⁶ Historic preservation could create unintended beneficial results, such as increasing the “pull” factors of cities and neighborhoods.⁶⁷

*C. The Landmarks Preservation Commission’s Composition,
Legal Mandate, and Processes*

Landmarking decisions in New York City are made by the Landmarks Preservation Commission. The LPC was established by New York City’s Landmarks Preservation and Historic District Law of 1965.⁶⁸ The Landmarks Law is a local law passed under the New

Tom Topousis, *Jane Hotel Cite to Behold*, N.Y. POST, Oct. 7, 2009, http://www.nypost.com/p/news/local/manhattan/item_RYCpYEUgwKkSnPvfKJRjiO (describing use of landmark law violations in response to neighborhood complaints about bar at Jane Hotel). The use of landmark status as an ideological or anti-development tool has been a longstanding problem with the Landmarks Law. See David W. Dunlap, *Change on the Horizon for Landmarks*, N.Y. TIMES, Apr. 29, 1990, at 32.

63. While precise numbers are not necessarily relevant to this point, it is worth noting that economists have attempted to account for some of the harder to quantify or strictly qualitative benefits. For a summary of some of these methods, see Mason, *supra* note 25, at 16–18.

64. Rose, *supra* note 24, at 497.

65. See John Costonis, *Law and Aesthetics: A Critique and Reformulation of the Dilemmas*, 80 MICH. L. REV. 355, 392–94 (1982). See generally Rose, *supra* note 24. Both sources are cited in ROBERT C. ELLICKSON & VICKI L. BEEN, *LAND USE CONTROLS: CASES AND MATERIALS* 496 (3d ed. 2005).

66. *But see* Leichenko et al., *supra* note 50, at 1984 (mentioning potential displacement of residents).

67. There is a significant debate over which level of government is most appropriate for the redistribution of public goods, but this debate is beyond the scope of this Note. For a discussion of the strengths and weaknesses of centralized and decentralized regimes of government in the context of public good distribution, compare Roderick M. Hills, Jr., *Compared to What? Tiebout and the Comparative Merits of Congress and the States in Constitutional Federalism*, in *THE TIEBOUT MODEL AT FIFTY: ESSAYS IN PUBLIC ECONOMICS IN HONOR OF WALLACE OATES* 239 (William A. Fischel ed., 2006), with Clayton P. Gillette, *The Tendency to Exceed Optimal Jurisdictional Boundaries*, in *THE TIEBOUT MODEL AT FIFTY: ESSAYS IN PUBLIC ECONOMICS IN HONOR OF WALLACE OATES* 264 (William A. Fischel ed. 2006) (commenting on Hills). While both of these essays focus on the distinction between the federal and state levels of governance, the arguments are applicable in the context of state and local governance as well.

68. Barron, *supra* note 3.

York State Municipal Laws Enabling Act.⁶⁹ The purpose of the law was to protect “landscape features, as herein defined, having a special character or a special historical or aesthetic interest or value”⁷⁰

The conventional narrative is that the creation of the LPC was largely driven by concerns that emerged from the demolition of Penn Station in October of 1963.⁷¹ In reality, landmarking had been a significant and controversial area of city politics for quite some time before the destruction of Penn Station.⁷² In 1941, Robert Moses’ attempted destruction of Castle Garden⁷³ prompted a resolution at an angry meeting of the New-York Historical Society and the American Scenic and Historic Preservation Society,⁷⁴ and in 1961, Mayor Robert Wagner created the Committee for the Preservation of Structures of Historic and Esthetic Importance.⁷⁵ Regardless of the historical reasons underlying its creation, the LPC has been the primary force for preservation in New York City since 1965.

69. N.Y. GEN. MUN. LAW § 96-a (2006) (“[T]he governing board or local legislative body of any . . . city . . . is empowered to provide by regulations, special conditions and restrictions for the protection, enhancement, perpetuation and use of places, districts, sites, buildings, structures, works of art, and other objects having a special character or special historical or aesthetic interest or value.”).

70. N.Y.C. ADMIN. CODE § 25-301 (1992).

71. See, e.g., FURMAN CENTER REPORT, *supra* note 5, at 26; Barron, *supra* note 3. But see ANTHONY C. WOOD, PRESERVING NEW YORK: WINNING THE RIGHT TO PROTECT A CITY’S LANDMARKS 9 (2008) (referring to narrative as “myth”).

72. Rose, *supra* note 24, at 481–91 contains an overview of some of the thematic developments in the justifications for landmark preservation. WOOD, *supra* note 71, is a comprehensive treatment of the history of New York City’s landmark scheme.

73. ROBERT A. CARO, THE POWER BROKER: ROBERT MOSES AND THE FALL OF NEW YORK 639–87 (1975) contains a lengthy treatment of this dispute. Caro’s work, despite its relative age, remains the definitive study of the life and work of Robert Moses. Castle Garden was initially saved from destruction through the personal intervention of Eleanor Roosevelt, *id.* at 672, and was designated a national landmark in 1946. *Id.* at 686. It was ultimately placed on the National Register of Historic Places in 1966 under the National Historic Preservation Act. 16 U.S.C. § 470 (2006). While the site was not protected by local law, the dispute demonstrates that, prior to the enactment of the landmarks law, there were still preservationist mechanisms. The problem was not that historic buildings were never preserved but rather that there was no singular process for determining which ones ought to be preserved. For a review of the historical significance of Castle Garden, see generally BARRY MORENO, CASTLE GARDEN AND BATTERY PARK (2007).

74. WOOD, *supra* note 71, at 13–14.

75. *Id.* at 9.

The LPC is composed of eleven commissioners,⁷⁶ which must include “at least one resident of each of the five boroughs,”⁷⁷ three architects,⁷⁸ one “historian qualified in the field,”⁷⁹ a city planner or landscape architect,⁸⁰ and a realtor.⁸¹ None of these terms is defined,⁸² but the mayor is permitted to consult with “the fine arts federation of New York and any other similar organization” when appointing architects, historians, or city planners.⁸³ The commissioners are appointed by the mayor⁸⁴ and serve three-year, staggered terms.⁸⁵ The mayor also designates a chair and vice-chair for the commission.⁸⁶ With the exception of the chair, the commissioners are not paid for their work on the commission, but they do receive reimbursement for necessary expenses.⁸⁷ There is also a panel “independent of the commission” with five members “appointed by the mayor with advice and consent of the council” who hear appeals from denials of a variety of exemptions from LPC decisions.⁸⁸

The LPC has the legal authority to declare sites as individual (or exterior) landmarks, interior landmarks, scenic landmarks, and historic landmarks.⁸⁹ The commission can act on its own recommendation⁹⁰ or in response to a public petition.⁹¹ The commission can also amend designations⁹² and approve or disapprove proposed modifications or alterations to existing landmarks.⁹³ Within 120 days of the designation of a historic site, the city council can “modify or disapprove” the designation by majority vote.⁹⁴ The mayor

76. N.Y.C. CHARTER § 3020(1) (1989).

77. *Id.*

78. *Id.*

79. *Id.*

80. *Id.*

81. *Id.*

82. *See* N.Y.C. ADMIN. CODE §§ 1-112, 25-302 (1992); N.Y.C. CHARTER § 3020 (1989).

83. N.Y.C. CHARTER § 3020(2)(b).

84. *Id.* § 3020(2)(a).

85. *Id.*

86. *Id.* § 3020(4).

87. *Id.* § 3020(3).

88. *Id.* § 3020(10)(a).

89. N.Y.C. ADMIN. CODE § 25-303(a) (1992); *see supra* note 4.

90. ADMIN. § 25-303.

91. *Id.*

92. *Id.* § 25-303(c).

93. *Id.* § 25-303(d). The power to control construction and alterations of landmarks is more fully addressed in § 25-305.

94. N.Y.C. Charter § 3020(9) (1989).

can veto council resolutions relating to LPC decisions, subject to override by a two-thirds vote of the council.⁹⁵

Once a site has been designated a landmark, its owner loses the right to modify or use the property in certain ways.⁹⁶ While a variety of waivers and exemptions exist,⁹⁷ the loss of rights in the property and the concomitant obligations to maintain the landmark⁹⁸ have been described as onerous⁹⁹ and are enforced by both civil penalties¹⁰⁰ and criminal sanctions, including imprisonment and fines.¹⁰¹ Under certain conditions, the LPC is even authorized to condemn and seize landmarked property in order to protect it¹⁰² and, in extraordinary cases, may file lawsuits against property owners to force repairs or levy further fines.¹⁰³

The “rescission” provision¹⁰⁴ states that “[t]he commission shall have the power, after a public hearing, to adopt a resolution proposing rescission, in whole or in part, of any designation or

95. *Id.*

96. ADMIN. § 25-305.

97. *See, e.g., id.* § 25-306 (discussing “certification of no effect on protected architectural features”); *infra* Part II.A.

98. ADMIN. § 25-311.

99. *Penn Cent. Transp. Co. v. New York City*, 438 U.S. 104, 139 (1978) (Rehnquist, J., dissenting).

100. ADMIN. § 25-317.1.

101. *Id.* § 25-317. In 2009, approximately seven hundred enforcement actions and more than one thousand warning letters were issued. *See* 2009 MAYOR’S REPORT, *supra* note 7, at 108. As a practical matter, courts sometimes stay the enforcement of these violations in order to allow violators to “cure” the problem. *See, e.g.,* 259 West 12th, LLC v. Grossberg, 836 N.Y.S.2d 504 (City Civ. Ct. 2007) (allowing defendant found to be in violation of her lease because of failure to comply with LPC regulations a ten-day stay in order to obtain approval for changes made to her apartment). *But see* Craig Karmin, *City: Take Off Top Floor of Townhouse*, WALL ST. J., May 4, 2010, <http://online.wsj.com/article/SB10001424052748704342604575222053129473006.html>.

102. ADMIN. §25-309(g) (describing procedures in conjunction with certificate of appropriateness); *see also* Stewart, *supra* note 50, at 170.

103. *See, e.g.,* Mike McLaughlin, *City Sues Homeowner Over Crumbling Historic Buildings in Cobble Hill*, N.Y. DAILY NEWS, Apr. 22, 2010, http://articles.nydailynews.com/2010-04-22/local/27062355_1_vacant-buildings-historically-significant-buildings-buildings-department. Even leaving aside enforcement powers, the LPC’s statutory authorization is quite broad, because the power to approve or disapprove alterations within the historic district gives the LPC effective control over a variety of decisions that are probably better suited to the zoning board. *See, e.g.,* Glaeser, *supra* note 6 (discussing how LPC limited height of building proposed on corner of 91st Street and Madison Avenue, despite the explicit statutory limitation on LPC’s authority to regulate height of buildings).

104. ADMIN. § 25-303(h).

amendment or modification”¹⁰⁵ The provision describes the procedure for rescission, which includes a report from the city planning commission¹⁰⁶ and allows for city council and mayoral veto.¹⁰⁷ While this process closely mirrors the designation process, it is infrequently used.¹⁰⁸ There is no specific requirement that there be an additional public hearing, and there are no additional procedural requirements for acting upon a petition. When the rescission provision is used, the LPC typically issues a brief and largely conclusory decision with virtually no explanation of its reasoning.¹⁰⁹ Also missing is any rule governing how much time must elapse between designation and rescission or amendment of the designation.¹¹⁰ Finally, application for rescission does not prevent an owner from later seeking another avenue for removal of the landmark status.¹¹¹

105. *Id.* §25-303(h)(1).

106. *Id.* §25-303(h)(2).

107. *Id.* §25-303(h)(3). The city council and mayor can each approve through silence. *Id.*

108. *See infra* note 145.

109. *See, e.g.*, N.Y.C. LANDMARK PRESERVATION COMMISSION, *Landmark Site of Former New Brighton Village Hall*, Dec. 12, 2006, at <http://nyc.gov/html/records/pdf/govpub/2782nbvillagehallrecis.pdf>. The entirety of the “findings,” which are part of a scant six-page document, states that:

On the basis of a careful consideration of the history, the architecture, and other features of this Landmark and Landmark Site, the Landmarks Preservation Commission finds that the site of the former New Brighton Village Hall no longer possess special character or special historic or aesthetic interest or value as part of the development, heritage, and cultural characteristics of New York City.

