ARTICLES

Seeing the State: Transparency as Metaphor.....................Mark Fenster  617

A History of the Military Authority Exception in the
Administrative Procedure Act.......................... Kathryn E. Kovacs  673

“Your Results May Vary”: Protecting Students and
Taxpayers Through Tighter Regulation of
Proprietary School Representations.................. Aaron N. Taylor  729

New Governance, Financial Regulation, and Challenges
to Legitimacy: The Example of the Internal Models
Approach to Capital Adequacy Regulation.......... Robert F. Weber  783

COMMENT

In Name Only: Employee Participation Programs
and Delegated Managerial Authority
after Crown Cork & Seal.................................Joseph D. Richardson  871

RECENT DEVELOPMENTS

Congress and the Courts Close Their Eyes:
The Continuing Abdication of the Duty to Review
Agency’s Noncompliance with the
Congressional Review Act............................Sean D. Croston  907

The Economics of Railroad
“Captive Shipper” Legislation............................Russell Pittman  919
## Volume 62  Summer 2010  Number 3

### Editor in Chief
Talitha Macharia

### Executive Editor
Amy Gaither

### Managing Editor
Peter J. White

### Senior Editors

<table>
<thead>
<tr>
<th>Field</th>
<th>Editors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Senior Articles Editors</td>
<td>Jonathan Harnois, Cristopher D. Jones, Jonathan T. Zimmer</td>
</tr>
<tr>
<td>Senior Note &amp; Comment Editors</td>
<td>Francis Massaro, Lauren C. Clark, Kirsten Soto</td>
</tr>
<tr>
<td>Senior Symposia Editor</td>
<td>Jesse S. Sommer</td>
</tr>
</tbody>
</table>

### Senior Staff

<table>
<thead>
<tr>
<th>Name</th>
<th>Position</th>
</tr>
</thead>
<tbody>
<tr>
<td>Camden D. Burton</td>
<td></td>
</tr>
<tr>
<td>Kara M. Clunk</td>
<td></td>
</tr>
<tr>
<td>Ryan Duffy</td>
<td></td>
</tr>
<tr>
<td>Jessica A. Flores</td>
<td></td>
</tr>
<tr>
<td>Cyrus M. Gerammarye</td>
<td></td>
</tr>
<tr>
<td>Gregory Glofak</td>
<td></td>
</tr>
<tr>
<td>Jonathan Wang</td>
<td></td>
</tr>
</tbody>
</table>

### Junior Staff

<table>
<thead>
<tr>
<th>Name</th>
<th>Position</th>
</tr>
</thead>
<tbody>
<tr>
<td>Katherine Aljinovic</td>
<td></td>
</tr>
<tr>
<td>Christina M. Andreen</td>
<td></td>
</tr>
<tr>
<td>Alexander Bard</td>
<td></td>
</tr>
<tr>
<td>Emily Baver</td>
<td></td>
</tr>
<tr>
<td>David Van Fleet Bloys</td>
<td></td>
</tr>
<tr>
<td>Jessica Buonaccorsi</td>
<td></td>
</tr>
<tr>
<td>Shannon Cain</td>
<td></td>
</tr>
<tr>
<td>Kristan Callahan</td>
<td></td>
</tr>
<tr>
<td>Lauren Caplan</td>
<td></td>
</tr>
<tr>
<td>Robert D. Cardina</td>
<td></td>
</tr>
<tr>
<td>Sarah Chiang</td>
<td></td>
</tr>
<tr>
<td>Amanda Dodd Lutz</td>
<td></td>
</tr>
</tbody>
</table>

### Faculty Board

<table>
<thead>
<tr>
<th>Name</th>
<th>Position</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gary J. Edles</td>
<td></td>
</tr>
<tr>
<td>Jeffrey S. Lubbers</td>
<td></td>
</tr>
<tr>
<td>William J. Snape, III</td>
<td></td>
</tr>
<tr>
<td>Paul F. Figley</td>
<td></td>
</tr>
<tr>
<td>Adeen Postar</td>
<td></td>
</tr>
<tr>
<td>Anthony E. Varona</td>
<td></td>
</tr>
<tr>
<td>Lewis A. Grossman</td>
<td></td>
</tr>
<tr>
<td>Jamin B. Raskin</td>
<td></td>
</tr>
<tr>
<td>Robert G. Vaughn</td>
<td></td>
</tr>
<tr>
<td>Benjamin Leff</td>
<td></td>
</tr>
<tr>
<td>Stephen J. Wermiel</td>
<td></td>
</tr>
</tbody>
</table>

### Coordinator, Law Reviews
Sharon E. Wolfe
AMERICAN UNIVERSITY
WASHINGTON COLLEGE OF LAW FACULTY

Administration
Claudio M. Grossman, B.A., J.D., S.J.D., Dean
Anthony E. Varona, A.B., J.D., LL.M., Associate Dean for Faculty and Academic Affairs
Christine Haight Farley, B.A., J.D., LL.M., J.S.D., Associate Dean for Faculty and Academic Affairs
Trishana E. Bowden, B.S., Assistant Dean for Development and Alumni Relations
Lewis Grossman, B.A., J.D., Ph.D., Associate Dean for Scholarship
David B. Jaffe, B.A., J.D., Associate Dean for Student Affairs
Khalid R. O. Khalif, B.A., A.M., Assistant Dean for Finance and Administration
Rebecca T. Davis, B.S., M.A.T., Assistant Dean for Academic Services and Registrar
D. Akira Shirama, B.S., J.D., Assistant Dean for Admissions and Financial Aid
Billie Jo Kaufman, B.S., M.S., J.D., Associate Dean of Libraries and Information Resources
David Aaronson, B.A., M.A., LL.B., LL.M., Ph.D., Director of the Trial Advocacy Program
Robert Dinestien, A.B., J.D., Director of the Clinical Program
Daniel Bradlow, B.A., J.D., LL.M., Director of the International Legal Studies Program
Teresa Godwin Phelps, B.A., M.A., M.S.L., Ph.D., Director of the Legal Rhetoric and Writing Program
Jamin B. Raskin, B.A., J.D., Director of the LL.M. Program in Law and Government
Ann C. Shaleck, A.B., J.D., Director of the Women and the Law Program

Full-Time Faculty
David E. Aaronson, B.A., M.A., Ph.D., The George Washington University; LL.B., Harvard University; LL.M., Georgetown University, B.J. Tenney Professor of Law and Director of the Trial Advocacy Program
Evelyn G. Abrahame, A.B., J.D., Case Western Reserve University, Professor of Law
Padideh Ala'i, B.A., University of Oregon, J.D., Harvard University, Professor of Law and Acting Director of the International Legal Studies Program
Kenneth Anderson, B.A., University of California at Los Angeles; J.D., Harvard University, Professor of Law and Director of the JD/MBA Dual Degree Program
Isaiah Baker, B.A., Yale University; M.A., DePaul University; M.B.A., J.D., Columbia University; LL.M., Harvard University, Associate Professor of Law
Jonathan B. Baker, A.B., J.D., Harvard University; Ph.D., Stanford University, Professor of Law
Susan D. Bennett, B.A., M.A., Yale University; J.D., Columbia University, Professor of Law and Director of the Community and Economic Development Law Clinic
Daniel Bradlow, B.A., University of Witwatersrand, South Africa; J.D., Northeastern University; LL.M., Georgetown University, Professor of Law and Director of the International Legal Studies Program
Pamela Bridgewater, B.A., Florida Agricultural and Mechanical University; J.D., Florida State University; LL.M., University of Wisconsin, Professor of Law
Barlow Burke, A.B., Harvard University; LL.B., M.C.P., University of Pennsylvania; LL.M., S.J.D., Yale University, Professor of Law and John S. Myers and Alvin Reckman Myers Scholar
Susan D. Carle, A.B., Bryn Mawr College; J.D., Yale University, Professor of Law
Michael Carroll, A.B., University of Chicago; J.D., Georgetown University, Professor of Law and Director of the Program on Information Justice and Intellectual Property
David F. Chavez, B.S., Michigan State University; J.D., University of California at Berkeley, Professor of Law
Janie Chuan, B.A., Yale University; J.D., Harvard University, Assistant Professor of Law
Mary Clark, A.B., Bryn Mawr College; J.D., Harvard University, Assistant Professor of Law
John B. Corr, B.A., M.A., John Carroll University; J.D., Georgetown University; Ph.D., Kent State University, Professor of Law
Angela Jordan Davis, B.A., Howard University; J.D., Harvard University, Professor of Law
Robert D. Dinerstein, A.B., Cornell University; J.D., Yale University, Professor of Law and Director of Clinical Programs
Lynda Dost, J.D., Yale Law School; Ph.D., Politics, Princeton University, Assistant Professor of Law
Walter A. Edwards, B.A., Princeton University; J.D., Harvard University, Professor of Law
Mary D. Fan, B.A., University of Arizona; M.Phil., University of Cambridge; J.D., Yale Law School, Assistant Professor of Law
Christine Haight Farley, B.A., Binghamton University; J.D., University at Buffalo Law School; LL.M., J.S.D., Columbia Law School, Associate Dean for Faculty and Academic Affairs and Professor of Law
Amanda Frost, A.B., J.D., Harvard University, Associate Professor of Law
David Gantz, A.B., Harvard College; J.D., J.S.M., Stanford Law School, Visiting Professor of Law
Anna Gelpern, A.B., Princeton University; M.Sc., London School of Economics and Political Science; J.D., Harvard University, Associate Professor of Law
Robert K. Goldman, B.A., University of Pennsylvania; J.D., University of Virginia, Professor of Law and Louis C. James Scholar
Claudio M. Grossman, Licenciado en Ciencias Juridicas y Sociales, Universidad de Chile, Santiago, Doctor of the Science of Law, University of Amsterdam, Dean, Professor of Law and Raymond I. Gershon Scholar in International and Humanitarian Law
Lewis A. Grossman, B.A., Yale University; J.D., Harvard University; Ph.D., Yale University, Associate Dean for Scholarship and Professor of Law
Heather Hughes, B.A., University of Chicago; J.D., Harvard Law School, Associate Professor of Law
David Hunter, B.A., University of Michigan; J.D., Harvard University, Associate Professor of Law
Darren L. Hutchinson, B.A., University of Pennsylvania; J.D., Yale Law School, Professor of Law
Special Faculty Appointments

Nancy S. Abramowitwitz, B.S., Cornell University; J.D., Georgetown University. Practitioner in Residence

Jamie Abrams, B.A., Indiana University at Bloomington; J.D., American University Washington College of Law. Legal Rhetoric Instructor

David Bahaatte, B.S., Brown University; J.D., American University Washington College of Law. Practitioner in Residence

Elizabeth Boas, B.S., Virginia Polytechnic Institute and State University; J.D., George Mason University. Assistant Director of the Trial Practice Program

Mary Kathryn Garrett Bunker, B.S., University of Maryland; J.D., The George Washington University. Practitioner in Residence

Gary Edles, B.A., Queens College; J.D., New York University; LLM, S.J.D., The George Washington University. Fellow in Administrative Law

Paul Figley, B.A., Franklin & Marshall College; J.D., Southern Methodist University School of Law. Legal Rhetoric Instructor

Sean Flynn, B.A., Pitzer College - Claremont; J.D., Harvard Law School. Associate Director, Program on Information, Justice and Intellectual Property

Horacio Grigera Naon, LL.B., J.D., School of Law of the University of Buenos Aires; S.J.D., LLM., Harvard Law School. Distinguished Practitioner in Residence, Director of the International Arbitration Program

Mentall Jain, B.S., Georgetown University; J.D., Hastings College of the Law. Practitioner in Residence and Human Rights Clinic

Elizabeth Keith, B.A., University of North Carolina at Chapel Hill; J.D., George Mason University School of Law. Legal Rhetoric Instructor

Daniela Kraiem, B.A., University of California at Santa Barbara; J.D., University of California at Davis. Associate Director of the Women and the Law Program

Jaimie Lee, B.A., Yale University; J.D., Harvard Law School. Practitioner in Residence

Jerome Levinson, B.A., L.L.B., Harvard University. Distinguished Lawyer in Residence

Adrienne Lockie, B.A., University of North Carolina; J.D., Georgetown University Law Center. Practitioner in Residence, Women's and the Law Clinic

Jeffry S. Lubben, A.B., Cornell University; J.D., University of Chicago Law School. Professor of Practice in Administrative Law

Daniel Marcus, B.A., Brandeis University; L.L.B., Yale Law School. Fellow in Law and Government

Claudia Martin, Law Degree, Universidad de Buenos Aires; LLM, American University Washington College of Law, Professorial Lecturer in Residence

Corrine Parve, Dip. PT., B.P.T., McGill University; J.D., American University Washington College of Law. Practitioner in Residence and Executive Director, Health Law Project, Program on Law and Government

Victoria Phillips, B.A., Smith College; J.D., American University Washington College of Law. Professor of the Practice of Law

Heather Ridensour, B.B.A., Texas Women's University; J.D., Texas Wesleyan School of Law. Director of Legal Analysis Program and Legal Rhetoric Institute

Diego Rodríguez-Pinzón, J.D., Universidad de los Andes; LLM, American University Washington College of Law; S.J.D., The George Washington University. Professional Lecturer in Residence and Co-Director, Academy on Human Rights and Humanitarian Law

Susana SaíConto, B.A., Brown University; M.A.L.D., The Fletcher School of Law and Diplomacy; J.D., Northeastern University Law School. Professional Lecturer in Residence and Director, Vet Crimes Research Office

Joshua Sarnoff, B.S., Massachusetts Institute of Technology; J.D., Stanford Law School. Professor of the Practice of Law

William Snape, B.A., University of California at Los Angeles; J.D., George Washington University. Fellow in Environmental Law

David Spratt, B.A., The College of William and Mary; J.D., American University Washington College of Law. Legal Rhetoric Instructor

Richard Ugelow, B.A., Hobart College; J.D., American University Washington College of Law; LLM, Georgetown University University. Practitioner in Residence

Rangeley Wallace, J.D., American University Washington College of Law; LLM, Georgetown University Law Centers. Practitioner in Residence

Stephen Wernmier, A.B., Tufts University; J.D., American University Washington College of Law. Fellow in Law and Government and Associate Director of the Saper Institute on Law and Government

Robin Westbrook, B.A., Smith College; J.D., Yale Law School. Practitioner in Residence, Tax Clinic

Bill Yeomans, B.A., Trinity College; J.D., Boston University Law School; LLM, Harvard Law School. Fellow in Law and Government
AMERICAN BAR ASSOCIATION
SECTION OF ADMINISTRATIVE LAW AND REGULATORY PRACTICE

OFFICERS AND COUNCIL

Officers
Chair
Chair-Elect
Vice Chair
Secretary
Budget Officer
Section Delegates
Last Retiring Chair

Jonathan J. Rusch*
Michael E. Herz*
James W. Conrad, Jr.*
Anna Williams Shavers*
Ronald L. Smith*
Randolph J. May*
Hon. John M. Vittone*
William V. Laneburg*

* Executive Committee Member

ABA Board of Governors Liaison

Peter A. Winograd

Council
Member 2011
Daniel Cohen
James P. Gerkis
Linda C. Lasley
Steve Vieux

Member 2012
Cynthia R. Farina
Jeffrey B. Litwak
Fiona Agnes Philip
Jeffrey Rosen

Member 2013
Linda D. Jellum
Suedeen G. Kelly
Nina A. Mendelson
Jason Schlosberg

Ex-Officio
State Administrative Law
Executive Branch
Judiciary
Legislative Branch
Administrative Judiciary
Young Lawyers Division
Law Student Division

Edward J. Schoenbaum
Michael A. Fitzpatrick
Brett M. Kavanaugh
James Park
Jodi B. Levine
Richard Raiders
ARTICLES

Seeing the State: Transparency as Metaphor..................Mark Fenster 617

A History of the Military Authority Exception in the
Administrative Procedure Act.......................... Kathryn E. Kovacs 673

“Your Results May Vary”: Protecting Students and
Taxpayers Through Tighter Regulation of
Proprietary School Representations.................. Aaron N. Taylor 729

New Governance, Financial Regulation, and Challenges
to Legitimacy: The Example of the Internal Models
Approach to Capital Adequacy Regulation.......... Robert F. Weber 783

COMMENT

In Name Only: Employee Participation Programs
and Delegated Managerial Authority
after Crown Cork & Seal.................... Joseph D. Richardson 871

RECENT DEVELOPMENTS

Congress and the Courts Close Their Eyes:
The Continuing Abdication of the Duty to Review
Agencies’ Noncompliance with the
Congressional Review Act..................Sean D. Croston 907

The Economics of Railroad
“Captive Shipper” Legislation........................... Russell Pittman 919
ARTICLES

SEEING THE STATE: TRANSPARENCY AS METAPHOR

MARK FENSTER*

Introduction ............................................................ 617
I. Transparency as Populist Metaphor ..................................... 624
   A. Transparency as Metaphor ........................................... 624
   B. Transparency and the Democratic Wish .......................... 628
   C. The Impossibility of Transparency ................................ 632
II. Exposing the Organizational State ................................... 636
   A. Constitutional Transparencies ...................................... 637
   B. Statutory Transparencies ............................................ 643
   C. Private Transparencies ............................................. 647
   D. The Impossibility of Organizational Exposure .................. 652
      III. Exposing the Physical State ................................. 653
         A. Distance ......................................................... 653
         B. Enclosure ....................................................... 658
         C. The Impossibility of Containment ......................... 660
             1. Containing Meetings ....................................... 660
             2. Containing Documents .................................... 664
         D. The Impossibility of Physical Exposure ................... 668
Conclusion: The Panopticized State ..................................... 668

INTRODUCTION

Early in his memoir Secrets, Daniel Ellsberg recalls the moment he first surreptitiously accessed top-secret government information, an experience

* UF Research Foundation Professor and Associate Dean for Faculty Development, Levin College of Law, University of Florida. The author offers thanks to David Pozen for comments and suggestions, and to Trysh Travis for conversations and guidance. You may contact the author at fenster@law.ufl.edu.

617
that would lead him, ultimately, to become the most famous liberator of
classified documents in American history. Ellsberg was then a young, rising
Pentagon bureaucrat who had been hired away from his previous position
as a research analyst at Rand, a private think tank that served as a
consultant to the Pentagon’s efforts fighting the Vietnam War, to work for
John T. McNaughton, Assistant Secretary of Defense for International
Security Affairs.¹ In the course of his duties, McNaughton received
classified documents that Ellsberg lacked sufficient security clearance to
read. The binder in which those documents were filed sat on a rolling
bookstand in McNaughton’s office. Every evening, the bookstand was
rolled into a secure, locked closet. Ellsberg could see the binder but was not
allowed to look inside, despite its promise of invaluable information that
could divulge the secrets of the unfolding drama in Vietnam. Ellsberg
narrates the event of one fateful evening:

It was too much for me. There came a night—I can’t remember how many
weeks it was after [McNaughton] had directed my attention to this forbidden
binder—when I did pull it out of the row of files and open it. . . . The office
was dark; the light was coming from inside the closet. I was in the process of
putting the rolling stand away for the night. I looked inside the thick binder
and rifled through the contents. It was like opening the door on Ali Baba’s
treasure. . . . At a glance I could see that what I held in my hand was
precious. Reading just a few paragraphs here and there was, for me, like
breathing pure oxygen. My heart was pounding.²

Witness the tension and expectation as Ellsberg—who would later
illegally release to United States newspapers what would be famously
referred to as the “Pentagon Papers”—describes the ecstasy of access and
anticipates what would soon become his troubled, infamous relationship to
secret documents. The records that he was forbidden to view almost
commanded that he view them. They offered him new, important
information, and therefore revelation—the purest form of “oxygen” an
analyst like Ellsberg requires to survive and prosper. But their access had
been strictly limited. Not only were they removed from the public, which
was ignorant of their existence, they were even kept separate from someone
like Ellsberg, a Harvard-trained wunderkind specifically hired to assist the
government agency that forbid him access. Ellsberg was forced to violate
the law that prohibited him from viewing the documents, to cross both the
legal line and physical boundary that placed this binder beyond his view.
His heroism, to those who see it as such, began when he traversed that well-

1. Daniel Ellsberg, Secrets: A Memoir of Vietnam and the Pentagon Papers
2. Id. at 81.
guarded (but not well-guarded enough!) threshold into the sacred space where the most privileged information is secured. Only then could he imagine freeing that information from its physical constraints; only then could he imagine educating the public of the policies and actions that were being undertaken in its name.

For Ellsberg and those committed to the expansion and strict enforcement of open government laws, the antidote to the wrong of excessive governmental secrecy is greater transparency. Without access to the government, the public can neither evaluate the government’s performance in the past, nor hold the government accountable in the present, nor deliberate over the government’s future representatives or policies. As Ellsberg’s description vividly reveals, transparency suggests both visibility—these documents exist, and powerful government officials can see them—and a distance that makes that visibility difficult to achieve—you can’t see them, and you don’t even know they exist. The young bureaucrat would only become the (in)famous Daniel Ellsberg by allowing the public to view the information that was kept secret and secure.

When applied as a foundational concept for federal and state administrative laws mandating some form of open government, transparency assumes the existence of a gap that arises naturally between the state and its public. Its underlying logic works as follows. Government institutions operate at a distance from those they serve. To be held accountable and to perform well, the institutions must be visible to the public. But in the normal course of their bureaucratic operation, public organizations—sometimes inadvertently, sometimes willfully; sometimes with good intent, sometimes with unethical or illegal intent—create institutional impediments that obstruct external observation. These

obstructions must be removed in order for the institutions to be visible and, ultimately, transparent. The dictionary definition of the word transparency makes this dynamic plain: something that is transparent has “the property of transmitting light, so as to render bodies lying beyond completely visible; that can be seen through . . . ”. A transparent window, for example, enables one to see inside from outside or vice versa, rendering visible to each other those that are on either side, despite their separation.

Employed in this way, the term transparency simultaneously describes both an aspirational goal—full openness to the public—and the core problem that must be overcome in order for that goal to be met—the separation between the state and public. Judges, policy advocates, academics, and legislators frequently deploy the concept’s metaphorical authority when adjudicating, advocating, and legislating transparency. “Democracies die behind closed doors,” a federal appellate court declared when finding that the First Amendment prohibits the government from closing immigration hearings to the public and press without an individualized showing of justification. “Sunlight” or “sunshine,” when it is allowed to shine through previously darkened, secretive places, provides the best of “disinfectants,” Louis Brandeis famously contended when he decried the corrupt trusts of the early twentieth century.

Information must be set free from its bureaucratic constraints, as Congress declared in the name of its act requiring executive branch agencies to disclose information. Deep secrets—those state secrets that the public does not know that it does not

5. The same dynamic exists even when a commentator complicates the concept by substituting “translucent” for “transparency” in recognizing the inevitable limitations on public access to government information. See, e.g., Adam M. Samaha, Government Secrets, Constitutional Law, and Platforms for Judicial Intervention, 53 UCLA L. REV. 909, 923, 969–76 (2006).
7. LOUIS D. BRANDEIS, OTHER PEOPLE’S MONEY AND HOW THE BANKERS USE IT 92 (Augustus M. Kelley 1986) (1914) (“Sunlight is said to be the best of disinfectants; electric light the most efficient policeman.”); see also infra note 28 (identifying the influence this metaphor has on legal academic writings).
know because they are hidden below the public’s view—pose the greatest danger in liberal constitutional democracy, two important recent law review articles have persuasively argued. Transparency thus serves as more than a mere technical concept that provides the basis for constitutional, legislative, and regulatory rules. It also acts as a powerful metaphor that drives and shapes the desire for a more perfect democratic order.

Ideally, of course, there would be no distance between observer and observed, between the governed and those institutions that govern. The metaphor, in other words, would accurately diagnose the problem and set an agenda for the cure. Under a strong form of transparency, government doors should never be closed; government should not operate in the darkness; all government information should be available to the public; and in the rare instance when they must be kept from the public, government secrets should not be so deep that their existence is unknown. How else can citizens make up their minds independently of government officials and media gatekeepers, and advise elected officials as to the wisest course of action? A weaker conception of transparency conceives the need to balance transparency’s beneficial effects and normative value against the state’s need to withhold a limited amount of information whose disclosure would cause identifiable harm. As a metaphor, transparency suggests two solutions: allow the public to view the state directly, or require the state to make its work available for the public to review. Open government laws rely on both of these solutions by requiring certain government entities to hold open meetings, trials, and deliberations, and by mandating that


government records be made public routinely or in response to a public request. 13 Both the strong and weak conceptions of transparency assert that the legal order imposed by such laws—and other efforts by the state, urged on by the public, to impose openness—can unveil the state, eradicating or at least mitigating its distance from its citizens through mandates and obligations placed on government institutions and officials.

And yet, the regular, ritualistic outpouring of public complaints about the weakness of such laws and the power and dangers of a secretive government suggests that transparency’s metaphorical ideal in fact does not prevail. 14 The state remains distant and unseen, perhaps even concealed. In an earlier article, I explored the conceptual reasons why this disappointment seems endemic to transparency. 15 In this article, I explain how transparency’s metaphoric dimensions—the problem it identifies and the goal it sets—impede our ability to understand and address the complexities of the modern administrative state.

The public prefers a proximate, comprehensible, responsive bureaucracy, one that fulfills the “ democratic wish” of a directly accountable government. 16 Populist and progressive reforms and political campaigns endeavor to take the nation back from the present crisis caused by an autocratic, secretive “ other” ensconced in Washington and state capitols. 17 They promise that by revealing the state’s operations, transparency’s metaphorical understanding can enable the public to control the state. The transparency movement, which came of age as part of what Richard Stewart called the “ reformation” of American administrative law in the 1970s and after, suggests that the state must and can be made visible. 18

Administrative reform cannot, however, deliver on transparency’s metaphoric promise. The state’s large, organizationally and physically dispersed public bureaucracies perform a variety of functions and make a staggering number of decisions of varying importance, not all of which can

15. See generally Fenster, supra note 11.
17. See infra text accompanying notes 36–47.
be viewed before the fact or even easily reviewed later. The state is too big, too remote, and too enclosed to be completely visible. The very nature of the state, in other words, creates the conditions of its obscurity. It can never be fully transparent, at least not in the sense that the term and its populist suspicions of the state require. Overinvestment in transparency as a metaphor leads open government advocates to lament insufficiently effective administrative laws, while the debate over how best to make the government open too often focuses on how to make the state permanently and entirely visible rather than on devising means to improve public oversight and education. Transparency’s fear of a secret, remote government—like its promise of a visible, accessible one—heightens the concept’s salience even as it obscures the limits of its enforceability as an administrative norm.

Transparency is a means to achieve the end of a more responsive state that more effectively achieves democratically agreed-upon ends. Transparency’s symbolic pull, its ability to grab the public’s imagination, leads us to fetishize means at the cost of ends. My underlying assumption is that bureaucracy is necessary to carry out the tasks required in a complex society and economy. As the public administration scholar Donald Kettl has argued, “society has yet to discover anything that works better in coordinating complex action” than public bureaucracies. The public must certainly know about the government’s operations, but obtaining that knowledge is not a costless transaction. Simplistic understandings of the state’s operations and the potential of imposing equally simplistic

---

19. Cf. Cary Coglianese, The Transparency President? The Obama Administration and Open Government, 22 Governance 529, 537 (2009) (distinguishing between “fishbowl” transparency, which focuses on the maximal release of government data, and “reasoned” transparency, which more effectively requires government officials to provide “sound reasons for their decisions”); Mark Schmitt, Transparency for What?, Am. Prospect, Mar. 2010, at A10 (criticizing efforts to require the release of government data and praising legislative enactments that instead focus on increasing public understanding).


21. Donald F. Kettl, Public Bureaucracies, in The Oxford Handbook of Political Institutions 366, 373 (R.A.W. Rhodes et al. eds., 2006); see also Kenneth J. Meier & Gregory C. Hill, Bureaucracy in the Twenty-First Century, in The Oxford Handbook of Public Management 51, 51 (Ewan Ferlie et al. eds., 2005) (“Large-scale tasks that government must perform . . . will remain key functions of governments in the twenty-first century and . . . bureaucracies, likely public but possibly private, will continue to be the most effective way to do these tasks.”).
understandings of transparency can lead to imperfect, costly measures to disclose information and less effective governance.

This Article proceeds as follows: Part I explores transparency’s metaphoric work within American law, politics, and culture, and identifies its dual role as both a powerful, populist metaphor and a set of imperfect technocratic tools. It introduces the argument that Parts II and III then develop: transparency’s obsessive concern with visibility and the effort that this concern inspires to contain the state ultimately fail and disappoint because of the state’s inevitable organizational and geographic distance from the public. The technocratic tools of open government cannot in fact meet the demands that transparency’s force as a political and administrative symbol animates. Part II focuses on the state’s organizational complexity, both as a matter of form and function, and describes the various constitutional and statutory mechanisms that simultaneously establish an intricate institutional network and impose a limited, variable set of transparency commands. Part III describes the physical impediments to transparency caused by the vast territory of the American state, the complexity of its jurisdictional units, and the physical structures that house government offices. Both Parts II and III explain the impediments to the state’s visibility and the imperfect means that have been developed to overcome them.

A final, concluding part posits that the ultimate technocratic tool that could successfully contain the state and make it visible would reverse Jeremy Bentham’s Panopticon, rendering the state a prisoner of the public’s gaze. The impossibility of this solution demonstrates the limits of transparency as a symbol and suggests that the way forward is to understand transparency’s limited usefulness as a term for achieving both an effective and accessible state. Nevertheless, this Article concludes, transparency’s prevalence as a political concept requires reform efforts to balance delicately technocratic efficacy with populist demands.

I. TRANSPARENCY AS POPULIST METAPHOR

A. Transparency as Metaphor

Among other things, Barack Obama’s 2008 presidential campaign pledged to reverse the Bush administration’s penchant for secrecy and its general opposition to transparency norms, proclaiming on its campaign website that if elected Obama would “Shine the Light on Washington Lobbying” as well as on federal contracts, tax breaks, and earmarks, and
“Bring Americans Back into their Government.” Although it is difficult to ascertain what role Obama’s transparency pledge played in his victory, it was one among many issues that constituted his campaign’s narrative of Obama as an agent of change. Obama’s message was not an idiosyncratic one. The Democratic Party’s 1976 campaign platform, when Jimmy Carter defeated Gerald Ford in the first post-Watergate presidential election, offered quite similar calls for “responsive” and “competent” government that would end the “remote government” whose “secretive and unresponsive” approach the Nixon–Ford presidency had established.

Both campaigns featured self-proclaimed outsiders who touted their promises to reform a corrupt and secretive Washington and to make government accessible and visible to the public. Elect me and you will have your government back, their campaigns vowed. Underlying this partisan political discourse are the notions that the government you fear operates behind a veil of secrecy while the government you want operates in the open, and that no amount of secrecy is warranted while no amount of transparency is too great. These campaigns described a fallen world in which the state is remote and apart from its citizenry, operating corruptly and out of the public’s view. At the same time, they promised a government that would be close, visible, trustworthy, and transparent.

Such rhetoric is in fact quite common when an organization or writer advocates on behalf of transparency. “America is a nation of secrets,” one


recent popular book warns, “an increasingly furtive land where closed doors outnumber open ones . . . ” A large, international network of nongovernmental organizations that seek to expand public rights to information attempt to aid journalists and members of the public by pressuring governments to “free” information, operate in the open and in the sunshine, and make government data constantly and immediately available on an on-demand, real-time basis. The image pervades the academic literature on transparency as well, with definitions and introductory sections that imaginatively and provocatively present the government as a closed, isolated entity with shuttered windows and locked doors. One academic definition of transparency states that the term “refers to the degree to which information is available to outsiders that enables them to have informed voice in decisions and/or to assess the decisions made by insiders.” Following Brandeis’s dictum, hundreds of law review articles assert that “sunlight” offers a solution that can “disinfect” bad government and corruption. Some authors cast information as a substance that in a proper democracy must flow freely out of the government’s clutches and into the waiting arms of the public.


29. See, e.g., Aftergood, supra note 28, at 399 (“[T]he free flow of information to interested members of the public is a prerequisite to their participation in the deliberative process and to their ability to hold elected officials accountable.”); Michael Herz, Law Lags Behind: FOIA and Affirmative Disclosure of Information, 7 Cardozo Pub. L. Pol’y & Ethics J. 577 (2009) (arguing for the relevance of understanding information as needing to be free as
Compare this rhetoric to the far more fanciful depictions of a corrupt, secretive state in popular culture, which vividly and imaginatively harness the same imagery for dramatic effect. The dénouement of the first season of *The X-Files* reveals the locked Pentagon repository where the government sequesters the most prized, awful secrets from an ignorant public—the files that contain evidence of alien life and government conspiracy and that sit locked in a secured vault, accessible only to the few pernicious bureaucrats that know of the vault’s existence.\(^{30}\) The film adaptation of *All the President’s Men* memorably depicts the only place where the intrepid Woodward and Bernstein can obtain crucial government information about the illegal activities of the Nixon White House: the dark, obscure garage where they meet their anonymous source, Deep Throat. In one famous scene, the reporters sift through a huge stack of paper slips in order to find evidence of the administration’s malfeasance. The camera tracks steadily upward towards the library’s very high ceiling in a shot that captures the plight of two private citizens who attempt, against all odds, to pierce the informational haze that a complex but coordinated state can create. They are small and insignificant, forced to piece together a crucial story from obscure bits of evidence made only partially available, if at all, within the state’s cavernous, intimidating architecture.\(^{31}\)

The series of paired terms upon which transparency proponents and filmmakers rely—open and closed, transparent and secret, sunshine and darkness, inside and outside, and the like—works powerfully and metaphorically to give some normative, symbolic bite to an administrative norm. Films and television shows, political campaigns, and popular political discourse generally present secrecy and conspiracy as political commonplace, and suggest that the lone individual—as in Daniel Ellsberg’s leak of the Pentagon Papers and Woodward and Bernstein’s reporting on what became known as the Watergate scandal\(^ {32}\)—must save us from official corruption and perfidy.\(^ {33}\) Indeed, the political reforms that followed the Vietnam War and Watergate depended in part on popular disgust with government secrets,\(^ {34}\) as well as on Ellsberg’s and Woodward and

---

\(^{30}\) *The X-Files: The Erlenmeyer Flask* (Fox Television broadcast May 13, 1994).

\(^{31}\) *All the President’s Men* (Warner Bros. 1976).

\(^{32}\) CARL BERNSTEIN & BOB WOODWARD, *All the President’s Men* (1974).


Bernstein’s deification as heroic actors exposing government deceitfulness and treachery.  

Transparency thus operates simultaneously in two ways. It constitutes a technical concept that, when properly implemented in law and regulation, produces goods deemed essential for a democratic society: an effective administrative state; a knowledgeable citizenry that can hold the government accountable; and an active, deliberative polis. In implementing this understanding of the concept, constitutions and legislatures impose transparency through legal and administrative commands and institutional design, all of which require the intricate drafting of provisions and the delicate balancing of interests. At the same time, transparency also offers a highly charged metaphor of a corrupt, secretive state that must be made visible. The metaphoric understanding of transparency animates deeply held beliefs about the state’s legitimacy, escalating to the level of a preeminent democratic imperative the technocratic legal issue of how best to make the official administrative bureaucracy accessible.

B. Transparency and the Democratic Wish

Transparency’s two understandings, the technical or technocratic and the metaphorical, can work to mutual advantage. The Obama administration, for example, is attempting to meet the vivid rhetorical promises made in the Obama campaign with bureaucratic and technological reforms—small bore, technocratic efforts to change the bureaucratic culture of the federal government and to make government data more easily accessible. But they can also conflict. Each time the Obama administration has failed to take the most pro-transparency positions—on state secrets, photos of prisoners taken at the Abu Ghraib prison, and congressional negotiations over health care reform legislation, for example—critics from various points on the political spectrum have

(1999).


36. See Fenster, supra note 11, at 895–902.

37. See Memorandum from Peter Orszag, Director, Office of Mgmt. and Budget, on Open Government Directive to the Heads of Executive Departments and Agencies, (Dec. 8, 2009) http://www.whitehouse.gov/omb/assets/memoranda_2010/m10-06.pdf (announcing the directive to federal agencies to increase government information available online, improve the quality of government information, and “create and institutionalize a culture of open government”); see also Cogliano, supra note 19, at 533–35 (describing the Obama administration’s early efforts to expand transparency).
asserted that the President has failed to meet his campaign promises. In such instances, the metaphorical understanding of transparency overwhelms its technocratic understanding by creating a set of expectations that legal and regulatory reforms cannot fulfill. By invoking transparency’s symbolic meanings, a candidate or political movement may fire a drive for comprehensive solutions that rejects or minimizes the importance of technical, incremental efforts and that will accept nothing less than a perfectly accessible and visible state. Even as it reforms executive branch compliance with open government laws and norms, the Obama administration will continually frustrate transparency advocates, leftist reformers skeptical of the administration’s centrism, and conservative political opponents who characterize every refusal to disclose information or open government as another victory by a closed, secretive bureaucracy over the people’s will.