The Commission further finds that the New Brighton Village Hall has been demolished, and that the site has been cleared of all structures. Accordingly, pursuant to the provisions of Chapter 74, Section 3020 of the Charter of the City of New York and Chapter 3 of Title 25 of the Administrative Code of the City of New York, the Landmarks Preservation Commission rescinds the designation of the Landmark of the New Brighton Village Hall and Landmark Site, which consists of Borough of Staten Island Tax Map Block 71, Lot 117.

This may be an exceptional case as the building was itself demolished, so there may have been little that the LPC could have said about it. Since rescission decisions are infrequent, an examination of some recent designation reports may be instructive. For example, while there is a lengthy recitation of historical facts, the “findings” in a report on the Brill Building are less than one page long. N.Y.C. LANDMARKS PRESERVATION COMMISSION, *The Brill Building*, Designation List 427, LP 2387 (Mar. 23, 2010), available at <http://www.nyc.gov/html/lpc/downloads/pdf/reports/brill.pdf>.

110. *See ADMIN.* §25-303(h)(1).

111. *See, e.g.*, N.Y.C. LANDMARKS PRESERVATION COMMISSION, DISPOSITION OF HARDSHIP APPLICATIONS 3 (2008) (discussing application and re-application by

*D. Institutional Concerns About the Landmarks
Preservation Commission*

The Landmarks Preservation Commission is not democratically accountable: the commissioners are appointed, not elected, and serve staggered terms,¹¹² so a mayor only has a chance to appoint the full commission by the final year of his or her first term of office. Even if the mayor had the power to appoint a full commission, there is little reason to believe that the voting public would care enough about landmarks to make this its single voting issue.¹¹³ Additionally, with the exception of the chair, the commissioners do not draw a salary,¹¹⁴ so even the threat of removing them from their position is weaker than it would be for ordinary city officers.¹¹⁵

Further, the LPC is vulnerable to “capture”¹¹⁶ by stakeholders in the process. An agency is typically thought to be vulnerable to capture if there is an interest group, or a small number of interest groups, with a disproportionate stake in the work of the agency relative to the general public.¹¹⁷ The two most likely participants in the landmark process are developers and preservationist associa-

owners of 74-86 Greene Avenue, Brooklyn) (internal document acquired through records request to LPC) [hereinafter “LPC HARDSHIP DISPOSITIONS”].

112. N.Y.C. CHARTER § 3020(2)(a) (1989).

113. A search of *New York Times* articles for the phrase “landmarks preservation commission” in the six months preceding the 2005 mayoral election reveals sixty-one articles, while the same search for the six months before that reveals forty-one. A search of the same six-month period for 2004 revealed thirty-eight articles. Most of the discrepancy between the results can be explained by one landmark controversy in particular, the “Lollipop” building at 2 Columbus Circle. *But see* Gutman, *supra* note 22, at 432 n.29 (“The destruction of historic landmarks became an issue during the 1961 mayoral election campaign.”).

114. N.Y.C. CHARTER § 3020(3).

115. Of course, there may be reputational constraints on commissioners’ behavior, but insofar as commissioners are likely to come from preservationist backgrounds, those constraints would likely lead to more, not less, preservation. *But see* Robin Pogrebin, *An Opaque and Lengthy Road to Landmark Status*, N.Y. TIMES, Nov. 25, 2008, <http://www.nytimes.com/2008/11/26/arts/design/26landmarks.html> (quoting preservationists as stating that chair of LPC is not committed to preservation and wields too much power over LPC’s agenda).

116. “Regulatory capture is generally understood to refer to the undue influence of a regulated party over the regulator.” Shi-Ling Hsu, *A Two-Dimensional Framework for Analyzing Property Rights Regimes*, 36 U.C. DAVIS L. REV. 813, 858 n.171 (2003) (citing, *inter alia*, IAN AYRES & JOHN BRAITHWAITE, *RESPONSIVE REGULATION: TRANSCENDING THE DEREGULATION DEBATE* ch. 3 (1992) (stating that scope of agency capture in general is not entirely clear)).

117. *See* Michael E. Levine & Jennifer L. Forrence, *Regulatory Capture, Public Interest, and the Public Agenda: Toward a Synthesis*, 6 J.L. ECON. & ORG. 167, 185 (1990) (discussing vulnerability to capture in terms of “slack” given to regulators by combination of high monitoring costs and impacted special interest groups).

tions, both of whom can be characterized as repeat players or special interest groups. Because landmarks operate as a public good in private hands, both participants have incentives, whether monetary or ideological, to try to win disputes in front of the LPC rather than attempt to reach negotiated solutions with one another. If a developer successfully fights off an attempt to landmark his property, he can maintain the unrestricted use without having to provide a public benefit from which he will be unable to capture the full return. On the other hand, a preservationist group seeking to save a building will not have to pay to maintain the site once it is declared a landmark.

Despite these concerns, there are reasons to believe that the LPC is not a captured agency. The first is that the composition of the commission is designed to balance competing interests.¹¹⁸ As noted above, six of the eleven seats on the commission have professional requirements and are split between architects, historians, city planners, and developers.¹¹⁹ On the other hand, the LPC uses a simple majority requirement for most matters, so the fact that six seats are split between different professions may be insufficient to prevent any one subset from gaining undue influence over the process.

Another reason why there may not be capture of the LPC is that, even if the landmarking process is structurally vulnerable to capture, as suggested above, the parties involved may have more to gain from negotiation than from capture of the commission. While high-profile landmark disputes, such as the recent fight over St. Vincent's Hospital, tend to create extremely bitter feelings on both sides,¹²⁰ most landmark decisions involve much lower stakes. This produces an incentive on both sides to reach an accommodation. Voluntary associations have often been involved in preservation efforts outside of the formal landmark process but in consultation with the LPC. There are also examples of partnerships in which both preservationists and private developers have worked in concert to preserve buildings while developing the area.¹²¹ Still, the

118. See N.Y.C. CHARTER § 3020(1) (1989).

119. *Id.*

120. See Glenn Collins, *Landmarks Panel Approves St. Vincent's Tower*, N.Y. TIMES CITY ROOM BLOG (Mar. 10, 2009, 5:30 PM), <http://cityroom.blogs.nytimes.com/2009/03/10/landmarks-panel-approves-new-tower-for-st-vincent/> (describing "yearlong battle" over O'Toole Building); John Del Signore, *Preservationists Sue to Save St. Vincent's O'Toole Building*, GOTHAMIST (Mar. 9, 2009, 3:27 PM), http://gothamist.com/2009/03/09/preservationists_file_lawsuit_to_sa.php.

121. See, e.g., David Gonzalez, *A Neglected Bronx Landmark Gets a New Life*, N.Y. TIMES CITY ROOM BLOG (Aug. 25, 2008, 10:50 AM), <http://cityroom.blogs.nytimes.com/2008/08/25/a-neglected-bronx-landmark-gets-a-new-life/>.

fact that exterior landmarks remain the most common form of landmark in New York City creates a situation in which “negotiated regulation” is possible, as preservationists work with developers to find ways to adapt interiors to new uses while preserving landmarked exteriors.¹²²

Finally, the two primary groups that would be interested in capturing the LPC, developers and preservationists, may simply balance each other out. This is possible but seems unlikely because developers may suffer from a collective action problem.¹²³ For any given landmark, there will tend to be one interested developer, whereas preservationists can seek to fight every major decision.¹²⁴

Another potential institutional issue that may impact the LPC’s behavior is the internal dynamic between the eleven commissioners. Both commissioners and private citizens can initiate the designation process.¹²⁵ But each of the commissioners may come to the process with a personal agenda.¹²⁶ Presumably, a historian is likely

com/2008/08/25/a-neglected-bronx-landmark-gets-a-new-life/; Press Release, National Trust for Historic Pres., *Drugstores: A Success Story in Westfield* (Sept. 12, 2002), available at <http://www.preservationnation.org/resources/case-studies/chain-drugstores/drugstores-westfield.html>.

122. See J. Peter Byrne, *Regulatory Takings Challenges to Historic Preservation Laws After Penn Central*, 15 *FORDHAM ENVTL. L. REV.* 313, 330–31 (2004) (arguing that “historic preservation law is well suited for this kind of negotiated regulation,” and using rezoning of SOHO lofts in the 1970s as primary example).

123. The collective action problem is that general lobbying requires everyone to get involved, but most developers only care to the extent their building is affected. Trade associations may be somewhat able to overcome this issue, especially in terms of systemic reform and lobbying. Ultimately, each individual designation is extremely valuable to the individual property owner, and of limited importance to the profession in general.

124. While both sides have cost concerns, those costs should be roughly the same for each, so for the purposes of discussing the relative strengths and weaknesses of each side, cost is something of a non-issue. In addition, some (although certainly not all) of the preservationist groups in New York City are extremely well funded. It is also worth mentioning that several developers have made significant donations to preservationist groups, such as the Municipal Art Society. See, e.g., MUNICIPAL ART SOCIETY, ANNUAL REPORT 2008–2009 11, 34, available at http://mas.org/images/media/original/MAS_AnnRep_2009_jul13_final.pdf. While further evidence would be needed, this suggests that it may be the case that developers feel it is easier to work with preservationists than lobby the LPC.

125. A private citizen can begin the process by filing a “Request for Evaluation” with the LPC. See N.Y.C. LANDMARKS PRESERVATION COMMISSION, <http://www.nyc.gov/html/lpc/html/propose/landmark.shtml>.

126. Though it is beyond the scope of this Note, there is a lurking question as to whom the LPC will favor. Though the LPC was created, in part, to avoid the unseemly influence peddling that dominated landmark decisions prior to its enactment, it might still reflect the power dynamic of city politics writ large, thus impli-

to be sympathetic to claims of historic value, while an architect is likely to seek out aesthetic monuments for preservation. Whenever a commissioner wants a particular building landmarked, the incentive structure suggests that the rational strategic choice for another commissioner would be to approve that landmark. Because the costs of the decisionmaking process (i.e. the hearings, reports, and other statutory requirements) are the same regardless of whether the commission chooses to designate a site, these costs can be considered neutral. Once the site is landmarked, there are enforcement costs, but these tend to be relatively low.¹²⁷ If any individual commissioner wants a site designated, other commissioners are likely to approve that choice in order to maintain goodwill and have their own preferences respected.

While the LPC does not keep searchable databases of votes, a look at some of the voting agendas posted on their website bears this out.¹²⁸ The agenda for December 8, 2009, contains twenty-three items, seventeen of which were brought to a vote.¹²⁹ In these

cating questions of race and class. Also troubling, the commission may perpetuate the biases of those people who tend to sit on the city's artistic boards, which are not even vaguely representative of the city's population. These issues can affect both judgments about the impact of a landmark decision on the neighborhood as a whole, and decisions about what to landmark, in the sense that the history of a particular group may be undervalued. *See, e.g.,* Rose, *supra* note 24, at 478 (“[P]oor black families might be displaced as middle class whites moved into spruced-up ‘historic’ neighborhoods . . . and . . . it wasn’t *black* history that the preservationists had in mind.”) (citing Michael deHaven Newsom, *Blacks and Historic Preservation*, 36 LAW & CONTEMP. PROBS. 423, 423–24 (1971)); Tipson, *supra* note 22, at 309 (“It is only recently that working-class neighborhoods have been appreciated as historic districts.”). There is some evidence that this tendency is changing. *See, e.g.,* N.Y.C. LANDMARK PRESERVATION COMMISSION, RIDGEWOOD NORTH HISTORIC DISTRICT DESIGNATION REPORT 1–2 (2009) (describing the significance of model tenements built in the early 20th century), *available at* <http://www.nyc.gov/html/lpc/downloads/pdf/reports/rnhd.pdf>.

127. Of the LPC’s total budget of a little under five million dollars in 2009, less than one million dollars was spent on all actions besides “personal services.” The category of “personal services” includes the entire hearing process and general work of the commission, while “other expenditures” include all other work. THE CITY OF NEW YORK, EXPENSE REVENUE CONTRACT 146E (2010), *available at* http://www.nyc.gov/html/omb/downloads/pdf/erc6_09.pdf.

128. The archive of LPC calendars can be found at WORKING WITH LANDMARKS: CALENDAR ARCHIVE, http://www.nyc.gov/html/lpc/html/working_with/calendar_archive.shtml. A few recent agendas were chosen more or less at random for this purpose.

129. *Public Hearing of the Landmark Preservation Commission* (Dec. 8, 2009), *available at* http://www.nyc.gov/html/lpc/downloads/pdf/calendar/12_08_09.pdf.

votes, there was one dissent and one abstention.¹³⁰ The April 13, 2010 agenda contains eleven items, all of which reached a vote.¹³¹ There were zero votes in opposition and zero abstentions.¹³² While there are occasional contested votes, such as the approval of St. Vincent's Hardship Application,¹³³ the posted agendas contain mostly unanimous or nearly unanimous votes.

To explain the general lack of dissent within the LPC, a look at some of the scholarly work done on collegial courts may be instructive.¹³⁴ The phenomenon of "dissent aversion,"¹³⁵ in which judges engaged in a cooperative structure will actively avoid leaving a majority opinion,¹³⁶ may explain why votes on the LPC tend not to produce 6–5 decisions. This is especially true because, like appellate judges but unlike university faculties or lower civil servants, the commissioners are not chosen by "a stable, uniform management layer."¹³⁷ Rather, they are appointed by a democratically elected mayor, which would tend to create a greater likelihood of dissent aversion between commissioners.¹³⁸

Furthermore, the LPC may be vulnerable to informational cascading.¹³⁹ Insofar as commissioners are willing to defer to the opinions of other commissioners on matters concerning home boroughs, or in areas of individual commissioner's expertise, there

130. *Id.*

131. *Public Meeting of the Landmark Preservation Commission* (Apr. 13, 2009), available at http://www.nyc.gov/html/lpc/downloads/pdf/calendar/04_13_10.pdf.

132. *Id.*

133. *Public Meeting of the Landmark Preservation Commission* (May 12, 2009), available at http://www.nyc.gov/html/lpc/downloads/pdf/calendar/05_12_09.pdf; see also LANDMARKS PRESERVATION COMMISSION, DETERMINATION OF THE APPLICATION FOR A CERTIFICATE OF APPROPRIATENESS OR NOTICE TO PROCEED TO DEMOLISH A DESIGNATED BUILDING PURSUANT TO SECTION 25-309 OF THE LANDMARKS LAW (May 12, 2009), available at http://www.gvshp.org/_gvshp/preservation/st_vincent/doc/notice-to-proceed05-12-09.pdf [hereinafter "St. Vincent's Hardship Approval"].

134. RICHARD A. POSNER, *HOW JUDGES THINK* 32 n.30 (2008), contains a brief review of literature discussing the impact of dissenting opinions on collegiality in the appellate courts.

135. *Id.*

136. *Id.*

137. *Id.* at 33.

138. See *id.*; see also CASS R. SUNSTEIN, *WHY SOCIETIES NEED DISSENT* 166–68 (2003) (discussing tendencies of three-judge panels based on political affiliation of the appointing president).

139. *Id.* at 55 ("In an informational cascade, people cease relying, at a certain point, on their private information or opinions. They decide instead on the basis of the signals conveyed by others.")

is already some evidence that cascades are occurring.¹⁴⁰ Even more to the point, the commissioners are not full-time government workers, so they are likely to be even more vulnerable to social pressure from other members of the commission.¹⁴¹ And because commissioners tend to be preservationists, there is a further risk of group polarization towards a more extreme position.¹⁴²

There is at least one plausible alternative explanation for the paucity of dissenting votes on the LPC agenda: it could be the case that there is frequently disagreement among the commissioners, but they negotiate among themselves to reach a consensus before calendaring the vote. This possibility, however, does not limit the problems seen above. If anything, it simply exacerbates them by making the process entirely opaque, which is contrary to the spirit of holding public meetings on the subject.

E. Landmark Status Should Not Necessarily Be Perpetual

It is sometimes clear that landmark status is no longer appropriate or efficient. The cultural importance of a given site may seem much stronger at some point in the past than it does in the present or future. The Landmarks Preservation Commission might make a mistake in assigning the initial designation, either by protecting a building that lacks aesthetic or historical merit, or by protecting exemplars of an historical trend that has been reevaluated. Alternatively, a landmark designation may impede the development of an area so significantly that the costs of maintaining the landmark designation significantly outweigh any cultural benefit. Additionally, the circumstances of the area surrounding the landmark may have changed since the designation was issued, either in terms of an areas development needs or cultural needs. Finally, the value of the landmark relative to other landmarks may change over time. In each of these circumstances, rescission of the landmark designation might be potentially appropriate, yet the current law merely provides that any order of the LPC can be rescinded without providing further guidance or addressing these possibilities.

140. Informal conversations with members of the LPC's legal staff have confirmed this, but unfortunately neither hard data nor public statements are readily available.

141. SUNSTEIN, *supra* note 138, at 79–80.

142. *Id.* at 112. No hard data is available for this proposition, both for questions of what and how to measure, but this is supported by informal conversations with involved parties and the tendency, discussed throughout this Note, for the LPC to approve landmarking decisions and oppose rescissions.

While owners of landmarked property can seek permission to modify the property,¹⁴³ removing the designation altogether is extremely difficult. There is a process for the rescission of a landmark designation,¹⁴⁴ but it is infrequently used.¹⁴⁵ The LPC last rescinded landmark status in 2006, and that case was somewhat unusual in that the site itself had been destroyed.¹⁴⁶ Some of this is probably explained by the composition of the LPC; many of the commissioners are likely to have an interest in preservation,¹⁴⁷ and there are reasons why owners might hesitate to seek rescission, including concerns about negative publicity. This does not suggest that there are no circumstances in which rescission would be desirable, nor that rescission should be as difficult as it currently stands. Rather, this implies that, even if there were a more robust mechanism for rescission, fears of an under-supply of landmarks are unfounded.

The next section will address the existing mechanisms within the Landmarks Law for modifications to landmarked property and show why they are inadequate to address the problem of landmark designations that have outlived their usefulness.

143. *Infra* Part II.B. Critics of the current Landmarks Law have also criticized the difficulties of making some of these modifications. See, e.g., Barron, *supra* note 3 (quoting critic and professor of architecture Paul Goldberger as stating that

“[preservation fundamentalism] is particularly egregious in New York, where the insistence that everything be exactly as it was, and that historic districts cannot continue to evolve, that they cannot contain modern buildings within them and that the only way to show proper care for an individual landmark building is to make sure its appearance never changes—these attitudes seem altogether inconsistent with our nature and identity as a city. Respecting and preserving landmarks does not have to mean treating them like hothouse orchids.”)

The consequences of ignoring these regulations can also be severe; in at least one case, the LPC ordered a homeowner to tear down the addition of an entire floor. See Karmin, *supra* note 101.

144. N.Y.C. ADMIN. CODE § 25-303(h) (1992).

145. The LPC does not aggregate its records, so it is virtually impossible to find hard data on the actual frequency of its use, but some analysis of the LPC's agendas indicates that it is rare. There are a few cases in which buildings in historic districts were granted permits for demolition, but these are slightly different types of cases in that they usually involve non-historic buildings in generally protected areas and are typically handled by “certificates of no effect.” See *infra* Part II.A.

146. N.Y.C. LANDMARK PRESERVATION COMMISSION, *supra* note 109.

147. *Supra* Part I.D.

II. CURRENT MECHANISMS FOR DE-DESIGNATION: REASONABLE RETURN AND THE HARDSHIP EXCEPTION

Though outright rescission is rare, the Landmarks Law contains other provisions that allow owners to modify their properties in certain ways. This section will briefly review the existing provisions that allow for modification of properties and will then address the reasonable return standard and both the statutory and judicial hardship exceptions, which allow for more drastic changes, including demolition of the landmarked property.

In addition to rescission, there are other potential avenues for owners of landmarked property to pursue. There are three potential certificates that the owner of the property can petition for in order to modify a landmarked property: a “certificate of no effect,”¹⁴⁸ a “certificate of appropriateness,”¹⁴⁹ and a “certificate of insufficient return.”¹⁵⁰

A. *Certificates of No Effect and the Certificate of Appropriateness*

A certificate of no effect allows a property owner the opportunity to demonstrate that a proposed modification to the landmarked property would not “change, destroy or affect” the protected elements,¹⁵¹ or in the case of a historic district, that the proposed modification would be “in harmony with the external appearance” of the district.¹⁵² This is by far the most common certificate issued by the Landmarks Preservation Commission,¹⁵³ although this likely does not demonstrate the ease with which one can obtain relief from landmark status, but instead reflects the expansive regulatory power of the LPC: even the most minor altera-

148. ADMIN. § 25-306.

149. *Id.* § 25-307.

150. *Id.* §25-309(a)(1). Technically, this is also a “certificate of appropriateness,” wherein the owner petitions for a certificate of appropriateness, asking for permission to demolish, alter, or renovate on the specific grounds of “insufficient return.” In this Note, it is referred to as a certificate of insufficient return or a claim under the reasonable return standard, for the sake of clarity.

151. *Id.* § 25-306(a)(1).

152. *Id.*

153. In 2009, the LPC issued 3466 such certificates and 848 expedited petitions, after receiving 8929 applications. 2009 MAYOR’S REPORT, *supra* note 7, at 107–08.

tions to landmarked exteriors or buildings in historic districts require approval.¹⁵⁴

The second certificate an owner might seek is a certificate of appropriateness.¹⁵⁵ Unlike a certificate of no effect, in which a property owner asserts that the protected elements will remain unchanged, a certificate of appropriateness¹⁵⁶ covers situations in which the owner concedes that some element will be modified but argues that those modifications do not undermine the purposes of the Landmarks Law.¹⁵⁷ This certificate is mostly used for additions or renovations that are set back from the street in some way or are otherwise inconspicuous. Unlike the standards for rescission proposed in this Note¹⁵⁸ this certificate is used for a narrow class of renovations in which the owner is able to preserve the historic values of the site, despite making some changes.

B. *Reasonable Return and the Statutory Hardship Exception*

Finally, a certificate of insufficient return allows commercial property owners to demonstrate that the landmark designation of their property prevents them from obtaining a “reasonable return” on their property because of the landmark status.¹⁵⁹ Reasonable return is defined as “six percent of the valuation of an improvement parcel.”¹⁶⁰ A similar provision, the statutory hardship exception, exists for nonprofit owners of landmarked property.¹⁶¹ The Landmarks Preservation Commission can modify the landmark designation in order to assist with a sale or long-term lease of the property if the nonprofit owner can demonstrate that the landmark status makes the site unusable for its current purpose and for which it was used when the nonprofit acquired the property, and demonstrate that, were the owner a commercial enterprise, it would be unable to obtain a reasonable return.¹⁶² Neither provision is frequently used by the LPC, nor is the hardship provision often sought.¹⁶³ Since 1967, there have been sixteen applications for

154. There are provisions for extremely minor work, but the general rule is that any work that would require a Department of Buildings Permit would also require LPC approval.

155. ADMIN. § 25-307.

156. *Id.*

157. *Id.* § 25-307(a).

158. *Infra* Part III.C.

159. ADMIN. § 25-309(a)(1).

160. *Id.* § 25-302(v)(1).

161. *Id.* § 25-309.

162. *Id.* § 25-309(a)(2).

163. LPC HARDSHIP DISPOSITIONS, *supra* note 111.

hardship dispositions.¹⁶⁴ Of these applications, fifteen asked for the right to demolish the structure in whole or in part.¹⁶⁵ Eight of the fifteen were granted outright, and three were denied.¹⁶⁶ In the remaining cases, either the request was withdrawn, or the LPC helped the applicant find a buyer.¹⁶⁷ In sum, landmark status is rarely rescinded either by a rescission decision made by the LPC or through one of the exceptions.

*C. The Regulatory Takings Doctrine and the
Judicial Hardship Exception*

The remaining ways in which an owner can remove landmark status are through a challenge under the regulatory takings doctrine or, in the case of nonprofit owners, the judicial hardship exception.

If the Landmarks Preservation Commission denies an owner's application for a certificate of insufficient return, the owner can bring a claim in New York State Supreme Court with an Article 78 petition.¹⁶⁸ The burden of proving a lack of reasonable return falls to the owner of the property,¹⁶⁹ and judicial scrutiny of these decisions is based on case-by-case analysis¹⁷⁰ under a deferential standard of review.¹⁷¹ Owners who lose the Article 78 proceeding can still challenge the designation of the property as a taking or argue that the decision to landmark was arbitrary and capricious and therefore violated due process.¹⁷² Nonprofit owners can claim that

164. *Id.*

165. *Id.*

166. *Id.*

167. *Id.*

168. N.Y. C.P.L.R. § 7801 (McKinney 2010). Article 78 is the primary mechanism for challenging administrative decisions in New York State. Generally speaking, it encompasses the common law writs of mandamus, prohibition, and certiorari. The LPC can also apply the "judicial" hardship exception, despite its lack of statutory authorization.