The paired terms upon which transparency relies thus establish openness as a metonym for democracy—an element of a representative government that appears to stand for its entirety. An engaged, informed populace can control a transparent state, but a distant, secretive bureaucracy rules the nontransparent state. In this sense, transparency offers a deeply populist account of politics and the administrative state in which an unresponsive state can and ultimately will obstruct and oppose inquisitive private individuals. By “populist,” I mean both the historical populist movements in the United States and, more particularly, the populist rhetoric and logic that suffuse American politics. Populism simplifies complex political

---


39. I leave aside for purposes of this Article the precise nature of the historical relationship between populism and transparency’s metaphoric understanding and whether, for example, it represents an aspect of what historian Richard Hofstadter described as the “paranoid style” in American politics, or whether it is a more recent and more rational response to the expansion of the executive branch since the Great Depression and especially following World War II. See Richard Hofstadter, THE PARANOID STYLE IN AMERICAN POLITICS (2d prtg., Alfred A. Knopf 1966). My purpose here is merely to note the relationship and to assert that the rhetoric of strong-form transparency advocacy and that implied by the term’s underlying metaphor clearly align with the rhetoric of American populism.

40. On populism as a flexible, rhetorical mode of persuasion in politics as well as an historical survey of populist movements, see generally Michael Kazin, THE POPULIST PERSUASION (1995). On the populist logic in American political culture, see Mark Fenster, CONSPIRACY THEORIES: SECRECY AND POWER IN AMERICAN CULTURE 84–89
alignments and issues within a stark, symbolic dichotomy between “the people” at one pole and “the other”—the power bloc in charge—at the absolute opposite. Populist appeals identify threats to the national identity and claim to speak on behalf of an identifiable collective “we,” a people who are rising up to challenge and resist the concentrated interests that hold power and the seemingly dangerous ideas and values those interests represent. Populism drifts left and right, with no necessary connection either to an institutional party or ideology. It can appear conservative (in the anticommunism of the 1950s and early 1960s), liberal (in the New Deal of the 1930s), or thoroughly independent (in the Populist campaigns of the late nineteenth century)—in each instance it identifies some concentration and combination of state and private power that threatens the people.

Populism plays a recurring role in the inevitable fight over the institutional processes of democratic political and social order. Because democratic representational politics relies on a gap between the public and its elected representatives that is mediated by established political institutions, populist rhetoric claims to offer some more direct or authentic means of representation in the name of the people when those institutions appear illegitimate, whether as a result of substantive or procedural irregularities. As Jack Balkin has explained, populist approaches to law and government commit to two basic preferences: popular participation and regular rotations of authority and power. Each preference envisions a state that is proximate and thoroughly visible to the citizens that control it. Thus, self-proclaimed populist or popular constitutional theorists in the legal academy embrace a vision of the constitutional order that they claim would prove more responsive to the popular will and less capable of elite manipulation.

(rev. and updated ed. 2008).

41. See ERNESTO LA CLAU, ON POPULIST REASON 18 (2005).
42. Margaret Canovan, Trust the People! Populism and the Two Faces of Democracy, 47 POL. STUD. 2, 4–5 (1999).
43. KAZIN, supra note 40, at 192–93.
Critics or skeptics of populism, especially those tied to what Balkin has called the progressive category or strain of public law, decry the retrograde and conservative implications of understanding the complex contemporary state in such simplistic terms. For progressives committed to the regulatory intervention into market activity provided by the administrative state, the government cannot rely on direct democratic rule, but must instead utilize expert, public agencies that deliberate rationally and are protected from direct political control and popular sentiment. Populist ideals can thus constitute a barrier to good, progressive governance. In Edward Rubin’s terms, they rely on an inherited set of symbols and metaphors that “produce a sense of dissonance or incongruity, a grinding of intellectual gears, when applied to a modern administrative state.” The progressivism of the regulatory state supports open government, but as a tool for improved governance rather than as a democratic end in itself.

Transparency thus operates somewhat uneasily and ironically at the conjunction of legal and political populism and progressivism. Its populism pursues what James Morone has called the “democratic wish” for direct democracy, consensus, and localism that generates and assembles a popular will to create a more perfectly accessible and instrumental state. Its mobilization around the ideal of the visible state proceeds restlessly and endlessly, driven by the unsatisfactory nature of the corrupt present. At the

---

51. See id. at 140 (noting criticisms of various federal open government laws, but ultimately approving of them as means by which administrative agencies interact with the public); see also infra notes 142–45 (discussing ambivalence of “new public governance” scholars towards transparency). Political leftists and progressives may espouse a strong commitment to transparency. See, e.g., Ellen Miller, Obama at One, NATION, Feb. 1, 2010, at 21 (contribution by the Sunlight Foundation Executive Director to a progressive magazine’s forum both praising and criticizing Obama’s record on transparency in his first year in office); Greenwald, supra note 38 (leftist writer condemning Obama’s poor commitment to transparency). In doing so, they espouse a left populism analogous to that of the popular constitutional theorists identified above, many of whom would also identify themselves as progressives or leftists. See supra note 48 and accompanying text.
52. MORONE, supra note 16, at 5–9.
same time, its progressive cast—its commitment to legal rules and institutions that can constrain the state and make it visible—attempts to address and manage popular discontent through a bureaucratic apparatus, one that has grown steadily at the federal, state, and local levels since the nation’s founding. The state’s bureaucratic apparatus executes legal rules and regulations and is itself controlled by an evolving and expanding set of laws.

This produces a cyclical, ironic dynamic: the populist demand for popular control of the state in turn leads to a more expansive state that in turn creates a larger bureaucratic organization that in turn leads to calls for more popular control. The Jacksonian era illustrates this dynamic quite well. Swept into power on a wave of populist sentiment that sought to wrest power away from what they characterized as a ruling Federalist elite and replace it with egalitarian, popular control of the state, Andrew Jackson and Jacksonian Democrats remade and expanded the federal bureaucracy, recasting the emerging American bureaucracy as one based on offices and rules rather than individuals and privilege. In this instance and others throughout American political and administrative history, the effort to make the state more accessible and accountable to the people also has led to an expanded administrative state. The narrower contemporary populist call to create a more visible state creates a similar dynamic. Forced to impose its will on a complex, decentralized set of governmental institutions created to meet its citizens’ substantive demands for public goods, benefits, and regulatory programs, efforts to create a more open government must rely on complex combinations of procedural laws, regulations, and institutions. The democratic wish for transparency may (or may not) lead to a more visible state, but it will certainly produce more of the state to make visible.

C. The Impossibility of Transparency

As a result of the populist dynamic that at once fears and expands the state, transparency has proven and will continue to prove impossible to achieve as an administrative norm in its strongest, metaphorical form.

From its beginnings, the new United States faced a dire organizational problem: how and whether to create a federal government out of a disparate set of colonies spread over a large territory while still addressing the popular demand for a direct, accessible government. The effort to do so spawned anxious commentary from proponents of the new constitution and angry condemnations by their critics. In The Federalist Number 37, James Madison worried about the “arduous” task facing the constitutional convention in “marking the proper line of partition between the authority of the general and that of the State governments,” and suggested that the issue was so complex, and its solution so difficult to derive, that the resulting lines drawn in the constitutional convention were the necessary result of human estimation and political compromise.\(^56\) The Anti-Federalists, meanwhile, characterized the task as impossible rather than merely arduous, and dismissed the resulting constitution as fatally flawed. Writing as Cato in The New-York Journal in 1787 (in a letter later collected as part of The Antifederalist Papers), New York Governor George Clinton warned against the “consolidation or union” of states that comprise an “immense extent of territory” “into one great whole”:

> [W]hat can you promise yourselves, on the score of consolidation of the United States into one government? Impracticability in the just exercise of it, your freedom insecure, even this form of government limited in its continuance, the employments of your country disposed of to the opulent, to whose contumely you will continually be an object. You must risk much, by indispensably placing trusts of the greatest magnitude, into the hands of individuals whose ambition for power, and aggrandizement, will oppress and grind you. Where, from the vast extent of your territory, and the complication of interests, the science of government will become intricate and perplexed, and too mysterious for you to understand and observe; and by which you are to be conducted into a monarchy, either limited or despotic; the latter, Mr. Locke remarks, is a government derived from neither nature nor compact.\(^57\)

In response to such arguments, Alexander Hamilton conceded that those who lived closer to the seat of power would enjoy greater access to the state than those who lived far away, but he argued that the proper institutional design of government, combined with the development of an active civil society and independent press, would produce a functional, accountable state.\(^58\) The Hamiltonian belief that organization can correct the structural


\(^{58}\) The Federalist No. 84, supra note 56, at 516–17 (Alexander Hamilton); cf. The Federalist No. 10, supra note 56, at 83 (James Madison) (arguing that a republic
problems caused by a large territory and complex federal system has remained prevalent throughout the twentieth century, most notably in repeated efforts to reorganize and tame what are seen as fragmented, haphazardly structured executive branches of both the federal and state governments. Bureaucratic organization has its “ups and downs” in modern democracies, in organizational theorist Johan Olsen’s terms, but its hold remains “tenacious” and its history marked by theoretical and political arguments over how best to design institutions and rules that might improve or perfect governmental operations.

These anxieties and arguments about the state originate in two distinct obstructions to the public’s ability to view it. The first barrier is organizational. If, in Madison’s terms, it has proven difficult to draw lines among the various levels and agents of government that wield state authority, then, in the Anti-Federalists’ terms, the state will appear “intricate and perplexed, and too mysterious” to monitor. Visibility requires simplicity because complexity creates opacity. The second barrier is spatial. Hamilton argued that the state could manage its offices and officers across vast distances through the formal and informal relationships among federal, state, and local governments, and by the diligent work of an alert press and public. He assumed that a complex organization of governmental institutions and civil society would develop, built in large part on the public’s agents in the press and federal and state capitals that would promote the national and public interest. The Anti-Federalists, by contrast, predicted that the vast post-colonial territory—itself having a small footprint compared to the current United States—would frustrate the

encompassing a larger territory, and therefore a larger population, would include more distinct parties and interests that would result in more factions that would check each other’s tendency to dominate.


61. The Antifederalist Papers No. 14, supra note 57, at 37 (George Clinton).
development of a functional national government and cohesive civil society.

If transparency abhors the distance between the state and public and requires immediacy, then efforts to make the government’s operations fully visible must overcome the organizational and spatial distances that arise naturally from the size and complexity of the American state. Writing in the early twentieth century, Max Weber predicted the development of this conflict between an expanding territory and state on the one hand and the populist American desire for an accessible government on the other. “It is obvious,” Weber declared, “that technically the large modern state is absolutely dependent upon a bureaucratic basis. The larger the state, and the more it is a great power, the more unconditionally is this the case.” 62 He foresaw that the United States, which was then “not fully bureaucratized,” would likely become so as the nation faced “greater . . . zones of friction with the outside and . . . more urgent . . . needs for administrative unity at home.” 63 The relatively young nation’s expanding size—both in population and space—would propel the American state from a relatively small, directly accountable democracy toward becoming the administrative state required to perform the functions citizens demand. 64

Thus would the government bureaucracy, a key element of what Weber famously characterized as the antidemocratic, authoritarian, and instrumental rationality of modernity’s “iron cage,” enmesh the United States. 65 Its vastly expanded administrative apparatus, which collects and preserves vast quantities of data in its everyday operation, would take advantage of the informational asymmetry that bureaucracies typically enjoy over the public. 66 “Bureaucratic administration,” Weber wrote, “means fundamentally domination through knowledge”—domination made possible by the bureaucracy’s ability to hoard knowledge and keep its intentions secret. 67 To the extent that a state’s large territory dictates a larger and more powerful administrative apparatus, then, a state the size of the United States, with its necessary bureaucratic rule, would inevitably attempt to protect itself from the public’s view. It would, in sum, make

---

62. See MAX WEBER, ECONOMY AND SOCIETY 971 (Guenther Roth & Claus Wittich eds., 1968).
63. Id.
64. Id. at 949–52 (discussing the limits of direct democracy).
66. See 1 WEBER, supra note 62, at 218–23.
67. Id. at 225; 3 WEBER, supra note 62, at 992.
transparency an impossible goal to attain.

Parts II and III explore these organizational and territorial issues in greater detail and identify the variable, imperfect measures developed in an effort to make the state visible. They assert that the vast territory of the United States, along with its citizenry’s expectations of both an expansive but also accessible and accountable government, have increased the demand for transparency even as they have made it more difficult to meet.

II. EXPOSING THE ORGANIZATIONAL STATE

As a result of its framers’ quite conscious intent, the United States Constitution inaugurated a prototypically modern, complex organization. It sets forth in its articles a range of roles (legislator, executive, administrator, judge) and institutions that would shape the behaviors of those who would assume official positions, simply by virtue of the organizational scheme. Contemporary government agencies, many of them subject to additional organizational mandates by their state constitutions, carry on this tradition. Their official organizational charts graphically represent how they delegate their institutional authority and tasks, again under the assumption that the correct organization and hierarchy will produce the correct official behavior, which will in turn result in the optimal kind and extent of governance. If linked together, all such governmental charts—those of the co-equal branches of the federal government and their agencies, committees, and respective hierarchies, as well as of state governments and their multitudinous municipal governments and administrative agencies—constitute a formal atlas of American government, a great chain of the state’s being.

Such maps seem to inscribe a spatial logic that plots the division of labor and allocates authority within units and positions. As the maps expand and proliferate—down within branches of a particular level of government, and across federal, state, and local levels—they seem to form a never-ending, bewildering series of Leviathans rather than a comprehensible single state. Under a strong conception of transparency that would require a continually


71. See Wolin, supra note 68, at 351–52.
visible state, such complexity constitutes a significant problem. If the state is to be visible and perceptible, it ought to be visible in its entirety as a whole and as constituent parts—from the federal top of the President, the Congress, and the Supreme Court, down to the lowest-level service provider of the local government. To implement transparency’s inherent promise, public access laws must thus attempt to bridge or collapse the vast organizational distance the state creates so that the public, as citizens, subjects, and clients, can know the government that ultimately, and at least theoretically, serves it. Perhaps a Nozickian “night-watchman state” could be so flat and simple that it proves thoroughly and perfectly visible. But even the relatively simple modern government envisioned by the United States Constitution allocates tasks and authorities in a complex system that strains the public’s capability to view and comprehend the state—especially once the regulatory state, nascent from the colonial period through the early twentieth century, began to grow.

Below I consider three distinct legal authorities that either create or reflect this complexity: a constitutional order that imposes only minimal and quite variable openness requirements on the various branches and levels of government; an executive branch whose evolving size and complexity limit Congress’s efforts to impose statutory openness obligations on it; and the blurred lines between the government and the private entities with whom it collaborates and to whom it outsources operations that challenge the reach of open government laws.

A. Constitutional Transparencies

The Constitution’s initial distribution of authority between the federal and state governments and among the federal government’s branches blocks the creation of a uniform, comprehensive approach to public access. Consider the first four Articles in turn. Although the framers engaged in spirited debates about the need for the proposed legislative branch to be open to the public, the Constitution imposes no structural, uniform openness requirement upon Congress. Instead, it requires certain and


73. See infra Part II.A.

limited disclosure practices, and allows only Congress to impose procedural rules upon itself. Notably, when Congress saw fit to place disclosure and other procedural requirements on executive branch agencies through the Freedom of Information Act (FOIA), it imposed no such requirements on itself.

The Constitution makes even fewer openness demands of the executive, requiring only that the President “from time to time give to the Congress Information of the State of the Union,” a minimal command that has resulted in an annual speech that ritualistically offers self-selected information deemed politically important to the President’s agenda and

75. U.S. CONST. art. I, § 5, cl. 3 (requiring Congress to keep and publish “from time to time” a journal of its proceedings and its members’ votes, while also allowing Congress to except “such Parts as may in [its members’] Judgment require Secrecy”; id. § 9, cl. 7 (requiring Congress to publish “a regular Statement and Account of the Receipts and Expenditures of all public Money”; id. § 7, cl. 2 (“[T]he Names of the Persons voting for and against the Bill shall be entered on the Journal of each House respectively.”).


popularity. The only additional transparency requirements made of the presidency and executive branch are those that Congress mandates or that are self-imposed. The most prominent general statutory mandates placed upon executive branch agencies are largely uncontroversial in the abstract. The Freedom of Information Act requires the disclosure by executive branch agencies of certain documents, the Government in the Sunshine Act requires executive branch agencies to hold open meetings, the Federal Advisory Committee Act places open government requirements on certain types of committees created by the executive branch, and the Presidential Records Act requires the President to retain records and make them available to the public after he or she leaves office. Each statute imposes a particular openness requirement on a limited universe of entities, most typically those defined by the respective statutes as agencies and advisory committees.

But as the history of these statutes demonstrates—especially the history of the FOIA—both the extent of their applicability and the specific


80. This was not always the case. President Johnson did not support the original statute, and President Ford vetoed the 1974 amendments to the FOIA that strengthened its disclosure obligations. See Foerstel, supra note 34, at 39–48. Prior to his confirmation as a judge on the D.C. Circuit, Antonin Scalia wrote a blistering critique of the statute in the *American Enterprise Institute’s journal* in 1982. See Antonin Scalia, *The Freedom of Information Act Has No Clothes*, REGULATION, Mar./Apr. 1982, at 14. Today, however, no elected official would propose repealing any of the existing open government laws, and efforts to strengthen them frequently have bipartisan support. See, e.g., Daniel J. Metcalfe, *The Cycle Continues: Congress Amends the FOIA in 2007*, ADMIN. & REG. L. NEWS, Spring 2008, at 11 (noting the bipartisan effort to enact amendments to the FOIA in 2007). In addition, nongovernmental organizations (NGOs) supporting the FOIA are either nonpartisan or range across the political system. Of the NGOs cited supra note 26, Judicial Watch is avowedly conservative, while others are nonpartisan. See Judicial Watch, About Us, http://www.judicialwatch.org/about-us (last visited Aug. 5, 2010).

requirements they impose have proven hotly contested.\textsuperscript{85} As they have
grown more vigorous and coercive, congressional mandates on the
executive branch’s openness have approached constitutional common law
limits on inter-branch interference, most notably through the tangled
doctrine of executive privilege and the more generalized concept that the
President should be free from constraint in seeking advice and counsel from
close advisors.\textsuperscript{86} At the same time, presidential administrations have varied
in their commitment to transparency in general and in their willingness to
interpret these statutes broadly or narrowly,\textsuperscript{87} while agency compliance

\textsuperscript{85} See Foerstel, supra note 34 [history of the FOIA]; Suzanne J. Piotrowski,
Governmental Transparency in the Path of Administrative Reform 21–24 (2007)
history of the FOIA, focusing on discontent with its shortcomings); Peter L. Strauss,
Todd D. Rakoff & Cynthia R. Farina, Gellhorn & Byse’s Administrative Law 762–
66 (rev. 10th ed. 2003) [history of the Government in the Sunshine Act and discussion of
criticisms of its effects on agency deliberations]; Mark Fenster, Designing Transparency: The
9/11 Commission and Institutional Form, 65 Wash. & Lee L. Rev. 1239, 1253–58 (2008) [history of
the FAC, focusing on discontent with its shortcomings and controversies over its
constitutioality]; Jonathan Turley, Presidential Papers and Popular Government: The Convergence
of Constitutional and Property Theory in Claims of Ownership and Control of Presidential Records, 88
Cornell L. Rev. 651, 666–77 (2003) [history and controversies surrounding the
Presidential Records Act of 1978].

\textsuperscript{86} On the current general state of the doctrines of executive privilege, state secrets,
and presidential prerogatives over information bearing on national security and foreign
affairs, see Pozen, supra note 9, at 321–22. On the constitutional issues surrounding FACA’s
limitations on the President’s ability to seek advice, see Fenster, supra note 85, at 1254–56.

\textsuperscript{87} The Attorney General typically issues a memorandum to the federal branch
agencies declaring its interpretation of the FOIA and how the Department of Justice plans to
litigate contested cases. They tend to vary with each change of party control of the White
House—with a Democratic president, the memo tends to favor disclosure, and with a
Republican president, it tends to favor nondisclosure. Compare Memorandum from Eric
Holder, Attorney General, on the Freedom of Information Act to the Heads of Executive
Departments and Agencies (Mar. 19, 2009), http://www.justice.gov/ag/foia-mem-
march2009.pdf [withdrawing memorandum from Attorney General Ashcroft and
announcing “a clear presumption: In the face of doubt, openness prevails.”] (quoting
Memorandum from President Barack Obama on the Freedom of Information Act to the
Heads of Executive Departments and Agencies (Jan. 21, 2009),
http://www.whitehouse.gov/the-press-office/freedom-information-act], and Memorandum
from Janet Reno, Attorney General, on the Freedom of Information Act to the Heads of
Vol_XIV_3/page3.htm (“The Department [of Justice] will no longer defend an agency’s
withholding of information merely because there is a ‘substantial legal basis’ for doing so.
Rather, in determining whether or not to defend a nondisclosure decision, we will apply a
presumption of disclosure.”), with Memorandum from John Ashcroft, Attorney General, on
the Freedom of Information Act to the Heads of all Federal Departments and Agencies
requests and decide to withhold records, in whole or in part, you can be assured that the
Department of Justice will defend your decisions unless they lack a sound legal basis . . .”).
with the FOIA mandates varies considerably. Significantly, the Constitution’s lack of any general openness requirement permits such variance among administrations.

Some constitutional doctrines force a degree of openness on the federal and state judiciary. The Sixth Amendment rights to “a speedy and public trial, by an impartial jury of the State and district wherein the crime shall have been committed” require that at least a proportion of the work performed by courts must be public and include a degree of public participation, while the First Amendment also requires public access to criminal trials. But there is no constitutional requirement for open judicial deliberation and conferences, and the tradition of published judicial opinions is just that—a tradition, rather than a constitutional requirement. Some federal district courts and circuit courts of appeal allow cameras in the courtroom, as do some state courts, but no federal constitutional requirement or right binds courts, and no systematic approach prevails. At the same time, modern criminal and civil procedural rules place significant emphases on pretrial procedures and alternative dispute resolutions that undercut the relatively simple and abstract constitutional provisions regarding an open judicial process.

89. U.S. CONST. amend. VI.
93. See Stephanos Bibas, Transparency and Participation in Criminal Procedure, 81 N.Y.U. L. REV. 911 (2006) (decrying lack of public access to discretionary governmental decisions in the criminal process, especially in the plea bargain process); Kenneth Feinberg, Transparency and Civil Justice: The Internal and External Value of Sunlight, 58 DEPAUL L. REV. 473 (2009) (former Special Master of the September 11th Victim Compensation Fund of 2001 discussing the incomplete progress of and prospects for greater transparency in civil litigation); Hamilton & Kohnen, supra note 91, at 293–97 (noting the existence of general rules of judicial and court access, as well as the various exceptions and limiting principles to
Constitution’s lack of a general, expansive right or requirement for judicial transparency allows federal and state courts significant leeway in opening or closing their operations to public view.

Because the United States Constitution fails both to command states to be transparent and to provide individual rights that would allow individuals to impose administrative openness, individual state constitutions and governments have been free to devise their own open government mandates. Shaped by idiosyncratic institutional designs, states take relatively diverse approaches that mix statutory and constitutional requirements and impose different degrees of openness. Transparency advocates frequently express frustration at the variability and relative rigor of state laws. A 1993 survey, for example, found wide variation in the form and substance of state open meeting laws. A 2007 report issued by two nongovernmental organizations used a variety of criteria to evaluate state constitutional and statutory provisions and declared that thirty-eight states had failing laws. Compounding the problem, state officials and judges exhibit varying degrees of commitment to and compliance with their respective open government laws; nongovernmental organizations and media groups in many states that have performed audits of state and municipal government agencies’ response to open record requests variably decry and hesitantly applaud agencies’ performances.


federalist system in an area unregulated by federal constitutional rights and commands thus results in a wide-ranging degree of transparency across states and municipalities.

Rather than imposing transparency’s ideal of a constantly and thoroughly visible state, the constitutional scheme sets forth some limited, variable transparency requirements to individual federal branches, while it restrains the ability of any branch to impose further requirements on another.\(^\text{98}\) The Constitution created a decentralized complex of government institutions without a uniform standard or set of commands that would make the state as a whole and in its parts fully visible to its public. It also leaves to individual states the authority to establish their own governmental structure and administrative norms (within constitutional constraints) and limits the federal government from interfering with state governance. The idiosyncratic nature of each branch and level—its different tasks, its distinct history, and the conditions under which each of its bureaucracies works—renders an organizational map that resists transparency as an abstract and absolute norm, especially as each branch and level expands to engage more complex and demanding tasks. The Constitution’s organizational plan, then, not only fails to create a transparent state—it affirmatively stands in the way of creating one.

**B. Statutory Transparencies**

Like the Constitution, congressional efforts to impose openness obligations on the executive branch have also failed to establish a general, uniform legal norm, again in part because of the complex organization of government institutions. Congress’s intent in enacting the FOIA, the most prominent of Congress’s open government enactments, as well as language within the statute itself suggested that it would sweep broadly across the federal government.\(^\text{99}\) Those entities subject to its mandates are required to make certain information available as a matter of course,\(^\text{100}\) and must also

\(^{98}\) *Cf.* Samaha, *supra* note 5, at 948–49 (describing the constitutional regime for public access to information as “Unsatisfying”).


respond to public requests for documents not subject to those requirements.\textsuperscript{101}

The FOIA does not, however, apply uniformly across the federal government, as it explicitly does not apply to the judiciary or to Congress itself.\textsuperscript{102} Indeed, it does not even apply to all entities within the executive branch. It only affirmatively applies to “[e]ach agency,”\textsuperscript{103} a term that the FOIA defines in an enumerated list.\textsuperscript{104} Congress has granted certain agencies, most notably the CIA, broad exemptions from disclosure.\textsuperscript{105} The Supreme Court has held that the FOIA’s legislative history makes clear that Congress intended to exclude the Office of the President, the President’s immediate personal staff, and units in the Executive Office of the President whose sole function is to advise and assist the President.\textsuperscript{106} It remains unclear how broadly that exception sweeps. The Court has yet to provide an authoritative interpretation of it,\textsuperscript{107} while lower federal courts have developed an indeterminate multifactor test to ascertain whether the FOIA

\textsuperscript{101} Id. § 552(a)(3). Some documents are exempted based either on their content, their status as inter- or intra-office memoranda, or specific exemptions created by other statutes. Id. § 552(b)(1)-(9).

\textsuperscript{102} Id. § 551(1)(A), (B) (exempting the Congress and federal courts from the definition of “agency”).

\textsuperscript{103} Id. § 552(a).

\textsuperscript{104} Id. § 552(f)(1) (defining “agency” as “any executive department, military department, Government corporation, Government controlled corporation, or other establishment in the executive branch of the Government (including the Executive Office of the President), or any independent regulatory agency”).

\textsuperscript{105} See, e.g., 50 U.S.C. § 403-1(b)(1) (2006) (directing the CIA to “protect intelligence sources and methods from unauthorized disclosure”); id. § 403g (exempting the CIA from any law requiring “disclosure of the organization, functions, names official titles, salaries, or numbers of personnel employed by the Agency”). See generally CIA v. Sims, 471 U.S. 159, 167–68 (1985) (applying statutory exemption to CIA). The third exemption of the FOIA, 5 U.S.C. § 552(b)(3) (2006), provides that the FOIA does not apply to matters that are “specifically exempted from disclosure by statute,” so long as the statute meets certain requirements.


\textsuperscript{107} The Court considered this issue briefly in Kissinger, but did no more than resolve the issue that Kissinger was acting in his capacity as National Security Adviser when the documents in controversy were created, and therefore, the documents were not considered the records of an agency under the FOIA. See 445 U.S. at 156. The Court made no effort to develop a test for lower courts to apply in more difficult cases.
applies to nontraditional and advisory entities that the President or executive branch agencies created within the Executive Office of the President. The factors include whether the entity exercises “substantial independent authority” and has been granted sufficiently broad delegated power such that it has “less continuing interaction with the President,” whether the entity’s “sole function [is] to advise and assist the President,” and it is “close operationally” to the President, and “whether it has a self-contained structure.” The more independent the entity seems, the more likely a court will deem it an agency and subject it to the FOIA’s disclosure regime; while the closer the entity is to the President, the less likely the FOIA will apply.

This standard leads to seemingly random results. Among the entities found to be agencies under the FOIA that were sufficiently removed from the President and that possessed sufficient independent authority are the Office of Science and Technology (1971), the Office of Management and Budget (1978), and the Council on Environmental Quality (1980). Among those found not to be agencies because they are too close to the President, have insufficient independent authority, or both, are the Council of Economic Advisers (1985), White House Counsel (1990), the

108. See Citizens for Responsibility & Ethics in Wash. v. Office of Admin., 566 F.3d 219, 222–23 (D.C. Cir. 2009) (reiterating the series of tests). On the complexity of the Executive Office of the President (EOP) and the fact that presidential decisionmaking exempt from the FOIA is in fact decisions made by executive branch bureaucrats, not by the President him- or herself, see Peter L. Strauss, Overseer, or “The Decider”? The President in Administrative Law, 75 GEO. WASH. L. REV. 696, 753 (2007).


110. If so, then it is an “agency” subject to the FOIA. Meyer v. Bush, 981 F.2d 1288, 1293–94 (D.C. Cir. 1993).

111. Soucie, 448 F.2d at 1075.

112. Meyer, 981 F.2d at 1293.

113. Id. at 1295.

114. In this way, the FOIA’s definition of agency implicitly recognizes constitutional limits on Congress’s authority to regulate the presidential deliberative process, which also turns in part on the relative position of the advisor. As the D.C. Circuit has held, communications made between presidential advisers, but not directly to the President, can be protected under the privilege for presidential communications only if the advisors are not too “remote and removed from the President,” and at minimum must be within the staff of a White House adviser rather than an executive branch agency. See In re Sealed Case, 121 F.3d 729, 751–52 (D.C. Cir. 1997).

115. Soucie, 448 F.2d at 1078–79.


President’s Task Force on Regulatory Relief (1993), the Executive Residence of the White House (1995), the National Security Council (1996), the Smithsonian Institution (1997), and the Office of Administration within the Executive Office of the President (2009).

*Meyer v. Bush*, a 2–1 decision in one of the D.C. Circuit’s most influential efforts to parse the definition of *agency*, illustrates this confusing indeterminacy. The issue before the court was whether the FOIA applied to the President’s Task Force on Regulatory Relief, a cabinet-level entity created by President Ronald Reagan to lead his administration’s efforts to reduce federal regulation. For Judge Lawrence Silberman, joined by fellow Reagan appointee Judge David Sentelle in the majority, the Task Force served as an advisory body that offered nothing more than guidance to the Office of Management and Budget (OMB) regarding regulatory rules and programs. It “was positioned between the OMB” and the President, and thus “only a hair’s breadth from the President,” and its members, many of whom (including the Vice President) were also agency heads or cabinet members in their own right, were also “the functional equivalents of assistants to the President.” Therefore, it was not an agency under the FOIA. For Judge Patricia Wald, a Carter appointee writing in dissent, the Task Force was a “separate functional establishment within the Executive Office of the President to which the President delegated some of his executive powers,” and therefore a powerful cohesive unit with direct supervisory control over agencies below it in the hierarchical chain of executive branch authority.

Both arguments seem plausible under the D.C. Circuit’s test, and no essential, consistent logic emerges from the test’s application in *Meyer v. Bush* or in the related case law. The executive branch has proven too amorphous and confusing for a thorough and uniform legislative transparency regime. When the President or Congress creates a new entity within the executive branch that does not clearly constitute an agency, we will not know its obligations under the FOIA without an extensive, fact-specific survey on the messy organizational map of the federal government.

126. Id. at 1294.
127. Id. at 1298, 1307, 1313 (Wald, J., dissenting).
unless Congress clearly exempts it from or clearly subjects it to the FOIA in its organic statute.\textsuperscript{126} Indeed, presidential administrations create such entities regularly, especially to oversee or advise politically significant and controversial programs. Examples include the taskforce created to oversee deregulatory efforts during the Reagan administration, as seen in \textit{Meyer v. Bush};\textsuperscript{129} the Task Force on National Health Care Reform on health care reform during the Clinton administration, headed by first lady Hillary Clinton, to which the Federal Advisory Committee Act was held not to apply;\textsuperscript{130} and the National Energy Policy Development Group in the George W. Bush administration, headed by Vice President Cheney, to which the Federal Advisory Committee Act was also held not to apply.\textsuperscript{131} These entities played key roles in devising and implementing policy for the presidents who created them, and their creators designed and placed them within the executive branch in a way that limits public access to their proceedings and records.

\textbf{C. Private Transparencies}

The American state has long used private entities to perform seemingly public functions,\textsuperscript{132} and it has long delegated to or worked closely with private actors when it has engaged in law- and regulation-making.\textsuperscript{133} Indeed, these relationships are so longstanding and embedded in public governance that no clear boundary separates the state from the private entities with which it works to regulate and deliver services.\textsuperscript{134} Should seemingly public information produced or possessed by private entities be made public? Similarly, should information produced by or concerning

\begin{footnotesize}
\begin{enumerate}
\item\textsuperscript{126} Congress occasionally exempts new, innovative agency-like entities from FOIA obligations. See, e.g., 15 U.S.C. § 7215(b)(5)(A) (2006) (exempting the Public Company Accounting Oversight Board, created as part of the Sarbanes–Oxley Act, from the FOIA).
\item\textsuperscript{129} See supra text accompanying notes 125–27.
\item\textsuperscript{130} See Ass’n of Am. Physicians & Surgeons, Inc. v. Clinton, 997 F.2d 898, 916 (D.C. Cir. 1993).
\item\textsuperscript{131} In re Cheney, 406 F.3d 723, 731 (D.C. Cir. 2005) (en banc). See also Judicial Watch, Inc. v. Dep’t of Energy, 412 F.3d 125, 131–32 (D.C. Cir. 2005) (holding that employees of the Department of Energy, whose work for that agency would be subject to the FOIA, produced work that was not “agency records” subject to the FOIA when they were detailed to the National Energy Policy Development Group, which was not subject to the FOIA).
\item\textsuperscript{133} See Louis L. Jaffe, \textit{Law Making by Private Groups}, 51 HARV. L. REV. 201 (1937).
\end{enumerate}
\end{footnotesize}
private entities with which the state collaborates or transacts be made public? Under the populist understanding of transparency, the private information that the government possesses or could or should possess, or that private actors produce or disclose while participating in or negotiating with government, becomes public information and therefore should be made available to the public. If the state is to be visible, then all of its parts, including private individuals and entities that actively interact with or serve as adjuncts to the state, should be visible.

This proposition has not, however, prevailed. Consider first the dynamics at play over the disclosure of information the government gathers about private individuals and entities through its lawmaking, rulemaking, and law enforcement activities. Federal law requires the federal government to protect the privacy of private individuals from and about whom it collects information in some contexts, while the FOIA excepts from disclosure the privileged or confidential commercial data the government collects in order to encourage those it regulates to continue to share information. Federal law also protects some information submitted by owners and operators of “critical infrastructure” from disclosure on the grounds that the release of such information might threaten national security. The state’s intimate and ongoing relationship with individuals and those it regulates limits the extent to which current law allows it to

135. The Privacy Act prohibits disclosure of routine personal information except to the person to whom the record pertains, or with that person’s permission. 5 U.S.C. § 552a(b), (d) (2006). In addition, the FOIA’s exemptions include privacy protection. See id. § 552(b)(6) (exempting files on individuals for which disclosure “would constitute a clearly unwarranted invasion of personal privacy”); id. § 552(b)(7)(C) (exempting records or information compiled for law enforcement only to the extent that their disclosure “could reasonably be expected to constitute an unwarranted invasion of personal privacy”).

136. Id. § 552(b)(4). Indeed, corporations engage in extensive “reverse-FOIA” litigation in order to preempt efforts by their competitors to use FOIA requests to obtain their trade secrets and other valuable information. See, e.g., Chrysler Corp. v. Brown, 441 U.S. 281, 293–94, 317–18 (1979) (approving of reverse-FOIA litigation by finding a private right of action under the APA to seek injunctive relief prohibiting the disclosure of information submitted to the government that plaintiffs claim to be commercially sensitive). For agencies, the reverse-FOIA process has proven costly, as regulated corporate entities use litigation to secure their information from competitors. See David C. Vladeck, Information Access—Surveying the Current Legal Landscape of Federal Right-to-Know Laws, 86 Tex. L. Rev. 1787, 1817 n.197 (2008). Nevertheless, for industry representatives, the reverse-FOIA process proves relatively indeterminate and not a guarantee to protect against disclosure. See James W. Conrad, Protecting Private Security-Related Information from Disclosure by Government Agencies, 57 Admin. L. Rev. 715, 729–32 (2005).

disclose private information that it controls. Whether as a matter of personal privacy, corporate function and commercial property, or national security, private entities are not treated simply as part of the state, even if the state collects information about them or uses them to perform important state or state-like functions.