169. *See* 400 East 64/65th Street Block Ass'n v. City of New York, 583 N.Y.S.2d 452, 454-55 (App. Div. 1992).

170. *Soho Alliance v. N.Y.C. Bd. of Standards & Appeals*, 703 N.Y.S.2d 150, 158 (App. Div. 2000).

171. *Shubert Org., Inc. v. Landmarks Pres. Comm'n of City of New York*, 570 N.Y.S.2d 504, 507 (App. Div. 1991) (applying "arbitrary and capricious" standard of review).

172. *See, e.g., Rudey v. Landmarks Pres. Comm'n of New York*, 627 N.E.2d 508 (N.Y. 1993). As a procedural matter, an owner must seek a certificate of appropriateness before seeking either no-return or hardship relief, and must exhaust these options before challenging them in court. *See* N.Y.C. ADMIN. CODE §§ 25-307, 25-309 (1992); C.P.L.R. § 7801.

they qualify for the judicial hardship exception, created by New York State Appellate Division in *Trustees of Sailor's Snug Harbor in the City of New York v. Platt*,¹⁷³ while commercial owners typically proceed under the regulatory takings standard.¹⁷⁴ This section discusses each standard and then argues that neither is appropriate nor adequate for the problem of albatross landmarks.

1. Regulatory Takings

Governments are permitted to make decisions for aesthetic reasons.¹⁷⁵ In *Penn Central Transportation Co. v. New York City*,¹⁷⁶ the United States Supreme Court held that a government designation of a site as a landmark does not constitute a taking.¹⁷⁷ Penn Central, the same railroad company that had destroyed Penn Station, also owned Grand Central Terminal.¹⁷⁸ The station, built in 1913, was and is a prized example of Beaux-Arts architecture.¹⁷⁹ The Terminal had been declared a landmark in 1967.¹⁸⁰ Penn Central wanted to build a fifty-three-story office building atop the Terminal and applied for permission to do so from the Landmark Preservation Commission, submitting two separate plans.¹⁸¹ The LPC denied the applications, and Penn Central filed suit in New York State court, alleging violations of the Fifth and Fourteenth Amendments to the Constitution.¹⁸² Specifically, it asserted that the application of the Landmarks Law to Grand Central Station was a taking of its property without just compensation¹⁸³ and that the landmark status

173. 288 N.Y.S.2d 314 (App. Div. 1968). While this obviously pre-dates the decision in *Penn Central*, any discussion of the judicial response to the concerns in *Penn Central* must begin with *Snug Harbor*.

174. While the two tests have been discussed separately in this Note, one open question is the extent to which the regulatory takings standard and the hardship variance standard are co-extensive. In other words, if the judicial hardship exception is read as stating that the denial of the hardship variance would cause taking, then is the nonprofit entitled to the variance? However, whether the *Snug Harbor* test governs or the regulatory takings doctrine controls, the criticisms made in this Note as to the inappropriateness of these tests for certain decisions still hold.

175. *Berman v. Parker*, 348 U.S. 26, 32–33 (1959).

176. 438 U.S. 104 (1978).

177. *Id.* at 138.

178. *Id.* at 115.

179. *Id.*

180. *Id.* at 115–16.

181. *Id.* at 116–17.

182. *Id.* at 117, 119.

183. *Id.*

“arbitrarily deprived them of their property without due process of law.”¹⁸⁴

The New York State Supreme Court granted an injunction to Penn Central, but the Appellate Division reversed,¹⁸⁵ and the New York Court of Appeals affirmed the Appellate Division.¹⁸⁶ The New York Court of Appeals cursorily rejected the Fifth Amendment takings argument because the City had merely restricted certain uses of the property, rather than transferring ownership.¹⁸⁷ After more extensive discussion, the court also rejected Penn Central’s claim that its substantive due process rights had been violated, because the landmark designation, *inter alia*, still allowed a “reasonable return” on its investment interest in the property.¹⁸⁸

The United States Supreme Court affirmed the judgment of the New York Court.¹⁸⁹ Rather than address the substantive due process argument, the Court certified the question for appeal as whether there had been a taking under the Fifth Amendment and, if so, whether the transferable development rights that had been granted to Penn Central were sufficient as “just compensation.”¹⁹⁰ Because it found no taking, the Court never addressed the second question.¹⁹¹ Although the Court explicitly rejected the idea that a taking required the transfer of physical control of the property,¹⁹² it ruled that the Landmarks Law “[had] not effected a ‘taking’ of [Penn Central’s] property.”¹⁹³

The Court’s ruling made a number of points that are important for the purposes of contemporary landmark determinations. It rejected the argument that landmark decisions are fundamentally arbitrary because they rely on the subjective judgment of institutions like the LPC.¹⁹⁴ While the Court did not adopt a rule requir-

184. *Id.*

185. Penn Cent. Transp. Co. v. City of New York, 377 N.Y.S.2d 20, 30 (App. Div. 1965).

186. Penn Cent. Transp. Co. v. City of New York, 366 N.E.2d 1271, 1279 (N.Y. 1977).

187. Penn Cent. Transp. Co. v. New York City, 438 U.S. at 120–21.

188. *Id.* at 121.

189. *Id.* at 138.

190. *Id.* at 122. The New York Court of Appeals had granted leave to Penn Central to present further evidence that they could not make a reasonable return on their property, and thus could still have a viable Fourteenth Amendment claim, but as the company decided not to pursue that in lieu of appeal to the Supreme Court, the Court did not address that claim. *Id.*

191. *See id.*

192. *Id.* at 123 n.25.

193. *Id.* at 1–38.

194. *Id.* at 132.

ing a physical transfer of ownership or intrusion for a finding of a regulatory taking,¹⁹⁵ neither did it apply a rule requiring compensation whenever there is a diminution of value caused by government action.¹⁹⁶ Finally, the Court re-affirmed the test for regulatory takings¹⁹⁷ established in *Pennsylvania Coal Co. v. Mahon*¹⁹⁸ and explicitly applied it to the landmark context.¹⁹⁹

Justice Rehnquist, joined by Chief Justice Burger and Justice Stevens, dissented.²⁰⁰ Justice Rehnquist explicitly justified his dissent with the language of public goods discourse²⁰¹ and identified two important aspects of landmarking that distinguish it from other zoning actions. Unlike zoning, which might limit land uses for all owners in an area, a landmark designation only restricts the owners of a particular site.²⁰² In addition, the landmark designation creates “an affirmative duty to *preserve* his property *as a landmark* at his own expense.”²⁰³ According to the dissent, the individual and affirmative nature of the burdens created by a landmark designation were sufficient to create a claim under the Fifth Amendment.²⁰⁴

Despite these concerns, the majority’s ruling in *Penn Central* remains intact.²⁰⁵ Yet the recognition by the Court that the concerns of the dissenters should inform the use of an ad hoc test has had consequences for litigation surrounding landmarks.

195. *Id.* at 123 n.25. That actions other than a transfer of title can amount to a takings has been established since at least the nineteenth century. Gary Lawson, Katharine Ferguson & Guillermo A. Montero, “*Oh Lord, Please Don’t Let Me Be Misunderstood!*”: *Rediscovering the Mathews v. Eldridge and Penn Central Frameworks*, 81 NOTRE DAME L. REV. 1, 25 (2005).

196. *Penn Cent. Transp. Co. v. New York City*, 438 U.S. at 131.

197. For a brief yet helpful summary of the evolution of the regulatory takings doctrine up to *Penn Central*, see Lawson, Ferguson & Montero, *supra* note 195, at 24–30.

198. 260 U.S. 393 (1922).

199. *Penn Cent. Transp. Co. v. New York City*, 438 U.S. at 136 (“We now must consider whether the interference with the appellants’ property is of such a magnitude that ‘there must be an exercise of eminent domain and compensation to sustain [it].’”) (quoting *Pennsylvania Coal Co. v. Mahon*, 260 U.S. at 413).

200. *Id.* at 138.

201. *Id.* at 139 (Rehnquist, J., dissenting) (“The question in this case is whether the cost associated with the city of New York’s desire to preserve a limited number of ‘landmarks’ within its borders must be borne by all of its taxpayers or whether it can instead be imposed entirely on the owners of the individuals properties.”).

202. *Id.* at 139–40.

203. *Id.* at 140 (emphasis in original).

204. *Id.*

205. See, e.g., *Palazzolo v. Rhode Island*, 533 U.S. 606, 616–17 (2001).

2. The Judicial Hardship Exception

In *Trustees of Sailor's Snug Harbor in the City of New York v. Platt*,²⁰⁶ a sailors' home on Staten Island challenged the designation of four of their dormitories as landmarks.²⁰⁷ While the Landmarks Law includes an exception for nonprofits that is somewhat similar to the reasonable return status,²⁰⁸ it only applies when the nonprofit intends to sell or lease the property in question and the sale is to a commercial enterprise.²⁰⁹ The plaintiffs in *Snug Harbor* only wanted to modify their property.²¹⁰ In *Snug Harbor*, the Appellate Division reversed the decision for the plaintiffs in the lower court and remanded for reconsideration in light of the rule it announced.²¹¹ The *Snug Harbor* test states that nonprofit owners can successfully challenge the refusal to grant a variance as a taking when "maintenance of the landmark either physically or financially prevents or seriously interferes with carrying out the charitable purpose."²¹²

Despite the announcement of a major new constitutional rule in *Penn Central*, which, as a takings clause case addressing New York's Landmarks Law, might have superseded the New York State cases, the basic framework for hardship variance applications for nonprofits in New York retains the *Snug Harbor* test. In *Lutheran Church in America v. City of New York*,²¹³ the New York Court of Appeals explicitly applied the *Snug Harbor* test²¹⁴ in order to declare a

206. 288 N.Y.S.2d 314 (App. Div. 1968).

207. *Id.* at 315.

208. *Supra* at Part II.B.

209. N.Y.C. ADMIN. CODE § 25-309(a)(2) (1992).

210. *See Snug Harbor*, 288 N.Y.S.2d 314, 316 (App. Div. 1968).

211. *Id.* at 317.

212. *Id.* at 316. It is worth noting, in light of later arguments in this section, that *Snug Harbor* was essentially a case on physical hardship, not financial hardship. *See Stewart, supra* note 50, at 180.

213. 316 N.E.2d 305 (N.Y. 1974). Cases involving religious institutions, which are frequently organized as nonprofit organizations, also introduce a First Amendment dimension to the jurisprudence. *See, e.g.*, *Church of St. Paul & St. Andrew v. Barwick*, 496 N.E.2d 183, 185–86 (N.Y. 1986) (dismissed on ripeness grounds). While this problem is beyond the scope of this Note, a general treatment of the issue can be found in Catherine Maxson, Note, "Their Preservation is Our Sacred Trust"—Judicially Mandated Free Exercise Exemptions to Historic Preservation Ordinances Under *Employment Division v. Smith*, 45 B.C. L. REV. 205 (2003). *See also* Steven P. Eakman, Note, *Fire and Brownstone: Historic Preservation of Religious Properties and the First Amendment*, 33 B.C. L. REV. 93 (1991) (written before Supreme Court's decision in *City of Boerne v. Flores*, 521 U.S. 507 (2000)).