At times the government does more than merely collect information about private entities and individuals; it also collaborates or negotiates with them in regulatory programs sometimes referred to as “new” or “new public” governance.138 Departing from a traditional top-down command-and-control approach, in which an identifiable state agency requires an identifiable private entity to comply with mandatory practices or regulatory targets or face punishment, a state entity adopting a new governance approach to achieve a particular outcome works closely with private actors to develop and implement a program or programs that can best achieve its goal. The federal Negotiated Rulemaking Act139 has created the most formalized and congressionally authorized model for new governance processes, allowing an administrative agency to negotiate openly with regulated entities and interested parties through a chartered committee that observes the openness requirements of the Federal Advisory Committee Act.140 Additional “tools” developed by new governance advocates offer a much wider spectrum of public–private coordination than the formal negotiation process, including some that are significantly less formalized.141

The blurring of government authority in new governance efforts raises significant concerns about a resulting program’s accountability and visibility to the public.142 Government delegation of some degree of regulatory authority to private or hybrid public–private entities may increase the state’s organizational complexity and may thereby decrease the state’s visibility to the public. Some degree of privacy may be essential to

140. See id. § 564(a) (requiring notice of regulatory negotiations in the Federal Register); id. § 565(a) (requiring formal chartering of committees); id. § 566(d), (g) (requiring that committees keep meeting minutes and records consistent with the FACA).
the process, however. If private entities that collaborate with the government would thereby become subject to open government laws, they may be less willing to engage directly with the government. Their reluctance would in turn undermine the collaborative approach that new governance seeks to promote. At the same time, to the extent that current law limits the FOIA’s applicability to new governance efforts, then the new governance approach appears significantly less than perfectly transparent.

Proponents argue, however, that collaborative governance offers a more “dynamic accountability” than conventional top-down regulatory programs; new governance, they argue, imposes measures like peer review and reporting requirements that provide as much if not more government oversight than traditional public governance. In addition, some new governance programs themselves enhance information disclosure, targeting particular kinds of data whose release to the public can inform individuals and positively shape their behavior. Thus, proponents argue, new governance results in better, more effective regulation, although perhaps it allows less openness according to traditional conceptions of public disclosure and transparency. Again, a vigorous populist approach to transparency would protest against those aspects of new governance programs that offer less than full disclosure—protests that, if made into law, might conflict with and undermine whatever gains this less traditional form


144. See ARCHON FUNG, MARY GRAHAM & DAVID WEIL, FULL DISCLOSURE: THE PERILS AND PROMISE OF TRANSPARENCY 1–7 (2007); Janet A. Weiss, Public Information, in TOOLS OF GOVERNMENT, supra note 141, at 227–33.

145. See Orly Lobel, The Renew Deal: The Fall of Regulation and the Rise of Governance in Contemporary Legal Thought, 89 MINN. L. REV. 542, 455–57 (2004) (arguing that transparency and increased access to information do not themselves improve regulation, and that the state may need to be less than perfectly transparent in order to develop more effective regulatory programs). In a volume of essays intended to serve as a guide to new governance, the only essay that mentions and seems to embrace open-ended public transparency appears as the twentieth of twenty-three chapters and includes the topic as one among many “policy tools” that further democratic ends. Steven Rathgeb Smith & Helen Ingram, Policy Tools and Democracy, in TOOLS OF GOVERNMENT, supra note 141, at 565, 579. Furthermore, the same volume’s introduction concedes that for new governance to succeed in producing a more effective regulatory state, “classical notions of democratic accountability may need to be loosened and more pluralistic conceptions developed,” while the introduction fails to include transparency as one of its criteria for evaluating particular new governance tools. Salamon, supra note 141, at 23–24, 38.
of governance offers.

The state frequently does more than collaborate with private entities—it often and explicitly contracts out or privatizes government services. This longstanding tradition of American governance offers, so its proponents say, a more efficient and effective means to deliver services that the government has performed in the past or can perform. In an especially poignant example, the federal government has begun outsourcing to private firms not only the digital storage of its information, but also its handling of FOIA requests, for the stated reason that private information management companies can provide better, more reliable, and less expensive service in these activities than the federal civil service. Proponents argue that outsourcing not only improves government services, but it makes the resulting smaller government leaner, more efficient and flexible, and more responsive—a type of reform that enjoys bipartisan support. For transparency proponents and critics of privatization alike, the public’s need to view the state’s operations does not disappear merely by virtue of a contractual agreement with a private entity. When the law extends open government obligations to private entities, however, it threatens to undercut the instrumental and political advantages of privatization and new forms of governance. Unsurprisingly, given this

146. See Gillian E. Metzger, Privatization as Delegation, 103 COLUM. L. REV. 1367, 1369–71 (2003). The academic literature on the privatization of public services is vast; a useful citation to it is in Sagers, supra note 134, at 43–48 & nn.14–38.


150. See David G. Frederickson & H. George Frederickson, Measuring the Performance of the Hollow State 21 (2006); Light, supra note 134, at 6. This claim is widely contested, particularly in terms of the actual size of government and the limits placed on government control and management of contractors’ work. See Frederickson & Frederickson, supra, at 20–21; Light, supra note 134, at 176–79.


conflict, federal and state laws have taken halting, uncertain steps to impose transparency norms on private entities with whom the state is contracting or governing.\textsuperscript{153} The issue pervades all national governments with freedom of information laws, and as Alasdair Roberts has explained, it has caused a “conceptual muddle” regarding how “to determine where the boundaries of government lie” and how best to draft rules that can force disclosure upon private entities that “appear governmental.”\textsuperscript{154}

These conflicts between the gains of public–private collaboration, and the limits such collaboration place on the state’s visibility, illustrate the inevitable and pervasive barriers to making the government thoroughly transparent. In order to meet the public’s expectations for the range and quality of services it must perform, the state must work with private entities; but that work may as a result either make the state less transparent, or may provoke an effort to treat private entities as state actors that will in turn undercut the range and quality of services the state can offer. If the state must be visible, its efforts to provide effective regulation and services are likely to suffer, at least to some extent.

\textbf{D. The Impossibility of Organizational Exposure}

In all of its various complexities, the contemporary state organization of the United States poses great challenges to any effort to impose visibility. The complexities are both endogenous—reflecting historical, path-dependent decisions about institutional design made at the nation’s founding and throughout its history—and exogenous—the result of governmental adaptations to social and economic development in civil


society. Transparency cannot simply be imposed on such a massive network of institutions and individuals; legal, regulatory, and normative projects to make the state more visible must grapple with design, implementation, and enforcement issues across a broad, diverse range of levels, branches, and webs.

III. EXPOSING THE PHYSICAL STATE

Two of the state’s most basic physical characteristics impede its visibility to the public. The first is the state’s territorial size and political-geographic complexity. For the state to be thoroughly transparent and reduce or collapse its distance from the public, its operations and personnel must be identifiable and made available for public inspection, no matter their location. The immense size and intricate overlap of government entities in the United States frustrate any effort to achieve such perfect or even near-perfect visibility. The second impediment is architectural. The thoroughly transparent state must be capable of allowing the public to view where and how government employees work: the physical spaces of the built bureaucratic environment. Government buildings have standard architectural elements—walls, ceilings, doors, and windows—that serve naturally to exclude the public and obscure the state. Even if it were physically possible either to enable the public to see through the structures that house the state or to invite the public into these structures at all times, the effort can prove so intrusive and costly as to make the work of public officials difficult if not impossible. The first two sections of this Part offer a more detailed account of these geographic and architectural issues, while the third section describes two instances in which access to information laws confront, and ultimately fail to respond coherently to, the state’s spatial and physical complexity.

A. Distance

The federal government is sovereign over a significant amount of well-populated territory. Its three branches may all have their headquarters in Washington, but their decisions and administration also occur in agency and congressional offices and federal courthouses scattered throughout Washington as well as the fifty states. The federal government shares sovereignty over the same territory with state governments, and both the federal and state governments overlap municipal governments. Many state and local governments preside over extraordinary amounts of territory from their capitols and city halls—heavily populated Los Angeles County,
for example, occupies more than 4,000 square miles of land, while sparsely populated Alaska sits on over 570,000 square miles. Enabling the public to view such diffuse Leviathans proves a difficult challenge, as does enforcing any general edict for openness upon officials in geographically scattered organizations.

Both the Hamiltonian faith in administrative and structural means to manage government and enable democracy across vast distances, and the Weberian warning that such solutions would lead to an imperfect modern state ruled by an information-hoarding bureaucracy, foreshadowed ongoing arguments and anxieties about the state’s operations in an expansive American territory. Weber correctly predicted the expansion of the American administrative state, while Hamilton anticipated systematic efforts to control it, efforts that began almost immediately in the federalist period of the early Republic. A larger and more diverse nation than even Hamilton’s Anti-Federalist opponents feared, coupled with an administrative apparatus that Weber foresaw but that Hamilton could not have anticipated, has made Hamilton’s confident forecast of private collective actions to control the administrative state appear naive at best. His general prescription for public and private institutional checks and balances, however, survives in the federal and state laws that attempt to provide uniform controls over vast and far-flung bureaucracies. In the present day, federal and state administrative laws impose standard procedural rules, including requirements for public access to information, equally on the operations of agencies’ headquarters and its offices. At the same time, federal courthouses, enforcing federal law and using uniform federal rules of civil and criminal procedure and providing equal levels of openness, were dispersed across the nation in the twentieth century in an

156. See supra text accompanying notes 58–61.
159. Nevertheless, the role of the press in checking government misdeeds—one of the roles that Hamilton hoped it would play—has remained the strongest justification for First Amendment protections against prior restraint. See Near v. Minnesota ex rel. Olson, 283 U.S. 697, 719–20 (1931) (asserting that press liberties are necessary as a means to protect against corrupt officials who take advantage of the increasingly complex administration of government).
effort to extend both federal authority and federal rights. At least as a
formal matter, then, the American state appears to have proven Hamilton
correct by successfully addressing the territorial concerns of the Anti-
Federalists.

As a matter of practice, however, these formal commands are not self-
enforcing. Central authorities have limited control over their dispersed
organizations, and not all branches and agencies of the individual units are
equally visible to their citizens. Even assuming that those at the center of
authority want their inferior officers to be visible to the public—a desire
that appears to vary among agencies and executive administrations, given
the variability of their levels of compliance—the periphery can resist
central commands, as Michael Lipsky observed in his study of “street-level
bureaucrats” and the “relative autonomy from organizational authority”
enjoyed by front-line government officials. Police officers on the street
and teachers in the classroom, as well as public information officials and
FOIA officers removed from an agency’s central command, inevitably have
significant discretion to make substantive and administrative decisions both
as a means of responding to the particular context in which they find
themselves and because they cannot in fact be controlled. Physical
distance, whether counted in miles, in feet, or by the floors of an office
building, limits the extent to which superiors can monitor and exercise
authority. If administrative discretion increases across space, and Weber’s
assertion that bureaucracies prefer to hoard information is correct, then
efforts to impose transparency on large, far-flung agencies will be doomed
to failure—or at least to incomplete success. The geographical dispersal
of authority thus limits both the state’s ability to supply bureaucracies that the
public can see and the law’s ability to command them to be seen.

The government’s size and dispersal across the territory it governs is one
obstacle to achieving a populist ideal of transparency; the state’s
jurisdictional complexity is an additional one that can hinder the public’s
view of the state. “By its very nature,” the political geographer John Short
has written, “the nation-state is a spatial phenomenon,” one that manifests

161. See supra text accompanying note 97.
163. James Q. Wilson, Bureaucracy 327–29 (1989). For a recent reconsideration of
Lipsky’s concept of the street-level bureaucrat, see Simon Halliday, Nicola Burns, Neil
Hutton, Ferguson McNeill & Cyrus Tata, Street-Level Bureaucracy, Interprofessional Relations, and
Coping Mechanisms: A Study of Criminal Justice Social Workers in the Sentencing Process, 31 Law &
Policy 405 (2009).
itself most clearly in the frontiers and borders between nations and in a nation’s internal division into such administrative subdivisions as regional, state, and local governments and their sub-agencies. This might suggest that a geographical map, which visualizes a series of logical—if somewhat haphazardly arranged—nested centers and peripheries, would provide a blueprint for political order and behavior. Like an organizational chart that claims to offer a hierarchical rendering of coordinated government entities, a map of the United States implies that political power is dispersed across a territory: the nation, with its federal capitol; the states, with their state capitals; and metropolitan regions, with their city halls, urban cores, and suburban and exurban peripheries. Where authority is dispersed logically, the public can view, comprehend, and hold accountable those officials it can find in the cores of the respective (federal, state, and local) jurisdictional bodies.

As Richard Thompson Ford has noted regarding local governments, however, we cannot assume that territory and the maps that record it accurately reflect an essential, authoritative sovereign power, nor can we assume that a hierarchical relationship among political divisions subordinates the smallest and lowest subunit. A governmental unit’s authority, jurisdictional reach, and public accessibility are never as fixed or stable as a map suggests. Federal, state, and local authorities whose territorial jurisdictions overlap any particular location frequently confuse the public. How can the public see a state when they cannot discern


165. On maps’ representational ideal, see Michael R. Curty, Shelf Length Zero: The Disappearance of the Geographical Text, in Space and Social Theory 88, 90 (Georges Benko & Ulf Strohmayer eds., 1997). As representations, maps are not natural but are instead the result of efforts to produce a visual representation of the social world. See Henri Lefebvre, The Production of Space 84–85 (Donald Nicholson-Smith trans., 1991).


which government entities are sovereign over a particular piece of land? Two examples: At the local level, especially in major metropolitan service areas, city and county governments frequently overlap or have shifting boundaries, requiring regional or crossjurisdictional coordination and governance and making regulatory responsibility difficult to pinpoint.\(^{170}\)

Secondly, lakes and rivers often traverse state boundaries and are overseen (or, sometimes, are not overseen and are therefore the site of significant conflict) by complex regional agreements or government authorities.\(^{171}\) These liminal spaces—parts not of one but of numerous jurisdictions, with no clear or obvious boundaries—render efforts both to govern and to view governance difficult if not impossible. Moreover, the modern state’s sovereignty has long extended beyond its mere territory and been shaped and challenged internally not only by its citizens but by other states, nongovernmental organizations, transnational corporations, supranational institutions, and the global flows of economic trade and capital.\(^{172}\) To the extent that different levels of government might cooperate with, ignore, or contest each other’s jurisdiction and policies, the public will struggle to identify the particular government entity or entities from which they need to seek information.\(^{173}\)
B. Enclosure

Government buildings and offices enable public employees to perform their tasks by housing the spaces where officials, managers, and civil servants work, converse, officially meet, and store and protect official records. By containing state activity within built structures, buildings and offices also enclose that activity within walls and ceilings, and control access and visibility to it via doors and windows. As a result of making it possible for officials to work and to sort and protect the records that they collect and produce, government buildings inevitably separate officials from the public that they serve. Accordingly, allowing the public to view and enter government buildings is at once an issue of design and practice: can the public see and navigate its way into the building, and is the public in fact invited or allowed to enter?174 The competing concerns of design and public policy help determine the extent of public access to officials and to government information.

Public architecture aspires to more than the simple, utilitarian goal of housing offices and allowing or limiting public access, however. It also attempts to shape the affective relationship between the state and its public.175 It works iconically and symbolically to establish an identity for the national, state, or municipal governmental unit or units that a building hosts.176 A public building’s size, architectural design, and location


175. Two additional, secondary purposes that government buildings attempt to further, which this Article does not discuss, are the broader sense of community and social capital they can create in dense urban locations, see JANE JACOBS, THE DEATH AND LIFE OF GREAT AMERICAN CITIES 179–86 (1961), and their ability to help deter criminal activity, see Neal Kumar Katyal, Architecture as Crime Control, 111 YALE L.J. 1039 (2002). Neither purpose directly furthers public access, while efforts to achieve them may in fact limit the state’s visibility.

176. See CHARLES T. GOODSELL, THE AMERICAN STATEHOUSE: INTERPRETING
announce the state’s existence and indicate its occupant or occupants’ relative prominence. In doing so it may invite the polis to enter or intimidate them and discourage their entry. The interior design and features of public buildings can also communicate openness or its opposite as they either foster or inhibit interaction among government actors and between the state and the public. Public architectural design may consider the public visibility of and access to government officials’ work as a significant end, but it may not.

Transparency laws must therefore attempt to address and mitigate the physical obstructions that walls and ceilings place before the public’s ability to view and access state operations. They can succeed, at least to an extent. The Americans with Disabilities Act (ADA), which requires that public buildings, and public accommodations generally, be “readily accessible to and usable by individuals with disabilities,” is one notable example. The ADA has significantly improved access to government offices and officials.


for a population that previously faced barriers to enter public buildings.

Open government laws attempt to mitigate the enclosure problem for the entire public in two primary ways: under the aegis of so-called “open meeting” or “sunshine” laws, government officials are required to make certain meetings accessible for public viewing, while open records laws (including the FOIA and its state analogues) require that agencies open their files to members of the public. In addition, video recordings and broadcast of government meetings via C-SPAN and state and local cable television and webcasting channels make otherwise public meetings more widely available. None of these efforts provide unlimited physical or visual access to all public buildings and offices at all times, however. As the next section explains, such transparency that they do provide is limited, either as a legal or practical matter, to certain preplanned public events or to files over which the government has initial control. The physical enclosure that walls and ceilings provide almost inevitably offer cover for the state from the public’s gaze, and transparency obligations cannot fully overcome or compensate for enclosure’s distance.

C. The Impossibility of Containment

The state’s geography and built environment thus pose significant barriers to government visibility and accessibility. Unsurprisingly, legislatures and courts struggle with these issues, and it proves difficult to shine light on the government and to free its information, especially when officials and documents refuse to stand still across the state’s vast territory and public employees work within their offices or other interior spaces where their actions cannot so easily be viewed. As ever, the law can handle easy cases—most government documents are in fact housed in government offices and can be requested and found, while official meetings regularly occur in official meeting halls with public access. But more difficult cases—private documents that are held in government offices or government documents that are held in private spaces, or public officials’ interactions in private locations outside formal meeting halls and government offices—test the limits of open government laws and challenge efforts to force compliance with the symbolic dimensions of transparency.

1. Containing Meetings

Although the constitutional framers met behind closed doors, in chambers (presumably) limited in sunlight though surely not infected, federal and state legislatures have long allowed the public to view their
formal meetings, whether by constitution or custom. These comprehensive open meeting laws emerged in the states largely during the post-World War II period, and especially in response to revelations of the Nixon Administration’s abuses of power (when Congress enacted the Government in the Sunshine Act). These laws extended the openness obligation to administrative bodies and local governments. Current statutory and state constitutional laws, frequently named “sunshine” laws (like the federal Act), require such meetings to be open and accessible to the public, thus echoing transparency’s broader emphasis on visibility and presence, as do the court decisions interpreting them.

The public is not invited to view everything the government does, however. By definition, the only event that these laws make thoroughly visible is the official occasion of a “meeting,” a term whose meaning is


185. See, e.g., Regents of Univ. of Cal. v. Superior Court, 976 P.2d 808, 826 (Cal. 1999) (Brown, J., concurring) (“There is rarely any purpose to a nonpublic premeeting conference except to conduct some part of the decisional process behind closed doors.”); Town of Palm Beach v. Gradison, 296 So. 2d 473, 477 (Fla. 1974) (declaring that Florida’s “government in the sunshine law” barred instances when a city engages in its “decisional process behind closed doors”); Atlanta Journal v. Hill, 359 S.E.2d 913, 914 (Ga. 1987) (describing Georgia’s Open Meetings Act as intended “to protect the public—both individuals and the public generally—from ‘closed door’ politics”); Okla. Ass’n of Mun. At’ys v. State, 577 P.2d 1310, 1313–14 (Okla. 1978) (“If an informed citizenry is to meaningfully participate in government or at least understand why government acts affecting their daily lives are taken, the process of decision making as well as the end results must be conducted in full view of the governed.”).

186. See, e.g., 5 U.S.C. § 552b(a)(2) (defining meeting as “the deliberations of at least the number of individual agency members required to take action on behalf of the agency where such deliberations determine or result in the joint conduct or disposition of official agency business”).
not self-evident. 187 How far along in a decisionmaking body’s consideration of a matter does a gathering of its members constitute an official meeting? 188 Does an open meeting mandate apply only to the formal conferences that officials hold in an agency’s official meeting space, or does the definition of meeting extend outside the official enclosure, to other rooms in government buildings, or even to gatherings and encounters held in restaurants and homes? And if the latter, more capacious definition applies, can officials be required to provide notice and public access to informal meetings that occur by chance or appointment—in which case, such meetings cannot as a practical matter take place within the ambit of the law? Do the government’s operations and transparency’s reach extend infinitely across the territories that its officials travel?

Consider the following case. Two elected members of a collegial body (e.g., a local legislature or hospital board) spontaneously decide to dine together with the general manager of a public agency overseen by the body. The two elected members alone do not constitute a quorum of the body, and they had no intent to circumvent the statutory open meeting requirement in the relevant state. 189 Nevertheless, at dinner they could discuss matters that are currently before the body or that could conceivably come before the body at a later date, while the public would be unable to monitor the conversation or even know the conversation took place. Is this a meeting that would require the members to give advance notice of their meal and to invite the public to join them?

Most open meeting statutes reach only formal meetings, defined as those that would adopt final actions, or at which a majority or quorum is in attendance. 190 This approach assumes that a meeting occurs in the normal course of a government entity’s operations, at a scheduled time, most

187. SCHWING, supra note 183, § 6.6 (discussing various definitions of meeting in open meeting law, and describing it as “[t]he most telling single element to determine whether an open meeting act is strong and encompassing or weak and limited in scope”).

188. See David A. Barrett, Note, Facilitating Government Decision Making: Distinguishing Between Meetings and Nonmeetings Under the Federal Sunshine Act, 66 Tex. L. Rev. 1195, 1205–06 (1988) (distinguishing among stages in which a body is engaged in “collective inquiry” into the existence of and facts surrounding an issue, deliberation over a narrow range of proposals, or when the officials are deciding about a particular proposal).


typically though not necessarily in the entity’s office or in an official public meeting room. Therefore, the majority of jurisdictions would allow the dinner meeting to take place without public notice or access because of its small size and the informal nature of the gathering, even if it results in a discussion by the members of an issue before the body. A small number of jurisdictions would bar the meeting, however. Interpreting their state statute, Florida courts and attorney general opinions would view the case as a violation of Florida’s sunshine law unless the public is provided notice and access; to do otherwise, an intermediate appellate court has held, would allow members to “gather with impunity behind closed doors and discuss matters on which foreseeable action may be taken by that board or commission in clear violation of the purpose, intent, and spirit of the Government in the Sunshine Law.”

The issue maps the spatial and architectural problems the state creates onto the private lives and dual identities of public officials who are at once government officers and private individuals. Any space an official occupies, even a private restaurant, can be transformed into a government office and meeting room by virtue of the official’s discussion of public business with colleagues. A populist understanding of transparency would not allow officials to avoid their duty to be visible to the public by escaping into their private lives and identities because, as a California appellate court asserted,

191. See, e.g., GA. CODE ANN. § 50-14-1(a)(2) (2009) (defining meeting as a “gathering of a quorum of the members . . . at a designated time and place” to discuss or take action on official business); N.C. GEN. STAT. § 143-318.10(d) (2009) (defining meeting as “a meeting, assembly, or gathering together at any time or place . . . of a majority of the members of a public body for the purpose of conducting hearings, participating in deliberations, or voting upon or otherwise transacting the public business within the jurisdiction, real or apparent, of the public body”).

192. Hough v. Stembridge, 278 So. 2d 288, 289 (Fla. Dist. Ct. App. 1973); see also Fla. Op. Att’y Gen. 2000-08 (2000), http://myfloridalegal.com/ago.md/Opinions/EBDA5F9E248932DA85256B000523870 (opining that the Sunshine Law “is generally applicable to any gathering where two or more members of a public board or commission discuss some matter on which foreseeable action will be taken by that board or commission,” including a forum for all county fire commissioners where on some occasions more than one commissioner from a specific district may attend the same meeting). For an extended critique of this approach to Florida’s law arguing that it is inconsistent with the statute’s text and legislative history, see Seed, supra note 189. Other states take a similar approach. See, e.g., VA. CODE ANN. § 2.2-3701 (2008) (defining meeting to include the “informal assemblage of (i) as many as three members or (ii) a quorum, if less than three, of the constituent membership”); Mayor of El Dorado v. El Dorado Broad. Co., 544 S.W.2d 206, 207–08 (Ark. 1976) (holding that state Freedom of the Information Act applies to informal meetings of less than a quorum of members).

193. See SCHWING, supra note 183, § 5.74 (discussing how state open meeting laws consider the public or private character of the government’s meeting place).
“[a]n informal conference or caucus permits crystallization of secret decisions to a point just short of ceremonial acceptance.” If an official can conduct public business out of the public’s sight, and public business includes nearly any action that could lead to an official government act, then any enclosure and any space that the official occupies must be made open to the public when necessary. Understood this way, the state can be everywhere, and the public must be able to view its officials everywhere across the state’s territory and in any building where the public’s business takes place. Taken to its logical end, however, this view would allow no space that an official occupies to be securely private—including his or her home (from where the official can make calls and send e-mails via private phone lines, computers, and e-mail accounts). The fact that federal law and the vast majority of states refuse to extend their open meeting laws to this degree suggests that legislatures and courts have been hesitant to make the state thoroughly and constantly visible. Their unwillingness to adopt the populist approach suggests either a failure of will or a recognition that the state’s visibility can and should be sacrificed to other interests, including the practical limits of transparency’s enforcement and the private interests of public officials.

2. Containing Documents

Government agencies regularly possess in their facilities documents they did not create; conversely, records produced by the government frequently end up in the hands of individuals and institutions and are housed in buildings that are not themselves part of the government. Open government laws struggle to resolve the issue of whether an agency must disclose a record that it does not possess, and whether it should be required to release a record that it possesses but that originated with another part of the government. Do freedom of information statutes cover records that are not in government offices or on government property? Can they tame the tendency of documents to move across the government and into the file


cabinets (and hard drives) of private individuals? Under the FOIA, the issue turns on whether a document is an “agency record,” which the statute fails to define, and whether an agency has the duty to obtain and retain records, which the statute fails to specify.

The answer, according to the Supreme Court, is that to be subject to disclosure under the FOIA, a record must either be born governmental—it must have, as its provenance, a governmental pedigree—or be adopted by the government—that is, the government must willingly take possession of it. This definition has a spatial dimension to it: the record must be produced within the government’s domain, or later incorporated within it. Consider, for example, the case of Henry Kissinger’s telephone notes. Kissinger served as both National Security Advisor (from 1969 until 1975) and Secretary of State (between 1973 and 1977) under Presidents Nixon and Ford. Throughout his service, he regularly recorded his telephone conversations, and the resulting tapes were then transcribed and stored in documentary form in his personal files within the Department of State. In October 1976, after obtaining a legal opinion from the Legal Adviser of the Department of State concluding that the transcribed notes constituted personal papers rather than agency records and were therefore his to keep after he left office, Secretary Kissinger arranged to remove the files to the private estate of Vice President Nelson Rockefeller. By a later agreement, Kissinger deeded the notes to the Library of Congress with restrictions on public access to the materials prior to the death of the parties to the phone conversations. When journalists and public interest groups subsequently filed requests to view the documents, the Department of State claimed that it no longer had possession of the files.

The issue before the Court in *Kissinger v. Reporters Committee for Freedom of the Press*, as Justice Brennan highlighted in his dissent, was the extent to which the FOIA restrains an agency’s authority to move documents—especially if a requester claims that the agency intended the documents’ removal to make them inaccessible—and the effect that physical location

196. See U.S. Dep’t of Justice v. Tax Analysts, 492 U.S. 136, 144–46 (1989). Note that this only speaks to the question of whether a record was improperly withheld, not to the question of whether it is an “agency record” subject to the FOIA. The latter issue is complicated by the organizational question of which entities are in fact subject to the FOIA, an issue discussed *supra* Part II.B.


198. Id. at 140.

199. Id. at 140–41.

200. Id. at 141–42.

201. Id. at 142–43. Some of the requests were filed before the files’ removal. Id.
has on their public access.\textsuperscript{202} If the FOIA extends only to physical control by and within the state’s facilities, and the law does not require an agency to disclose all of the records it considered in its decisionmaking process (no matter if the agency ever gained possession of them),\textsuperscript{203} then a document’s location outside of the state not only matters but is outcome determinative—a document not within the state’s control cannot be made available under the FOIA. A majority of the Supreme Court took this more limited approach to the issue in \textit{Kissinger}, holding that a document that an agency does not possess has not been “withheld” under the FOIA.\textsuperscript{204} If an agency does not possess a document, even if it has allowed the document to leave its possession, then its failure to retrieve it does not violate the law.\textsuperscript{205} Because Secretary Kissinger’s telephone records were no longer housed within Department of State offices and under the agency’s control, the Department of State did not violate the FOIA by failing to release them.\textsuperscript{206} To be an agency record, a document must be physically located within the state.\textsuperscript{207}

The reverse situation creates what appears to be an odd result that further confounds the populist understanding of transparency. Just as documents created but not retained by an agency are no longer subject to the FOIA when they leave the agency’s control, so documents controlled by an agency that is subject to the FOIA but created by a public or private entity that is not subject to the FOIA are also not subject to the FOIA. Thus, in \textit{Kissinger}, files that Kissinger created while he was a close advisor to the President (a role that does not fall within the FOIA’s ambit)\textsuperscript{208} and before he became Secretary of State (when documents he created would fall within the FOIA) did not become Department of State records when they

\begin{thebibliography}{9}
\bibitem{202} \textit{Id.} at 139 (Brennan, J., concurring in part and dissenting in part).
\bibitem{204} \textit{Kissinger}, 445 U.S. at 150–51.
\bibitem{205} Part of this limitation emanates from the FOIA’s limited reach. It does not require an agency to create or retain records; instead, the Federal Records Act, 44 U.S.C. §§ 2901–2910 (2006), and the Records Disposal Act, \textit{id.} §§ 3301–3324, govern how records are managed and disposed of, and neither statute provides for a private right of action. The FOIA thus does not itself oblige an agency to retrieve a document that it allowed to leave its possession. \textit{Kissinger}, 445 U.S. at 148–50.
\bibitem{206} \textit{Kissinger}, 445 U.S. at 155.
\bibitem{207} A companion case to \textit{Kissinger}, decided by the Court on the same day, came to a similar conclusion, holding that medical records produced by a private research organization under the aegis and with the funding of a federal agency are not subject to the FOIA because they were neither made nor received by a federal agency. \textit{Forsham}, 445 U.S. at 186.
\bibitem{208} \textit{See} \textit{Kissinger}, 445 U.S. at 156.
\end{thebibliography}
were moved to his new office. \(^\text{209}\) Similarly, the record of a secret congressional committee hearing did not become an agency record because it was possessed by the CIA; rather, it remained within congressional control and was thus not subject to the FOIA, even if it was housed within the CIA’s facilities. \(^\text{210}\) The D.C. Circuit’s current test for these types of cases, a two-part standard to determine whether documents created either by or for Congress but in an agency’s possession constitute agency records, inquires into whether Congress has in fact ceded control of the documents and whether the agency has gained over them full property rights, rather than simple possessory interests. \(^\text{211}\)

Kissinger’s result is the exact opposite of what an open government law that embraces the full implications of transparency would expect and demand. \(^\text{212}\) A document located outside the state, Kissinger held, is not subject to the FOIA. But a document located within the state is also not necessarily subject to the FOIA. If the state created it or controls it, a populist understanding of transparency would argue the document ought to be made available to the public. The state’s organizational and physical complexity should not keep it from being visible. The present state of the law appears to allow the government and its officials to move documents...

\(^{209}\) \textit{Id.} at 157.

\(^{210}\) Goland \textit{v.} CIA, 607 F.2d 339 (D.C. Cir. 1978), \textit{vacated in part on other grounds}, 607 F.2d 367 (D.C. Cir. 1979) (per curiam).

\(^{211}\) \textit{Id.} at 347. \textit{See, e.g.}, United We Stand Am., Inc. \textit{v.} IRS, 359 F.3d 595 (D.C. Cir. 2004) (holding records created by IRS for the congressional Joint Committee on Taxation were agency records because, other than in its initial request, Congress failed to show sufficient intent to retain control over them); Paisley \textit{v.} CIA, 712 F.2d 686, 695–96 (D.C. Cir. 1983), \textit{vacated in part on other grounds}, 724 F.2d 201 (D.C. Cir. 1984) (per curiam) (holding records created by the CIA to aid a congressional investigation were agency records subject to the FOIA because Congress did not manifest sufficient intent to retain control over them); Holy Spirit Ass’n for the Unification of World Christianity \textit{v.} CIA, 636 F.2d 838, 842–43 (D.C. Cir. 1980), \textit{vacated in part on other grounds}, 455 U.S. 997 (1982) (per curiam) (holding that documents created by the CIA for Congress, which were sent to Congress and then returned to the CIA, constituted agency records subject to the FOIA because Congress failed to retain control over them).

\(^{212}\) \textit{See, e.g.}, Feiser, \textit{supra} note 153, at 58 (criticizing Kissinger’s approach as “cramped” and arguing that “this approach would keep its records out of the public eye unless the FOIA agency actually possesses and uses the documents”); Samaha, \textit{supra} note 5, at 971–72 (criticizing Kissinger as exemplifying one of the FOIA’s main weaknesses: the ability of the government to avoid accountability to the public by moving or destroying documents); \textit{The Supreme Court, 1979 Term—Freedom of Information Act: Threshold Definitional Barriers to Disclosure}, 94 Harv. L. Rev. 232, 240 (1980) (characterizing Kissinger’s limited reading of the FOIA as “unsatisfactory”); Marie Veronica O’Connell, Note, \textit{A Control Test for Determining “Agency Record” Status Under the Freedom of Information Act}, 85 Colum. L. Rev. 611, 628–29 (1985) (attempting to read Kissinger broadly as part of a “control” theory that would make possession a non-determinative test for the FOIA’s applicability).
around its offices and territory in order to avoid disclosure.

D. The Impossibility of Physical Exposure

Geography and the built environment help define the state’s reach and presence. The American state encompasses a huge territory, and in its branches and levels occupies a vast number of buildings. Insofar as the state and its administrative apparatus have solidified their position at the core of an expansive and complex nation, their material scope and existence will continue to prove difficult to contain in a manner that will render them fully visible.

CONCLUSION: THE PANOPTICIZED STATE

The metaphoric understanding of transparency, which defines the accessible, accountable government as one that can be seen, faces innumerable obstacles in the complex and dispersed American state. Technology can ameliorate but not remove such obstacles, notwithstanding constructive efforts to improve the release and usefulness of government data—and then to claim those improvements as technological fixes to a secretive, likely corrupt state.213 Like the ongoing quest for legal and regulatory solutions to the problem of government opacity and unsatisfactory performance, the ongoing quest for technological fixes that make the state more accountable is itself symptomatic of the populist embrace of the visible state ideal. Information technology can make the state more visible, which will in turn force government officers to behave in ways that better comport with citizens’ expectations. If we cannot see the physical state, and if we cannot thoroughly force the state to be seen through law, perhaps we can see a digital one—or at least its informational traces—on the Internet or through a spreadsheet.

These efforts call to mind another technological fix for a significant social problem that requires the surveillance of a set of dangerously wayward actors. In all of its guises, the transparency metaphor urges the construction of an inverted panoptic penal facility, one that puts the public—or some subset thereof—in the position of the guard and that casts government officials as the incarcerated. Jeremy Bentham’s original design

---

213. See, e.g., David Robinson et al., Government Data and the Invisible Hand, 11 YALE J.L. & TECH. 160, 160 (2009) (claiming that the government should release reusable, rather than processed, data, which would “embrace the potential of Internet-enabled government transparency”); Sunlight Foundation, supra note 26 (characterizing itself as using “cutting-edge technology and ideas to make government transparent and accountable . . . [by] focus[ing] on the digitization of government data and the creation of tools and Web sites to make that data easily accessible for all citizens”).
for his Panopticon arranged and illuminated cells so that the inmates would be constantly visible to prison guards located securely in a central tower. Prisoners could see the tower but could not see into it, and could constantly be seen, despite being confined to a cell from which they could not escape.\textsuperscript{214} The prison’s enclosure would illuminate them, removing the darkness that offered them protection while it captured them for the supervisor’s eye. The Panopticon thus makes its subjects transparent to authority.

For the Panopticon’s effect to reach its “[i]deal perfection,” Bentham asserted, the subject should be unable to recognize when he is being watched, but should at all times “conceive himself to be so” scrutinized.\textsuperscript{215} Constant and unending, the belief that one is being watched would prove self-regulating as it was internalized by the prisoner; it would thereby be less difficult and costly to impose, and would require fewer guards to administer.\textsuperscript{216} The architecture of the Panopticon that creates the conditions of feeling under constant surveillance thereby shapes the prisoner and causes him to learn to shape himself, rendering through its physical design and organization a subject who considers himself to be the object of permanent surveillance. Such surveillance does not merely disincentivize resistance or thwart escape—it disciplines and organizes the behavior, thought, and desire of the surveilled. As Michel Foucault noted, Bentham brilliantly recognized that “[v]isibility is a trap.”\textsuperscript{217} Rather than an old-fashioned institution of power that banished certain undesirable activities and people—the criminal, the sick, the insane—the Panopticon could “carry the effects of power right to them” through “the calculation of openings, of filled and empty spaces, passages and transparencies.”\textsuperscript{218} It offers an architecture of “continuous observation made possible by technical arrangements.”\textsuperscript{219}

For Bentham, the panoptic model had clear implications for representative democracy. Throughout his political writings, Bentham


\textsuperscript{215} Id., supra note 214, at 40.

\textsuperscript{216} Id.

\textsuperscript{217} Michel Foucault, \textit{Discipline and Punish: The Birth of the Prison} 200 (Alan Sheridan trans., 1979).

\textsuperscript{218} Id. at 172.

\textsuperscript{219} Shoshana Zuboff, \textit{In the Age of the Smart Machine} 322 (1988).
emphasized the importance of allowing the public to view its political rulers. Publicity, he argued, would “constrain” the ruling assembly to perform its duty, allow it to secure the confidence of its public, and develop a more informed electorate. Bentham imagined mechanisms to achieve a state that was constantly under scrutiny, particularly through the concept of the “Public Opinion Tribunal,” a kind of societal committee or judiciary of the whole that would play a key role in a constitutional democracy. Specifically, it would gather facts and evidence regarding the performance of public institutions; express approval or disapproval of the state, as well as reward or punish representatives and officials; and propose reforms and new institutional arrangements. In this regard, his Tribunal, and his general understanding of publicity, imagined the public’s check on government behavior as analogous to the Panopticon, in which the informed, collectively organized public “attempts to serve as the all-seeing eye, casting its critical reforming gaze over the full spectrum of governmental (indeed public) activity.”

For Bentham, democracy’s foundation was built on the panoptic principle of an ever-vigilant public managing a captive state and rulers.

As with Bentham’s Panopticon, the populist metaphorical conception of transparency views its objects—government institutions and officers, rather than incarcerated prisoners—as requiring discipline. Both long to provide an institutional solution to the problem they identify, one that can develop in their objects the self-discipline that will transform them into proper subjects: rehabilitated citizens for Bentham, a more responsive and responsible state for transparency advocates. Strong-form transparency thus would reverse the Panopticon, placing the people in the lookout and recasting the state as the object of surveillance. The sentiment is populist, but the institutional apparatus that would enact the sentiment is decidedly progressive: a solution to a significant social problem that works through a state institution intended to shape human behavior.

The fly in transparency’s ointment is the same one that Bentham faced. As a practical matter, building a Panopticon proves difficult. Bentham could not persuade the various relevant authorities of his time—late

222. Id. at 111; see also Janet Semple, Bentham’s Prison 321 (1993) (“Bentham’s democracy is a structure full of light, as was the panopticon, but the light falls on those in authority.”).
eighteenth and early nineteenth century prison administrators, political leadership, and landowners—to allow him to build his model prison. Instead, the Panopticon has come to stand as what Foucault calls a “program” rather than a material, historical fact: one of the “diverse realities articulated onto each other” that produces a series of wide-ranging effects throughout society; most importantly, these technologies “crystallize into institutions, they inform individual behavior, they act as grids for the perception and evaluation of things.” The Panopticon serves as a metaphor for the modern institution, one that seeks to discipline its subjects by forcing them to internalize external authority, to develop the discipline of the self. It also represents the madness and excess of modernity, the pernicious but essential means by which the state could develop as the apex of the modern, rational civilization. It is impossible and horrifying to imagine a world in which one is perpetually under threat of observation. But it is also necessary as a metaphor to understand how the modern liberal state develops its subjects, and unsurprising therefore that one of the great liberal and utilitarian political philosophers—one whose writings on the role of publicity in a representative constitutional democracy remain filled with viable, relevant ideals—should have proposed it.

Viewing the boundless and endless desire to achieve a visible state in relation to the panopticized state model leads to two related conclusions. First, because the state cannot be made wholly visible, short of dismantling it or imposing a maddening (and likely impossible to construct) panoptic apparatus, such a desire will lead only to cycles of frustration. The popular will to see the state will ride an asymptotic line that approaches—but never reaches—the perfect and perfectly accountable and responsive government. Second, the will to see the state is so much a part of American democratic, populist political culture that is skeptical of the state that it cannot itself be wished away. Technocratic reform to provide incremental

227. I am for this reason skeptical of Edward Rubin’s efforts to purge political concepts of their popular and (what he sees as therefore) unhelpful resonances with historical references to a long-vanished state and ideological misrecognitions of the current one by employing uninteresting, uninformative, and naïve heuristics. See Rubin, supra note 50, at 16–17. As the legal realist Thurman Arnold argued regarding the conservative opposition
improvements to government performance—including but not limited to making government more open to the public—can neither ignore nor counteract populist demands for a fully visible state. Successful legislative, regulatory, and institutional interventions must recognize and respect the desire for a visible state while they also concede and grapple with the state’s inevitable push towards opacity. In the struggle over transparency, the populist will and the technocratic will cannot be separated.

to New Deal reform, which frequently expressed itself in legal formalist terms that attempted to thwart the administrative state, “[s]o long as our belief in rational moral government depends upon the law, it must continue to balance logically the contradictory ideals which that government must express.” THURMAN W. ARNOLD, THE SYMBOLS OF GOVERNMENT 69 (5th prtg. 1948). In other words, incremental reform that appears to be a substitute for a new age of transparency must nevertheless present itself as the next important step toward the dawn of a full transparency that can never be achieved.
A HISTORY OF THE MILITARY AUTHORITY EXCEPTION IN THE ADMINISTRATIVE PROCEDURE ACT

KATHRYN E. KOVACS*

Table of Contents

Introduction ............................................................................................................. 674
   I. Legislative Historiography ................................................................. 678
      A. Pre-APA Bills .............................................................................. 681
          1. 1929–1936 .......................................................................... 681
          2. The Walter–Logan Bill ......................................................... 683
          3. The War Years ...................................................................... 691
      B. The APA of 1946 ........................................................................ 696
          1. Introduction and Passage ....................................................... 696
          2. The Aftermath ...................................................................... 704
      C. 1966 and 1976 Amendments ....................................................... 708
      D. Summation ............................................................................... 709
   II. The Articles of War ............................................................................. 710
      A. “In the Field” ............................................................................ 712
      B. “Time of War” ........................................................................... 717
   III. Where the Courts and Commentators Have Gone Wrong ............ 720

Conclusion .......................................................................................................... 725

* The Author is an attorney in the U.S. Department of Justice, Environment & Natural Resources Division, Appellate Section, and is a graduate of Yale University and the Georgetown University Law Center. The views expressed in this Article are those of the Author and are not intended to represent the views of the United States. The Author owes a huge debt of gratitude to Professor George Shepherd for his painstaking history of the Administrative Procedure Act, infra note 34, and to Todd Aagaard, Allan Brabender, Richard Lazarus, and Daniel Ernst for reading drafts of this Article.

673
INTRODUCTION

The War on Terror ignited a firestorm of commentary on issues related to civil liberties and international law. It also sparked a debate about the environmental impacts of military action and whether existing statutory provisions are too restrictive or not restrictive enough. Few commentators, however, have mentioned the provision in the Administrative Procedure Act (APA) that exempts “military authority exercised in the field in time of war,” and none have examined it closely. Yet there can be no doubt about the importance of the APA. In 1945, the American Bar Association’s Special Committee on Administrative Law, which drafted the bill that eventually became the APA, predicted that it might “become the most important event in improving the administration of justice since the Judiciary Act of 1789.” Current commentators think the APA “is arguably the most important piece of legislation governing federal regulatory agency policy making[,]” and that “of all the administrative laws, none have been more significant than the federal Administrative Procedure Act of 1946 and the similar state-level administrative procedure acts . . . .”

The APA may provide an avenue to judicial review for individuals detained by the military. Although habeas corpus is available to those held

in “territory over which the United States exercises plenary and exclusive jurisdiction,” it may not be available to those held elsewhere. And the Alien Tort Act does not waive the government’s sovereign immunity. Hence, some persons challenging their detention by the military have relied on the APA’s waiver of sovereign immunity to gain access to the courts. Even if the APA is not the detainee’s road to the courthouse door, it may be relevant to the procedures used by military tribunals or in the judicial review of their decisions.

In environmental law, the APA is particularly significant because it provides the jurisdictional grounding for suits against the federal government under many, if not most, environmental and natural resources statutes. While some of the pollution statutes provide the waiver of sovereign immunity and cause of action required to sue a federal agency, the environmental statutes that are of the greatest current relevance to the military do not. In Winter v. Natural Resources Defense Council Inc., for example, the Supreme Court vacated in part an injunction that barred the Navy from using sonar during training exercises off the coast of California and clarified the standards for granting preliminary injunctive relief. The injunction at issue in Winter was premised primarily on the National Environmental Policy Act (NEPA), which requires federal agencies to examine the potential environmental impacts of their proposals before implementing them. NEPA, “the statute that launched the ‘environmental decade’ of the 1970s, has been hailed as one of the nation’s most important environmental laws.” Yet it is actionable only through

10. See Basaridh v. Obama, 612 F. Supp. 2d 30, 33 n.9 (D.D.C. 2009) (noting that “it could certainly be argued that the [Administrative Review Board] had violated the principles of the APA,” but that argument “would be met with the claim that the APA is inapplicable due to” the “military authority” exception); Bismullah v. Gates, 514 F.3d 1291, 1294–95 (D.C. Cir. 2008) (Ginsburg, C.J., concurring in denial of reh’g en banc) (concluding that Combatant Status Review Tribunals fall outside the scope of the APA); id. at 1303 n.3 (Randolph, J., dissenting from denial of reh’g en banc) (concluding that Combatant Status Review Tribunals “are military ‘functions’ the APA specifically exempts”).
12. See id. at 372–74.
the APA. The same is true of the Migratory Bird Treaty Act, under which the D.C. District Court enjoined Navy exercises on an island in the Pacific Ocean, and certain claims under the Marine Mammal Protection Act and the Endangered Species Act, which underlie many challenges to the Navy’s use of sonar in training.

The academic literature lacks any in-depth analysis of the military authority exception. Given the critical importance of the APA to military detainees and environmental plaintiffs, that silence is deafening. Part II of this article provides a much needed and long overdue historiographic analysis of the APA’s military authority exception. While a court interpreting this provision might limit itself to examining the statute’s text and official legislative history, the surrounding historical context of this enactment allows a fuller understanding of Congress’s probable intent. As Professor Eskridge observed, “statutory interpretation is all about words, but words are about much more than dictionaries and ordinary usage; they also involve policies chosen by the legislature and enduring principles suggested by the common law, the law of nations, and the Constitution.”


20. See, e.g., Cetacean Cmty v. Bush, 386 F.3d 1169, 1171 (9th Cir. 2004).

21. The term agency as it appears in the waiver of sovereign immunity in 5 U.S.C. § 702 must be read in favor of the government. Thus, any ambiguity in the definition of that term, including the “military authority” exception, should be read to exempt the military from judicial review. See Lane v. Pena, 518 U.S. 187, 192 (1996). Recently, however, the Supreme Court said that the “sovereign immunity canon” does not “displace[] the other traditional tools of statutory construction.” Richlin Sec. Serv. Co. v. Chertoff, 128 S. Ct. 2007, 2019 (2008); see also Gregory C. Sisk, The Continuing Drift of Federal Sovereign Immunity Jurisprudence, 50 WM. & MARY L. REV. 517, 575 (2008) (“So understood, the rubric of strict construction is not a substitute for careful attention to the statutory language and structure actually enacted by Congress or a basis for ignoring the manifest purpose of the statutory waiver.”). And the military authority exception cannot be given such an “unduly generous interpretation” that it “defeat[s] the central purpose of the statute.” Dolan v. U.S. Postal Serv., 546 U.S. 481, 492 (2006) (quoting Kosak v. United States, 465 U.S. 848, 853 n.9 (1984)); see also United States v. Nordic Vill., Inc., 503 U.S. 30, 34 (1992) (“We have on occasion narrowly construed exceptions to waivers of sovereign immunity where that was consistent with Congress’ clear intent . . . .”).

22. William N. Eskridge, Jr., All About Words: Early Understandings of the “Judicial Power” in
This article does not delve into the debate about whether extratextual sources should be examined to determine congressional intent, but instead presents the full historical context and leaves it to others to take what they will from the material presented.

Although the first bill to constrain administrative agencies was introduced in 1929, it was not until Roosevelt took office in 1933 and plowed ahead with the New Deal that the drive for administrative reform took root. In late 1940, a coalition of Republicans and conservative Democrats passed the Walter–Logan bill, a bill that was similar in many respects to the present-day APA, but which broadly exempted “any matter concerning or relating to the military or naval establishments, including . . . any other agency or authority hereafter created to expedite military and naval defense.” As Hitler occupied Paris and bombed London, President Roosevelt won reelection to a third term and vetoed the Walter–Logan bill, in part because he thought the military exemption was not broad enough. Administrative reform went into hibernation for several years during the war and reemerged after D-Day to blossom into the APA of 1946, which enjoyed wide support from Congress and the President. While the rulemaking and adjudication provisions of the APA contained broad exemptions for military functions, the judicial review provisions exempted only “military . . . authority exercised in the field in time of war or in occupied territory.” This article concludes that the history of that shift reveals much about Congress’s changing relationship with the military, but that further analysis is required to unearth the contemporary understanding of the particular phrases Congress chose to employ in the military authority exception.

Part III explores the provisions of the Articles of War from which the military authority exception was apparently drawn. At the time it enacted the APA, Congress would have understood the phrase “in the field” to be a term of art encompassing not just the locus of combat overseas, but any...
place “where military operations are being conducted,” including military training camps in the United States and military transport ships docked in the U.S. The phrase “time of war” is also a term of art that has long been understood not to require a congressional declaration of war. Although the courts have interpreted those phrases narrowly in the context of court martial jurisdiction over civilians, this Article asserts that such a narrow interpretation is not appropriate for the military authority exception, and that Congress’s understanding of those phrases in 1946 would have been somewhat broader than that of a modern reader.

This historical analysis of the military authority exception begins to bring the plain language’s ambiguity into focus and reveals that some courts’ and commentators’ assumptions about the military authority exception have been flawed. Those flaws are explored in Part IV. Most commonly, modern readers interpret the phrase “in the field” too narrowly and fail to recognize that, in the 1940s, that phrase would have been understood to reach well beyond the locus of combat to the high seas and domestic facilities and to activities with only a faint connection to combat operations. Those erroneous assumptions could lead courts to review a broader scope of military action than Congress might have intended. On the other hand, the history of the military authority exception calls into question the courts’ continued willingness to employ common law deference doctrines to avoid reviewing military action. Congress, implementing the lessons learned in World War II, mandated judicial review of a broad range of military actions. This Article will begin to define how broad that range actually is.

I. LEGISLATIVE HISTORIOGRAPHY

The term agency is central to the APA. Virtually all of the Act’s provisions apply to agencies: § 553 requires “the agency” to allow public participation in rulemaking;§ 554 requires “the agency” to give interested parties notice of adjudicatory hearings and an opportunity to participate; and § 702 provides a cause of action and waiver of sovereign immunity for claims challenging “agency action.” The term agency is defined identically in two separate provisions of the current APA, §§ 551 and 701, as, “each authority of the Government of the United States, whether or not it is within or subject to review by another agency.” Among the exceptions from that definition are Congress, the courts, the territories or possessions

29. § 554(b), (c).
30. § 702; Puerto Rico v. United States, 490 F.3d 50, 72 (1st Cir. 2007).
of the United States, courts-martial and military commissions, and, the subject of inquiry here, “military authority exercised in the field in time of war or in occupied territory.”

The APA passed the House and Senate in 1946 on voice votes with little debate. But that seemingly peaceful end was preceded by seventeen years of competing legislative proposals, fierce debate, a presidential veto, and finally, successful compromise. The debate was only in part “a search for administrative truth and efficiency.” Underlying that substantive dispute “was a pitched political battle for the life of the New Deal.” World War II interrupted the political battle temporarily and helped to shift the debate sufficiently to yield compromise. This historical context puts flesh on the bones of the military authority exception.

The American Bar Association (ABA) was central to the development and passage of the APA, the law that still imposes the primary statutory constraints on federal administrative agencies. That the Bar was so resistant to the rise of the administrative state may seem odd to us, given the significant number of lawyers currently employed in administrative practice. Professor Zeppos explains, however, that, in the “age of formalism” at the end of the nineteenth century, law was considered “a

32. §§ 551(1)(A), (B), (C), (F), 701(b)(1) (A), (B), (C), (F).
33. §§ 551(1)(G), 701(b)(1)(G). The exceptions at § 551(1)(E) through (H) do not apply to the public information requirements in § 552, § 551(1).
35. Id. Walter Gellhorn portrayed the debate leading to the APA’s passage, on its fortieth anniversary, as a battle between the reactionary, ideological, and rhetorically inflammatory American Bar Association (ABA) and the systematically analytical Attorney General’s Committee on Administrative Procedure. See generally Walter Gellhorn, The Administrative Procedure Act: The Beginnings, 72 VA. L. REV. 219 (1986). Gellhorn was research director and “intellectual leader” of the Attorney General’s Committee. Shepherd, supra note 34, at 1593; see also Martin Shapiro, APA: Past, Present, Future, 72 VA. L. REV. 447, 449 (1986) (“Gellhorn’s view is a conflation of New Deal ideology with good, common law lawyering. In fact, Professor Gellhorn was one of a cohort of New Deal lawyers who . . . created a body of administrative law that rationalized and legitimated the administrative state that the New Deal created and that the New Deal ideology defended.”). Employing positive political theory and econometrics, McNellis demonstrates that, in contrast to Gellhorn’s view of the APA as a unanimous effort “to enhance administrative efficiency and to extend individual rights through the establishment of procedural due process for federal agencies,” McNellis, supra note 5, at 206 (citation omitted), “political preferences over economic outcomes as well as prosaic political strategizing and coalition building played major roles” in the APA’s passage. Id. at 183.
36. See Nicholas S. Zeppos, The Legal Profession and the Development of Administrative Law, 72 CHI.-KENT L. REV. 1119, 1151 (1997) (“The puzzling question is why, given the obvious business opportunities presented by the rise of administrative law, the bar (or elite segments of the bar) was so slow (or reluctant) to fill this important new niche in legal services.”).
scientific, objective reasoning process,” separate “from the world of politics.”37 The Bar enjoyed a “privileged role . . . inextricably bound up with the power, independence, and prestige of the courts.”38 Then, in the first thirty years of the twentieth century, the number of federal administrative agencies doubled.39 Administrative adjudication “substitute[d] informal meetings presided over by a political actor for the formalized, structured, and ritualistic hearing before an independent judge.”40 That shift away from the formalist legal model breached the “boundary between law and politics,” raising “basic issues of due process and bias” and threatening the elite status and livelihoods of elite lawyers.41 The ABA’s early efforts at administrative reform thus focused on subjecting administrative decisionmaking to judicial review.42

Those early efforts were not successful. As Professor Schiller explains,

37. Id. at 1121–22 (citations omitted).
38. Id. at 1130.
39. Shepherd, supra note 34, at 1561 (citation omitted). “By 1940 there were over fifty federal administrative agencies compared to the eleven that existed at the beginning of the Civil War.” Reuel E. Schiller, Reining in the Administrative State: World War II and the Decline of Expert Administration, in TOTAL WAR AND THE LAW: THE AMERICAN HOME FRONT IN WORLD WAR II 185, 186 (Daniel R. Ernst & Victor J. E. eds., 2002); see also Daniel R. Ernst, The Ideal and the Actual in the State: Willard Hurst at the Board of Economic Warfare, in TOTAL WAR AND THE LAW: THE AMERICAN HOME FRONT IN WORLD WAR II, supra, at 149 (“[seventy] percent of the practice of major law firms in 1934 was before agencies that did not exist a generation earlier.”).
40. The rise of administrative bodies probably has been the most significant legal trend of the last century and perhaps more values today are affected by their decisions than by those of all the courts, review of administrative decisions apart. They also have begun to have important consequences on personal rights . . . . They have become a veritable fourth branch of the Government, which has deranged our three-branch legal theories much as the concept of a fourth dimension unsettles our three-dimensional thinking.


41. Zeppos, supra note 36, at 1127–29; see also Shepherd, supra note 34, at 1571 (“The increasing importance of administrative tribunals appeared to elite lawyers to threaten the lawyers’ livelihoods by diminishing the importance of lawyers and traditional lawyering.”). Professor Zeppos points out that, unlike the ABA, the Federal Bar Association and the National Lawyer’s Guild opposed the ABA’s proposals. Zeppos, supra note 36, at 1131 n.64. He thus posits that the ABA’s efforts at administrative reform may have been based in part on self-interested protection of large, industrial clients. Id. at 1133–37; see also Shepherd, supra note 34, at 1571 (“The lawyers feared that New Deal agencies threatened their clients.”).
42. Zeppos, supra note 36, at 1129.
progressive reformers early in the twentieth century expressed “a faith that properly trained experts could find objectively correct solutions to the myriad of social problems extant in a rapidly industrializing, increasingly fractious society.”\textsuperscript{43} That faith carried over to New Deal-era reformers, who believed that expert administration would solve the massive problems the Great Depression had caused.\textsuperscript{44} In the early 1930s, “liberals and progressives believed that administrative government was a scientific solution to an economic and social crisis of unparalleled proportions.”\textsuperscript{45} The three branches of government were seen as “insufficiently flexible” to solve such enormous and complex problems.\textsuperscript{46} Strict judicial review of expert administrative action had no place in that belief system.\textsuperscript{47} Invasive judicial review would “hobble governmental efficiency” and “defeat the purpose of creating expert agencies in the first place.”\textsuperscript{48} “Progressives had long viewed the American judiciary as a reactionary institution . . . .”\textsuperscript{49} The Supreme Court’s invalidation of early New Deal programs “heightened this suspicion” and made it “an article of faith among dedicated New Dealers” that the judiciary should have a limited role in reviewing agency action.\textsuperscript{50} It was not until those notions and the political balance of power began to shift in the late 1930s and early 1940s that administrative reform, and in particular the drive for judicial review of administrative action, picked up steam.

\textbf{A. Pre-APA Bills}

\textbf{1. 1929–1936}

Senator George Norris introduced “the first legislation for constraining administrative agencies” in 1929, four years before Franklin Delano Roosevelt took office.\textsuperscript{51} Norris’s bill would have established a Court of Administrative Justice to adjudicate claims against the United States.\textsuperscript{52}

\begin{itemize}
  \item \textsuperscript{43} Reuel E. Schiller, \textit{Free Speech and Expertise: Administrative Censorship and the Birth of the Modern First Amendment}, 86 V.A. L. REV. 1, 14 (2000).
  \item \textsuperscript{44} \textit{Id.} at 15.
  \item \textsuperscript{45} Schiller, \textit{supra} note 39, at 201.
  \item \textsuperscript{46} \textit{Id.} at 186.
  \item \textsuperscript{47} \textit{Id.} (“New Deal-era administrative law reflected these beliefs by allowing administrative agencies an exceptional amount of independence and flexibility.”).
  \item \textsuperscript{48} \textit{Id.} at 187.
  \item \textsuperscript{50} \textit{Id.} at 1404.
  \item \textsuperscript{51} Shepherd, \textit{supra} note 34, at 1566.
  \item \textsuperscript{52} S. 5154, 70th Cong. (1929).
\end{itemize}
Norris was a liberal Republican who later supported the New Deal and became an independent. Professor Shepherd posits that Norris “apparently introduced the bill in 1929 in order to control the excesses of Republican-controlled agencies.”\textsuperscript{53} Congress took no action on Norris’s bill, but those who sought administrative reform would remain focused on the administrative court concept for several more years.

In May 1933, the ABA’s Executive Committee established a Special Committee on Administrative Law, which spurred the debate over administrative procedure and played a pivotal role in passing the APA.\textsuperscript{54} The ABA Committee’s first report implies that the Committee was formed at that particular time in reaction to the first New Deal, which President Roosevelt had kicked off vigorously and immediately after his inauguration only two months earlier.\textsuperscript{55} Also in May 1933, Senator Mills Logan, a Kentucky Democrat and former chief justice of that state’s highest court, introduced a bill that was almost identical to Senator Norris’ administrative court bill of 1929.\textsuperscript{56} Like Senator Norris, Logan likely intended his bill as a “sincere, nonpolitical attempt to foster agency fairness and efficiency.”\textsuperscript{57} Congress took no action on the bill.

In 1935 and 1936, the ABA Committee proposed a draft bill nearly identical to those previously introduced by Senators Norris and Logan.\textsuperscript{58} Though the ABA withheld its full approval of that bill, the ABA Committee’s chairman, with the approval of only the Executive Committee, proposed a similar bill to Senator Logan, who introduced it in 1936; the bill died in committee.\textsuperscript{59} The administrative court proposals likely enjoyed little conservative support and hence saw no congressional

\textsuperscript{53} Shepherd, supra note 34, at 1567.
\textsuperscript{54} Id. at 1569–70.
\textsuperscript{55} 58 REPORT OF THE FIFTY-SIXTH ANNUAL MEETING OF THE AMERICAN BAR ASSOCIATION 407 (1933); see also John Dickinson, Administrative Procedure Act: Scope and Grounds of Broadened Judicial Review, 33 A.B.A. J. 434, 434 (1947) (“[the] Committee came into existence simultaneously with a mass of early so-called ‘New Deal legislation’… statutes which called into play a vast extension of administrative powers.”). The Twentieth Amendment moved inauguration day from March 4th to January 20th; U.S. CONST. amend. XX, § 1 (ratified January 23, 1933).
\textsuperscript{56} S. 1833, 73d Cong. (1933).
\textsuperscript{57} Shepherd, supra note 34, at 1569.
\textsuperscript{58} Id. at 1575; 60 REPORT OF THE FIFTY-EIGHTH ANNUAL MEETING OF THE AMERICAN BAR ASSOCIATION 136–43 (1935); 61 REPORT OF THE FIFTY-NINTH ANNUAL MEETING OF THE AMERICAN BAR ASSOCIATION 756 (1936) (“[T]he proposal as worked out by the committee is neither revolutionary nor particularly novel in character or purpose but, on the contrary, has been constructed upon a foundation already laid by bills introduced in earlier Congresses and upon study which was made in connection with them.”).
action before 1937 because “the Supreme Court’s rejection of New Deal programs made political attacks unnecessary.” So long as Republicans dominated the federal judiciary, legislative efforts to reign in New Deal agencies were not a high priority for congressional conservatives. Not surprisingly, given Congress’s growing isolationism and its focus on domestic issues at the time, none of these early proposals mentioned the military.

2. The Walter–Logan Bill

Starting in 1937, when the Supreme Court began to approve New Deal programs, President Roosevelt had been weakened by the failure of his Court-packing plan, and recession set in, the drive for administrative reform strengthened. Republicans joined conservative Democrats in support of administrative reform proposals. In the mid-term elections of 1938, Republicans picked up eighty-one seats in the House and eight seats in the Senate, which left Congress firmly in Democratic hands, but nonetheless enhanced the coalition to reign in administrative agencies.

By that time, the liberal faith in expert agencies to solve the nation’s problems began to be overshadowed by a fear of those agencies’ totalitarian tendencies. In the mid-1930s, “the true dimensions of European totalitarianism forced themselves into the American consciousness” with Stalin’s Show Trials, Hitler’s Kristallnacht, and Mussolini’s invasion of Ethiopia, among others. By the late 1930s, many Americans began to fear that “Roosevelt’s ambitions” and “economic desperation” could lead

---

64. Shepherd, supra note 34, at 1580–82.
65. Id. at 1580.
66. Schiller, supra note 40, at 424 n.141.
67. Schiller, supra note 39, at 188; see also Schiller, supra note 43, at 77.
to totalitarianism at home. Administrative agencies “would be a vehicle for fostering its growth.” Although conservatives had long attacked the absolutist tendencies of administrative agencies, by the end of the 1930s, those fears had spread even to New Deal supporters. Professor Schiller posits that Americans’ exposure to the abuses in Europe, coupled with their fear of “administrative absolutism” and Roosevelt’s dictatorial tendencies at home, made them “less and less inclined to trust legislators or administrative experts to look after their civil liberties.” Instead, “Americans came to expect the judiciary . . . to protect individuals and minorities from the deadly tide of totalitarianism that seemed to be infecting” Europe. That shift put wind in the sails of administrative reform in Congress. Despite the military involvement in the atrocities in Europe, however, Congress gave the U.S. military a wide berth in the years leading up to the war.

The 1937 ABA Committee proposal was much stricter than its previous administrative court proposals; it provided for formal administrative hearings, required regulations to be preceded by notice and public hearings, and provided for judicial review. For the first time at the ABA’s annual meeting in September 1937, just weeks after Japan began the Second Sino-Japanese War by attacking China in what some consider the first battle of World War II, the Committee proposed to exempt from its bill “the conduct of military and naval operations in time of war or civil insurrection.” The bill also would have exempted foreign affairs, the Federal Reserve Board, the Comptroller of the Currency, the Federal Deposit Insurance Corporation, and internal revenue, patent, and customs matters, among others. The Committee chair explained that some of those agencies and matters were exempted because the agencies themselves or members of the bar objected or because judicial review was already

68. Schiller, supra note 39, at 188–89.
69. Id. at 189.
70. Schiller, supra note 43, at 85–86, 88; Schiller, supra note 39, at 189.
71. Schiller, supra note 43, at 75, 85–86.
72. Id. at 75–76.
75. 62 REPORT OF THE SIXTIETH ANNUAL MEETING OF THE AMERICAN BAR ASSOCIATION 794, 850 (1937).
76. Id.
available, but he did not explain the military exemption.77 The ABA Board of Governors approved the proposed bill in 1939.78 The Committee, chaired for that one year by “a particularly cranky” Roscoe Pound,79 former Dean of Harvard Law School, began to employ more inflammatory rhetoric in support of its proposals, likening Roosevelt’s administration to the Soviet Union’s Marxist dictatorship.80

Senator Logan and Democratic Representative Emanuel Celler of Brooklyn, New York, introduced the ABA’s proposed bill in 1939, the year Hitler invaded Poland, and Great Britain and France declared war on Germany.81 Pennsylvania Democratic Congressman Francis Walter reintroduced the ABA bill a few months later.82 The bill came to be known as the Walter–Logan bill.83

The bill’s proponents in Congress, both Republicans and Democrats, picked up on the ABA’s rhetoric. The floor debates were “riddled with comparisons of the administrative state to fascist and communist governments and accusations that administrative agencies were [being] used to advance [Roosevelt’s] totalitarian ambitions.”84 Ohio Republican Congressman White endorsed the bill as a “vitaly important” means of counteracting Roosevelt’s ceaseless “greed for power.”85 Michigan

77. Id. at 271, 284–85. The ABA’s archivist informed the Author that the extant committee papers in the ABA archives do not reveal the source of the military exemption.
78. Shepherd, supra note 34, at 1588; 63 REPORT OF THE SIXTY-FIRST ANNUAL MEETING OF THE AMERICAN BAR ASSOCIATION 334 (1938).
79. Schiller, supra note 43, at 86; see also Schiller, supra note 39, at 197 (characterizing Roscoe Pound as “dyspeptic”).
80. 63 REPORT OF THE SIXTY-FIRST ANNUAL MEETING OF THE AMERICAN BAR ASSOCIATION 340, 343 (1938); Schiller, supra note 40, at 422–23; Schiller, supra note 43, at 86 (“The Committee’s bitter critique of the administrative state was suffused with accusations of totalitarianism.”); Shepherd, supra note 34, at 1590–91; Gellhorn, supra note 35, at 221 (“In 1938 . . . [t]he decibel count rose markedly. This was the heyday of Congressman Martin Dies and the House Committee on Un-American Activities. . . . Consideration of administrative law moved perceptibly to the level of ‘the good guys against the bad [guys].’”).
81. S. 915, 76th Cong. (1939); H.R. 4236, 76th Cong. (1939). In 1939, Sen. Logan and Rep. Celler also introduced administrative courts bills listing particular agencies within its jurisdiction, not including the military. S. 916, 76th Cong. (1939); H.R. 4235, 76th Cong. (1939). Those bills were never reported out of committee.
82. H.R. 6324, 76th Cong. (3d Sess. 1939).
83. Shepherd, supra note 34, at 1598.
84. Schiller, supra note 43, at 87 (citation omitted).
85. 86 CONG. REC. 4668 (1940); see also Schiller, supra note 40, at 424–25 (“[T]he congressional debates were full of . . . accusations that administrative agencies were being used to advance the totalitarian ambitions of . . . Franklin Roosevelt.”); Shepherd, supra note 34, at 1606, 1609–10 (explaining that “supporters explicitly intended the bill to control Roosevelt’s authority”).
Republican Congressman Michener, quoting Congressman Walter, said that the rise of administrative government paralleled “the developments in Europe, where control of governments by men who usurped the laws has culminated in the dictatorships which now hold much of that continent in their grasp.”\textsuperscript{86} He further asserted that allowing agencies to issue regulations that have the force of law without also providing for judicial review “rapidly approach[es] the totalitarian state.”\textsuperscript{87} South Dakota Republican Congressman Mundt topped it off when he said that “[n]o one interested in genuine self-government and the liberal concepts of the American system whereby the individual citizen is safeguarded from discrimination and dictatorial acts by powerful interests, political or economic, can fail to support the Walter–Logan bill.”\textsuperscript{88} Opponents of the bill answered in equally hyperbolic terms,\textsuperscript{89} if not as frequently, accusing the bill’s proponents of creating a “judicial fascist”\textsuperscript{90} and of being supported by “the utilities fascists, the most deadly enemy to economic democracy this country has ever seen.”\textsuperscript{91} The congressional pugilists in this “bruising political brawl”\textsuperscript{92} found perhaps their only repose in their common desire to avoid hampering military readiness, but they could not agree on the extent to which the bill would do so, despite the ultimate


\textsuperscript{87} Id. Texas Democratic Congressman Summers doubted that “either Hitler or Mussolini would consent to have their acts reviewed by a court proceeding under the provisions of law and the rules of evidence;” Id. at 13,811 (1940). Wisconsin Republican Congressman Hawks said that the Securities and Exchange Commission’s authority to investigate statutory and regulatory violations “call[s] to the mind stories of the tyrannies of the Gestapo of Germany, or the Russian [OGPU, the Soviet Union’s internal security force between 1922–1934].” Id. at 4603 (1940). And Utah Democratic Senator King branded the bill’s opponents socialists. Id. at 13,662 (1940). The Senate Judiciary Committee’s Report identified the bill’s purpose as reversing the nation’s drift into parliamentarism which, if it should succeed in any substantial degree in this country, could but result in totalitarianism with complete destruction of the division of governmental power between the Federal and State Governments and with the entire subordination of both the legislative and judicial branches of the Federal Government to the executive branch wherein are included the administrative agencies and tribunals of that Government.

S. Rep. No. 76-442, at 5 (1939); see also id. at 9; H.R. Rep. No. 76-1149, at 6, 7 (1939) (espousing the bill as necessary to prevent autocracy).

\textsuperscript{88} 86 Cong. Rec. 13,813 (1940).

\textsuperscript{89} Shepherd, supra note 34, at 1592; see also id. at 1593 (“[B]oth sides in the administrative reform debate expressed real fears of dictatorship and communism”); id. at 1611.

\textsuperscript{90} 86 Cong. Rec. 4530 (1940) (statement of Rep. Rankin (D-Miss.)).

\textsuperscript{91} Id. at 4595.

\textsuperscript{92} Shepherd, supra note 34, at 1596.
inclusion in the Walter–Logan bill of a broad military exemption.