214. *Lutheran Church*, 316 N.E.2d at 311 (citing *Snug Harbor*, 288 N.Y.S.2d 314).

landmark designation “confiscatory.”²¹⁵ In *Society for Ethical Culture v. Spatt*,²¹⁶ the first hardship exception case decided after *Penn Central*, the Appellate Division applied *Snug Harbor*²¹⁷ to uphold the designation of a building with an estimated sale price of four million dollars.²¹⁸ In that case, the court was careful to articulate that the hardship exception is evaluated with reference to the present uses of the property, not the potential financial gains that could be made by other uses of the site.²¹⁹

Later cases have re-affirmed the use of the *Snug Harbor* test and given a great deal of discretion to the Landmark Preservation Commission in applying the test. In *1025 Fifth Avenue, Inc. v. Marymount School of New York*,²²⁰ neighbors of a school challenged the issuance of a certificate of appropriateness to renovate a rooftop gymnasium located within the Metropolitan Museum Historic District.²²¹ Despite a finding that the modification was inappropriate,²²² the LPC had granted the certificate because Marymount had demonstrated insufficient returns on the property, given their charitable mission.²²³ The court, in upholding that decision, held that the LPC was correct to apply the *Snug Harbor* test, and thus the court could only reverse if the decision of the LPC was arbitrary or capricious.²²⁴

In the wake of *Board of Estimate v. Morris*,²²⁵ the City Charter Revision Commission considered and rejected several changes to the way in which the Landmarks Law deals with nonprofit organizations.²²⁶ While several changes, such as the creation of an appeals

215. *Id.* at 310–12. This was also a physical, not financial, hardship case. Stewart, *supra* note 50, at 181.

216. 416 N.Y.S.2d 246 (App. Div. 1979).

217. *Id.* at 251.

218. *Id.* at 248. This case was somewhat unusual, as the nonprofit owner challenged the designation under the hardship exception immediately after the designation. The court ruled that the challenge was not premature but still ruled against the owners. *Id.* at 253.

219. *Id.* at 252; see Stewart, *supra* note 50, at 184.

220. 475 N.Y.S.2d 182 (Sup. Ct. 1983).

221. *Id.* at 183–84. An additional issue in this case was that the Marymount building itself was not landmarked but fell within a historic district. The petitioners claimed that this meant that the issuance of the certificate was therefore ultra vires, but the court found that under *Snug Harbor*, it was entitled to create common law rules for “situations not foreseen in the statute.” *Id.* at 185.

222. *Id.*

223. *Id.* at 184.

224. *Id.* at 186–87.

225. 489 U.S. 688 (1989) (declaring structure of Board of Estimate unconstitutional).

226. Gutman, *supra* note 22, at 440–45.

board, would have made it easier for such groups to challenge LPC decisions,²²⁷ other changes would simply have transferred authority that had previously been held by the Board of Estimate²²⁸ to the city council.²²⁹

D. The Regulatory Takings Doctrine and the Judicial Hardship Exception Do Not Address the Problem of Albatross Landmarks

While the *Penn Central* decision and the ad hoc test it created have been the subject of significant criticism,²³⁰ they have remained the dominant framework for regulatory takings analysis.²³¹ The *Penn Central* decision has also largely protected historic preservation statutes from constitutional attack on Fifth Amendment grounds.²³² Despite this, the takings doctrine and the related judicial hardship exception are at present the most viable means for landmark rescission. The problem is that these doctrines were not designed to deal with albatross landmarks, and when applied to them, the doctrines create distortions and inefficiencies.

Both the exceptions within the Landmarks Law and the judicially created hardship exception create perverse incentives for owners of landmarked property.²³³ Landmark status can give own-

227. *Id.* at 445.

228. The New York City Board of Estimate was a governing body composed of the mayor, comptroller, city council president, and the five borough presidents. The mayor, comptroller, and city council president each got two votes, while the borough presidents each had one vote. The structure was declared unconstitutional under the doctrine of “One Person, One Vote” in *Board of Estimate v. Morris*, 489 U.S. 688 (1989), and most of the Board’s authority was given to the city council by the new city charter.

229. Gutman, *supra* note 22, at 446.

230. See, e.g., Andrew S. Gold, *The Diminishing Equivalence Between Regulatory Takings and Physical Takings*, 107 DICK. L. REV. 571, 571 (2003); Stewart E. Sterk, *The Federalist Dimension of Regulatory Takings Jurisprudence*, 114 YALE L.J. 203, 205 (2004). See generally Gideon Kanner, *Making Laws and Sausages: A Quarter Century Retrospective on Penn Central Transportation Co. v. City of New York*, 13 WM. & MARY BILL RTS. J. 679 (2005).

231. A long line of cases have reinterpreted or applied the *Penn Central* test. See, e.g., *Palazzolo v. Rhode Island*, 533 U.S. 606 (2001); *Lucas v. South Carolina Coastal Council*, 505 U.S. 1003, 1015–19 (1992); *Loretto v. Teleprompter Manhattan CATV Corp.*, 458 U.S. 419, 434 (1982); *Kaiser Aetna v. United States*, 444 U.S. 164, 175 (1979). The basic framework from *Penn Central* is still the standard in regulatory takings cases.

232. Byrne, *supra* note 122, at 316. In certain cases, constitutional attacks on designations have been made under the First Amendment. *Supra* note 213.

233. In general, landmark regimes may create perverse incentives for owners to build unremarkable buildings, although this may be mitigated through other incentives built into the landmark system. Rose, *supra* note 24, at 500–01.

ers a reason to take subpar care of their buildings.²³⁴ While fines and other penalties can help prevent this, it is sometimes more costly to repair the landmark than to pay the fine. The ultimate recourse, condemning and seizing the property, is difficult and does not necessarily solve the problem.²³⁵ Because rescission of a landmark is difficult, the easiest way to get a landmark designation removed entirely or modified may be to demolish the property through neglect²³⁶ or, short of that, to demonstrate that the designation impedes one's ability to see a profit on the property.

This creates the possibility of strategic behavior on the part of owners of landmarks. Commercial owners may be incentivized to less than fully exploit the landmark site, in order to demonstrate that it is impossible to receive a reasonable return on the site. While there is some case law stating that this is not an acceptable way to obtain the exception,²³⁷ such behavior may be difficult to detect. Similarly, nonprofit owners may be disincentivized from improving the services which they provide in order to demonstrate that the landmark status interferes with their charitable purposes.

Beyond these perverse incentives, the doctrines have a number of problems. First, the tests are relatively vague.²³⁸ Although there is a nominally objective standard defining "reasonable return" as "a net annual return of six percent of the valuation of an improvement parcel,"²³⁹ it is not clear what values should be used for calcu-

234. Shawn G. Kennedy, *Landmarking's Double-edged Sword: Sometimes It Works Against Preservation*, N.Y. TIMES, Jan. 13, 1991, § 10 (Real Estate), at 1 (cited in ELLICKSON & BEEN, *supra* note 65, at 500 n.5).

235. See, e.g., Lisa Selin Davis, *Neglect, Lack of Funds Damage NYC Landmark*, WALL ST. J. REAL ESTATE JOURNAL (2004), available at <http://lisaselindavis.com/articles/Sea%20View%20Hospital.pdf> (describing how city agencies could not be forced to repair property through standard landmarks enforcement regime). While the LPC can fine owners who fail to maintain landmarked property, enforcement is sometimes lax both because of issues of funding and staff shortages and because of administrative incompetence. See HON. BILL PERKINS, COMMITTEE ON GOVERNMENTAL OPERATIONS, FORTY YEARS OF NEW YORK CITY LANDMARKS PRESERVATION: AN OPPORTUNITY TO REVIEW, REFLECT AND REFORM 4, 8 (2005) [hereinafter "COUNCIL REPORT"].

236. See, e.g., Lana Bortolot, *amNY's Endangered NYC special report: The city's crumbling past*, AMNY, Jan. 3, 2010, <http://amny.com/urbanite-1.812039/amny-s-endangered-nyc-special-report-the-city-s-crumbling-past-1.1682392>.

237. *Seawall Assocs. v. City of New York*, 534 N.Y.S.2d 958, 963–64 (App. Div. 1988).

238. Moy, *supra* note 3, at 458; Stewart, *supra* note 50, at 163 ("New York City's nonprofit landowners allege, however, that the current historic preservation system employs vague, arbitrary and overly strict eligibility criteria for hardship relief . . .").

239. N.Y.C. ADMIN. CODE § 25-302(v) (1992).

lating the six percent figure. Even with a defined standard of “six percent,” it is extremely difficult to properly account for the failure of the owner to realize a reasonable return. For example, the owner’s losses may be a result of mismanagement, rather than the landmark designation. Courts have stated that a reasonable return is not the present return compared to the most economically efficient use of the property,²⁴⁰ but, especially when property has been landmarked after being acquired by the present owners, it is not clear why other potential uses of the property should not be considered.²⁴¹

Furthermore, the Landmarks Preservation Commission may well be best-suited to decide the historical or aesthetic value of a particular site,²⁴² but may not be equally adept at evaluating the financial implications of maintaining a landmark designation. The problem essentially repeats itself in the context of nonprofit owners; the LPC is designed to assess architectural and historic merit, not evaluate the books of myriad nonprofit enterprises.²⁴³ And if the conventional rationale for maintaining a deferential standard of review is administrative competence,²⁴⁴ then that rationale is insufficient to maintain the current deference in landmark decisions.²⁴⁵

Another potential issue is that the constitutional standards governing the various exceptions are traceable to *Penn Central*, which dealt with an obviously important, especially valuable, highly visible, and centralized structure.²⁴⁶ While early commentators on the *Penn*

240. 400 East 64/65th Street Block Ass’n v. City of New York, 583 N.Y.S.2d 452, 454–55 (App. Div. 1992).

241. The *Penn Central* test includes analysis of the owner’s “investment-backed expectations.” *Penn Cent. Transp. Co. v. New York City*, 438 U.S. 104, 124 (1978). While this may have relevance to the constitutional issue of takings, insofar as a plaintiff cannot claim the government took something they never thought they had, as this Note argues, the question of future uses should be relevant to the question of maintaining a landmark designation. *Infra* Part III.C.

242. Stewart, *supra* note 50, at 172.

243. *Id.* at 171 (“The LPC’s essential task in [the hardship relief] context is to distinguish between landmark owners who truly need hardship relief and those who seek relief merely for easy gain.”).

244. See, e.g., *Chevron U.S.A., Inc. v. Nat’l Res. Def. Council, Inc.*, 467 U.S. 837, 865 (1984).

245. This is not to suggest that courts should, without authorization from the legislature, apply a higher level of scrutiny to landmarking decisions but rather that the New York State legislature would do well to revisit this particular scheme.

246. Richard A. Jaffe & Stephen Sherrill, Comment, *Grand Central Terminal and the New York Court of Appeals: “Pure” Due Process, Reasonable Return, and Betterment Recovery*, 78 COLUM. L. REV. 134, 152–53 (1978) (published in January 1978, before the Supreme Court granted a writ of certiorari in *Penn Central*).

Central decision argued that the decision might have little practical import²⁴⁷ or that restrictions on less deserving sites would be harder to constitutionally maintain,²⁴⁸ the decision has proven to be versatile and enduring. But many landmarks are nowhere near as important or obvious as Grand Central. It is possible that the decision, made after the very same company that had destroyed Penn Station had attempted to destroy Grand Central, should not have the same persuasive force when applied to other less significant sites. The ad hoc test of *Penn Central* allows courts to balance the importance of the government's interest in the designation against the magnitude of the imposition on private property owners, but the Supreme Court has interpreted that as a balance between the regulatory scheme in general (i.e. the government's interest in having any landmark scheme at all) and the imposition on an individual owner, rather than a balance between the government's interest in a particular landmark and the imposition on the owner.²⁴⁹ In *Penn Central*, the Court explicitly stated that it approved of the government interest in maintaining landmarks.²⁵⁰ By approving the use of landmark laws in order to balance the public interest in historic preservation in general against the burden on particular private property owners, the Court did not ask if there might be individual cases where the general scheme of landmarking would be too burdensome given the actual structure being preserved.

247. *Id.* at 152.

248. *Id.* at 153.

249. See *Lingle v. Chevron U.S.A., Inc.*, 544 U.S. 528, 542–546 (2005); Christopher T. Goodin, *The Role and Content of the Character of the Governmental Action Factor in a Partial Regulatory Takings Analysis*, 29 U. HAW. L. REV. 437, 442 (2007); Michael B. Kent, Jr., *Construing The Canon: An Exegesis of Regulatory Takings Jurisprudence After Lingle v. Chevron*, 16 N.Y.U. ENVTL. L.J. 63, 93 (2008); see also *Yee v. City of Escondido*, 503 U.S. 519, 522–23 (1992) (interpreting *Penn Central* as stating that a regulatory taking occurs “only if considerations such as the purpose of the regulation or the extent to which it deprives the owner of the economic use of the property suggest that the regulation has unfairly singled out the property owner to bear a burden that should be borne by the public as a whole”). The Court never reached the regulatory takings argument in *Yee*, stating that it had not been properly certified for appeal. *Id.* at 537–38.