Logan’s bill, as originally submitted, required publication of regulations in the Federal Register, provided for administrative appeals boards in executive agencies and judicial review in the courts of appeals, and enunciated a standard of review that is not dissimilar from that of the current APA. Like the ABA’s proposal, the bill originally exempted “any matter concerning or relating to the conduct of foreign affairs; the conduct of military or naval operations in time of war or civil insurrection,” and a laundry list of agencies and specific types of administrative decisions.\footnote{93} Rep. Celler later explained that “[t]hose bureaus that yelled most loudly got their answer in exemptions, and those bureaus that did not yell too loudly did not.”\footnote{94}

The War Department complained that the bill would be “gravely subversive of military discipline in all components of the Army, destructive of efficiency in the performance of the functions of the War Department, both military and non-military, obstructive to progress in preparedness for national defense, and generally disastrous from the viewpoint of the public interest.”\footnote{95} In particular, the Department objected that the bill would allow military personnel to challenge orders “on any occasion except in time of war or insurrection.”\footnote{96} The Department therefore suggested that “all matters concerning or relating to the operations of the War Department and the Army” be exempted.\footnote{97}

Logan submitted an amended version of the bill in May 1939, which defined agency for the first time, removed the exemption for foreign affairs, and amended the military exemption to delete the reference to “time of war or civil insurrection,” leaving a broad exemption for “the conduct of military or naval operations.”\footnote{98} The Senate Judiciary Committee approved the amended bill unanimously;\footnote{99} its report did not mention the military

\footnote{93. S. 915, 76th Cong. § 6(b) (1939).}
\footnote{94. 86 Cong. Rec. 4547 (1940). McNollgast points out that the list of exempted agencies “included nearly all agencies created before 1933 under Republican administrations and thus more likely to be serving interests favored by Republicans.” McNollgast, supra note 5, at 197.}
\footnote{95. Letter from Harry H. Woodring, Sec’y of War, to Rep. Hatton W. Summers, Chairman, Judiciary Comm. (May 6, 1939), reprinted in Hearings on H.R. 4236, H.R. 6198, and H.R. 6324: Bills to Provide for the More Expeditions Settlement of Disputes with the United States, and for Other Purposes Before the Subcomm. No. 4 of the H. Comm. on the Judiciary, 76th Cong. 102 (1939).}
\footnote{96. Id. at 102.}
\footnote{97. Id.}
\footnote{98. S. 915, 76th Cong. §§ 1(3), 7(b) (1939).}
\footnote{99. See Shepherd, supra note 34, at 1601 (explaining why “[e]ven the committee’s eleven Northern Democrats approved the bill”).}
exemption.100 Logan’s bill passed the Senate a few days later while its opponents were absent,101 but a motion to reconsider passed the following day, and debate was scheduled for early 1940.102

Like Logan’s Senate bill, Walter’s House bill exempted “the conduct of military or naval operations.”103 The House Judiciary Committee reported Walter’s bill favorably on July 13, 1939.104 The majority report did not mention the military. The minority report, however—authored by Congressman Celler, who had introduced the ABA’s bill, but changed his mind about its wisdom and ultimately voted against it105—criticized the bill for failing to include a broader exemption for “other activities” of the Departments of War and the Navy “which highly affect public interest and the national defense, such as river and harbor improvements, and purchase of munitions and supplies.”106

The House debated the bill from April 15–18, 1940, just days after Hitler invaded Denmark. On the third day of debate, some congressmen suggested that the military exemption was too narrow.107 On the final day of debate, Congressman Walter proposed to amend the exemption for military “operations” to include “strictly military and naval activities of the War and Navy Departments.”108 He explained that the War Department had requested the amendment “because the word ‘establishments’ has a well-known technical meaning.”109 In Congressman Walter’s view,

100. S. REP. NO. 76-442 (1939).
101. Shepherd, supra note 34, at 1603.
102. Id.
103. H.R. 6324, 76th Cong. § 7(b) (3d Sess. 1939).
105. H.R. 4236, 76th Cong. (1939) (introducing the ABA bill); 86 Cong. Rec. 4744 (1940) (voting against the ABA bill); see also Shepherd, supra note 34, at 1604 (“Celler’s transformation is consistent with the Roosevelt administration’s having convinced him that the Walter–Logan bill would hinder New Deal programs; Celler otherwise firmly supported the New Deal.”).
106. H.R. REP. NO. 76-1149, pt. 2, at 5 (1940) (Minority Rep.); see also id. at 6 (“[I]t would be manifestly inappropriate to require the War Department to conduct hearings on Army regulations.”).
107. See 86 Cong. Rec. 4653 (1940) (statement of Rep. McGranery) (suggesting that the military exemption would allow military officers to contest promotion decisions in the courts of appeals and substitute the court’s judgment for that of the Army or Navy). One of the bill’s supporters, Republican Congressman Gwynne, defended the bill’s exemptions, including the military exemption. Id. at 4649 (statement of Rep. Gwynne) (“We know that under the Constitution Congress declares war, but the actual conduct of the armies and the navies is an executive function and Congress and the courts have very little, if anything, to do with it.”).
108. Id. at 4725 (statement of Rep. Walter).
109. Id.; see also id. at 4726 (statement of Rep. May) (accepting this amendment, but also suggesting express inclusion of the War and Navy Departments); id. at 4727 (statement of
however, the exemption, even as amended, would not cover the military’s “civil operations.” 110 Other members of the House argued strenuously that all operations of the War and Navy Departments, including work on rivers and harbors, should be excluded from the bill, and Rep. Celler expressed his view that the term establishments would do so. 111 Others agreed with Celler’s interpretation. 112 Rep. Hobbs then stated that Walter “certainly did not intend to give the impression that the word ‘establishments’ was not all-inclusive of the present activities of the War and Navy Departments, nor even of those that are not primarily and wholly military or naval functions.” 113 The amendment then passed. 114 Rep. Keller’s proposal to title the bill “[t]he lawyers’ emergency relief bill to end unemployment in the legal profession, and for no other purpose” was defeated, and the bill passed by a vote of 282 to 96. 115

In the summer of 1940, war became a more pressing concern as Hitler marched into Paris and began the London Blitz. Roosevelt won reelection to a third term in November. 116 Shortly after election day, the Senate Committee took the military amendment a step further and deleted “the conduct of” such that the military exemption provided: “Nothing contained in this Act shall apply to or affect any matter concerning or relating to the military or naval establishments.” 117 On November 26, 1940, the Senate accepted the Committee’s amendment without discussion. 118 The Senate also accepted without discussion Senator Hatch’s amendment adding to the military exemption “any other agency or authority hereafter created to expedite military and naval defense.” 119 The bill passed the same day by a vote of 27 to 25, much closer than the House’s lopsided vote in the spring. 120

Rep. Hobbs (arguing that establishments covers “the complete functioning of all parts of the Military and Naval Establishments of Uncle Sam”).

110. Id. at 4725 (statement of Rep. Walter).

111. Id. at 4726–27 (statements of Reps. Celler, May, and Bulwinkle).

112. Id. at 4727 (statement of Reps. Hobbs and May).


114. Id. at 4728.

115. Id. at 4742, 4743–44 (1940). “[A]ll but two Republicans, 83% of Southern Democrats, and 41% of Northern Democrats” voted for the bill. Shepherd, supra note 34, at 1619.

116. Id. at 1622.

117. H.R. 6324, 76th Cong. § 7(b) (3d Sess. 1940).

118. 86 Cong. Rec. 13,746–47 (1940).

119. Id. at 13,747.

120. Id. at 13,747–48. Every voting Republican senator and ten conservative Democrats, both Northern and Southern, voted in favor of the bill. Shepherd, supra note 34, at 1622.
Back in the House on December 2, 1940, Representative Cochran said that Senator Hatch’s amendment was “a legislative afterthought . . . designed to prevent the Walter–Logan bill from hampering the national-defense program” that “falls far short of accomplishing this purpose” because it would not cover civilian agencies “performing functions which are indispensable to the workings of our defense program.”\footnote{86 Cong. Rec. 13,810 (1940) (statement of Rep. Cochran).} Congressman Sumners, on the other hand, believed that the military exemption included not just the military, but “covers everything that may be done by any agency concerning or relating to the Military and Naval Establishments.”\footnote{Id. at 13,811 (1940) (statement of Rep. Sumners).} The House concurred in the Senate amendments by a vote of 176 to 51.\footnote{Id. at 13,815–16. Apparently, many of the 203 House members who did not vote, including Rep. Celler, were absent. Id.}

President Roosevelt vetoed the Walter–Logan bill on December 18, 1940.\footnote{Id. at 13,942–43; Shepherd, supra note 34, at 1625.} On one hand, he did not think the bill went far enough in enabling administrative agencies to resolve disputes so as to avoid litigation. He saw the bill as “one of the repeated efforts by a combination of lawyers who desire to have all processes of Government conducted through lawsuits and of interests which desire to escape regulation.”\footnote{Id. at 1625.} On the other hand, he felt that the bill imposed too much of a burden on national defense. He noted that affected agencies, “including many whose activities have an important collateral effect on the defense program, have pointed out serious delays and uncertainties which would be caused by the present bill.”\footnote{Franklin Delano Roosevelt, The President Vetoes the Bill Regulating Administrative Agencies, Note to the House of Representatives, Dec. 18, 1940, in THE PUBLIC PAPERS AND ADDRESSES OF FRANKLIN D. ROOSEVELT 619 (1941).} Roosevelt acknowledged that the bill exempted “agencies engaged in National Defense functions,” but found the bill flawed in that it would subject other agencies, like the Maritime Commission and the Departments of Commerce and Treasury, to delay when engaging in defense-related functions.\footnote{Id. at 620.} “Quite apart from the general philosophy of this Bill,” he concluded, “its unintentional inclusion of defense functions would require my disapproval at this time.”\footnote{Id. at 620–21.} Later that day, the House failed to override the President’s veto.\footnote{Id. at 621.}

\footnote{86 Cong. Rec. 13,953 (1940). While 113 of 115 Republicans voted to override, “many conservative Democrats now defected from their coalition with Republicans.” Shepherd, supra note 34, at 1630.}
3. The War Years

As he mentioned in his veto of the Walter–Logan bill, President Roosevelt had asked the Attorney General to form a committee to study administrative reform in 1939, about a month after Logan introduced the ABA bill in the Senate. The President’s suggestion in his veto message that legislation should await the committee’s report “may have swayed some members of Congress” who voted for the bill to vote against an override. The Attorney General’s Committee on Administrative Procedure submitted its report on January 22, 1941, about a month after the override vote failed. Despite President Roosevelt’s concern that the Walter–Logan bill, even with its broad exemption for “any matter concerning or relating to the Military or Naval Establishments,” would have hampered military readiness, the Attorney General’s Committee did not recommend special treatment for the military. Instead, the Committee said that its recommendations “relating to delegation; . . . administrative information; . . . informal methods of adjudication; and . . . rule-making procedures are applicable to the War Department.” Indeed, the only substantial discussion of the military in the Committee’s report concerned the War Department’s civil jurisdiction over navigable waterways and toll bridges.

The report included two draft bills, one favored by the eight-member liberal majority and the other favored by the four-member conservative minority. Senator Hatch introduced both bills in the Senate on January 29, 1941. While the majority bill “imposed little restraint on agencies,” the minority bill “would have controlled agencies substantially, but not as strictly as the Walter–Logan bill.” Interestingly, it was the minority bill that provided broader exemptions for the military. Both bills would have

131. Shepherd, supra note 34, at 1594.
132. Id. at 1631 (citation omitted).
133. COMM. ON ADMINISTRATIVE PROCEDURE, Administrative Procedure in Government Agencies, S. Doc. No. 77-8 (1941).
134. H.R. 6324, 76th Cong. § 7(b) (3d Sess. 1940); S. Doc. No. 77-8, at 155.
135. Id. at 155–57.
136. Shepherd, supra note 34, at 1632.
137. S. 675, 77th Cong. (1941); S. 674, 77th Cong. (1941). See also S. 918 and H.R. 3464, 77th Cong. (1941), which “combined the most restrictive sections of the Walter–Logan bill and the minority bill.” Shepherd, supra note 34, at 1636. Hatch introduced S. 918 on February 13, 1941, and Congressman Walter introduced H.R. 3464 on February 18, 1941. S. 918, 77th Cong. (1941); H.R. 3464, 77th Cong. (1941). Section 900(b) of these bills exempted “the conduct of the Military or Naval Establishments” from all of their provisions. S. 918 § 900(b); H.R. 3464 § 900(b).
138. Shepherd, supra note 34, at 1633–34.
created an Office of Federal Administrative Procedure to review practices and procedures of executive agencies, required promulgation and publication of regulations, and mandated procedures for administrative hearings. Both defined “agency” to include executive branch agencies, but neither exempted the military from that definition. Instead, both bills exempted from their adjudication provisions “the conduct of the Military or Naval Establishments, or the selection or procurement of men or materials for the armed forces of the United States.”\(^\text{139}\) The bills’ commonalities end there. The majority bill included no military exemption from its rulemaking provisions; the minority bill, on the other hand, authorized the President to temporarily suspend any of the act’s provisions under certain circumstances.\(^\text{140}\) In addition, Title II of the minority bill, which required notice and comment rulemaking “whenever practicable,”\(^\text{141}\) provided:

> Whenever expressly found by an agency to be contrary to the public interest, the provisions of this title, in whole or in part, shall not apply to (a) the conduct of military, naval, or national-defense functions, or the selection or procurement of men or materials for the armed forces of the United States\(^\text{142}\)

In the spring of 1941, as Germany prepared to invade the Soviet Union, the Senate Judiciary Committee held hearings.\(^\text{143}\) Captain Karl R. Bendetson from the Office of the Judge Advocate General appeared for the War Department. He opined that that the military “could not properly function” if it had to comply with the proposed procedural requirements and that the bills’ military exemptions were “not sufficiently broad to provide a complete exemption.”\(^\text{144}\) The Secretary of War’s written statement even questioned Congress’s constitutional authority to impose on the President’s “command function” statutory requirements “which would gravely impair” military efficiency.\(^\text{145}\) The word conduct, Bendetson said, was “restrictive,” and the term military establishment could be read to not cover the War Department.\(^\text{146}\) Bendetson objected in particular to the provisions requiring publication of rules, which “might conceivably cover

\(^{139}\) S. 675, 77th Cong. § 301(d) (1941); S. 674, 77th Cong. § 301(e) (1941).

\(^{140}\) S. 674, 77th Cong. § 111 (1941).

\(^{141}\) Id. § 208.

\(^{142}\) Id. § 201.

\(^{143}\) Administrative Procedure: Hearings on S. 674, S. 675, and S. 918 Before a Subcomm. of the S. Comm. on the Judiciary, 77th Cong. (1941).

\(^{144}\) Id. at 36.

\(^{145}\) Id. at 48.

\(^{146}\) Id. at 46.
any type of regulation whether adopted in the field or not.”147 The War Department wanted Congress not just to clarify the exemption of military functions, but also to provide a “full exemption” for the War Department’s civil functions,148 specifically its “civil jurisdiction over navigable waters” which Bendetson said “is so closely allied to the national defense that it partakes of the same character.”149 Bendetson proposed amending the bills to provide a “complete exclusion” of “every function” of both the War and Navy Departments by specifying that they “shall have no application to” the War and Navy Departments, including the Army, the Marine Corps, and the Coast Guard “when serving under the jurisdiction of the Navy Department.”150

Other witnesses rejected Bendetson’s suggestion. The Chairman of the ABA Committee responded that the military “should not be completely free of judicial review” and in fact “has not been since 1853 when the Court of Claims was established.”151 The ABA committee’s proposed bill at that time exempted from its judicial review provisions “any case involving military or naval operations in time of war.”152 Assistant Secretary of State and Chairman of the Attorney General’s Committee Dean Acheson opined that the majority bill’s exceptions for “the military service, the armed forces, or the selection and discharge of employees” were “very clear.”153 Carl McFarland, a member of the Attorney General’s Committee and proponent of the stricter minority bill, objected to the “exemption of any agencies, as such, since they almost all exercise certain functions which, as a matter of principle, should be governed by the mild requirements of these proposals.”154 In particular, McFarland disagreed with the War Department’s request that it be “bodily exempted . . . even beyond the present exemption of its military functions.”155

The United States declared war on Japan on December 8, 1941, and work on administrative reform took a back seat. Several changes during

147. Id. at 37.
148. Id. at 38.
149. Id. at 45.
150. Id.
151. Id. at 961.
152. Id. at 995; see also 66 Annual Report of the American Bar Association 451 (“This committee believes that there can be no compromise on the necessity and desirability of there being in the courts ultimate and final judicial authority to fully review any and all administrative decisions of whatever character which the Congress does not specifically exempt from such review for reasons of state or military policy.”).
154. Id. at 1349.
155. Id.
the war years “paved the path to the APA.”156 Both sides of the administrative reform debate became more flexible during the war years, and “[v]ituperation . . . [went] out of style.”157 Congressional Democrats grew politically weaker during the war. In the 1942 midterm election, Republicans gained nine seats in the Senate, leaving Democrats with a twenty-one-seat majority, but they picked up forty-seven seats in the House, leaving Democrats with only a nine-seat majority.158 President Roosevelt remained enormously popular,159 but retreated from the New Deal “to ensure industrialists’ cooperation” in the war effort.160 That retreat may have deflated the drive for strict administrative controls among the anti-New Deal conservatives who had supported the Walter–Logan bill. In addition, the President’s judicial appointments had shifted the federal bench to the left.161 The new liberal judiciary was less likely to strike down New Deal programs, making broad judicial review of administrative decisions less attractive to conservatives and less feared by liberals.162 The ABA backed off of its previously “combative approach” and gave “an olive branch to Roosevelt” by appointing Carl McFarland chair of its Special Committee on Administrative Law.163 Francis Biddle, who had served with McFarland on the Attorney General’s Committee on Administrative

---

156. Shepherd, supra note 34, at 1641.
157. See Gellhorn, supra note 35, at 229–30 (“Seemingly, all concerned heard the message that agitated advocacy was no longer appropriate.”).
158. Shepherd, supra note 34, at 1643.
159. Roosevelt’s approval rating hit 84% immediately following Pearl Harbor and remained high until his death. See Roper Center, Job Performance Ratings for President Roosevelt, http://webapps.ropercenter.uconn.edu/CFIDE/roper/pres presidency/presidential_rating_detail.cfm?allRate=True&PresidentName=Roosevelt (last visited August 1, 2010).
160. Shepherd, supra note 34, at 1643.
161. See Schiller, supra note 40, at 404 (“[B]y the end of the 1930s there had been a dramatic change in the relationship between courts and the administrative state. Courts were placed in a position frankly subservient to the administrators whose task it was to rationalize and reform the failing economy through the application of scientific expertise.”).
162. Shepherd, supra note 34, at 1643–44; McNollgast, supra note 5, at 183, 191. Roosevelt made 204 judicial appointments. See Judiciaries Appointments by President, http://www.uscourts.gov/JudgesAndJudgeships/Viewer.aspx?doc=/uscourts/JudgesJudgeships/docs/appshypres.pdf (last visited June 21, 2010). Although that number is far less than modern presidents like Reagan, Clinton, and Bush II, it represents a significant proportion of the then-smaller federal judiciary. When Roosevelt took office in 1933, there were 211 Article III judges. See Authorized Judiciaries, http://www.uscourts.gov/JudgesAndJudgeships/Viewer.aspx?doc=/uscourts/JudgesJudgeships/docs/allauth.pdf, at 5 (last visited August 1, 2010). When he died in 1945, the judicial ranks had swelled to 262 judges. Id. at 6. Thus, Roosevelt replaced a large proportion of the federal judiciary in his years in office. In 2009, there were 864 Article III judgeships. Id. at 8.
163. Shepherd, supra note 34, at 1645–46.
Procedure, became Attorney General in 1941. Professor Shepherd explains that McFarland and Biddle “continued a cooperative relationship that they had established on the Attorney General’s Committee.”

In addition, the size of the federal bureaucracy increased dramatically, as twenty-six new agencies were created to guide the nation through the war. Federal agencies gained broad new powers and sometimes “blundered,” revealing to the public “the abuses and irritations that agencies could cause.” For example, the Office of Price Administration, which had the authority among other things to fix prices and control rents, rationed over 90% of consumer goods by the end of the war. Inflation and chronic shortages were blamed on federal agencies. The war demonstrated that “agencies could be inefficient, incompetent, bullying, and perhaps even captured by the interests they were supposed to regulate.” The ABA Committee’s 1943 report explained, “[w]ar has complicated and aggravated the problems of administrative law, particularly federal administrative law. The impact of administrative regulation has vastly increased in both degree and in scope.” The Committee believed that the war, “fought for freedom and the dignity of the individual,” had made citizens “more keenly aware” that their “freedom and rights lie under the pall of a war emergency reflected in operations of the federal administrative establishment.”

Professor Schiller posits that the war bureaucracy “weakened Americans’ faith in expertise.” At the same time, the belief that administrative power could pave the road to totalitarianism gained prominence “across the political spectrum and had even entered mainstream culture.” Those factors, coupled with growing economic prosperity, led to a shift in liberal reformers’ belief that government regulation would solve society’s

164. *Id.* at 1647.
165. *Id.*
166. Schiller, *supra* note 39, at 190.
167. Shepherd, *supra* note 34, at 1641–42.
170. Schiller, *supra* note 39, at 195; *see also* *Id.* at 201 (“Too often American wartime agencies had the appearance of incompetent bullies, captured by special interests, acting with autocratic disregard of due process.”).
171. 68 *ANNUAL REPORT OF THE AMERICAN BAR ASSOCIATION* 249 (1943).
172. *Id.* at 250.
173. Schiller, *supra* note 43, at 93; *see also* *Id.* at 95 (“[A]dministrative expertise, once a powerful rationale for exempting agencies from judicial oversight, became nothing more than an excuse for frightening excesses of governmental power.”); Schiller, *supra* note 39, at 188.
problems.\textsuperscript{175} Consequently, after World War II, “[t]he claim that imposing
the rule of law on agency behavior could protect Americans from an
administrative state run amok . . . was increasingly heard across the
political spectrum.”\textsuperscript{176}

The Walter–Logan bill demonstrated that, at least in Congress, the
military was generally exempt from these concerns before the war. But that
changed during the war. By the time Congress passed the APA in 1946, it
was still hesitant to impose procedural constraints on the military, but its
willingness to leave the military free of judicial oversight had waned,
perhaps due to the militaristic regimentation of civilian life or exposure to
the abuses of Europe’s fascist armies.\textsuperscript{177}

B. The APA of 1946

1. Introduction and Passage

Two weeks after D-Day in 1944, Senator McCarran and Representative
Summers introduced the bill that would eventually become the APA.\textsuperscript{178}
The ABA Committee designed the bill as a compromise between the
Attorney General’s Committee’s majority bill, which the ABA Committee
believed was “designed to confirm administrative practices,” and the
minority bill, which the ABA Committee believed was “too long and
prolix.”\textsuperscript{179} The bill included public information, notice and comment
rulemaking, and adjudication provisions; it specified parties’ rights in
formal rulemakings and adjudications and provided for judicial review.\textsuperscript{180}
“The introduction of these bills brought forth a volume of further
suggestions from every quarter.”\textsuperscript{181}

McCarran and Summers revised and reintroduced the bill in January

\textsuperscript{175} Id. at 75–76, 84, 95.

\textsuperscript{176} Schiller, supra note 49, at 1405; see also id. at 1409 (“[B]y the end of World War II, a
consensus had developed that the judiciary should take a more active role in policing
agencies than it had during the New Deal.”); Schiller, supra note 43, at 102 (“It was events
that occurred in the 1940s—particularly the rise of fascism in Europe and the
disillusionment with the administration in the United States—that resulted in the judiciary
taking on an institutional role as the protector of civil liberties.”); Schiller, supra note 39, at
190.

\textsuperscript{177} See Schiller, supra note 39, at 189.

\textsuperscript{178} S. 2030, 78th Cong. (1944); H.R. 5081, 78th Cong. (1944).

\textsuperscript{179} 68 ANNUAL REPORT OF THE AMERICAN BAR ASSOCIATION 256 (1943); Shepherd,
supra note 34, at 1649–50.

\textsuperscript{180} See Shepherd, supra note 34, at 1650–52.

\textsuperscript{181} S. REP. NO. 79-752 (1945), reprinted in ADMINISTRATIVE PROCEDURE ACT:
LEGISLATIVE HISTORY, 1944–46, at 190 (1946) [hereinafter LEGISLATIVE HISTORY].
1945. Like their original bill, § 3 and § 4 of the revised bill exempted “any military, naval, and diplomatic function of the United States” from its public information and rulemaking provisions. The bill also included a blanket exemption for temporary wartime functions: “functions which by law expire on the termination of present hostilities, within any fixed period thereafter, or before July 1, 1947.” The adjudication and judicial review provisions contained no military exemption.

Henry L. Stimson, the Secretary of War, wrote to Pat McCarran, chair of the Senate Judiciary Committee, on February 15, 1945, in response to the Committee’s request for comments on the revised bill. Stimson noted that the “military function” exceptions had “no precise statutory meaning.” He also expressed concern that courts-martial might be subject to the adjudication and judicial review provisions and that various other provisions in the bill “would be ruinous if made applicable to the War Department and the Army.” Accordingly, Stimson suggested adding to the bill a blanket exemption for the War Department, the Army, the Navy, and “the selection or procurement of personnel or materiel for the armed forces of the United States.”


183. The rulemaking provisions of § 4 included the clause: “Except to the extent that there is directly involved any military, naval, or diplomatic function of the United States . . . .” The public information provisions of § 3 included the same clause with the following addition at the end: “requiring secrecy in the public interest.” H.R. 1203, 79th Cong. §§ 3, 4 (1945); S. 7, 79th Cong. §§ 3, 4 (1945). Congressman Gwynne introduced an earlier, slightly stricter version of the same bill several months before McCarran and Summers introduced their bill. See Shepherd, supra note 34, at 1633. Gwynne’s bill included the same military exemption in its public information and rulemaking provisions. See H.R. 4314, 78th Cong. §§ 2, 6 (1944). Congressman Smith also introduced a stricter version of the bill in 1944, which, “[r]eflecting popular anger at the agencies that had arisen during the war . . . eliminated the McCarran-Summers bill’s exemption of wartime agencies.” Shepherd, supra note 34, at 1633–34. Yet even Smith’s bill included the same military exemption in its public information and rulemaking provisions. See H.R. 5237, 78th Cong., §§ 2, 3 (1944).

184. H.R. 1203, 79th Cong. § 2(a) (1945); S. 7, 79th Cong. § 2(a) (1945).

185. Also introduced in 1945 were H.R. 104, 339, 1117, 1206, and 2602, 79th Cong. (1945), all of which contained some sort of military exemption.


187. Id. at 2.

188. See id. at 2–3.

189. Id. at 4.

190. Stimson recommended adding the following language:

The provisions of this statute shall not apply to the War Department, the Army of the United States, the Navy Department, or the United States Navy (including the
President Roosevelt died on April 12, 1945, three months into his fourth term, and Harry Truman became President. Professor Shepherd explains that “Truman supported administrative reform with marginally greater fervor than had Roosevelt.”191 Truman had an in-depth understanding of the United States’ war effort, in part from his experience as chair of the Senate Special Committee to Investigate the National Defense Program, which examined the nation’s war preparations,192 but he was politically weaker than Roosevelt.193 Indeed, when Congress passed the bill in the spring of 1946, Truman was engaged with the national railroad strike. “Truman could devote neither attention nor political resources to the APA.”194 Meanwhile, public support for administrative reform continued to grow after the war ended in Europe on May 8, 1945 and in Japan on August 14, 1945, as federal agencies’ reconversion to a peacetime economy faltered.195

Roosevelt’s death may also have given New Deal Democrats in Congress “the incentive to consolidate the gains of the New Deal thus far” as they “realized that their prospects for retaining the presidency were growing increasingly dim.”196 McNollgast explains that the prospect of a Republican executive led congressional Democrats to “favor procedural restraints on agency action” because such constraints “would blunt any republican president’s ability to dismantle or shift the regulatory policies of the New Deal.”197 For their part, congressional Republicans may have continued to support the bill because it “would slow the adoption of further New Deal regulations” and could be altered if Republicans won the presidency and majorities in Congress in the next election.198

United States Marine Corps and the United States Coast Guard when operating under the control of the Navy], or to the selection or procurement of personnel or materiel for the armed forces of the United States.  

Id. at 4.  
191. Shepherd, supra note 34, at 1658.  
192. See S. Res. 71, 77th Cong. (1941); 87 Cong. Rec. 1615 (1941).  
193. Shepherd, supra note 34, at 1658. Although Truman’s approval rating was high when he became President, it dipped rapidly during his first year in office. By the time he signed the APA in June, 1946, his approval rating was 45% and dropping. See Roper Center, Job Performance Ratings for President Truman, http://webapps.ropercenter.uconn.edu/CFIDE/roper/presidential/webroot/presidential_rating_detail.cfm?allRate=True&PresidentName=Truman [last visited Aug. 1, 2010]. Republicans took control of both houses of Congress in the 1946 mid-term election.  
194. Shepherd, supra note 34, at 1659.  
195. See id. at 1658.  
196. McNollgast, supra note 5, at 182–83; see also id. at 190–91.  
197. Id. at 192; see also id. at 203.  
198. Id. at 194–95. Alan Schwartz criticizes McNollgast for failing to explain why Republicans did not stall until the 1948 election and posits that “the contribution of the
Less than a month after Roosevelt died, on May 5, 1945, H. Struwe Hensel, formerly a partner at Milbank, Tweed in New York, then the Navy’s first General Counsel and, at the time, Assistant Secretary of the Navy, wrote to Senator McCarran as Acting Secretary that the bill “is so drafted as to cause difficulties to this Department which are wholly disproportionate to the intended benefits to the public.”199 “To the extent that it is possible to construe the bill with any reasonable degree of certainty,” he continued, “it appears to be, in many instances, affirmatively prejudicial to the operations of the Navy Department.”200 In Hensel’s view, the “fundamental difficulty” with the bill was the lack of certainty as to which of the Navy’s functions would be covered.201 In particular, Hensel was unsure whether the exceptions for naval functions would cover all of the Navy’s activities or “only those directly related to Navy ships.”202 Hensel also expressed concern that the adjudication provisions would impact courts-martial and the judicial review provisions would subject the Navy’s actions to de novo review in court.203 Unlike the War Department, the Navy did not suggest any amendments to fix those problems, but instead “urgently” recommended against the bill’s enactment.204

Following private negotiations with representatives of the Attorney General, the ABA, and others, the Senate Judiciary Committee issued a proposed revision of the bill in May 1945.205 The Senate Committee proposed retaining the military exemptions in the public information and rulemaking provisions in §§ 3 and 4; adding an exemption to the adjudication provision in § 5 for “the conduct of military, naval, or foreign affairs functions”; moving the exemption for temporary wartime functions from § 2 to a new § 13; and adding to that exemption the Selective Service Act and other specific statutes.206 The Senate Committee received further comments and, in June 1945, issued a second committee print with columns showing the provisions of the original bill, the Committee’s

---

200 Id.
201 Id.
202 Id. at 2.
203 See id. at 3, 5, 10.
204 See id. at 10.
205 See LEGISLATIVE HISTORY, supra note 181, at 11, 191.
206 Id. at 15, 17, 21, 43.
proposed revisions, explanations of the provisions, and a summary of comments received.\footnote{207} The Secretary of War wrote to Senator McCarran again on June 13, 1945, supporting the addition of a “military function” exception to the adjudication provisions, but again expressing concern about the imprecision of those terms.\footnote{208} Stimson’s “special concern,” however, was the lack of a military exception in § 10’s judicial review provisions.\footnote{209} He suggested that “at least” § 10 should include a “military function” exception similar to those in the other sections,\footnote{210} but reiterated the proposal in his February letter that the War Department, the Army, and the Navy be exempted from the bill entirely.\footnote{211} The Senate Committee declined that proposal.

Among the other comments the Committee received and responded to in the June 1945 Committee Print was a recommendation that it add to § 13 “courts[martial], military or naval authority exercised in the field in time of war or in occupied territory.”\footnote{212} The Committee commented that “[t]his may properly be done to remove any question of the application of the measure to purely military functions.”\footnote{213} Although the legislative history does not indicate who made that suggestion, a handwritten note on a copy of the Committee Print in the bill file at the National Archives in Washington, D.C., indicates that this amendment may have come from the War Department.\footnote{214} Since much of the bill’s development took place behind closed doors, it would not be surprising if the War Department suggested that language as a compromise. Another commenter “strongly urged” that “there should be no exemption of war functions except those relating to courts-martial and the authority of the Army and Navy”—in other words, that the exemption in § 13 be limited to “courts[martial],

\begin{footnotes}
\footnote{207} Id. at 11–44, 191.
\footnote{208} See Letter from Henry L. Stimson, Sec’y of War, to Hon. Pat McCarran, Chairman, S. Comm. on the Judiciary 1 (June 13, 1945) (on file at Record Group 46, Records of the U.S. Senate, 79th Cong.,Senate 79A-E1 Box 4, National Archives, Washington, D.C., and on file with the author).
\footnote{209} Id. at 1–2.
\footnote{210} Id. at 2.
\footnote{211} See id. at 3; Letter from Henry L. Stimson, Sec’y of War, to Sen. Pat McCarran, Chairman, S. Comm. on the Judiciary (Feb. 15, 1945) (on file at Record Group 46, Records of the U.S. Senate, 79th Cong., Senate 79A-E1 Box 4, National Archives, Washington, D.C., and on file with the author).
\footnote{212} LEGISLATIVE HISTORY, supra note 181, at 44.
\footnote{213} Id.
\footnote{214} See S. COMM. ON THE JUDICIARY, 79TH CONG., AGENCY COMMENTS ON REVISED TEXT OF COLUMN 2 22 (Comm. Print 1945) (on file in Record Group 46, Records of the U.S. Senate, 79th Cong., Senate 79A-E1 Box 6, National Archives, Washington, D.C.).
\end{footnotes}
military or naval authority exercised in the field in time of war or in occupied territory” and not include temporary wartime agencies. 215 The Committee explained that “war agencies functions” were exempted in § 13 because “it would take at least a year” for an agency to revise its practices. 216 Presumably, the temporary wartime agencies would have been close to expiring at that point. 217 With regard to the exception in §§ 3 and 4 for “any military, naval, and diplomatic function of the United States” and the newly proposed exemption to the adjudication provision in § 5, the Committee commented that these exemptions were “self-explanatory” 218 and could be further clarified in committee reports. 219 Unfortunately, the committee reports did not clarify those terms.

The House Judiciary Committee held hearings in June 1945, 220 but those hearings were “a side show” to the “main act,” which entailed private negotiations with the Attorney General. 221 The only discussion of the military at the hearings concerned the exemption for temporary wartime agencies, which ABA Committee chairman Carl McFarland assured the Committee would “not be unduly injured in any way.” 222 After future Supreme Court Justice Tom Clark became Attorney General on July 1, 1945, negotiations reopened between the Senate Judiciary Committee and the interested agencies, parties, and organizations whose views were received informally rather than at public hearings. 223 By August 1945, the

215. LEGISLATIVE HISTORY, supra note 181, at 43–44. With regard to the bill’s adjudication provisions, one commenter suggested that “military functions—particularly courts—require specific exemption.” Id. at 36. The Committee responded that “[this is dealt with in the comment to Section 13.” Id.

216. Id. at 43.

217. See id. at 313 (statements of Sen. McCarran, Chairman, S. Comm. on the Judiciary) (“The pending bill does, however, in Section 2(a), exempt war agencies, because they are presumably self-liquidating, and it was deemed unwise to attempt to cover them at this late date.”).

218. Id. at 15, 17, 22.

219. See id. at 22.


221. Shepherd, supra note 34, at 1659–60.


223. See Shepherd, supra note 34, at 1661 (citing S. Rep. No. 79-752 (1945), reprinted in LEGISLATIVE HISTORY, supra note 181, at 191); see also id. at 1663 (“The bill had sprung not from public debate in Congress, as other bills had, but from months of private, off-the-record negotiations.”); 70 ANNUAL REPORT OF THE AMERICAN BAR ASSOCIATION 271 (1945) (“The Senate Judiciary Committee has a subcommittee at work on the bill and is proceeding by executive sessions, consultations, and written submittals of views.”).
Navy had backed off of its opposition to the bill somewhat. Hensel chose instead to interpret the “naval authority” exceptions to encompass “all operations under the jurisdiction of the Navy Department” and the judicial review provisions in § 10 to “create no new methods of review.”

The Senate Judiciary Committee issued another draft of the bill on October 5, 1945, § 2(a) of which “excluded from the operation of this Act . . . (2) courts[-]martial and military commissions, (3) military or naval authority exercised in the field in time of war or in occupied territory,” temporary wartime agencies, and several specific statutes. The draft also retained an independent exemption in the rulemaking provisions of § 4 for “any military, naval, or foreign affairs function of the United States”; included an exemption from the adjudication provisions in § 5 for “the conduct of military, naval, or foreign affairs functions”; deleted the military exemption from § 3’s public information requirements; and deleted the Committee’s proposed § 13.

The Senate Judiciary Committee’s report yields little insight into the intended meaning of the military authority exemption in § 2(a); the Committee’s few military-related comments concern the exemption for temporary wartime agencies. The Committee explained that § 2(a) “[e]xpressly exempted from the term ‘agency’ . . . defined war authorities including civilian authorities functioning under temporary or named statutes operative during ‘present hostilities’” and that “[t]he exclusion of war functions and agencies, whether exercised by civil or military personnel, affords all necessary freedom of action for the exercise of such functions in the period of reconversion.” The Committee further explained, however, that it exempted “functional classifications,” rather than “administrative agencies by name. Thus, certain war and defense functions [were] exempted, but not the War or Navy Departments in the performance of their other functions.” The Committee also said that it used the term “authority” in the definition of “agency” to include “whatever persons are vested with powers to act (rather than the mere form of agency organization such as department, commission, board, or bureau).

---

225. Legislative History, supra note 181, at 186.
226. Id. at 218.
227. Id. at 219, 223.
228. Id. at 196.
2010]  

**HISTORY OF MILITARY AUTHORITY EXCEPTION IN THE APA**  
703

because the real authorities may be some subordinate or semi[-]dependent person or persons within such form of organization.”  

The Committee formally submitted that draft to the Attorney General for his review. Not surprisingly, the Attorney General supported the bill, as did the ABA. The Senate Judiciary Committee passed the bill unanimously on October 29, 1945. On March 12, 1946, the Senate passed the bill on a voice vote with much back-patting, but little debate. Like the Judiciary Committee’s report, the Committee’s chairman, Senator McCarran, emphasized that the bill “followed the undeviating policy of dealing with types of functions as such and in no case dealing with administrative agencies by name.” Thus, “certain war and defense functions [were] exempted under the bill, but there [was] no exemption of the War or Navy Departments in the performance of their other functions.”

The House Judiciary Committee reported the bill unanimously, with a few minor amendments, which the Senate Judiciary Committee and the Attorney General approved. The House Report of May 3, 1946, reiterated the Senate Report’s explanation of the meaning of authority in the definition of agency: “Whoever has the authority is an agency.” The term authority was thus intended to encompass “those who have the real power to act.” The Report stated that the exemption for “war functions,” apparently referring to the “rapidly liquidating” temporary wartime agencies, was “self-explanatory,” but did not mention § 2(a)'s exemption of “military authority exercised in the field in time of war or in occupied territory.” On May 24, 1946, the House passed the bill on a voice vote,

---

231.  See id. at 223; 92 Cong. Rec. 2148 (1946); Shepherd, supra note 34, at 1661.
235.  Shepherd, supra note 34, at 1668; *Legislative History*, supra note 181, at 344.
236.  *Legislative History*, supra note 181, at 302.
237.  Id.
238.  *Legislative History*, supra note 181, at 233–91, 347; Shepherd, supra note 34, at 1669 (citation omitted).
239.  92 Cong. Rec. 5647 (1946), reprinted in *Legislative History*, supra note 181, at 349; Shepherd, supra note 34, at 1670.
241.  Id.
242.  Id.
again with little debate. On May 27, 1946, the Senate concurred on a voice vote. President Truman signed the bill on June 11, 1946.

2. The Aftermath

In the end, the definition of agency in § 2(a) narrowly “excluded from the operation of this Act . . . military or naval authority exercised in the field in time of war or in occupied territory,” and temporary wartime functions. The military authority exemption did not apply to § 3 of the Act, which concerned public information and independently exempted “any function of the United States requiring secrecy in the public interest.” The rulemaking provisions in § 4 also retained a separate exemption for “any military . . . or foreign affairs function of the United States,” and the adjudication provisions in § 5 exempted “the conduct of military, naval, or foreign affairs functions.” The broad exclusions of functions “requiring secrecy in the public interest” from § 3 and of “military functions” from §§ 4 and 5 underscore the narrowness of the exemption in § 2(a) for “military or naval authority exercised in the field in time of war.” Indeed, since §§ 6, 7, 8, 9, and 11 largely supplemented the rulemaking and adjudication provisions in §§ 4 and 5, the narrow military authority exemption in § 2(a) primarily related to the judicial review provisions in § 10. Thus, the military was largely exempted from the APA’s rulemaking and adjudication requirements, but only a narrow slice of military action was exempt from judicial review under the Act.

243. Shepherd, supra note 34, at 1670–74. The only comment that is notable here is Rep. Walter’s statement that “[p]urely military and naval functions should obviously be exempt.” LEGISLATIVE HISTORY, supra note 181, at 355. Yet, again, he appeared to direct that comment to “defined war authorities functioning under temporary or named statutes,” not the broader “military authority” exemption. Id.

244. LEGISLATIVE HISTORY, supra note 181, at 423.


246. Id. § 3.

247. Id. § 4.

248. Id. § 5.


250. If Congress believed at the time that the exception for “agency action [that] is by law committed to agency discretion,” § 10(2), 60 Stat. at 243, would insulate the military from judicial review, the legislative history does not reflect it. Congress believed this exception would codify the status quo. The Senate Report explained that the “agency discretion” exception “would apply even if not stated at the outset” where, for example, the terms of a statute are so broad that “there is no law to apply.” LEGISLATIVE HISTORY, supra note 181, at 212. On the other hand, “where statutory standards, definitions, or other grants of power deny or require action in given situations or confine an agency within limits as required by the Constitution, then the determination of the facts does not lie in agency
Although the ABA President and the Senate Judiciary Committee Chairman denied it publicly, all parties involved in the APA’s passage understood that the Act was a compromise, and none were fully satisfied.\textsuperscript{251} Several years after its enactment, the Supreme Court said that the APA “represents a long period of study and strife; it settles long-continued and hard-fought contentions, and enacts a formula upon which opposing social and political forces have come to rest. It contains many compromises and generalities and, no doubt, some ambiguities.”\textsuperscript{252} Indeed, the bill’s ambiguity was essential to its passage.\textsuperscript{253} As Professor Vermeule observes, the Congress that passed the APA employed ambiguity to reconcile conflicting desires.\textsuperscript{254} On one hand, “a major purpose of the APA was to retrench the administrative state and to reassert legislative and judicial control over administrative action.”\textsuperscript{255} On the other hand, “the drafters of the APA had just lived through a global hot war and were on the verge of a global cold one,” and “[e]xecutive power was, perhaps, near a kind of local maximum.”\textsuperscript{256} Thus, “[t]he framers of the APA quite deliberately left escape hatches from the administrative code of legal liberalism, recognizing that unforeseen and emergency circumstances would inevitably arise, and that no code of administrative law and procedure could hope to specify, in advance, what to do about those circumstances.”\textsuperscript{257} It was neither possible
discretion” but is subject to judicial review. Id.; see also id. at 275 (H.R. Rep. No. 79-1980 (1946) reprinted in LEGISLATIVE HISTORY, supra note 181, at 275). The Attorney General agreed that this exception “declares the existing law concerning judicial review.” LEGISLATIVE HISTORY, supra note 181, at 229; see also U.S. DEPT OF JUSTICE, ATTORNEY GENERAL’S MANUAL ON THE ADMINISTRATIVE PROCEDURE ACT 94–95 (1947), available at http://www.law.fsu.edu/library/admin/1947ix.html [hereinafter ATTORNEY GENERAL’S MANUAL] (restating the same assertion and discussing the scope of § 10). Hence, the “agency discretion” exception would not insulate military action from judicial review where that action was governed by sufficiently specific statutory or constitutional standards. Moreover, if the agency discretion exception encompassed “military authority exercised in the field in time of war,” there would have been no need to include the latter exception.
\textsuperscript{251} See McConnell, supra note 5, at 206 (“A grand coalition emerged in support of the APA, . . .; not because everyone agreed that this was the best form of procedures [but] because that bill altered the status quo in the direction preferred by everyone . . .”).
\textsuperscript{253} Schiller, supra note 39, at 199; Shepherd, supra note 34, at 1665.
\textsuperscript{254} Adrian Vermeule, Our Schmittian Administrative Law, 122 HARV. L. REV. 1095, 1130 (2009).
\textsuperscript{255} Id.
\textsuperscript{256} Id.
\textsuperscript{257} Id.; see also id. at 1139 (“The reasons that the APA’s enactors created the black and grey holes were quite pragmatic, including the inability to formulate comprehensive and precise rules that would apply to the sprawling diversity of the administrative state and its problems, and a lively appreciation of the inevitability of emergencies and unforeseen
nor desirable under the circumstances to make the statute any clearer. Further specificity may have doomed the bill’s chances of passage or its prospects for survival in the courts.

Rather than attempting to resolve their disagreements, the parties to this compromise set their sights on convincing the courts to adopt their interpretation of the Act. To that end, the parties ensured that the legislative history included their various interpretations. In addition to the House and Senate Reports and scant debate on the floor of the two chambers, the Senate Report included a letter from Attorney General Clark in which he expressed his “belief that the provisions of the bill can and should be construed reasonably and in a sense which will fairly balance the requirements and interests of private persons and governmental agencies.” The Attorney General’s letter included an appendix analyzing the bill’s provisions. For example, the Attorney General interpreted the term naval as including “defense functions of the Coast Guard and the Bureau of Marine Inspection and Navigation.” The Attorney General did not comment on the military authority exemption in §2(a) at that time. Representative Hobbs also introduced a memorandum from the Justice Department interpreting the bill, which did not mention the military.

Following its signing, the Attorney General issued a monograph interpreting the APA that “was intended primarily as a guide to the agencies in adjusting their procedures to the requirements of the Act.” The Attorney General’s Manual on the Administrative Procedure Act did not discuss the “military authority” exemption in §2(a). The only relevant comment on §4’s exemption from the rulemaking provisions for “any military, naval, or foreign affairs function” was that the exemption was “not limited to activities of the War and Navy Departments but covers all military and naval functions exercised by any agency,” including the Coast Guard and the Federal Power Commission. Similarly, the Attorney General interpreted §5’s exemption from the adjudication provisions of “military, naval, or foreign affairs functions” as including not

---

258. See Shepherd, supra note 34, at 1673 (noting how parties “sought to create favorable legislative history”).

259. LEGISLATIVE HISTORY, supra note 181, at 224.

260. Id. at 225.

261. Id. at 415.

262. ATTORNEY GENERAL’S MANUAL, supra note 250, at 6.

263. See id. at 11 (listing the exceptions).

264. Id. at 26.
2010] HISTORY OF MILITARY AUTHORITY EXCEPTION IN THE APA
707

just the War and Navy Departments, but also “any other agency . . . to the extent that the conduct of military or naval affairs is involved.”

Professor Shepherd explains that Congress “would have preferred a stronger bill” and thus “interpreted the bill’s ambiguous provisions as imposing strict new controls on agencies.” The Attorney General, in contrast, interpreted the bill “as merely restating existing case law” and “suppressed to a minimum the bill’s limits on agencies.” Attorney General Clark’s efforts to shape the courts’ interpretation of the APA paid off. The Supreme Court has “often found persuasive” and “given some deference,” if not “great weight,” to the Attorney General’s Manual, rather than to the contrasting congressional statements in the legislative history. Commentators have criticized the Court’s adoption of the administration’s thesis as the definitive interpretation of the Act. Congress’s and the Executive’s interpretations of the military authority exemption, however, do not conflict. Both appear to agree that, in contrast to the Walter–Logan bill, the APA’s exemptions were intended to cover functions, not particular agencies, and that the term authority was intended to cover whichever entity has the power to act. In fact, the military

265. Id. at 45.
266. Shepherd, supra note 34, at 1663; see also Dickinson, supra note 55, at 516; Pat McCarran, Improving “Administrative Justice”: Hearings and Evidence, Scope of Judicial Review, 32 A.B.A. J. 827, 827–28 (1946); Shepherd, supra note 34, at 1673.
267. Shepherd, supra note 34, at 1663. The ATTORNEY GENERAL’S MANUAL, a “highly political document designed to minimize the impact of the new statute on executive agencies, shrewdly characterized the APA provisions governing judicial review as merely a ‘restatement’ and thereby invited courts and the bar to treat the Act as something less than a statute, as subservient to judge-made doctrine.” John F. Duffy, Administrative Common Law in Judicial Review, 77 TEX. L. REV. 113, 119 (1998).
268. Shepherd, supra note 34, at 1666; see also Comment, The Federal Administrative Procedure Act: Codification or Reform?, 56 YALE L.J. 670, 673 (1947) (“It is not surprising that those sponsoring the legislation should favor a construction that would establish new standards of law; and conversely those whom the Act is designed to control would favor making it as weak as possible.”).
272. See, e.g., Duffy, supra note 267, at 132 (“In fact, most of the passages from the legislative history cited by Clark to bolster his restatement thesis were letters sent to Congress by Clark himself.”); Roland M. Frye, Jr., Restricted Communications at the United States Nuclear Regulatory Commission, 59 ADMIN. L. REV. 315, 340 n.100 (2007); Shepherd, supra note 34, at 1682–83.
exemptions generated almost no debate and hence little contemporaneous
discussion. Thus, the question of which branch’s view of the APA should
garner more judicial respect need not be answered here.

C. 1966 and 1976 Amendments

In 1966, the APA was recodified and included in Title 5 of the U.S.
Code.273 The definition of agency, with the military authority exemption,
previously appeared only in § 2 of the Act, now codified at 5 U.S.C. § 551.
With the 1966 recodification, provisions concerning the Administrative
Conference of the United States bisected the original provisions of the
APA. Hence, Congress repeated the definition of agency at the beginning of
the provisions governing judicial review, previously § 10 of the Act, now
codified at 5 U.S.C. § 701(b). In both sections, Congress retained the “time
of war” exemption, but omitted “or naval” as included in “military.”274
The House and Senate Reports state that no substantive change was
intended.275

The factors that led Congress, after years of deliberation, to add a waiver
of sovereign immunity to 5 U.S.C. § 702 in 1976276 have been addressed
elsewhere.277 Congress did not change the military authority exception at
that time, but stated that § 702’s new waiver of sovereign immunity would
be limited by the exceptions to the definition of agency, including the
military authority exception. The House and Senate Reports stated:

274. See S. REP. NO. 89-1380, at 28, 33 (1966); H.R. REP. NO. 89-901, at 11, 16 (1965)
(listing what was omitted). Congress also “omitted as executed” the exemption in § 2(a) for
wartime functions when it recodified the APA in 1966. See S. REP. NO. 89-1380, at 28, 33
(1966); H.R. REP. NO. 89-901, at 11, 16 (1965).
277. See C. Stanley Dees, The Executive Branch as Penelope: Preserving the Tapestry of Sovereign-
Immunity Waivers for Suits Against the United States, 71 GEO. WASH. L. REV. 708, 709 (2003);
Gregory C. Sisk, The Tapestry Unravels: Statutory Waivers of Sovereign Immunity and Money Claims
Against the United States, 71 GEO. WASH. L. REV. 602, 616–17 (2003); Sharon J. Kronish,
Comment, Sovereign Immunity: A Modern Rationale in Light of the 1976 Amendments to the
Since the amendment is to be added to 5 U.S.C. Section 702, it will be
applicable only to functions falling within the definition of ‘agency’ in 5
U.S.C. Section 701. Section 701(b)(1) defines ‘agency’ very
broadly . . . except for a list of exempt agencies or functions: . . . courts-
martial and certain other military, wartime and emergency functions. 278

The fact that the waiver of sovereign immunity was “subject to the other
limitations of the Administrative Procedure Act” was “an important factor”
in the Justice Department’s support for the bill, 279 and presumably in its
passage. 280 More important than the military authority exception,
however, may have been the preclusion of judicial review of agency actions
that are “committed to agency discretion by law.” 281 The Administrative
Conference of the United States thought it “fanciful to suppose that
abolition of sovereign immunity will allow the courts to decide issues about
foreign affairs, military policy, and other subjects inappropriate for judicial
action.” 282 Of course, if the “agency discretion” exception covered
quintessentially military actions, the military authority exception would not
have been necessary. 283

D. Summation

The Walter–Logan Bill’s broad exemption for “any matter concerning
or relating to the military or naval establishments” 284 was too narrow, in
President Roosevelt’s eyes, for a nation on the brink of war. Yet, shortly
after the war ended, Congress willingly subjected military actions to judicial

statement in the House and Senate Reports was drawn almost verbatim from the hearing
testimony of Professor Roger Cramton, who authored the report of the Administrative
Conference of the United States’ Committee on Judicial Review. See Sovereign Immunity:
Hearing on S. 3568 Before the Subcomm. on Administrative Practice and Procedure of the S. Comm.
on the Judiciary, 91st Cong. 32 (1970) [hereinafter Sovereign Immunity Hearing]; see also 122 Cong.
affect other limitations on judicial review . . . .”); 122 Cong. Rec. 22,014 (1976) (statement
of Sen. Hruska) (“The Administrative Conference felt it important that the waiver of
immunity . . . be subject to the other limitations of the Administrative Procedure Act . . . .”);
scope of judicial review, which remains limited by the Administrative Procedure Act.”).
Antonin Scalia); S. Rep. No. 94-996, at 26 (1976) (letter from Assistant Attorney General
Antonin Scalia).
280. For a discussion of failed attempts to amend the APA, see William H. Allen, The
282. Sovereign Immunity Hearing, supra note 278, at 133.
283. See supra note 250.
284. H.R. 6324, 76th Cong. § 7(b) (1940).
review unless those actions were the result of “military authority exercised in the field in time of war” or fell within one of the APA’s other exceptions. The history of that shift lends insight into how World War II affected Congress’s and the President’s view of the need to control the Fourth Branch, specifically through judicial review of military action. That history, however, exposes little about Congress’s intended meaning in the military authority exception that the plain language does not already reveal. Congress had at its disposal language that would have insulated a wide swath of military actions from judicial review, language it employed in other provisions of the Act, but it denied the military’s pleas and opted to employ much more specific language when it came to judicial review. The history does show that military authority was not intended to be synonymous with the Departments of War and the Navy but to cover “defense functions,” regardless of which agency or individual was responsible.  

To unearth the meaning of the exception’s other key terms—*in the field* and *time of war*—more digging is required.

II. THE ARTICLES OF WAR

When Congress enacted the APA in 1946, the key phrases *in the field* and *time of war* had well established meanings. If Congress was aware that those terms had long been used in the Articles of War that governed the Army, however, the legislative history gives no indication of it. As mentioned above, the legislative history does not reveal the source of the military authority exception, but the War Department may have proposed it as a compromise when the Senate Judiciary Committee declined to exempt the military from the bill entirely. The lineage of the phrases *in the field* and *time of war* in the Articles of War reinforce that possibility. Thus, the contemporaneous understanding of these phrases in the Articles of War may assist in deciphering the military authority exemption in the APA.

The Second Continental Congress first enacted the Articles of War in 1775. Though they were revised many times, they did not change

285. Legislative History, supra note 181, at 191, 196.

286. William B. Aycock & Seymour W. Wurfel, Military Law Under the Uniform Code of Military Justice 9–10 (1955); Office of the Judge Advocate Gen. of the Army, Military Laws of the United States 163 (6th ed. 1940); William Winthrop, Military Law and Precedents 17 (2d ed. William S. Hein & Co. 2000) (1920). Colonel Winthrop, the “Blackstone of Military Law,” Reid v. Covert, 354 U.S. 1, 19 n.30 (1957), published the first edition of his treatise in 1886 and addressed Article 65 of the 1874 Articles of War: “All retainers to the camp, and all persons serving with the armies of the United States in the field, though not enlisted soldiers, are to be subject to orders, according to the rules and discipline of war.” Winthrop, supra at 991. At the time, the phrase “in the field” was “deemed clearly to indicate that the application of the Article [was]
significantly between 1806 and their replacement in 1950 by the Uniform Code of Military Justice. The version of the Articles that governed the Army during World War II was enacted in 1920 along with other “changes in details which the lapse of time and the experience of the [First World War] had shown to be necessary.” Article 2(d) provided for court-martial jurisdiction over:

All retainers to the camp and all persons accompanying or serving with the armies of the United States without the territorial jurisdiction of the United States, and in time of war all such retainers and persons accompanying or serving with the armies of the United States in the field, both within and without the territorial jurisdiction of the United States, though not otherwise subject to these articles.

The Uniform Code of Military Justice (UCMJ), enacted in 1950, replaced the Articles of War with a code that still governs all of the Armed Forces. Article 2 of the UCMJ, which was “taken from” Article 2(d) of the Articles of War, continued to subject to court-martial jurisdiction “all persons serving with, employed by, or accompanying the armed forces without the continental limits of the United States” and “[i]n time of war, all persons serving with or accompanying an armed force in the field.”

Civilians were thus subject to military jurisdiction at the time Congress enacted the APA if they served with or accompanied the Army “without the territorial limits of the United States” such that they were “beyond the territorial jurisdiction of state and federal civil courts.” In a “time of war,” however, civilians were subject to military jurisdiction either abroad

confined both to the period and pendency of war and to acts committed on the theatre of the war.” Id. at 101.

289. Pub. L. No. 66-242, 41 Stat. at 787 (emphasis added); AYCOCK & WURFEL, supra note 286, at 54. These phrases also appear together in Article 40(b), which authorized the dismissal of an officer below a certain grade, and Article 40(d), which authorized the death sentence for murder, rape, mutiny, desertion, and spying “in time of war . . . upon confirmation by the commanding general of the Army in the field.” 41 Stat. at 796-97. The phrase time of war appears unaccompanied by “in the field” in numerous other provisions as well. The Articles of War were amended substantially in 1948. Selective Service Act of 1948, Pub. L. No. 80-759, § 201, 62 Stat. 604, 627. Article 2(d) was not amended, § 202, 62 Stat. at 628, and the legislative history does not discuss the phrases in the field or time of war.
293. AYCOCK & WURFEL, supra note 286, at 57.
or within the United States if they served with or accompanied the Army “in the field.”

A. “In the Field”

Before World War I, the phrase in the field was interpreted somewhat narrowly as referring to “the theatre of military operations.” The Supreme Court later said, referring primarily to pre-World War I authorities, that “[c]ompetent experts on military law, the Judge Advocate General and the Attorney General have repeatedly taken the position that ‘in the field’ means in an area of actual fighting.” Consequently, an event could occur “in the field” only during a “time of war.” And even during a time of war, “in the field” did not apply “to any portions of the territory of the United States in which military operations were not being carried on against the public enemy.” During the World Wars, however, the phrase came to be interpreted more broadly as “not by any means limited to overseas service, but including service within the territorial limits of the United States.” Accordingly, when Congress enacted the UCMJ, the committee reports observed that “[t]he phrase ‘in the field’ has been

294. Id. During World War II, “[t]he trial by military tribunals of civilian employees of the military establishment in overseas areas . . . added substantially to the number confined by military authority.” Id. at 314. On January 1, 1950, there were more than 2,500 persons “serving civilian type felony sentences imposed by military tribunals.” Id. Although federal courts lack appellate jurisdiction over courts-martial, they may entertain “collateral attack[s] upon [their] determinations.” Shaw v. United States, 209 F.2d 811, 812–13 (D.C. Cir. 1954). Habecas corpus, for example, gives federal courts authority to examine whether the military tribunal had jurisdiction, but “[b]eyond this, [federal courts] need not look into the record.” Ex parte Reed, 100 U.S. 13, 23 (1879), aff’d, Hiatt v. Brown, 339 U.S. 103, 111 (1950).

295. GEORGE B. DAVIS, A TREATISE ON THE MILITARY LAW OF THE UNITED STATES 52 (3d ed. rev. 1915); see also id. at 52 n.2, 478; Merriam Webster’s New International Dictionary 941 (2d ed. 1939) (defining field as “a place where a battle is fought”); Winthrop, supra note 286, at 101.

296. Reid v. Covert, 334 U.S. 1, 34 n.61 (1957); see also 14 Op. Att’y Gen. 23 (1872) (“These words imply military operations with a view to an enemy.”).

297. See Davis, supra note 295, at 52 (“[T]he statute is restricted in its operation to persons accompanying armies in the field in time of war, and in the actual theatre of military operations”); EDGAR S. DUDLEY, MILITARY LAW AND THE PROCEDURE OF COURTS-MARTIAL 413–14 (2d ed. rev. 1906); Winthrop, supra note 286, at 101 (“[T]he application of the article is confined both to the period and pendency of war and to acts committed on the theatre of the war . . . . [T]his article is operative only in and for a time of war . . . .”).

298. Davis, supra note 295, at 52.

299. JULIAN J. APPLETON, MILITARY LAW FOR THE COMPANY COMMANDER 14 n.* (1944); see also Ex parte Jochen, 257 F. 200, 208–09 (S.D. Tex. 1919) (declaring that a civilian serving with troops patrolling the Texas-Mexico border was “in the field”).
construed to refer to any place, whether on land or water, apart from permanent cantonments or fortifications, where military operations are being conducted.”

For example, the Fourth Circuit in *Hines v. Mikell* held that Camp Jackson near Columbia, South Carolina, which “was established for the training of the military forces of the United States for service in the theater of operations overseas,” was “in the field” under Article 2(d). A civilian auditor was arrested and held in a military prison until the district court granted his petition for a writ of habeas corpus, holding that the phrase *in the field* was limited to “a place of contact with the enemy.” The court of appeals reversed. Looking to military statutes and regulations which indicated that domestic cantonments could be “in the field,” the court of appeals held that “there is a required distinction between the term ‘in the field’ and the places of contact with the enemy.” Whether a particular location was *in the field* was “not to be determined by the locality in which the army may be found, but rather by the activity in which it may be engaged at any particular time.” The court considered *in the field* to be a term of art that encompassed troops “engaged in training and preparing for service on the firing line overseas.” Hence, “[i]n time of war, with some exceptions, practically the entire army is ‘in the field,’ but not necessarily ‘in the theater of operations.’” Service members at Camp Jackson “took the first step which was to lead them to the firing line, and they were then as much ‘in the field’ in pursuit of such training as those who were encamped on the fields of Flanders awaiting orders to enter the engagement.” The court emphasized that a contrary decision “would handicap the military authorities, and greatly hinder and delay military


301. 259 F. 2d, 29, 35 (4th Cir. 1919) (Camp Jackson was established “for the purpose of training preparatory for service in the actual theater of war”). Article 2(d) of the 1916 Articles of War at issue in *Hines* was identical to Article 2(d) in the 1920 recodification. See id. at 30.

302. Id. at 29, 32.

303. Id. at 29–34. For example, in 1918 Congress provided that commissioned officers were entitled to retain quarters for dependents while they were “on duty in the field, or on active duty without the territorial jurisdiction of the United States.” Id. at 33 (quoting Pub. L. 65-129, 40 Stat. 330 (1918)).

304. Id. at 32.

305. Id. at 34.

306. Id. at 33.

307. Id.; see also id. at 31 (“[I]n case of war, when the army leaves the post and moves in the direction of the enemy, or to some intermediate point where they may temporarily stop for training, would it not be more reasonable to say that they were then ‘in the field?’”).

308. Id. at 33.
operations.”

Civilians serving on ships transporting troops or military supplies were also “in the field” under Article 2(d).

For example, *In re Berue* held that a civilian merchant seaman serving on a ship bound for Casablanca carrying military supplies was “in the field.”

The court noted that the ship “was in waters infested by submarines and other naval craft of the Axis Nations” and hence was in “actual combat zones.”

Similarly, in *Ex parte Gerlach*, the court held that a civilian employee of the United States Shipping Board who was returning from Europe on an army transport ship that faced “peril from submarines” was “in the field.”

The court said that “[t]he words ‘in the field’ do not refer to land only, but to any place, whether on land or water, apart from permanent cantonments or fortifications, where military operations are being conducted.”

Indeed, civilians on ships could be “in the field” even when docked in the United States.

Court-martial jurisdiction over civilians “in the field” was “spARINGLY exercised” and “reserved for cases where the individual is... engaged in

---

309. *Id.* at 55. *Hines v. Mikell* is consistent with other cases holding civilians subject to military jurisdiction during World War I:

- Port of Brooklyn storage office chauffeur and laborers; ... cook employed by quartermaster at New Mexico camp; scout in Texas;
- quartermaster civilian employee laborers on docks at ports of embarkation and at Camp Upton, New York; ... clerks employed by the Quartermaster overseas and at Camp Meade, Maryland; ... merchant seamen on Army transports at sea or in English, French, or American ports,

*Aycock, supra* note 286, at 55 (citations omitted), and World War II:

- decoding experts employed by the Signal Corps at installations in the United States; electricians employed by the Corps of Engineers in ... Alaska; ... employees of post exchanges at camps in the United States; ... firefighters at an air base within the United States; ... policemen and guards at installations in the United States important to the prosecution of the war.

*Id.* at 56–57 (citations omitted).

310. *Id.* at 58.


312. *Id.* at 253.

313. 247 F. 616, 617 (S.D.N.Y. 1917); *see also* McCune v. Kilpatrick, 53 F. Supp. 80, 84–85 (E.D. Va. 1943) (civillian cook on merchant ship transporting troops to battle zones was “in the field”).


315. *Ex parte Falls*, 251 F. at 416 (civillian cook on ship transporting supplies for the Army was “in the field” when he attempted to desert just before the ship sailed from Brooklyn, New York).
work essential to the war effort.” But a civilian did not have to be working directly on a military operation to fall within Article 2[a]’s jurisdiction. Perlstein v. United States concerned the habeas corpus petition of a civilian mechanic employed at a military base in Eritrea by a private contractor engaged in a salvage operation to raise enemy ships and docks that had been scuttled. The Third Circuit held that court-martial jurisdiction was properly exercised under Article 2[d], even though the petitioner was not directly involved in the salvage operation. The court concluded that, although the petitioner “did not physically assist in the raising of the ships and docks, his association with that program was as close as if he had and his contribution to its successful completion was of considerable importance.” Specifically, given the “very hot and humid” weather, his work on air conditioning and refrigeration “was an integral part of the whole endeavor and as such was within the intendment of Article 2[d].”

In the 1957 case of Reid v. Covert, the Supreme Court held that civilians accompanying the military overseas were not subject to court martial jurisdiction and, in the course of the opinion, also returned to a narrower interpretation of the phrase in the field. Reid v. Covert involved the habeas corpus petitions of two civilian women who were tried in military tribunals for the murder of their service-member husbands on military bases in Japan and England. The military asserted jurisdiction under UCMJ Article 2[11], which covers persons serving with or accompanying the military outside the United States. The Supreme Court initially held that the military trials were constitutional, but reheard the case the following term and concluded to the contrary that the defendants were entitled to the protections of the Bill of Rights. Although the government had not asserted jurisdiction

316. Aycock & Wurfel, supra note 286, at 57–58.
318. Id. at 168.
319. Id. On the other hand, “personnel at industrial establishments in the United States; employees of an independent contractor engaged in construction work on the Inter-American highway supervised by a few Army Engineer officers but where no troops were present; and War Department clerical employees in Washington or in a field office in the United States not located at a military camp” were not subject to court martial jurisdiction.
321. Id. at 3, 6.
322. Id. at 21 (“Every extension of military jurisdiction is an encroachment on the jurisdiction of the civil courts, and, more important, acts as a deprivation of the right to jury trial and of other treasured constitutional protections.”).
under Article 2(10), Justice Black, writing for a four-member plurality, distinguished cases upholding military jurisdiction over “civilians performing services for the armed forces ‘in the field’ during time of war,” including \textit{Hines, Perlstein}, and \textit{In re Berne}, as having “rest[ed] on the Government’s ‘war powers.’”

In the face of an actively hostile enemy, military commanders necessarily have broad power over persons on the battlefront. From a time prior to the adoption of the Constitution the extraordinary circumstances present in an area of actual fighting have been considered sufficient to permit punishment of some civilians in that area by military courts under military rules.

Since no “active hostilities were under way” in England or Japan at the time the defendants committed their crimes, the Court held that the military tribunals lacked jurisdiction. The Court thus linked the phrase “in the field” once again to the existence of a “time of war” and the proximity of combat.

The Supreme Court has continued not only to treat \textit{Reid v. Covert} as viable precedent, but also to follow its holding that military tribunals for non-service members must have some nexus to actual combat. In \textit{Hamdan v. Rumsfeld}, the Court opined that military commissions must be supported by “military necessity.” The Court held that the petitioner, who was being detained at Guantánamo Bay, could not be tried by a military commission because, among other things, his commission was appointed by “a retired major general stationed away from any active hostilities” and

\begin{enumerate}
\item Id. at 34 n.61.
\item Id. at 33. Justices Frankfurter and Harlan concurred in the result on the narrower ground that Article 2(11) did not permit court martial jurisdiction over civilians for capital offenses in times of peace. Id. at 64, 65 (Frankfurter, J., concurring).
\item Id. at 33 [plurality opinion].
\item Id. at 34.
\item See id. at 35 (rejecting “the Government’s argument that present threats to peace permit military trial of civilians accompanying the armed forces overseas in an area where no actual hostilities are under way”); see also \textit{McElroy v. United States ex rel. Guaudiero}, 361 U.S. 281, 284–86 (1960); Ian W. Baldwin, \textit{Comrades in Arms: Using the Uniform Code of Military Justice and the Military Extraterritorial Jurisdiction Act to Prosecute Civilian-Contractor Misconduct}, 94 \textit{Iowa L. Rev.} 287, 309–10 (2008) (“Reid again proves instructive by defining ‘in the field.’ A contractor is ‘in the field’ when that person is working ‘in an area of actual fighting’ at or near the ‘battlefront’ where ‘actual hostilities are under way.’”) (citations omitted). The Supreme Court extended \textit{Reid v. Covert} in later cases. \textit{Kinsella v. United States ex rel. Singleton}, 361 U.S. 234, 273–77 (1960) (extending \textit{Reid v. Covert} to noncapital trial of civilian dependent); \textit{Grisham v. Hagan}, 361 U.S. 270 (1960) (extending \textit{Reid v. Covert} to capital trial of civilian employee); \textit{McElroy}, 361 U.S. at 281 (extending \textit{Reid v. Covert} to noncapital offenses).
\item 548 U.S. 537, 612 (2006).
\end{enumerate}
none of the alleged unlawful acts “necessarily occurred during time of, or in a theater of, war.” Nonetheless, at the time Congress passed the APA, the phrase “in the field” was understood to reach more broadly to “any place . . . where military operations are being conducted.

B. “Time of War”

The interpretation of the phrase time of war has changed over time as well. Early military authorities state that a declaration of war was “not absolutely necessary to the legal existence of a status of foreign war.” Likewise, courts have long held that Congress need not issue a declaration of war for combat to qualify as “time of war.” “Although the United States has committed its armed forces into combat more than a hundred times, Congress has declared war only five times: the War of 1812, the Mexican–American War of 1848, the Spanish–American War of 1898, World War I, and World War II.” Nonetheless, “[s]ince the earliest

330. Id.; see also Chad DeVeaux, Rationalizing The Constitution: The Military Commissions Act and the Dubious Legacy of Ex Parte Quirin, 42 AKRON L. REV. 13, 30–31 (2009) (“Once removed from the exigency of the war zone, the legitimacy of military tribunals rapidly dissipates.”).


332. WINTHROP, supra note 286, at 668. Under the Articles of War, a war could commence when the United States’ territory or defenses were attacked and the President declared the existence of an insurrection. Id. “War” also included not just combat with foreign nations, but also civil war and “a state of active hostilities with an Indian tribe.” Id. at 86, 101.

333. See The Prize Cases, 67 U.S. 635, 666–70 (1862) [holding that the Civil War constituted a de facto state of war]; Bas v. Tingy, 4 U.S. 37, 41–42 (1800) (opinion of Washington, J.); id. at 43–44 (opinion of Chase, J.); id. at 45–46 (opinion of Paterson, J.) (all concluding that conflict with France constituted war); Campbell v. Clinton, 203 F.3d 19, 37–39 (D.C. Cir. 2000) (Tatel, J., concurring) (discussing cases); Kooli v. United States, 976 F.2d 1328, 1334 [9th Cir. 1992], cert. denied, 506 U.S. 960 (1993) (holding government not liable under the Federal Tort Claims Act for accidental downing of civilian aircraft during “tanker war” between Iran and Iraq); Mitchell v. Laird, 488 F.2d 611 (D.C. Cir. 1973) (holding that the President may wage “some types of war” without Congressional approval); Minns v. United States, 974 F. Supp. 500, 506 (D. Md. 1997) (holding that the Persian Gulf War constituted a “time of war” under the Federal Tort Claims Act).