250. *Penn Cent. Transp. Co. v. New York City*, 438 U.S. 104, 129 (1978) (“[A]ppellants do not contest that New York City’s objective of preserving structures and areas with special historic, architectural, or cultural significance is an entirely permissible governmental goal.”); see also *id.* at 134–35 (“Unless we are to reject the judgment of the New York City Council that the preservation of landmarks benefits all New York Citizens and all structures, both economically and by improving the quality of life in the city as a whole—which we are unwilling to do—we cannot conclude that the owners of the Terminal have in no sense been benefited by the Landmarks Law.”).

Finally, the Supreme Court has frequently referred to “investment-backed expectations” in the context of regulatory takings cases.²⁵¹ This makes perfect sense: an owner cannot reasonably claim that the government has taken property rights that no reasonable person would believe he or she possessed. But in the context of albatross landmarks, present and future uses that were not foreseen when the owner acquired the property might be persuasive reasons to eliminate the landmark designation. Requiring owners to show interference with their investment-backed expectations fails to address the possibility that a landmark designation that made sense at one point no longer provides the city with significant benefits. The next section of this Note proposes circumstances which would justify rescission of landmark designations.

III. PROPOSALS FOR RESCISSION

To recount, the problem can be stated as this: without the Landmarks Preservation Commission, there would likely be an undersupply of landmarks. But the LPC itself is vulnerable to capture and is likely to over-preserve. Even were the supply of landmarks at an optimal level,²⁵² there is still the possibility of mistaken landmark designations and the possibility that changed circumstances obviate the utility of a given landmark designation. At present, landmark status is virtually perpetual, both because of the internal dynamics of the LPC and because of the external constraints on its behavior. When this is combined with the courts’ likely increased scrutiny²⁵³ of decisions to rescind, there are powerful disincentives for the LPC to ever rescind landmark status.

The solution to the problem of albatross landmarks is to modify the statutory scheme for rescission. The scheme should explicitly delineate circumstances in which landmark status can be rescinded and establish the evidentiary burden that an owner must meet. In order to avoid the possibility of developers attempting to manipulate this system, it should also address the possibility of re-litigation.

251. See, e.g., *Palazzolo v. Rhode Island*, 533 U.S. 606, 617–18 (2001); *Lucas v. South Carolina Coastal Council*, 505 U.S. 1003, 1019 n.8 (1992) (quoting *Penn Cent. Transp. Co.*, 438 U.S. at 124); *Kaiser Aetna v. United States*, 444 U.S. 164, 175 (1979).

252. The question of the optimal level of supply of landmarks is an empirical and ideological question, which will exist regardless of the possibility of rescission. The point here is that the current system is likely to over-preserve and has no mechanisms in place to correct for this tendency.

253. *Infra* Part III.A.

A. The Current Rescission Provision's Inadequacy

One of the major problems with the rescission provision is that its own terms seem to undercut any significant possibility of rescission. The Landmarks Preservation Commission is charged with “safeguard[ing] the city’s historic, aesthetic and cultural heritage, as embodied and reflected in such improvements, landscape features and districts.”²⁵⁴ The very fact of an earlier designation suggests that there is significant cultural or historic value in the landmark. If it were clear that a site has simply lost historic value, as happened when the New Brighton Village buildings were destroyed,²⁵⁵ then perhaps rescission would be easy even in the absence of a clear standard or a set of factors to consider. Unfortunately, when rescission becomes appropriate, it is rarely as obvious.

Landmark disputes are generally controversial, and there is little reason to believe that rescission decisions would not be similarly hard-fought. This presents a disincentive for the LPC to rescind of its own accord. If the LPC reverses its own decision to protect a building, it risks increased scrutiny of that change.²⁵⁶ Were the LPC to rescind a designation, it would likely face a lawsuit from a preservationist group. Since there are no provisions delineating the circumstances in which rescission can be used, these decisions are, contrary to most decisions by the LPC,²⁵⁷ highly vulnerable to chal-

254. N.Y.C. ADMIN. CODE § 25-301(b) (1992).

255. *See supra* note 109.

256. *See, e.g.*, *Idlewild 94-100 Clark, LLC v. City of New York*, 898 N.Y.S.2d 808, 817 (Sup. Ct. 2010) (quoting *Matter of 2084-2086 BPE Assocs. v. State of New York Div. of Housing and Community Renewal*, 15 A.D.3d 288 (N.Y. App. Div. 2005)).

257. The standard for judicial review of determinations by the LPC, like other agencies, is whether the administrative determination is “warranted on the record and has a rational basis in the law.” *Mattone v. N.Y.C. Landmarks Pres. Comm’n*, No. 117604.03, 2004 WL 2567127 (N.Y. Sup. Ct. Sept. 24, 2004); *see also* *Teachers Ins. & Annuity Ass’n of America v. City of New York*, 623 N.E.2d 526, 528 (N.Y. 1993) (“A landmark designation is an administrative determination, ordinarily reviewable under article 78, that must be upheld if it has support in the record, a reasonable basis in law, and is not arbitrary or capricious.”) (citing *Lutheran Church v. City of New York*, 316 N.E.2d 305, 309 n.2 (N.Y. 1974)); *Shubert Org., Inc. v. Landmarks Pres. Comm’n*, 570 N.Y.S.2d 504, 507 (App. Div. 1991) (also applying “reasonable basis” and arbitrary and capricious standard of review). This is an extremely low requirement for the LPC to meet. One of the few examples of the LPC losing under this standard was a case in which the Commission set two different timetables for compliance with identical enforcement orders given to two residents of the same building, *Rudey v. Landmarks Pres. Comm’n*, 627 N.E.2d 508 (N.Y. 1993). Another court rejected an LPC determination, used by the City Planning Commission, that an increase of 4900 pedestrians “in a vicinity populated

lenge in an Article 78 proceeding. While judicial review of LPC decisions is rare,²⁵⁸ insofar as it is possible to extrapolate from a relatively small data set, courts seem generally unwilling to find the LPC's actions arbitrary or capricious.²⁵⁹ Moreover, the opacity of many LPC decisions makes it extremely difficult to demonstrate that they did not use improper factors.²⁶⁰ Owners seeking rescission of a landmark designation face significant additional obstacles.²⁶¹

The most likely result of this is that the Landmarks Preservation Commission will simply maintain landmarks, which appears to be what is happening under the status quo.²⁶² The LPC also has a variety of procedural mechanisms which can allow it to dispatch certain petitions without substantive rulings,²⁶³ and the LPC is not subject to particularly stringent procedural requirements in the first

daily by some 700,000 workers" would be a significant impact. 383 Madison Assocs. v. City of New York, 598 N.Y.S.2d 180, 182 (App. Div. 1993).

258. A search on Westlaw for "Landmarks Preservation Commission" for cases from either New York courts or federal courts sitting in New York reveals only 132 decisions, many of which are from different procedural phases of the same cases, and some of which pre-date the *Penn Central* decision. The absence of cases could also indicate that the threat of litigation constrains the behavior of the LPC. See Byrne, *supra* note 122, at 330.

259. See, e.g., Rudey v. Landmarks Pres. Comm'n, 627 N.E.2d 508 (N.Y. 1993).

260. See, e.g., Mattone v. N.Y.C. Landmarks Pres. Comm'n, No. 117604.03, 2004 WL 2567127 (N.Y. Sup. Ct. Sept. 24, 2004) (upholding LPC determination where petitioner failed to show that LPC had relied on one plan improperly).

261. The cases cited in this section contain several different kinds of obstacles, including challenges from third parties, which create their own dilemmas. For example, owners can challenge designations of their own property or the failure to grant a certificate, while third parties can challenge both designations and the failure to designate or the granting of the various certificates. One issue that sometimes occurs in cases is timing, as it is difficult to obtain a preliminary injunction compelling the LPC to either hold a hearing or stay a determination, see, e.g., Deane v. City of New York Dep't of Bldgs., 677 N.Y.S.2d 416 (Sup. Ct. 1998), but delaying litigation can allow a property owner to quickly make approved changes, which would lead to dismissal of a case against the LPC's decision as moot. See, e.g., Citineighbors Coal. of Historic Carnegie Hill *ex rel* Kazickas v. N.Y.C. Landmarks Pres. Comm'n, 811 N.E.2d 2, 4 (N.Y. 2004). Standing and timing are much bigger issues for third parties seeking to overturn decisions allowing demolition or modification of landmarks than for owners of property seeking relief.

262. Whether this is a direct result of the difficulty of obtaining rescission and the unavailability of judicial review or an indirect result of owners not seeking rescission because of these two factors is an empirical question.

263. See, e.g., Landmark West! v. Tierney, 807 N.Y.S.2d 342, 343 (App. Div. 2006) (holding that collateral estoppel prevented petitioner from raising challenge to LPC's procedures on appeal). The LPC had simply refused to calendar a hearing for the petition, which the petitioner had challenged below. The original decision is Landmark West! v. Tierney, No. 107387/05, 2005 WL 2108005 (N.Y. Sup. Ct. Sept. 1, 2005).

instance. For example, there is no requirement of any particular procedure to determine whether to conduct a hearing on a proposal brought to its attention.²⁶⁴

Many of the aforementioned problems are not unique to the LPC. Courts typically only examine agency decisions under arbitrary and capricious review.²⁶⁵ But the confluence of several factors makes this inappropriate for landmark decisions. The LPC is vulnerable to capture.²⁶⁶ When the agency decides to designate a landmark, the costs associated with that decision are relatively low.²⁶⁷ The only viable legal option for owners besides judicial review of the decision is a challenge under the regulatory takings doctrine, which is an extremely difficult standard to win under.²⁶⁸ Because of these factors, the owner of a landmark is more or less stuck with the designation. The problem is compounded by the lack of statutory language delineating when a designation should be rescinded. It is almost impossible to think of an example where the LPC would not meet the rationality standard when a prior group of commissioners had approved the landmark. If it is true that landmark status sometimes becomes inappropriate over time, or was simply sometimes mistaken when made, then there ought to be a way to re-visit the decision. As it is unlikely that the rules of judicial review for administrative decisions will change, the most practical way to alter the review of landmark designations is to add a statutory standard for rescission. This would improve the procedural route to rescission and create fixed parameters for the LPC's decision, providing courts with a standard against which to judge LPC determinations.²⁶⁹

B. Using "Change-Mistake" to Create a Framework for Rescission

At least one jurisdiction, Maryland, has created a doctrine to address areas of land use policy that have some common features with the problem of albatross landmarks. Under Maryland's

264. Robin Pogrebin, *New York Landmarks Agency Is Upheld on Appeal*, N.Y. TIMES, Mar. 4, 2010, http://www.nytimes.com/2010/03/05/arts/design/05arts-NEWYORKLANDM_BRF.html.

265. *Rudey v. Landmarks Pres. Comm'n of City of New York*, 627 N.E.2d 508 (N.Y. 1993).

266. *See supra* Part I.D.

267. *See supra* Part I.D.

268. *See supra* Part II.C.1.

269. Decisions under this standard would still be subject to a great deal of deference, but adding statutory factors would create additional questions of law, to which less deference is owed to the agency's interpretation. *See, e.g.,* *Belance v. Manhattan Beer Distribs.*, 861 N.Y.S.2d 797, 798 (App. Div. 2008).

“change–mistake” doctrine,²⁷⁰ in the context of “spot” zoning—where individual parcels within a comprehensive zoning plan are re-zoned—courts require the zoning board (or whoever is responsible for the decision) to justify the change under a higher standard than would ordinarily attach to an administrative decision. The rule is stated as follows: “[w]here property is rezoned, it must appear that either there was some mistake in the original zoning, or that the character of the neighborhood has changed to such an extent that such action ought to be taken.”²⁷¹ The rule is meant to create a presumption in favor of the initial zoning determinations,²⁷² because, unlike the comprehensive zoning plan, in a spot zoning decision, courts “probably suspect that regulations that favor (or disfavor) a particular tract serve private interests, not public purpose, and are therefore arbitrary and illegal.”²⁷³

This should be the basic framework for rescission in New York City. Landmark decisions, especially exterior designations, are effectively spot zoning decisions. In historic districts, the decision to rescind a landmark is also comparable. Courts should treat the initial landmark designation as presumptive evidence of historical or aesthetic merit. When an owner wants the designation rescinded,²⁷⁴ he or she should be allowed to overcome the presumption by presenting evidence to the Landmarks Preservation Commission based on factors suggested in the next section. The owner would have to demonstrate that some factor either not considered or not in existence at the time of designation now exists and that rescission is justified by reference to one of the enumerated standards. In essence, this should be a two-step process for owners: First, they should have to demonstrate that there has been a change in circumstances or a new potential use which justifies re-evaluating the decision based on a factor that was not considered during the initial

270. *Kracke v. Weinberg*, 79 A.2d 387, 391 (Md. 1951). See generally Barlow Burke, Jr., *The Change-Mistake Rule and Zoning in Maryland*, 25 AM. U. L. REV. 631 (1976).

271. *Kracke*, 79 A.2d at 391.

272. Philip J. Tierney, *Bold Promises But Baby Steps: Maryland's Growth Policy to the Year 2020*, 23 U. BALT. L. REV. 461, 483 n.154 (1994) (surveying cases on change-mistake rule and stating that, while rule has been successful, Maryland State Legislature should enable zoning boards to make changes in other circumstances).

273. John Mixon & Kathleen McGlynn, *A New Zoning and Planning Metaphor: Chaos and Complexity Theory*, 42 HOUS. L. REV. 1221, 1231 (2006).

274. In the somewhat unlikely event that the LPC decides to revoke landmark status on its own and there is a legal challenge, the current level of administrative deference is justifiable, as there are already safeguards against arbitrary action in place, and this is the kind of decision that is delegated to the LPC.

designation. Second, if they can do so, they should also have to demonstrate that the change is significant enough that landmark status is no longer appropriate.²⁷⁵ This would both provide the LPC with a way to avoid frivolous claims and allow owners a second chance to challenge the rulings of the LPC in court, thus increasing the transparency of LPC decisions by forcing them to present arguments to neutral arbiters. Permitting this kind of claim would add clarity to the current landmarks regime and further the goal of balancing historic preservation against the present needs of the city. Furthermore, it would provide standards against which a court could evaluate the LPC's decisions without modifying the existing standard of agency deference.

C. *Circumstances Justifying Rescission*

There are at least three circumstances in which rescission should be warranted.²⁷⁶ First, when there has been a significant change in the economic circumstances surrounding the landmark, owners should be entitled to present evidence that the initial designation as a landmark no longer makes sense. The kinds of changes for which this would be appropriate would primarily be economic.

Changes that should be considered include significant increases or decreases in the area's population²⁷⁷ and significant shifts in the economic activities of the neighborhood. Both of these situations would alter the economic impact of the landmark so significantly that a new look at the designation is justified. This provision would be especially useful in situations where the present use of the landmark is no longer useful to the area and the landmark designation impedes any other practical uses.

Landmark designations are justified, in part, by economic considerations.²⁷⁸ Therefore, it is incongruous to consider only the economic impact at the time of designation and not at the time of

275. The kinds of changes that would justify this should vary with the type of claim made. Changes that would qualify are addressed for each of the circumstances that would justify rescission.

276. One further circumstance, simple mistake in the original designation, has not been discussed, because it is uncomplicated, appears to be extremely rare, and is adequately covered by the existing provision. This Note should not be read as foreclosing the possibility that other circumstances could also justify rescission.

277. To be clear, this should not apply to shifts in the demographics of a neighborhood, including gentrification. Part of the reason for historic preservation is to maintain vestiges of the past; therefore, changes in demographics are insufficient to justify rescission.

278. N.Y.C. ADMIN. CODE § 25-301 (1992) (declaration of public policy, stating that purposes of law include, *inter alia*, "stabiliz[ing] and improv[ing] property

application for rescission. Allowing rescission in changed economic circumstances would strike a balance between the Landmarks Law's competing aims of preservation and development. The approach taken by courts in other types of land use disputes is instructive here. For example, Pennsylvania courts have refused to bar claims contesting denials of variances on the basis of *res judicata* when "there has been a subsequent substantial change in conditions incident to the land itself."²⁷⁹ Placing the burden on owners to show that such a change exists would limit the dangers of re-litigation and would still allow for reconsideration when appropriate.

Second, rescission should be granted when the landmark is one of the least valuable examples of a style of architecture that is over-protected in the city. The question should be: "How valuable is maintaining the landmark, given the existence of other protected exemplars of the same architectural style?" As architectural trends reach designation-eligible age, and as different commissioners identify particular architectural trends as being worthy of preservation, the Landmarks Preservation Commission can designate several exemplars of this style. For example, in 2009, the LPC committed to preserving more Modernist buildings and designated four examples in that fiscal year alone.²⁸⁰

While programmatic planning by the LPC may be desirable, there is also the possibility that, given the costs both to the owner and to the public of preserving a landmark, there may not be a need for every component of the scheme. Under the current system, landmarks are designated without reference to the broader program of protection, even when the LPC has targeted a particular style for protection.²⁸¹ This is the worst of both worlds: by programmatically identifying and then protecting sites in the first

values in [historic] districts . . . protect[ing] and enhanc[ing] the city's attractions to tourists and visitors . . . [and] . . . strengthen[ing] the economy of the city").

279. *City of Pittsburgh v. Zoning Bd. of Adjustment*, 559 A.2d 896, 901 (Pa. 1989). While that case rejected a variance on the basis of "mere economic hardship," the circumstances included "self-inflicted economic harm." *Id.* at 903-04. The point made in this Note is primarily procedural; thus, the Pennsylvania court's refusal to credit the economic harm is somewhat irrelevant. *Cf. Fisher v. City of Dover*, 412 A.2d 1024, 1027-28 (N.H. 1980) (reversing lower court's affirmation of zoning board's granting of variance because zoning board had not demonstrated change in circumstances).

280. 2009 MAYOR'S REPORT, *supra* note 7, at 107.

281. *See, e.g.*, N.Y.C. LANDMARK PRESERVATION COMMISSION, *Spring Mills Building*, Designation List 428, LP-2385 10 (Apr. 13, 2010), available at <http://www.nyc.gov/html/lpc/downloads/pdf/reports/SpringsMills.pdf> (describing site as "a particularly well-preserved example of a mid-20th century glass curtain wall skyscraper" without reference to any other extant examples).

instance without a way to later evaluate the landmark in the context of the scheme in its entirety later, there is a danger of both over-protection and of arbitrariness. If an owner can demonstrate that the significance of the structure within the overall preservation program was not considered when the landmark was designated, or that subsequent designations have significantly weakened the importance of preserving their building, or that there has been a major reevaluation of the importance of the protected attributes, then he or she should be able to apply for rescission.

Owners of landmarks should be granted rescission if they demonstrate that their property is protected primarily as an exemplar of certain kind of architecture and that that style of architecture is already adequately protected relative to its importance to the history of the city or the history of architecture.²⁸² The scarcity of buildings exhibiting a particular style of architecture is a valuable consideration, because it allows the city to find a balance between the public benefits of having historical buildings against the economic needs of the city. For example, if the Trylon Theater had been the last Art Deco movie house left in the city, then perhaps it would have been more crucial to save it.²⁸³ There are likely to be cases where it makes very little difference to the public which particular examples are preserved or where those exemplars are located, but the actual location of a particular landmark is crucial to a development plan or other advantageous use.

There are a number of ways in which the owner could demonstrate that the site is protected merely as an exemplar. The initial designation report would be useful for this, as it typically explains, at least briefly, the reasons why a site has been deemed to be important, and it often states that the building is typical of a particular style.²⁸⁴ If rescission on these grounds were allowed, however, the

282. There are other reasons, in addition to architectural style, for granting landmark status. The logic of this Note applies in these situations. For example, some sites are protected because of the historical importance of the period in which they were built, rather than because of the actual aesthetic merit of the buildings. When the argument for rescission centers on schematic concerns, it will likely be based on architectural style; therefore, I have referred solely to those circumstances for the purposes of clarity.

283. For a contemporary account of the dispute over the Trylon, see Jeff Vandam, *For an Art Moderne Theater, A Struggle Over Act II*, N.Y. TIMES, Sept. 18, 2005, § 14, at 8, available at <http://query.nytimes.com/gst/fullpage.html?res=9D0CE7D61E31F93BA2575AC0A9639C8B63>. The theater has since been destroyed.

284. For an example of one of the more comprehensive designation reports found while researching this Note, see LANDMARKS PRESERVATION COMMISSION, *130 West 57th Street Building*, Designation List 310, LP-2042 (Oct. 19, 1999), available at <http://www.nyc.gov/html/lpc/downloads/pdf/reports/130w57.pdf>.

LPC might simply add more reasons to the initial report in order to protect the designation from later challenge. Still, this would be preferable to the status quo. Forcing the LPC to provide more detailed justifications for landmark designations may limit the supply of landmarks and would create a written evidentiary trail, which would reduce the risk of arbitrary decisions.

One way to determine whether a landmark deserves continuing protection would be a ranking system. At least one state, Oregon, requires the state historic preservation office to survey all historic resources in a community prior to designation.²⁸⁵ While preservationists have taken issue with ranking systems because the rankings may make less significant structures seem “dispensable,” they have also praised their objectivity and responsiveness to community needs.²⁸⁶ As this Note has proposed, the less significant structures may in fact be dispensable, given the burdens on owners and the impediments to development. Furthermore, the Oregon system already uses rarity as one of its considerations even before designation.²⁸⁷ Similar ranking systems exist in San Francisco,²⁸⁸ Chicago,²⁸⁹ Boston,²⁹⁰ and England.²⁹¹ Using a ranking system, par-

285. See OR. REV. STAT. § 358.595 (2009); OR. STATE HISTORIC PRESERVATION OFFICE, 2005 OREGON HISTORIC PRESERVATION PLAN 7 (2005).

286. See, e.g. Bradford J. White & Richard J. Roddewig, *Elements of a Good Preservation Plan*, HISTORIC PRESERVATION LAW, SJ053 ALI-ABA 379, 386 (2004) (reprinted from American Planning Association, 1993).

287. See, e.g. CLACKAMAS COUNTY ZONING AND DEVELOPMENT ORDINANCE § 707.02(B)(5) (2008) (adopted in conformity with statewide standards for preservation).

288. San Francisco does not explicitly address the question of rarity, but it does divide buildings within the central city into five categories, assigning different limitations to each. See S.F., Cal., Planning Code, art. 11, §§ 1102, 1109–17.

289. Chicago has conducted at least one historical site surveys and ranks buildings in order of priority. See COMMISSION ON CHICAGO LANDMARKS, *Chicago Historic Resources Survey*, <http://webapps.cityofchicago.org/LandmarksWeb/chrs.do>. However, the ordinance establishing the Commission on Chicago Landmarks was recently struck down as unconstitutionally vague. *Hanna v. City of Chicago*, 907 N.E.2d 390 (Ill. App. Ct. 2009), *appeal denied*, 910 N.E.2d 1127 (Ill. 2009).

290. White & Roddewig, *supra* note 286, at 387; see also, e.g., THE DRUKER COMPANY, PROJECT NOTIFICATION FORUM: 350 BOYLSTON STREET 3-21 (2007), *available at* http://www.bostonredevelopmentauthority.org/DevelopmentProjects/PipeDocs/350%20Boylston%20Street/PNF/350%20Boylston%20Street_PNF.pdf (proposing development that would affect historic sites, and discussing applications of Boston ranking system).

291. Under the Planning Act of 1990, the English Secretary of State for Culture, Media, and Sport is required to list buildings for conservation status and categorize them according to three grades. Planning (Listed Buildings and Conservation Areas) Act, 1990 c. 9, *available at* <http://www.legislation.gov.uk/ukpga/1990/9/contents>.

ticularly for the rarity of what is being preserved, would have other benefits. For example, it would force the LPC to make explicit statements about their reasons for preserving a structure. This would, at the very least, give owners an administrative record to use if they sought to challenge a refusal to rescind as arbitrary. Another possible benefit would be greater flexibility for the LPC itself. Although designating a landmark as an exemplar will not necessarily create any great harm at the time of designation, the fact that landmark status is virtually permanent means that there is no opportunity to reevaluate the scheme in its entirety. If it becomes easier to rescind landmark designation, the LPC would be able to preserve a greater number of borderline landmarks while retaining the ability to reevaluate those decisions later.