334. Campbell, 203 F.3d at 29–30 n.6 [Randolph, J., concurring]. Moreover, since 1945 when “the United Nations Charter prohibited the use of force in international law, congressional declarations of war have disappeared from U.S. practice.” Catherine H. Gibson, Frankfurter’s Gloss Theory, Separation of Powers, and Foreign Investment, 36 N. KY. L. REV. 103, 125 (2009); see also Paul W. Kahn, War Powers and the Millennium, 34 LOY. L.A. L. REV. 11, 16–17 (2000) (“Since the advent of the United Nations (UN) Charter, war has been abolished as a category of international law. A declaration of war serves no purpose under international law; it can have no bearing on the underlying legal situation.”).
years of the nation, courts have not hesitated to determine when military
action constitutes ‘war.’”335 The Ninth Circuit, for example, interpreting
the phrase time of war in an exception to the waiver of sovereign immunity
in the Federal Tort Claims Act, held that “when, as a result of a deliberate
decision by the executive branch, United States armed forces engage in an
organized series of hostile encounters on a significant scale with the military
forces of another nation . . . a ‘time of war’ exists . . .”336 The court
observed that the military actions in Korea, Vietnam, Panama, Grenada,
Kuwait, and Iraq were not preceded by declarations of war, “[y]et no one
can doubt that a state of war existed.”337

The UCMJ includes the phrase “time of war” in numerous provisions,
but provides no definition of the phrase, and the legislative history contains
no discussion of whether a war requires a congressional declaration or not.338
After the adoption of the UCMJ, the military courts continued to
determine on a case-by-case basis whether a particular military engagement
constituted a “time of war.” In 1954, for example, the Court of Military
Appeals held that a service member who deserted in the continental United
States during the Korean Conflict could be charged with absence without
leave in time of war.339 Applying a “practical approach,”340 the court
emphasized that the phrase time of war appears in several Articles of the
UCMJ and its meaning “must be determined with an eye to the goal
toward which that Article appears to have been directed.”341 The court
observed that, during the Korean Conflict, “[c]ertainly . . ., from [the]
state-side, most of the attributes of a declared war were undeniably
present.”342

In 1969, the Court of Military Appeals shifted gears. In United States v.
Averette, the court held that a civilian employee of a military contractor in
Vietnam was not subject to court-martial jurisdiction under Article 2(10) of
the UCMJ.343 The court observed that, although “the fighting in Vietnam
qualifies as a war as that word is generally used and understood,” the
phrase in time of war in Article 2(10) means “a war formally declared by

335. Campbell, 203 F.3d at 37 (Tatel, J., concurring).
336. Koohi, 976 F.2d at 1335.
337. Id. at 1334.
however, does not define the term and its legislative history is not particularly
enlightening.”).
340. Id. at 222.
341. Id. at 227.
342. Id. at 224.
2010] HISTORY OF MILITARY AUTHORITY EXCEPTION IN THE APA 719

Congress.” 344 The court found that “recent guidance” from the Supreme Court, i.e., Reid v. Covert and its progeny, required “a strict and literal construction of the phrase ‘in time of war’” in Article 2(10) so as to avoid “the possibility of civilian prosecutions by military courts whenever military action on a varying scale of intensity occurs.” 345 Military courts continued to apply the “practical” approach in other contexts, 346 and following the terrorist attacks of September 11, 2001, the Averette decision began to receive harsh criticism. 347

In 2006, Congress amended Article 2(10) to overrule Averette. 348 Where the UCMJ previously provided for court martial jurisdiction over “persons serving with or accompanying an armed force in the field” only “[i]n time of war,” it now provides for such jurisdiction “[i]n time of declared war or a contingency operation.” 349 Apart from the Averette detour, however, and certainly at the time Congress enacted the APA, no declaration of war was required for combat to qualify as a time of war.

The courts’ narrow interpretations of the phrases “in the field” and

344. Id. at 365.
345. Id. at 364–65; accord Robb v. United States, 456 F.2d 768, 769, 771 (Ct. Cl. 1972) (holding that a civilian engineer employed by the Navy in Vietnam was not subject to court martial jurisdiction under Article 2(10) because the Vietnam conflict was not a formally declared war).
346. See United States v. Castillo, 34 M.J. 1160, 1163, 1166 (C.M.R. 1992) (holding, in case of marine charged with disobeying an order, that the Persian Gulf conflict was a “time of war” based on the “realities of the situation as distinguished from the legal niceties”) (citation omitted).
347. E.g., Joseph Romero, Of War and Punishment: “Time of War” in Military Jurisprudence and a Call for Congress to Define its Meaning, 51 Naval L. Rev. 1, 32 (2005) (commenting that Averette “may have to be described as the most divergent, inconsistent, and questionable decision to arise from the Court of Military Appeals with regard to UCMJ wartime provisions”); Lawrence J. Schwarz, The Case for Court-Martial Jurisdiction over Civilians under Article 2(a)(10) of the Uniform Code of Military Justice, 2002 Army Law. 31, 34–35 (2002) (identifying four weaknesses of the Averette decision’s definition of “in time of war”).
349. See Pub. L. No. 109–364, § 552, 120 Stat. 2083, 2217 (2006) (codified at 10 U.S.C. § 802(a)(10) (2006)). The term “contingency operation” is defined to include military operations designated by the Secretary of Defense or those resulting in “the call or order to, or retention on, active duty of members of the uniformed services . . . during a war or during a national emergency declared by the President or Congress.” 10 U.S.C. § 101(a)(13)(B) (2006).
“time of war” in Article 2(10) of the UCMJ do not compel a similarly narrow interpretation of those phrases in § 701 of the APA. The provision allowing court martial jurisdiction over civilians is appropriately construed narrowly to protect citizens’ constitutional rights.\textsuperscript{350} By the same token, the term \textit{agency} in the APA’s waiver of sovereign immunity should be interpreted narrowly to insulate agency action from judicial review.\textsuperscript{351} Any ambiguity in the meaning of that term that remains after employing “traditional tools of statutory construction” should be resolved in favor of the government’s immunity from suit.\textsuperscript{352} Likewise, any ambiguity in the definition of \textit{agency} should be resolved in favor of the government’s immunity. Hence, if the phrases \textit{in the field} and \textit{time of war} in § 701 of the APA are ambiguous, they should be interpreted broadly to insulate military actions from judicial review. The upshot is that, while the plain language of the military authority exception is certainly narrower than the military function exceptions, the contemporaneous understanding of the terms \textit{in the field} and \textit{time of war} in the 1940s was somewhat broader than the plain language might indicate to a modern reader.

\section*{III. Where the Courts and Commentators Have Gone Wrong}

Although courts mentioned the military authority exception over the years, it was not until 1979 that a court actually discussed it.\textsuperscript{353} Perhaps the lack of a clear waiver of sovereign immunity in the APA before 1976 provided the government with sufficient means for keeping APA claims against the military out of court. From 1976 until the start of the present conflicts in Iraq and Afghanistan, the United States may not have experienced a time of war of sufficient length to warrant invoking the military authority exception. Whatever other factors may have contributed to the dearth of case law discussing the exception in years past, the issue has arisen more often in recent years. Commentators have followed suit and begun to mention the exception, though without any in-depth inquiry. The above historiographic analysis demonstrates that courts and commentators have been reading the exception too narrowly.

The 1979 case that first analyzed the exception did not stray from

\begin{thebibliography}{9}
\bibitem{350} See Rohrb, 456 F.2d at 771 (“[T]he succession of Supreme Court decisions dealing with military jurisdiction over civilians leads to the conclusion that Article 2(10) is to be narrowly construed.”).
\bibitem{351} See supra note 21.
\bibitem{353} See Jaffee v. United States, 592 F.2d 712, 719–20 (3d Cir. 1979).
\end{thebibliography}
historical accuracy. In Jaffe v. United States, the plaintiff claimed that he developed cancer as a result of exposure to radiation during a nuclear test in Nevada in 1953 while he was serving in the Army. The district court dismissed the plaintiff’s class action claim, which sought to enjoin the government to warn all soldiers that had been present at the test of the medical risks of their exposure, for lack of a waiver of sovereign immunity. The court of appeals reversed on that point, holding that the Army is an agency and the military authority exception did not apply. The court held that, even if the explosion took place before the Korean War ended and the exception “could be interpreted to cover operations in Nevada,” the claim concerned “the Army’s failure to act in the years since the explosion,” which “was neither in the field nor in time of war.” The court was correct to assume that the Korean Conflict would qualify as a time of war and that the Nevada nuclear testing site could be considered in the field. As explained above, Congress in the 1940s would not have thought a declaration of war necessary for a time of war to exist, and weapons testing falls comfortably within the scope of the phrase in the field as interpreted in contemporary case law.

The D.C. Circuit’s 1991 decision in Doe v. Sullivan began the courts’ and commentators’ departure from historical accuracy. In Doe, a service member challenged a Food and Drug Administration regulation permitting the military to use “unapproved . . . drugs on military personnel[] in certain combat-related situations, [in this case, during Operation Desert Storm,] without . . . informed consent.” The D.C. Circuit held the military authority exception was not applicable essentially because the plaintiff was not challenging any military authority. The court said that the claim did not concern “military commands made in combat zones or in preparation for, or in the aftermath of, battle,” “judicial interference with the relationship between soldiers and their military superiors,” or “military strategy or discipline,” but rather “the scope of the authority Congress has entrusted to the FDA.” The court’s holding that the military authority exception would not reach the FDA’s regulation is consistent with the APA’s history, not because the FDA is not a military agency but because it probably was not exercising a “defense function” when it issued the

354. Id. at 714.
355. Id.
356. Id. at 719–20.
357. Id. at 720.
358. See supra text accompanying notes 332–349.
359. See supra text accompanying notes 299–331.
361. Id. at 1380.
challenged regulation. But the D.C. Circuit’s dicta did not contemplate the full reach of the military authority exception; the exception could reach beyond combat zones, preparation for battle, the relationship between service members and their superiors, military strategy, and military discipline. The Doe dicta is hard to square with some of the case law contemporaneous with the APA’s enactment that interpreted the phrase in the field in the UCMJ, like Ex parte Gerlach, which held that an Army transport ship bound for the United States was in the field simply because it might be attacked by a submarine, and Perlstein v. United States, which was decided a few weeks before the Senate added the military authority exception to the bill and held that a civilian air conditioning mechanic was in the field even though he was not directly involved in any military operation.

The Doe dicta has led some courts and commentators astray. The court in Vance v. Rumsfeld, for example, relied on Doe when it denied pending discovery the government’s motion to dismiss claims for the return of personal property filed by civilian employees of a private security firm in Iraq who had been detained by the U.S. military. The court was correct in its belief that the “military authority” exception was not intended to “exempt the military as a whole.” That is one of the few points the

362. See Legislative History, supra note 181, at 191, 196.
363. 247 F. 616, 617 (S.D.N.Y. 1917).
364. 151 F.2d 167, 168 (3d Cir. 1945).
365. No. 06 C 6964, 2009 WL 2252258, at *3 (N.D. Ill. July 29, 2009). The court in Vance also relied on Rosner v. United States, 231 F. Supp. 2d 1202 (S.D. Fla. 2002), in which the court denied pending discovery the government’s motion to dismiss claims filed by Hungarian Jews whose personal property was seized during World War II first by the pro-Nazi Hungarian government and later by the U.S. Army. Id. at 1203–04. Among other things, the plaintiffs brought claims alleging violations of international law and seeking nonmonetary relief under the APA, and the government invoked the military authority exception. Id. at 1209, 1211. The court rejected the plaintiffs’ argument that the military authority exception did not apply because the order for the Army to seize the property came from American soil, reasoning that the military authority exception “covers ‘military authority exercised in the field,’ without regard to where the underlying order to take military action arises” and that “virtually all military action will be traceable, at some level, back to United States soil.” Id. at 1211–12 & n.13. Instead, the court held that the exception did not apply because the plaintiffs asserted in their complaint that the Army had taken their property “after hostilities had ceased and peace was formally declared” and had engaged in “conduct that, although exercised by military personnel, is decidedly non-military in its nature.” Id. at 1212 n.14. On reconsideration however, the court ordered discovery on the underlying facts and noted that the plaintiffs bore the burden of proving that “the Government’s actions were non-military in nature.” Rosner v. United States, No. 01-1059-CIV, 2002 WL 31954453, at *2–3 (S.D. Fla. Nov. 20, 2002).
official legislative history makes quite clearly. The court departed from historical accuracy, however, in suggesting that the military authority exception only applies to “an exercise of authority from the field of battle” or to orders from “a commander in the field.” The court did not mention the cases that would have informed the understanding of the military authority exception in 1946, like Hines v. Militell, which held that “there is a required distinction between the term ‘in the field’ and the places of contact with the enemy.”

The Doe dicta also impacted Professor Masur’s article on judicial deference to executive branch factual determinations in wartime, in which he examines whether the substantive law governing judicial review in national security matters compels greater deference and concludes that it does not, in part because the military is an agency under the APA. Although he admits that the military authority exception’s “outer limits remain somewhat murky,” he contends that “[t]he judiciary has generally construed this exception narrowly and literally, constraining its application to genuine military operations in theaters of battle, as exemplified by the D.C. Circuit’s language in Doe.” As explained above, however, the dicta in Doe are not consistent with the full range of judicial decisions that would have informed contemporary understanding of the military authority exception in 1946. And aside from Doe, Professor Masur cites only three cases that did not discuss the military authority exception, but merely noted the exception’s obvious inapplicability. Thus, while Professor Masur is correct insofar as he suggests that military action that is reviewable under the APA would be examined under the arbitrary and capricious standard, a broader range of military actions might not be actionable under the APA than he contemplates.

Judge Randolph of the D.C. Circuit came closer to historic accuracy in his concurrence in Al Odah v. United States. Judge Randolph wrote the court’s opinion holding that it lacked jurisdiction to hear habeas corpus petitions filed by aliens held outside the United States at Guantánamo Bay Base, as the President had made a determination that the military commander at Guantánamo Bay Base “has the authority to conduct the military operations there . . . consistent with the law and treaties of the United States.”
Bay. He also filed a concurring opinion in which he said that the military authority exception would bar the plaintiffs’ APA claims because they were taken into custody and remain in custody “in the field in time of war.” Judge Randolph explained that historically the phrase “in the field” “was not restricted to the field of battle,” but “applied as well to ‘organized camps stationed in remote places where civil courts did not exist.’” He also believed that the lack of a congressional declaration of war did not prevent “the war against the al Qaeda terrorist network” from qualifying as a time of war within the meaning of the APA. Those observations were certainly consistent with the meaning of those terms when the APA was enacted. Whether Judge Randolph’s belief that the military authority exception was “meant to forbid” judicial review of “military decisions after those captured have been moved to a ‘safe’ location,” however, is less clear. Unfortunately, when the Supreme Court reversed, holding that habeas corpus jurisdiction extends to aliens held in “territory over which the United States exercises plenary and exclusive jurisdiction, but not ‘ultimate sovereignty,’” the Court did not address the “military authority” exception.

Professors Jinks and Sloss find Judge Randolph’s concurring opinion in Al Odah “unpersuasive.” In their article examining whether the President has the authority under domestic law to violate the Geneva Conventions in order to protect national security, they analyze, among other things, whether sovereign immunity would bar Guantánamo detainees’ claims that the executive branch breached the Geneva Conventions. In the course of that analysis, they assume that the military prison at Guantánamo Bay, Cuba cannot be in the field, even if that phrase is interpreted broadly,

374. Id. at 1144.
375. Id. at 1149–50. The district court had also concluded that it lacked jurisdiction, but held in the alternative that § 702 of the APA would not waive the government’s sovereign immunity because “the actions of the government in this case would be exempt” under the military authority exception. Rasul v. Bush, 215 F. Supp. 2d 55, 64 n.11 (D.D.C. 2002). Because the plaintiffs “were captured in areas where the United States was (and is) engaged in military hostilities pursuant to the Joint Resolution of Congress,” the court found that “[t]his situation plainly falls within” the exception. Id.
376. Al Odah, 321 F.3d at 1150 (quoting Kinsella v. United States ex rel. Singleton, 361 U.S. 234, 274 (Whittaker, J., concurring in part and dissenting in part)).
377. Id.
378. Id.
380. Jinks & Sloss, supra note 1, at 188.
381. Id. at 101.
because it is “thousands of miles away from the battlefield.”\textsuperscript{302} That assumption, however, is not consonant with the understanding of that phrase when Congress enacted the APA. The contemporary case law extended the field beyond the locus of combat to “any place . . . where military operations are being conducted,”\textsuperscript{303} including domestic training facilities.\textsuperscript{304} The history thus indicates that the military authority exception could well have been intended to encompass military prisons, regardless of where they are located. Accordingly, the criticism that “Judge Randolph’s interpretation of the statute is flawed because it deprives the phrase ‘in the field’ of any meaning whatsoever” is not well taken.\textsuperscript{305}

Recognizing the historical significance of Congress’s decision to use the phrases \textit{in the field} and \textit{time of war} in the APA and the contemporary meaning of those phrases in 1946 could narrow the circumstances in which the APA would pave an avenue to judicial relief for military detainees. Similarly, although no court has analyzed the military authority exception in the context of an environmental claim, acknowledging that the exception’s meaning in 1946 was somewhat broader than a modern reader would presume from the plain language might lead courts to decline to hear claims under environmental statutes that are actionable only through the APA. On the other hand, a court might limit its analysis to the plain language and official legislative history of the APA, thus paving a broader path to the courthouse door for detainees and environmental claimants.

\textbf{CONCLUSION}

The military authority exception is plainly narrower than the Walter–Logan bill’s exemption for “any matter concerning or relating to the military establishments” and the exemptions for military functions in §§ 4 and 5 of the APA. But Congress’s understanding of the phrases \textit{in the field} and \textit{time of war} may have been somewhat broader than a modern reader would suppose. Contemporary case law held that \textit{in the field} was not limited to the locus of combat and \textit{time of war} did not require a congressional declaration of war. Moreover, Congress was under pressure from the military to reduce the bill’s impact on military functions. As it did in many other provisions of the APA, Congress used ambiguous language in the

\textsuperscript{302} \textit{Id.} at 188 & n.473. Similarly, Professor Chesney assumes, without analysis, that decisions to transfer custody of Guantanamo detainees to nations that might subject them to torture are not made “in the field.” Chesney, \textit{supra} note 1, at 604 n.126.


\textsuperscript{304} \textit{See} Hines v. Mikell, 259 F. 29, 32–34 (4th Cir. 1919).

\textsuperscript{305} Jinks \& Sloss, \textit{supra} note 1, at 188.
“military authority” exception in order to compromise with competing interests and ease the bill’s passage. 386 Lacking any understanding of that historical context, courts and commentators have made assumptions about the exception’s scope based only on their modern interpretation of the plain language and have failed to take into account the contemporary understanding of the exception when Congress enacted the APA in 1946.

The historiographic analysis of the military authority exception presented in this article reveals questions that remain to be answered. Why has the exception not arisen in litigation more often? To what extent has the meaning of the exception changed since 1946 in response to shifts in fundamental doctrines of administrative law? And most importantly, is the history of the military authority exception, beyond the plain language and official legislative history, even relevant to a court’s analysis of its jurisdiction under the APA?

Litigation concerning persons held at military facilities or the environmental impacts of military actions may give the courts further opportunities to answer some of these questions. The Supreme Court, however, does not appear to be in any rush to discuss this provision. The Court declined the opportunity in *Bismullah v. Gates* in which the government’s rehearing petition in the D.C. Circuit generated five separate opinions, including a face-off between Judge Randolph and Judge Ginsburg about the relevance of the military authority exception to the procedures applied in petitions for review under the Detainee Treatment Act of 2005. 387


387. 501 F.3d 178 (D.C. Cir. 2007), *reh’g* denied, 503 F.3d 137, 138–39 (D.C. Cir. 2007), *reh’g en banc* denied, 514 F.3d 1291, 1293 (D.C. Cir. 2008) (per curiam), vacated and remanded, 128 S. Ct. 2960 (2008). The Detainee Treatment Act of 2005 gives the D.C. Circuit exclusive jurisdiction to review determinations by Combatant Status Review Tribunals that aliens are properly being detained as “enemy combatants.” *Id.* at 182. Among other things, the D.C. Circuit held that the record on review includes not just information submitted to the Tribunal, but all “reasonably available information in the possession of the U.S. Government bearing on the issue of whether the detainee meets the criteria to be designated as an enemy combatant . . . .” *Id.* at 139. Judge Randolph dissented from the denial of rehearing *en banc*. Although he acknowledged that the case was not “controlled . . . by the APA,” he opined that “the detention of enemy combatants, and the review processes related to them, are military ‘functions’ the APA specifically exempts.” 514 F.3d at 1303 n.3 (Randolph, J.). He attached as an addendum his concurrence in *Al Odeh* to further explain that point. *Id.* Judge Ginsburg responded in his opinion concurring in the denial of rehearing *en banc* that no court has ever held that a Combatant Status Review Tribunal is a “military authority exercised in the field in time of war” and “no party to this case has suggested as much.” *Id.* at 1294 n.3 (Ginsburg, C.J.), concurring in the denial of rehearing *en banc*). In Judge Ginsburg’s view, the Tribunal is “not a court[-], martial, not a military
The Supreme Court also has shown some willingness to defer to quintessentially military decisions, regardless of whether they fall within the scope of the military authority exception, and thus might not see a need to address the exception any time soon. In Winter, for example, the Supreme Court did not mention the military authority exception when it reviewed an order enjoining the use of sonar in Navy training, even though the courts’ jurisdiction rested in part on the APA. Instead, the Court reiterated that it would give “great deference to the professional judgment of military authorities concerning the relative importance of a particular military interest” and to the “essentially professional military judgments” concerning “the composition, training, equipping, and control of a military force.” The Court held that the “lower courts failed properly to defer to senior Navy officers’ specific, predictive judgments about how the preliminary injunction would reduce the effectiveness of the Navy’s . . . training exercises.”

The history of the military authority exception calls into question the courts’ continued willingness to employ common law deference doctrines to avoid reviewing military actions. Congress chose to use narrow language in the military authority exception, obviously intending for military actions


389. 129 S. Ct. at 377 (quoting Goldman v. Weinberger, 475 U.S. 503, 507 (1986)).

390. Id. (quoting Gilligan v. Morgan, 413 U.S. 1, 10 (1973)).

391. Id. at 378. Perhaps the Supreme Court was willing to defer so broadly in Winter because the APA expressly preserves the courts’ equitable discretion to fashion a remedy appropriate to the case, 5 U.S.C. § 702 (2006), because of the “longstanding consensus” that in some “domains affecting national security . . . executive action must proceed unhampered by even the threat of legal regulation and judicial review, no matter how deferential that review might be on the merits,” Vermeule, supra note 254, at 1133, or simply because the case concerned harm to marine mammals, not human beings. See Jonathan Cannon, Environmentalism and the Supreme Court: A Cultural Analysis, 33 Ecology L. Q. 363 (2006); Richard J. Lazarus, Thirty Years of Environmental Protection Law in the Supreme Court, 17 Pace Envtl. L. Rev. 1 (1999); Richard J. Lazarus, Restoring What’s Environmental About Environmental Law in the Supreme Court, 47 UCLA L. Rev. 703 (2000); Richard J. Lazarus, Environmental Law and the Supreme Court: Three Years Later, 19 Pace Envtl. L. Rev. 653 (2002). The Court probably declined to address the military authority exception because no party raised it.
falling outside the scope of the exception to be subject to judicial review (unless some other provision precludes review). The history of the APA’s enactment demonstrates that Congress believed judicial review of administrative action, including military action, was one of the keys to protecting individual liberties and avoiding totalitarianism. “With the enactment of the APA in 1946, the judicial method in most administrative law cases should have shifted to the task of interpreting the new statute, rather than continuing to formulate and apply judicially-created doctrines.”

Professor Masur demonstrates that military agencies are entitled to no greater deference than other expert administrative agencies and that the rule of law demands that courts exercise their authority to review military actions. Indeed, the plain language of the APA itself mandates judicial review in appropriate cases. The Supreme Court recently held that agency action may not be subjected to a stricter standard of review than that set forth in the APA. By the same token, it should not be subjected to a lesser standard of review, or excused from review altogether, based on a doctrine that has no “basis in the text of the statute.”

Whatever the reason for the Supreme Court’s continued willingness to defer broadly to the military, Winter and Bismullah indicate that those who desire a definitive interpretation of the military authority exception from the Court may have to be patient. In the interim, the historical context provided in this Article may lend some aid to those who are trying to bring Congress’s intentional ambiguity into focus.

392. Duffy, supra note 267, at 121.
393. Masur, supra note 370, at 493, 520.
394. 5 U.S.C. §§ 702 (“A person suffering legal wrong because of agency action... is entitled to judicial review thereof.”), 704 (“Agency action made reviewable by statute and final agency action for which there is no other adequate remedy in a court are subject to judicial review.”), 706 (“To the extent necessary to decision and when presented, the reviewing court shall decide all relevant questions of law, interpret constitutional and statutory provisions, and determine the meaning or applicability of the terms of an agency action.”); Duffy, supra note 267, at 130.
396. Id. at 1811.
397. Professor Babcock was somewhat prescient when she said that “[c]ourts have generally been protective of the military when confronted with a conflict between NEPA mandates and military needs.” Babcock, supra note 2, at 113–16. If she is also correct that “the circumstances of 9/11 and the continual state of war may change that posture,” id. at 116, perhaps the Supreme Court will address the “military authority” exception in the not-too-distant future.
“YOUR RESULTS MAY VARY”:
PROTECTING STUDENTS AND
TAXPAYERS THROUGH TIGHTER
REGULATION OF PROPRIETARY SCHOOL
REPRESENTATIONS

AARON N. TAYLOR*

TABLE OF CONTENTS

Introduction .......................................................... 730
I. Higher Education Misrepresentations and Fraud ...... 733
   A. Inflated Placement and Completion Rates ............. 735
   B. Inappropriate Compensation Arrangements .......... 736
   C. Entrance Test Improperities .......................... 739
   D. Cohort Default Rate Manipulation ....................... 740
II. The Commmotization of Higher Education .................. 742
   A. Knowledge-Based Economy .............................. 744
   B. New Technologies ........................................ 745
   C. Globalization ............................................ 746
   D. Neoliberal Financial Aid Policies ....................... 747
   E. Changing Student Populations ........................... 748
   F. Rising Tuition ............................................. 749
   G. Demands for Accountability .............................. 750
   H. New Higher Education Providers ......................... 751
   III. The Rise of Proprietary Colleges ......................... 752
       A. Title IV Expansion .................................... 753

* Assistant Dean for Admissions and Scholarships and Adjunct Professor of
  Education Law, University of Arkansas at Little Rock William H. Bowen School of Law,
  B.A., North Carolina A&T State University; J.D., Howard University School of Law; Ed.D.,
  Vanderbilt University. The Author wants to thank Lisa S. Taylor and Theresa M. Beiner
  for their helpful feedback and guidance. He thanks Latrenia Byrd and Elizabeth Davis for
  their research assistance.
INTRODUCTION

“Graduate with a career!”
“Make more cash!”
“Change your life!”
“It’s easy! Just pick up the phone!”

These were likely some of the exhortations Trina Thompson heard (between episodes of The Jerry Springer Show and Maury Povich) before deciding to enroll at Monroe College—a career-focused proprietary (for-
profit) college with campuses in New York and the Caribbean. Trina likely wanted to start a career, make more money, and change her life when she enrolled in Monroe’s bachelor of business administration program in April 2008. But after graduating without a job in April 2009, Trina realized that achieving her objectives was not as easy as she had been led to believe. According to Trina, her 2.7 GPA and her “good” attendance record should have resulted in job interviews and eventually employment. Shortly thereafter, she filed a lawsuit against Monroe alleging inadequate career-placement assistance and seeking a tuition reimbursement of $70,000.

Trina’s lawsuit, while laughable in some respects, is nonetheless reflective of the commoditization of higher education—a trend that promotes the endeavor simply as a means to an end instead of a complicated undertaking. In that vein, the lawsuit is also instructive of risks associated with the aggressive and often deceptive promotion of future benefits by colleges hawking their wares. Representations made by some colleges rival the most optimistic—and often unfounded—diet pill claims. The end result is thousands of “Trinas” entering higher education full of misguided optimism and leaving bitter, unfulfilled, and most of all, in debt. Trina is lucky—at least she earned a degree. Most others in her position do not.

Individuals who are induced to enroll in an institution based on misrepresentations are allowed little recourse to recoup damages they may incur. The courts have been very reluctant to recognize certain causes of action against higher education institutions. And regulatory safeguards are

3. See id. (stating that Trina is suing because Monroe has not helped her to secure a job).
5. See id. (noting that Trina seeks $2,000 for stress induced by her failed job search).
7. Patrick F. Linehan, Dreams Protected: A New Approach to Policing Proprietary Schools’ Misrepresentations, 89 Geo. L.J. 753, 754 (2001) (“Unfortunately, existing legal doctrine and regulatory regimes are ill-suited to protect proprietary school students from such predatory marketing practices.”).
8. See id. at 764–65 (describing how the academic abstention doctrine raises “a significant obstacle” to students seeking to recover against proprietary schools under tort
principally focused on protecting public, rather than individual, interests. As a result, these individuals are left to bear the brunt of the improper actions of others. And in spite of regulatory safeguards, taxpayers pay a heavy price as well.

This Article argues that there is an urgent need for tighter regulation of higher education recruitment and marketing, particularly among colleges in the proprietary sector. Specifically, colleges that promote future employment and financial benefits to induce enrollment should be subject to heightened disclosure requirements. Akin to the “triad” that monitors institutions’ Title IV financial aid eligibility,\(^9\) federal, state, and non-governmental entities should monitor disclosures. The goal of such oversight would be to prevent misrepresentations from being made to prospective students.

People who lack in-depth knowledge of higher education are frequent targets of higher education misrepresentations. They tend to be poor,\(^10\) thus rendering the idea of escaping poverty in a matter of months very appealing. They also tend to be poorly educated,\(^11\) coming from families with little, if any, higher education experience.\(^12\) Lastly, they tend to be older and further removed from their last educational experience than traditional students,\(^13\) and they are more likely to have experienced past educational difficulty.\(^14\) These characteristics make these individuals particularly susceptible to deceptive marketing and unfounded promises.\(^15\)

---

9. *Id.* at 783 (“Federal law envisions a highly complex and comprehensive bureaucratic ‘triad,’ with state licensing systems and accrediting agencies playing a significant complementary role alongside federal eligibility and certification requirements.”).


11. *See id.* at 8 (providing a quote from an admissions counselor stating that her former employer enrolled students who were ill-prepared to complete the program).

12. *See id.* at 46 (statement of Nick Glakas, President, Career College Association) (“[Seventy percent of proprietary school students] are the first in their families to attend college . . . .”).

13. *See U.S. Gov’t Accountability Office, GAO-09-600, Proprietary Schools: Stronger Department of Education Oversight Needed to Help Ensure Only Eligible Students Receive Federal Student Aid 7* (2009), available at http://www.gao.gov/new.items/d09600.pdf [hereinafter *GAO, Stronger Oversight*] (“[O]ver half of the student population at proprietary schools is comprised of ‘non-traditional’ students, such as students who are 25 years old and older.”).

14. *See Linehan, supra note 7, at 736* (“Most [proprietary school] enrollees . . . have previously experienced educational failure.”).

15. *See Anti-Fraud Hearings, supra note 10, at 22* (statement of Rep. Maxine Waters) (arguing that victims of proprietary schools’ misrepresentations “are less likely to complain,
Similarly, these characteristics put these individuals at higher risks of dropping out before program completion and eventually defaulting on student loans. The costs of higher education failure are high; therefore, the costs of higher education misrepresentations are high. Unfortunately, taxpayers, in addition to the victims, are saddled with these costs. As a result, effective oversight of higher education marketing and recruitment would not only protect individual students, but also contribute to the country’s fiscal health.

In making the case for better oversight, this Article describes, in Part I, the multifaceted nature of higher education misrepresentations and fraud. Part II discusses the commoditization of higher education. Part III chronicles the rise of proprietary colleges and describes their aggressive marketing and recruitment practices. Part IV argues that an imperative exists for tighter regulation of higher education marketing and recruitment. Part V asserts that current safeguards are inadequate in protecting students and taxpayers from misrepresentations. Lastly, Part VI presents proposals for regulating higher education representations in a manner that protects the public from misrepresentations without unduly restricting competition and protected speech.

I. HIGHER EDUCATION MISREPRESENTATIONS AND FRAUD

When 60 Minutes visited campuses of the Katharine Gibbs School and Brooks College to investigate proprietary school business practices, the

and when they do they are less effective, because they don’t know where to complain, or how to articulate their complaint, as they do not know the requirements of the law”.


17. Id. at 12 (“When students do not make payments on their federal loans and the loans are in default, the federal government and taxpayers assume nearly all the risk and are left with the costs.”).

18. Katharine Gibbs School consisted of two campuses in New York and one in Pennsylvania that were owned by Career Education Corporation (CEC). After an unsuccessful attempt to sell the campuses, CEC announced on February 15, 2008, that it would close one of its Katharine Gibbs campuses and convert the other to another brand within its corporation. Press Release, Career Educ. Corp., Career Education Corporation Announces Plans to Teach-Out Programs at Selected Schools Held for Sale (Feb. 15, 2008), http://phlx.corporate-ir.net/phoenix.zhtml?c=87390&g=p=irol-newsArticle_print&ID=1108811&highlight= [hereinafter CEC, Teach-Out].

producers found a virtual treasure trove of corruption. In its feature, *For-Profit College: Costly Lesson*, the newsmagazine documented recruiter misrepresentations and entrance exam improprieties, providing insider views into the high pressure world of proprietary school recruiting and the debilitating effects of unscrupulous recruitment tactics. These tactics have drawn the ire of former students, policymakers, governmental regulators, and consumer watchdogs, and have formed the bases of lawsuits, investigations, and congressional hearings. Proprietary schools, for all of their virtues, have a checkered collective past. The rate of investigations and sanctioning among the sector exceeds that of the nonprofit sector.

And even though corruption has been greatly reduced since the 1980s and early 1990s, proprietary schools are still tainted by improprieties that many believe are the direct result of their tuition-driven, profit-generating motives.

20. Rebecca Leung, *For-Profit College: Costly Lesson*, 60 MINUTES, Jan. 30, 2005, http://www.csbcnews.com/stories/2005/01/31/60minutes/main670479.shtml (“The admission counselors told 60 Minutes they were expected to enroll three high school graduates a week, regardless of their ability to complete the coursework. And if they didn’t meet those quotas, they were out of a job . . . . They all say the pressure produced some very aggressive sales tactics.”).

21. See id. (interviewing unemployed and underemployed graduates of Brooks College who stated that admission counselors induced their enrollment with promises of prestigious employment in the fashion industry upon graduation).


23. See Lisa K. Foster, CAL. STATE LIBRARY, CRB 04-010, FOR-PROFIT POSTSECONDARY EDUCATIONAL INSTITUTIONS: OVERVIEW OF ACCREDITATION AND STATE AND FEDERAL OVERSIGHT 14 (2004) (“During the 1970s and 80s, institutions operated with little or no oversight and few constraints in recruiting and training students. A large number of institutions did not provide the training advertised, did not comply with fair consumer practices, and mismanaged finances.”).

24. See Línehan, *supra* note 7, at 760 (“[A]lthough proprietary schools compose one-third of the approximately 6,000 schools eligible for federal student grants and loans, they recently accounted for three-fours of the Department [of Education]'s student loan fraud and abuse investigations.”).

25. See Catherine Elton, *Degrees of Difficulty: The Truth About Online Universities, Consumers Dig.*, Mar./Apr. 2009, at 20–21 (describing “a disturbing pattern” of aggressive recruiting tactics by for-profit schools that ensnare unqualified students who ultimately fail out or qualified students who receive little benefit from the noncompetitive learning environment).
Tighter Regulation of School Representations

A. Inflated Placement and Completion Rates

Proprietary school misrepresentations and other improprieties cost individual students and taxpayers. Misrepresentations typically pertain to job placement rates, a benchmark upon which proprietary schools market themselves. These rates can be misrepresented in terms of the number of students placed, the average salary offered by those placements, and the overall availability of quality placements. A prerequisite to placement, completion rates are often misrepresented or not disclosed as required. A common scenario is one experienced by the 60 Minutes producer who posed as a prospective student. She was told by a Katharine Gibbs representative that the school’s placement rate was 89%; however, Department of Education (DOE) data put the rate at 29%. A former Brooks recruiter captured the integrated nature of placement and completion rate misrepresentations when she summarized the essential elements of her deceptive sales pitch: “We are telling you that you are going to have a 95 percent [chance of getting] a job paying $35,000 to $40,000 a year by the time you are done in 18 months.”

26. See Linehan, supra note 7, at 757 (discussing how proprietary schools place “advertisements on daytime and late night television to reach the unemployed and those seeking new jobs”); see also Foster, supra note 23, at 12 (describing placement rates as “[t]he ultimate outcome measure”).

27. See Anti-Fraud Hearings, supra note 10, at 15 (statement of Rep. Maxine Waters) (“The biggest misrepresentations made to students that convince them to enroll are anticipated starting salary... and the placement rate... The starting salaries that prospective students are told are seldom true. Many schools tout a 90% plus placement rate. But these are self reported rates and not necessarily accurate.”); see also Linehan, supra note 7, at 759 (discussing the pressure that oversupplied labor markets place on proprietary schools); U.S. GEN. ACCOUNTING OFFICE, GAO/T-HEHS-96-158, HIGHER EDUCATION: ENSURING QUALITY EDUCATION FROM PROPRIETARY INSTITUTIONS 11–12 (1996), [hereinafter GAO, Ensuring Quality] (statement of Cornelia M. Blanchette), available at http://www.gao.gov/archive/1996/he96158t.pdf (discussing the financial ramifications of students incurring debt in order to train for jobs for which little demand exists). But cf. GAO, OVERSUPPLIED OCCUPATIONS, supra note 22, at 11 (discussing benefits conferred upon students who trained in oversupplied fields).