One possible problem with instituting a ranking system is that it might encourage a race among owners to get their buildings de-designated. Even assuming that an owner has no plans to demolish or modify his or her landmarked property and no significant upkeep costs resulting from the landmark designation, the value of having the freedom to make changes without the approval of the LPC may suffice to motivate the owner to seek de-designation. That said, there are two reasons why this is not a significant problem. First, when a landmark is part of an over-protected architectural movement, it may be good to encourage owners to seek rescission. Second, this would alert the LPC to owners who wish to be rid of their landmarks, giving it an opportunity to try to find other owners for the site. This would help ease some of the LPC's enforcement costs.²⁹²

The third situation in which an owner of landmarked property should be able to seek rescission would be when the costs of maintaining the landmark significantly outweigh the benefits that accrue to the city. While there would be some overlap between this provision and either changed circumstances or value of the landmark within the overall landmark scheme, there are circumstances which would fit neither of those two provisions but would nonetheless justify the rescission of a landmark designation. Examples of this would include situations where a potential use of the property that could be instituted without the landmark designation. One example of a change that could justify this type of rescission application would be a proposed use of the property that was not possible at the time of designation or a significant shift in the value of the site for some other potential use.

292. See COUNCIL REPORT, *supra* note 235, at 4–6.

There are also some similarities between this proposal and the current hardship relief and reasonable return standards. Under those standards, however, potential uses of the site are not considered when evaluating the impact of the landmark designation,²⁹³ the result of which may be that it impedes a future use of the land without significant benefit to the city. Permitting owners to present evidence of such a situation will allow the LPC to more fully address the question of costs and benefits of a particular designation.

The recent dispute over the O'Toole Building at St. Vincent's Hospital illustrates the need for this kind of provision.²⁹⁴ In that instance, the LPC ultimately permitted the demolition of the O'Toole Building.²⁹⁵ A Certificate of Appropriateness was granted on the basis of a hardship application.²⁹⁶ To reach this result, the LPC stated that demolition of the site was "inappropriate"²⁹⁷ but presented reasons why demolition was nevertheless desirable.²⁹⁸ But these arguments focused on two aspects of the proposed demolition: whether the site interfered with St. Vincent's charitable mission,²⁹⁹ and whether the judicial hardship exception applied.³⁰⁰ Although there is no doubt that these were the proper questions for the LPC to ask under existing law, neither of them get to the question that the LPC seemed to be trying to answer, which was

293. See *supra* Part II.D.

294. For a summary of the landmark dispute, see Matt Chaban, *The Hardest Choice*, THE ARCHITECT'S NEWSPAPER, Oct. 28, 2008, http://www.archpaper.com/e-board_rev.asp?News_ID=2916. For an account of the eventual closure of St. Vincent's Hospital, see Suzanne Sataline, *St. Vincent's Hospital to Close*, WALL ST. J., Apr. 8, 2010, <http://online.wsj.com/article/SB10001424052702303411604575168631569560638.html>.

295. St. Vincent's Hardship Approval, *supra* note 133, at 9.

296. *Id.* at 4–7.

297. *Id.* at 4.

298. *Id.* at 5–7. In order to prove this argument, the LPC recounted a variety of "standards" for hospital designs that the O'Toole building failed to meet.

299. *Id.* at 4–5.

300. *Id.* Because St. Vincent's did not seek to alienate the property, the statutory provision was unavailable, but the judicial hardship test only seemed to apply if the entirety of the "campus" of St. Vincent's was considered. As at least two future litigants would point out, the idea of a "campus exception" to the hardship test was an invention of the LPC. See, e.g., Brief Amici Curiae of The Municipal Art Society of New York City, as Amici Curiae Supporting Petitioners, at 39, *Protect the Village Historic District v. N.Y.C. Landmarks Pres. Comm'n* No. 102744/2009 (N.Y. Sup. Ct. 2009), available at http://www.gvshp.org/_gvshp/preservation/st_vincents/doc/amicus-curiae-11-04-09.pdf. For a summary of this brief, see Letter from David Schnakenberg, The Municipal Art Society of New York, to Participating Amici (Nov. 4 2009), available at http://www.gvshp.org/_gvshp/preservation/st_vincents/doc/amicus-summary-11-04-09.pdf.

whether the aesthetic or historical value of the building outweighed the cost to the City of preventing St. Vincent's from expanding. Based on this example, it seems as though the LPC is effectively engaging in the kind of cost-benefit analysis proposed in this Note, at least in the more extreme cases. Rather than issue a ruling to that effect, the LPC is forced to contort the existing law in ways that distort the incentives for owners to properly preserve property and manage their commercial or charitable missions properly. It would be better if they could state what they are in fact doing and allow owners to apply for rescission based on the substantive question as to whether maintaining the designation makes sense. Because the City Council can overturn decisions of the LPC,³⁰¹ a democratic check exists on the process, which would allow for more direct expression of voter preferences in these kinds of rescissions.

Cost-benefit analysis is never easy, especially when many of the benefits resist quantification.³⁰² However, almost all of the other land use regulatory schemes in New York City demand some form of cost-benefit analysis or environmental impact review, often with community involvement.³⁰³ There is no principled reason why this form of analysis should be limited to the initial designation, where at least a rudimentary form of cost-benefit analysis is used. Just as the circumstances surrounding the landmark may change, the potential uses of the site may change, and the needs of the city may evolve. Allowing property owners to present evidence that the restrictions on their property are inefficient would allow the LPC to create a better balance between preservation and other concerns.

The major concern with allowing owners to claim that there is simply a better use for the land than as a landmark is that an important goal of a landmarks law is to insulate historic sites from pressure from developers. Penn Station, which may have provided part of the impetus for the Landmarks Law, is an example. In that case, an unprotected historic site was destroyed by a coalition of developers, politicians, and construction unions.³⁰⁴ Were cost-benefit analysis grounds for rescission, there would be a concern that developers could more easily destroy historic sites. The point of a landmarks regime should be to find a balance between the desira-

301. N.Y.C. ADMIN. CODE § 25-303(g)(2) (1992).

302. See Mason, *supra* note 25, at 21.

303. See, e.g., N.Y. CITY CHARTER § 197-c (2004). See generally N.Y.C. CHARTER REVISION COMM'N, *Summary of Expert Testimony*, in FINAL REPORT OF THE 2010 N.Y.C. CHARTER REVISION COMM'N B-12 (2010), available at http://www.nyc.gov/html/charter/downloads/pdf/final_report_2010_charte_revision_9-1-10.pdf.

304. Wood, *supra* note 71, at 301.

bility of preservation and the economic and social needs of the city. These concerns can be mitigated in a number of ways. As discussed, placing the burden of demonstrating that some change has occurred on the owners of the property would limit their ability to rescind; the LPC is likely to be unsympathetic to claims that there is a better use for the property. Preservationists would still be able to challenge rescissions in court, so an additional level of oversight would remain. As discussed in greater depth below, limiting the number of such claims, would give owners a reason to only make claims that they are likely to win. Additionally, there are external constraints on the owners of landmarked property, which would limit the use of this provision to extreme cases. Furthermore, bringing such a claim would alert the LPC to the fact that the owner wished to be free of the designation; the LPC is uniquely positioned to bring together owners of landmarked property who wish to get rid of the designation and potential owners who would maintain it, and so this may provide an additional mechanism for collaboration.

IV. CONCLUSION

These proposals do not mean that rescission is a panacea to the public goods dilemma of landmarks. Without the Landmarks Preservation Commission, there is likely to be an undersupply of landmarks, yet the current system both oversupplies the good and places significant burdens on private actors. More aggressive use of rescission, in conjunction with the standards outlined above, can mitigate these two problems. Mechanisms exist within the current landmarks regime to address situations where landmark status may no longer be appropriate, but, as shown above, they are inadequate to the problem of oversupply, and they place a very high burden on individual owners.

Rescission is preferable to the status quo for a number of reasons. First, using rescission would allow the LPC to focus more closely on questions within its field of competence. Rather than have the LPC evaluate financial returns or potential for greater charitable uses, the LPC would be able to focus on the general programmatic goals of the landmarking regime. It would also force the LPC to justify its decisions in more scrupulous and transparent ways, because the initial determination that a site should be a landmark would likely create a strong presumption in favor of continued preservation in the eyes of reviewing judges. This would help ferret out developer interest and would expose the process to greater scrutiny and oversight.

The second reason to prefer rescission is that it can lead to greater accommodation between preservationists and developers.³⁰⁵ There would be less of a reason to fight landmark designations if they were not perpetual. Rescission claims would also alert the LPC and the preservationist community to the desires of the property owners prior to the developers having to formally submit a plan under the current certificate of appropriateness scheme. Because the current system requires a significant showing of a plan to modify or develop, of hardship, or no reasonable return, the proposed option could function as an early warning sign, allowing the LPC and preservationists to attempt to find buyers who would be willing to preserve the site. The LPC is uniquely positioned to bring together preservationists and developers and could potentially be useful in reducing transaction costs associated with efforts to move properties from the hands of those who wish to demolish into those who wish to preserve.³⁰⁶ This would have the added benefit of reflecting expressed preferences for landmarks, at least insofar as preservationists could raise money for their cause.³⁰⁷

While this system may encourage a certain amount of re-litigation, the problem can be mitigated in a couple of ways. In order to guard against re-litigation by owners of landmarks, a timing provision should be included, stating that rescission cannot occur until a fixed number of years have passed since the designation and permitting applications for rescission only once during a fixed period of years.³⁰⁸ This would have two benefits: it would limit the power of repeat players or wealthy owners of landmarked property to continually re-litigate their claims, and it would allow the LPC to make regular comparisons between properties over time, which could make cost-benefit analysis increasingly easy. Even if more of these claims are brought, the fact that this sort of litigation (bringing a

305. If owners of a landmarked property subject to demolition by decay or otherwise sub-optimal use are encouraged to alert the LPC that they no longer wish their property to be subject to the landmark designation, the LPC may be able to step in and assist them. *See supra* Part II.B.

306. For an example of the type of landmark dispute that this could help mitigate, see Noah Rosenberg, *Seek "Sensitive" Theater Buyer: Landmark Ridgewood Theater Not Protected Inside*, THE QUEENS COURIER, Apr. 14, 2010, available at http://www.queenscourier.com/articles/2010/04/14/news/top_stories/doc4bc608af57448981604477.txt.

307. Obviously, those unable to afford to participate in these deals would still be at risk of being underserved by landmarks, but this is the same under the present system.

308. The ideal length of this term is largely an empirical question, and should be determined in consultation with the relevant actors.

petition to the LPC) is likely to remain far cheaper than full-scale lawsuits under Article 78 may be cheaper in the aggregate for owners to come before the LPC and wait out the timing provision if they lose rather than challenge the decision in court. The attempt at rescission should also be tied to the property rather than to the owner, thus making it impossible to avoid this provision by conveying the property which would, over time, produce data on how valuable the possibility of rescission is to owners.

Forcing owners to demonstrate a change would lead to an increase in scrutiny by the courts. This would not require a revision of the current standard for administrative review, but rather a recognition that, were the LPC to rescind a designation, the record would need to demonstrate why the previous designation is no longer appropriate, which may require a more exhaustive evaluation than is currently used. This new process would therefore serve to mitigate some of the concerns surrounding the lack of transparency of the LPC.

Historic preservation has always been controversial in New York City. The rate at which the city evolves and the scarcity of space, particularly in Manhattan, and the rate at which the city evolves create significant tensions between the desire to preserve the past and the need to serve the present. The current landmarks regime is prone to oversupply landmarks, and, by failing to provide adequate mechanisms for rescission, does not create an ideal balance between these two goals. Given some of the past failures to preserve the city's landmarks, the urge to freeze aspects of the city is understandable but is not a sufficient reason to retain the status quo. There are circumstances in which maintaining landmark designations simply does not make sense, and the Landmarks Law should be changed to reflect that. The goal is not to return to the old days of Robert Moses, nor to suggest that the LPC should be less vigilant in protecting the City's cultural legacy. Rather, the City's Administrative Code should be amended to include a more realistic standard and procedure for the rescission of landmarks so that owners of landmarked property, as well as the public at large, can help strike the proper balance between the goals of development and preservation.

1000 NYU ANNUAL SURVEY OF AMERICAN LAW [Vol. 66:951