28. See, e.g., Leung, supra note 20 (reporting that representatives from the Katharine Gibbs School in New York lied about the school’s graduation rate).


30. Leung, supra note 20.

also been accused of misrepresenting the transferability of credits, \(^{32}\) programmatic content, \(^{33}\) and accreditation status. \(^{34}\) In addition, shareholders of publicly-traded schools have filed lawsuits alleging misrepresentations and omissions in annual reports and other required disclosures. \(^{35}\)

**B. Inappropriate Compensation Arrangements**

Improprieties also involve the fraudulent obtaining of Title IV financial aid funds. \(^{36}\) This fraud tends to be systematic in nature and often concerns the manner in which recruiters are compensated, the process by which students are enrolled and matriculated, and the manipulation of regulatory safeguards. Recruiter compensation is a frequent basis of lawsuits against proprietary schools. Title IV forbids schools from compensating recruiters based solely on the number of students they induce to enroll. \(^{37}\) This ban on “incentive compensation” is intended to protect students by lessening the pressure on recruiters to induce enrollment at all costs. \(^{38}\) The National

\(^{32}\) See, e.g., Till v. Delta Sch. of Commerce, Inc., 487 So. 2d 180, 182 (La. Ct. App. 1986) (alleging a breach of contract for several misrepresentations the school used to induce plaintiff to enroll, including the ability to transfer credits to other institutions).

\(^{33}\) See, e.g., Phillips Colls. of Ala., Inc. v. Lester, 622 So. 2d 308, 309 (Ala. 1993) (alleging that the school misrepresented the nature of its instruction).

\(^{34}\) See, e.g., Malone v. Acad. of Court Reporting, 582 N.E.2d 54, 55–56 (Ohio Ct. App. 1990) (suing for false claims of accreditation).


\(^{36}\) Title IV of the Higher Education Act governs the provision of federal financial aid funds for higher education. The statute governs mostly need-based programs, such as Federal Pell Grants, supplemental educational opportunity grants, payments to the States and institutions for need-based financial aid, and other special programs and projects. 20 U.S.C. § 1070(a) (2006).

\(^{37}\) See 34 C.F.R. § 668.14(b)(22)(i) (2009) (requiring that “an institution agrees that . . . [i]t will not provide any commission, bonus, or other incentive payment based directly or indirectly upon success in securing enrollments or financial aid to any person or entity engaged in any student recruiting or admission activities . . . .”). But see id. § 668.14(b)(22)(ii) (listing exceptions to this prohibition).

\(^{38}\) *H.R. 1992, The Internet Equity and Education Act of 2001: Hearing Before the Subcomm. on 21st Century Competitiveness of the H. Comm. on Education and the Workforce, 107th Cong. 14 (2001)* (statement of Lorraine Lewis, Inspector General of the United States), available at http://frwebgate.access.gpo.gov/cgi-bin/getdoc.cgi?dbname=107_househearings&docid=f:77908.pdf (“The prohibition was designed to protect students from the high pressure tactics used by recruiters to enroll students in programs for which they may not have been prepared or did not want. The students were saddled with unwanted debt at increased cost to the
Association for College Admission Counseling (NACAC) argues that “reducing the basis for compensation to the number of students enrolled in any circumstance introduces an incentive for recruiters to actively ignore the student interest in the transition to postsecondary education.” In spite of this ban, accusations against proprietary school compensation structures abound. In a 2003 audit, the DOE found that the University of Phoenix (UOP) had violated the ban. The audit described a system under which recruiter compensation was tied to “asses in classes” and a culture where a recruiter’s enrollment numbers could mean the difference between lucrative employment and unemployment. According to the DOE, the high pressure environment fostered by UOP’s compensation structure led to the very dangers that the ban is intended to prevent. Unqualified students and those facing unfavorable family or financial circumstances were pressured to enroll. Recruiters also pressured students to take out loans to pay tuition, and recruiters were encouraged to cease providing taxpayers.”


40. See, e.g., Loonin & Devanthery, supra note 29, at 1 (summarizing these accusations).

41. University of Phoenix (UOP) is “the nation’s largest private university, offering undergraduate and graduate degree programs at more than 200 locations, as well as online in most countries around the world.” University of Phoenix, History, http://www.phoenix.edu/about_us/about_university_of_phoenix/history.html (last visited Aug. 3, 2010).


43. The audit also referenced “butts in seats.” Id. at 10.

44. See, e.g., id. at 7–8 (listing a salary schedule showing that UOP recruiters could make upwards of $120,000 per year and discussing how the potential for high salaries often prompted employees in other departments to seek recruiter positions).

45. Id. at 12 (quoting a UOP enrollment director saying in a recruiter meeting: “My job is on the line. And I need you guys to perform. . . . if you’re not doing your job, you’re going to lose your job. And if you’re not hitting your goals, that’s how we’re going to measure if you’re doing the job. And . . . I don’t mean applications in. I mean starts.”).

46. Even when recruiters felt other educational options, such as community colleges, would be better for individual prospects, they were forbidden from making such recommendations. Id. at 24.

47. Recruiters were expected to complete financial aid documentation for students, and the forging of signatures by recruiters was commonplace. Id. at 25; see also Anti-Fraud
support to these students once their enrollment was credited for salary purposes.\textsuperscript{48} UOP ended up paying $9.8 million to settle the investigation but admitted no wrongdoing.\textsuperscript{49}

The issue of incentive compensation has spawned a spate of lawsuits filed pursuant to the federal False Claims Act (FCA).\textsuperscript{50} The \textit{qui tam}\textsuperscript{51} provisions of the Act permit private citizens to bring lawsuits alleging fraud against the government on behalf of the government.\textsuperscript{52} If the plaintiffs, referred to as relators, are successful in winning damages on behalf of the government, they share in the recovery.\textsuperscript{53} The law is intended to incentivize whistle-blowing by individuals with first-hand knowledge of fraud against the government.\textsuperscript{54} Proprietary schools are particularly susceptible to FCA

\textit{Hearings, supra note} 10, at 41 [statement of Paula L. Dorsey, former Director of Admissions, Bryman College] (discussing recruiters pressuring students to “improperly obtain social security numbers and signatures of other family members by whatever means necessary for the hopes of getting a ‘better’ financial aid package”).

\textsuperscript{48} See UOP \textit{PROGRAM REVIEW REPORT, supra note} 42, at 24 [illustrating how, to be considered enrolled for purposes of calculating recruiter salary, students had to “attend three nights of the first five-week course of a bachelors’ program or, for graduate students, attend two nights . . . and be scheduled to attend a second class. After the student has met these criteria . . . UOP requires [recruiters] to pursue new enrollments . . . ”].

\textsuperscript{49} See Apollo Group, Inc., The February 2004 Program Review Report Relating to the University of Phoenix was Fundamentally Flawed, http://www.apollolegal.com/prCritique.html [last visited July 31, 2010] (“[W]e believe that the terms of the Settlement Agreement between UOP and [the DOE] constitute a clear, albeit implicit, rejection of the [audit] and its alleged findings. The simple fact is that if the alleged findings in the [audit] had any merit—which they do not—[the DOE] would not and could not have settled the issues raised in the [audit] on the terms that it did. The terms of the Settlement Agreement are very favorable to UOP.”).


\textsuperscript{51} “\textit{Qui tam}” is short for the Latin phrase . . . “who pursues this action on our Lord the King’s behalf as well as his own.” \textit{Vt. Agency of Natural Res. v. United States ex rel. Stevens}, 529 U.S. 765, 768 n.1 (2000).

\textsuperscript{52} 31 U.S.C. § 3730(b)(1) (2006) (“A person may bring a civil action for a violation of section 3729 for the person and for the United States Government. The action shall be brought in the name of the Government.”).

\textsuperscript{53} Defendants found to have committed fraud under the FCA are assessed a fine ranging from $5,000 to $10,000 and must pay three times the government’s damages arising from the fraud. § 3729(a)(7). Plaintiffs can recover between 25% and 30% of the latter assessment. \textit{Id.} § 3730(d)(2).

\textsuperscript{54} See, e.g., 
Constitutionality of the \textit{Qui Tam} Provisions of the False Claims Act, 109 Op. Att’y Gen. 4–5 (1989) (stating that the 1986 Amendments to the FCA were the result of Congress being “dissatisfied with the way the executive branch was enforcing government procurement laws” and therefore desiring to “deputize” private citizens to ensure effective
lawsuits; plaintiffs have filed many lawsuits alleging fraud in the obtainment of Title IV funds, primarily arising from recruiter compensation arrangements. FCA actions of this type have proved fruitful. In 2007, Oakland City University paid $5.3 million to settle such a lawsuit. In late 2009, UOP settled an FCA suit for almost $80 million. In both cases, the whistle-blowers received millions of dollars for their efforts.

C. Entrance Test Improprieties

Proprietary schools have also faced accusations of improprieties relating to the enrollment of ineligible students for financial aid purposes. Entrance test improprieties and the falsification of attendance records commonly form the bases of these accusations. For federal financial aid purposes, students not possessing a high school diploma or General Equivalency Diploma (GED) must take and pass “an independently administered test of basic math and English skills, called an ‘ability-to-benefit’ or ATB test. The intent of the test is to measure whether students have the basic skills needed to benefit from higher education and succeed in school.” An investigation by the Government Accountability Office (GAO) documented test administrators giving out answers and changing answers to ensure passing scores. At the Katharine Gibbs School, the 60 Minutes producer intentionally failed the entrance exam, but was allowed to retake it and was told her second score was sufficient for admission purposes. Additionally,

---


56. See generally Hatch & Arnold, supra note 50, at B16 (providing an overview of FCA lawsuits brought against educational institutions).


58. Press Release, Apollo Group, Inc., Apollo Group, Inc. Resolves University of Phoenix False Claims Act Case (Dec. 14, 2009), http://phx.corporate-ir.net/phoenix.zhtml?c=79624&p=irol-newsArticle&ID=1365655 (“Under the terms of the agreement, the Company will pay $67.5 million to the United States. A separate agreement provides for the payment by the Company of $11 million in attorneys fees to the plaintiffs, as required by the False Claims Act.”).

59. GAO, STRONGER OVERSIGHT, supra note 13, at 9.

60. Id. at 22.

61. Her initial score was seven out of fifty. Upon retaking the test, the admissions recruiter said she got fourteen out of fifty answers correct, which was sufficient for enrollment. Leung, supra note 20.
recruiters alleged that schools forced them to enroll students even without required exam scores. A federal raid in 2004 of the headquarters of ITT Educational Services (ITT) and ten of its campuses is believed to have been related to allegations that the proprietary educational provider was overstating student enrollment in order to increase its federal financial aid revenue. The investigation found no wrongdoing among executives within the company; however, investigations of individual campuses remained open.

D. Cohort Default Rate Manipulation

Schools are commonly accused of manipulating cohort default rates. Cohort default rates provide schools with incentives to minimize defaults

62. See Anti-Fraud Hearings, supra note 10, at 41 (statement of Paula L. Dorsey, former Director of Admissions, Bryman College) (“There were students that had never taken the exams or who had failed the exams, sitting in class. I was instructed to clean up the files by whatever means necessary even if it meant backdating things.”).

63. See ITT Educational Services, Inc., Investor Relations Overview, http://www.ittesi.com/phoenix.zhtml?c=94519&p=irol-IRHome (“ITT Educational Services, Inc. . . provides accredited, technology-oriented undergraduate and graduate degree programs . . . to help students develop skills and knowledge they need to pursue career opportunities in a variety of fields. It owns and operates more than 120 ITT Technical Institutes and Daniel Webster College. [It] serves approximately 80,000 students at its campuses in 38 states and online.”).

64. See Press Release, ITT Educational Services, Inc., ITT Educational Services, Inc, Reports that It Has Been Served with a Search Warrant and Related Subpoenas from the U.S. District Court in Texas (Feb. 25, 2004), http://www.ittesi.com/phoenix.zhtml?c=94519&p=irol-newsArticle&ID=498922&highlight= (reporting that agents sought information pertaining to “placement figures and rates, retention figures and rates, graduation figures and rates, attendance figures and rates, recruitment and admissions materials, student grades, graduate salaries[3] and transferability of credits to other institutions”).


67. See GAO, STRONGER OVERSIGHT, supra note 13, at 10 (“[DOE] computes default rates for all schools with students who receive Title IV loans . . . by tracking whether borrowers in a cohort—a group of students who begin repaying their loans in a given fiscal year—at each school default on their federal student loans over a 2-year period.”).
among their student-borrowers and protect taxpayers from the costs of excessive defaults.68 Schools with default rates that exceed certain thresholds can have their federal financial aid eligibility stripped.69 Historically, proprietary schools have recorded higher default rates than nonprofit institutions,70 and numerous proprietary schools have been closed due to unacceptably high rates.71 Given the extent to which proprietary schools rely on Title IV aid, losing eligibility is akin to an institutional death sentence.72

In an effort to reduce cohort default rates, some proprietary schools unilaterally pay off loans obtained by students who later withdraw.73 While this may seem altruistic on its face, the true motivation is thus: Students who take out loans to pay for school, but withdraw before completing an academic program, are at high risk of defaulting on those loans;74 therefore, schools settle the loans for these students as a means of protecting their Title IV eligibility from likely defaults.75 The schools then engage in aggressive efforts to collect the debt from students, offering less favorable repayment terms than those available through Title IV.76 As a result, former students have brought lawsuits alleging contract-based causes of action.77


69. Schools with default rates of 25% or above for three years or above 40% for one year lose federal student loan eligibility for the remainder of the year after notification and for the subsequent two years, pending appeals and adjustments. See id. at 2.4-2 to -3 (discussing benefits conferred upon schools with default rates of less than 5% or 10%).

70. GAO, STRONGER OVERSIGHT, supra note 13, at 13, 15–17.

71. See, e.g., Linehan, supra note 7, at 760 (noting that the closures have resulted from “fraudulent misrepresentations and deceptive marketing”).


74. See, e.g., GAO, STRONGER OVERSIGHT, supra note 13, at 20 (providing an overview of six different research studies that showcase trends in students’ default rates).

75. Lederman, supra note 73.

76. Id.

77. See NACAC TESTIMONY, supra note 39, at 16 (summarizing a lawsuit against UOP filed by former students whose loans were paid by UOP without their permission).
The DOE has expressed disapproval of this behavior. In 2008, an inspector for the DOE found that a proprietary school in New York had improperly repaid Title IV loans or returned loan funds for 301 students who withdrew during their first semester of study.\textsuperscript{78} School officials stated that this “default prevention policy” was implemented due to past problems with the school’s default rate.\textsuperscript{79} As a result of its findings, the DOE made various recommendations to the school, including ceasing the practice of repaying loans and ending pending collection efforts resulting from that practice.\textsuperscript{80}

Proprietary school misrepresentations and improprieties are multifaceted and costly. However, for many of these institutions, their profit-generating motives make the allure of such behaviors irresistible. These motives are part of a larger trend of commoditization of higher education. Commoditization has introduced a market ethos into higher education that has changed the way all institutions operate—for better and worse.

\section*{II. The Commoditization of Higher Education}

The United States has the most market-oriented system of higher education in the world.\textsuperscript{81} Compared to its European counterparts, the American system has developed with little direct influence from the federal government.\textsuperscript{82} This freedom has spawned a vast, entrepreneurial expansion of higher education within the United States.\textsuperscript{83} Through much of the country’s history, new colleges were established “without restraint,”\textsuperscript{84} and the result has been a proliferation and democratization of higher

\begin{thebibliography}{9}
\item\textsuperscript{78} \textit{U.S. Dep’t of Educ., Office of Inspector Gen., ED-OIG/A02H0007, Final Audit Report: Technical Career Institutes, Inc.’s Administration of the Federal Pell Grant and Federal Family Education Loan Programs} 3 (2008), \url{http://www2.ed.gov/about/offices/list/oig/auditreports/6y2008/a02h0007.pdf}.
\item\textsuperscript{79} Id. at 4.
\item\textsuperscript{80} Id.
\item\textsuperscript{81} \emph{See, e.g.}, David D. Dill, \textit{Allowing the Market to Rule: The Case of the United States}, 57 \textit{Higher Educ.} Q. 136, 137 (2003) (discussing the increased “marketization” of higher education and its impact on the public interest).
\item\textsuperscript{83} \emph{See, e.g.}, Christopher J. Lucas, \textit{American Higher Education} 116–19 (2006) (discussing the proliferation of colleges in the early to mid-nineteenth century United States).
\item\textsuperscript{84} \emph{Id.} at 117.
\end{thebibliography}
education that stand in contrast to the elitist systems in Europe and other parts of the world.\textsuperscript{85} Historically, however, the higher education market has been largely protected from the sink-or-swim pressures that have characterized other industries. Higher education has enjoyed strong public support, and vast sums of public resources have flowed with virtually no strings attached.\textsuperscript{86} However, various economic and social trends have led to a diminishment of higher education’s protected market standing, and as a result, a capitalistic academic system has come to the fore.\textsuperscript{87}

The commoditization of higher education is a transcendent phenomenon. This multifaceted trend is the result of many interrelated factors, including demands of the knowledge-based economy, new technologies, increased globalization, neoliberal financial aid policies, changing student populations, rising tuition, demands for accountability, and the advent of new higher education providers. One result of commoditization is a higher education market rooted in capitalistic principles; thus, any discussion of the factors contributing to commoditization must begin with a discussion of academic capitalism.

Like capitalism in general, academic capitalism is about competition—competition for funding, students, and—for some schools—prestige.\textsuperscript{88} The primary competitors are institutions, which are embodied by the actors who operate therein: faculty, students, and administrators.\textsuperscript{89} Networks are central to viability within the academic capitalist system.\textsuperscript{90} As such, institutional actors seek to link institutions (and themselves) to the modern, knowledge-based economy.\textsuperscript{91} These links most often take the form of “new circuits of knowledge”—partnerships with the private sector, investments in marketing, product development and student services, and an expanded managerial core to handle these new demands.\textsuperscript{92} Fundamentally, the goal

\textsuperscript{85} See id. at 108 (arguing that republican ideals in the young nation’s history contributed to liberalization of higher education in the United States).

\textsuperscript{86} See, e.g., Trow, supra note 82, at 57 (calling the federal government’s approach to disbursing land grants under the Morrill Acts “extraordinarily permissive”).

\textsuperscript{87} See DAVID L. KIRP, SHAKESPEARE, EINSTEIN, AND THE BOTTOM LINE: THE MARKETING OF HIGHER EDUCATION 2 (2003) (“For better or worse—for better and worse, really—American higher education is being transformed by both the power and the ethic of the marketplace.”).

\textsuperscript{88} See, e.g., SHEILA SLAUGHTER & GARY RHODES, ACADEMIC CAPITALISM AND THE NEW ECONOMY: MARKETS, STATE, AND HIGHER EDUCATION 1 (2004).

\textsuperscript{89} Id.

\textsuperscript{90} See id. at 24 (providing examples of organizations and networks that act as intermediaries between the public and private sectors).

\textsuperscript{91} Id. at 1.

\textsuperscript{92} Id.
of institutions competing in this environment is to generate income, particularly from “alternative revenue streams,” with the assumption that robust, diversified funding will lead to greater prestige, better students, and increased viability.

A. Knowledge-Based Economy

In a knowledge-based economy, knowledge is a commodity that when exploited can reap tangible benefits upon the possessor. Higher education institutions are central to the knowledge-based economy because they are considered “a major source of alienable knowledge.” In other words, as creators, holders, and sellers of knowledge, these institutions hold the key to economic growth and the social cohesion that often accompanies such growth. As a result, one of the most pressing issues concerning the nature and function of higher education is how it can ensure that citizens can be productive participants within this economy. And just like the knowledge-based economy facilitates opportunity and success for possessors of vital knowledge, it “increasingly eliminates those without education and training beyond high school from employment opportunities that can support a middle-class standard of living.”

The relatively new emphasis on knowledge as a tool of economic vitality has created a market for education and has changed the motivations and

93. See id. at 11 (defining academic capitalism as “the pursuit of market and marketlike [sic] activities [by colleges and universities] to generate external revenues”).
94. See JAMES C. HEARN, AM. COUNCIL ON EDUC., DIVERSIFYING CAMPUS REVENUE STREAMS: OPPORTUNITIES AND RISKS 1 (2003) (identifying “alternative revenue streams” as those from sources other than state appropriations or tuition and fees).
95. Cf. id. at 5 (stating that institutions are being forced “to seek additional revenue sources” to meet the demands of “increased expectations” and the threats of “new providers and technologies”).
96. See SLAUGHTER & RHODES, supra note 88, at 15 (“[K]nowledge is a raw material to be converted to products, processes, or service.”).
97. Id.
99. See PATRICK M. CALLAN ET AL., THE NAT’L CTR. FOR PUB. POLICY & HIGHER EDUC., GOOD POLICY, GOOD PRACTICE: IMPROVING OUTCOMES AND PRODUCTIVITY IN HIGHER EDUCATION; A GUIDE FOR POLICYMAKERS 1 (2007), http://www.highereducation.org/reports/Policy_Practice/index.shtml (asserting that opportunities for higher education are not “as widespread as they need to be to place Americans in good jobs, fuel economic growth, promote social mobility and social justice, and sustain the country’s democratic ideals”).
mindsets of students. Students are increasingly viewing education as a product and themselves as consumers.  

101. Academic capitalism dictates that “[s]tudent consumers choose . . . colleges and universities that they calculate are likely to bring a return on educational investment.”  

102. The traditional student motivation—learning for learning’s sake—is making way for contemporary realities and pressures where education is increasingly seen as a private, rather than public, good.  

103. Therefore, as the value of a college education has skyrocketed,  

104. schools have been increasingly required to “reframe themselves as both education and business institutions.”  

B. New Technologies

Technology has changed “how students learn, how professors teach and conduct research, and how administrators manage institutions.”  

106. Today, many schools offer courses via distance learning frameworks. Professors are now able to analyze large datasets in seeking knowledge and are able to collaborate with colleagues from all over the world.  

107. Technology’s practical effect within the academic capitalist system has been to widen learning options for students and to increase competition among

101. See, e.g., SLAUGHTER & RHoades, supra note 88, at 12 (“[R]aising tuition . . . has heightened students’ and parents’ consumer consciousness about what they expect in terms of their educational experience . . . . These changed expectations reshape student identity from that of learner to that of consumer.”).

102. Id. at 1–2 (discussing how students “increasingly choose majors linked to the new economy, such as business, communications, [and] media arts”).

103. See, e.g., id. at 42–43 (“By the 1980s and 1990s, higher education was construed less as a necessary public or social good and more as an individual or private good, justifying ‘user pays’ policies.”) (citation omitted).

104. THOMAS J. KANE, THE PRICE OF ADMISSION: RETHINKING HOW AMERICANS PAY FOR COLLEGE 1 (1999) (noting that the difference in earnings between a high school graduate and a college graduate increased from 19% in 1980 to 52% in 1995). But see MICHAEL S. McPherson & MORTON OWEN SCHAPIRO, THE STUDENT AID GAME: MEETING NEED AND REWARDING TALENT IN AMERICAN HIGHER EDUCATION 40 (1996) (arguing that much of the increased economic differentiation between levels of education is mostly attributable to a decline in the value of a high school education, rather than an increase in the value of a college education).


107. See OECD POLICY ANALYSIS, supra note 98, at 18 (describing the impact of new digital technologies on higher education).
institutions for both students and faculty members. Technology has also led to greater institutional operating efficiencies, a trend that has allowed some institutions to better harness the academic capitalist system to their advantage.

C. Globalization

Broadly, globalization is “the flow of technology, economy, knowledge, people, values[,] and ideas ... across borders.” The phenomenon encompasses virtually all aspects of modern society, and the extent to which nations embrace it varies. As the world’s economies have become increasingly knowledge-based and integrated, demands for globalized trade in higher education have become more vocal. This trade embodies academic capitalism, as it “attract[s] foreign capital, invit[es] competition, and produc[es] a profit.” The globalized trade in higher education involves millions of people and billions of dollars.


109. See, e.g., CAROL A. TWIGG, THE NAT’L CTR. FOR PUB. POLICY & HIGHER EDUC., COURSE REDESIGN IMPROVES LEARNING AND Reduces Cost 1 (2005), available at http://www.highereduction.org/reports/pa_core/core.pdf (discussing a technology-based course redesign project that reduced the costs of offering these courses by an average of 37% for the participants).


112. See id. ("Globalization affects each country in a different way due to each nation’s individual history, traditions, cultures, resources[,] and priorities.") (quotation omitted).

113. See OECD POLICY ANALYSIS, supra note 98, at 103 ("There is growing interest in ways to build cumulative knowledge across the profession... ").


115. See id. at 11 ("In 2006, there were 2.7 million students studying abroad... [P]rojections... indicate that the demand... will increase to 7.2 million by 2025.").

Unsurprisingly, the principle motivation behind higher education globalization is economic. Individuals possessing foreign credentials can broaden employment prospects and demand higher salaries. Institutions can generate large sums of revenue by attracting foreign students. Governments can exploit the globalized knowledge of their workforce in dealings with foreign counterparts. In response, the World Trade Organization formally commoditized education under the General Agreement on Trade in Services (GATS) in 1995. Under GATS, education trade encompasses cross-border supply and consumption, as well as the presence of commercial providers in foreign countries.

Technology is the primary facilitator of globalization; as such, globalization has had many of the same effects on higher education as technology. Learning opportunities have been broadened, as have opportunities for academic collaboration.

D. Neoliberal Financial Aid Policies

Neoliberalism is premised on encouraging productivity through the empowerment of individuals as economic actors. Salient characteristics of neoliberal policies are “privatization, commercialization, deregulation, and reregulation.” The 1972 Amendments to the Higher Education Act

117. See VARGHESE, supra note 114, at 9 (noting that although economic growth depends on numerous factors, “human capital” has become an increasingly important economic driver). 118. See id. at 11 (explaining that cross-border education is mutually beneficial for providers and beneficiaries of education). 119. See, e.g., UNESCO, GLOBALIZED SOCIETY, supra note 111, at 8 (discussing the increased movement of students and scholars between countries “for commercial and for-profit purposes”). 120. See, e.g., OECD POLICY ANALYSIS, supra note 98, at 24 (“As the cost and multidisciplinary nature of research at the scientific frontier increase, countries will also increasingly need to draw on ideas generated abroad.”). 121. See VARGHESE, supra note 114, at 11 (“The GATS [General Agreement on Trade in Services] . . . represents a set of multilateral rules governing international trade in services.”). 122. Id. at 11–12. 123. UNESCO, GLOBALIZED SOCIETY, supra note 111, at 8 (listing new developments that globalization has helped facilitate in higher education). 124. See, e.g., SLAUGHTER & RHOADES, supra note 88, at 20 (“The neoliberal state focuses not on social welfare for the citizenry as a whole but on enabling individuals as economic actors. To that end, neoliberal states move resources away from social welfare functions toward production functions.”). 125. Id. at 21.
are considered landmark pieces of legislation. Among other things, the Amendments made federal student aid portable, meaning students could use their aid at the schools of their choice. The Amendments also “broadened the definition of which institutions were eligible to receive students with federal aid.” Specifically, proprietary and non-degree granting institutions could now collect Title IV aid from students. These shifts were classic neoliberal policies, in that they sought to encourage higher education efficiencies by empowering students and forcing institutions to engage in a “marketlike [sic] competition . . . for federally subsidized student tuition dollars.” The effects of neoliberal policies are apparent in other areas of higher education as well, including research.

E. Changing Student Populations

Broader access to higher education and the emergence of the knowledge-based economy have fueled higher education demand from new student markets. Students of color and nontraditional adult learners are seeking higher education in increasing numbers. Affirmative Action programs and the expansion of nonselective colleges and universities have helped spur this increased level of participation. The number of students of color undertaking higher education increased almost 49% between 1994 and 2004. Adult learners now comprise more than half of the college student

126. See, e.g., Trow, supra note 82, at 59 (asserting that the 1972 education amendments “established higher education as a national priority in its own right”).
127. See id. at 60 (explaining that student aid had previously been awarded directly to institutions in the form of block grants).
129. Id. at 6.
130. SLAUGHTER & RHODES, supra note 88, at 35.
131. See id. at 21 (discussing the effects of the Bayh-Doyle Act, which allows institutions to claim ownership of patents that are based on research conducted with federal funds).
132. See COUNCIL FOR ADULT & EXPERIENTIAL LEARNING, SERVING ADULT LEARNERS IN HIGHER EDUCATION: FINDINGS FROM CAEL’S BENCHMARKING STUDY 1 (1999), http://www.cael.org/pdf/publication_pdf/CAEL%20Benchmarking%20Findings%20Executive%20Summary.pdf (defining the adult learner as financially independent, with major responsibilities outside of school, and “whose principal identities have evolved beyond the role of full-time student”).
133. See Michael Kirst, Secondary and Postsecondary Linkages, in ECONOMIC INEQUALITY AND HIGHER EDUCATION: ACCESS, PERSISTENCE, AND SUCCESS 44, 44–46, 56 (Stacy Dickert-Conlin & Ross Rubenstein eds., 2007) (noting the substantial rise in college enrollment, especially in community colleges, and the fact that 80% of postsecondary students “attend postsecondary institutions that either accept all qualified applicants or are open enrollment”).
134. BRYAN J. COOK & DIANA L. CÓRDOVA, AM. COUNCIL ON EDUC., MINORITIES IN HIGHER EDUCATION: TWENTY-SECOND ANNUAL STATUS REPORT 3 (Supp. 2007),
population. Their primary motivation for undertaking higher education is job skills training and professional development. With these new consumers have come new needs and demands that vary from those of the "typical college student" of the past. And institutions seeking to exploit these new markets have been forced to adjust accordingly. Put simply, the academic capitalist system is prompting these institutions to respond to consumer demands.

F. Rising Tuition

State appropriations to higher education (as percentage of overall budgets) have declined steadily over the past three decades. As a result, institutions have had to generate more "market income" principally in the form of tuition. Between 1976 and 2005, the average cost of a public four-year institution increased 270%. Compounding the effects of rising tuition has been a decline in federal funding of need-based student aid, particularly Pell Grants. Twenty years ago, the maximum Pell Grant covered 60% of tuition at a typical public four-year institution; in 2006, that purchasing power had declined to 33%. The discretionary nature of

http://www.acenet.edu/AM/Template.cfm?Section=CAREE&Template=/CM/ContentDisplay.cfm&ContentID=23716.


136. See id. (discussing how economic volatility has fueled a "growing demand for continual learning and skill enhancement").

137. Id. (identifying the "typical college student" of the past as a "financially dependent, 18-year-old high school graduate who enrolls full time").

138. See, e.g., Peterson & Dil, supra note 108, at 18–19 (discussing how the rising popularity of non-degree and continuing education programs among adult learners has forced institutions to reassess their programmatic offerings).

139. See, e.g., SLAUGHTER & RHOADES, supra note 88, at 13 (noting that this trend occurred during the 1970s, 80s, and 90s).

140. ANCTIL, supra note 105, at 4 ("[M]arket income has increasingly substituted for public appropriations in higher education . . . .").

141. See, e.g., SLAUGHTER & RHOADES, supra note 88, at 12 (explaining that increasing tuition has led in turn to higher expectations from parents and students on the quality of education received).

142. See, e.g., Bridget Terry Long & Erin Riley, Financial Aid: A Broken Bridge to College Access?, 77 HARV. EDUC. REV. 39, 40 (2007) (noting that median family income only increased by 23% during the same period).

143. See, e.g., id. at 45 (noting that the maximum Pell Grant has decreased by 20% from 1975–1976 to 2003–2006 after accounting for inflation).

144. COLLEGE BOARD.COM, TOTAL PELL GRANT FUNDING DECLINES FOR FIRST TIME IN SIX YEARS (2006), http://www.collegeboard.com/prod_downloads/press/cost06/
higher education funding has made it an easy target for cuts, as entitled expenses such as health care increasingly strain state and federal budgets.\textsuperscript{145}

Students now finance a higher proportion of their tuition using loans. Between 1990 and 2004, the percentage of full-time students with loans rose from 36\% to 50\%.\textsuperscript{146} Between 1993 and 2004, cumulative debt levels for college students at public and private institutions rose 76\% and 57\% respectively.\textsuperscript{147} This shift has largely privatized the cost of higher education, thereby further entrenching the student-as-consumer mindset that has come to characterize academic capitalism.\textsuperscript{148}

\textbf{G. Demands for Accountability}

Against the backdrop of declining funding emerged the demand for educational accountability. Institutions are now being called upon to demonstrate their “value for money”—an expression used to denote an organization’s economy, efficiency, and effectiveness.\textsuperscript{149} In other words, institutions have had to provide evidence of successful outcomes, particularly as they relate to student learning.\textsuperscript{150} Institutional assurances are no longer sufficient, and assumptions are no longer freely granted. Fundamentally, policymakers want colleges and universities to behave more like private industry. In fact, the “resurgence of productivity and performance in American business” has been cited as an impetus behind the increased calls for educational accountability.\textsuperscript{151} Private sector watchwords like “performance,” “investment,” and “efficiency” have become part of the higher education lexicon.\textsuperscript{152} Accountability models first

\textsuperscript{145} See, e.g., Kane, supra note 104, at 65–66, 69–70 (illustrating an association between increases in state Medicare spending and increases in public four-year tuitions, and discussing the manner in which entitlement spending has affected federal spending on education, particularly grant programs such as Pell).

\textsuperscript{146} Long & Riley, supra note 142, at 47.

\textsuperscript{147} Id. at 47–48.

\textsuperscript{148} See, e.g., Slaughter & Rhoades, supra note 88, at 283 (noting that this trend has ignored the beneficial externalities of higher education).


\textsuperscript{150} Marilyn C. Kameen & Manuel J. Justiz, Using Assessment in Higher Education to Improve Success for Minority Students, 66 Peabody J. Educ. 46, 47 (1998) (asserting that state assessment of student outcomes “emerged as a remarkable new feature in American higher education” as far back as 1985).


\textsuperscript{152} See generally Joseph D. Creech, S. Reg’l Educ. Bd., Linking Higher Education
tested (and often discredited) in the private sector have found homes in higher education institutions. 153

Demands for accountability have manifested in various ways that reflect the capitalistic nature of education. Most prominently, states and accrediting agencies have required institutions to develop performance indicators and methods for assessing them. Many states have also tied institutional funding to performance. 154 While performance indicators take many forms, they are most often expressed numerically, 155 similar to private sector indicators. They can be internal in nature (e.g., graduation rates, research funds obtained, and teaching quality), external (e.g., employment rates of graduates), and operational (e.g., unit costs, class sizes, and course options). 156 The accountability movement has also contributed to a shift in how institutions present themselves to potential consumers. It is no coincidence that schools now tout employment rates and outcome-based indicators in advertisements to prospective students. The accountability movement has played a considerable role in fostering the spread of academic capitalism within higher education, particularly in terms of how institutional effectiveness is viewed internally and externally.

H. New Higher Education Providers

The advent of new providers exemplifies the interrelated nature of the trends contributing to the spread of academic capitalism. The knowledge-based economy has increased demand for higher education, which, along with neoliberal financial aid policies, has incentivized entry of new providers. Technological advancements and globalization have eased these

---

153. See Gary Rhoades & Barbara Sporn, Quality Assurance in Europe and the U.S.: Professional and Political Economic Framing of Higher Education Policy, 43 HIGHER EDUC. 355, 366 (2002) (explaining how accountability models “arrive at higher education’s doorstep” after initial trials in business that lead to these models being “discarded”) (quotation omitted).

154. See Dowd, supra note 151, at 109–10 (noting that by 2000, almost three-quarters of the states had performance funding systems in place and that the shift away from input-based funding to funding based on outcomes betokened a new emphasis on accountability).


156. Ball & Halwachi, supra note 149, at 401; see also Bob Barnetson & Marc Cutright, Performance Indicators as Conceptual Technologies, 40 HIGHER EDUC. 277, 278–79 (2000) (classifying indicators in terms of five organizational elements to which performance indicators are applied: (1) Inputs, e.g., faculty, facilities; (2) processes, e.g., teaching; (3) products, e.g., courses completed; (4) outputs, e.g., degrees awarded, grants secured; and (5) outcomes, e.g., employment rates).
providers’ entry into the market. Changing student populations and rising tuition among traditional institutions have fostered new markets and enhanced preexisting ones. Demands for accountability have introduced outcomes-based parlance and practices into the higher education industry—a shift upon which new providers have been able to capitalize.

Proprietary schools are the most salient new higher education providers. These providers have entered the market and, in many ways, adapted to new realities more effectively than their traditional peers. For starters, proprietary institutions have been successful at reframing themselves in response to the commoditization of education. In fact, it could be argued that they never needed to reframe in the first place. Their profit-generating motives already required them to respond to market demands in ways that nonprofit institutions did not. They have also adapted through an evolution of their own. The conventional mom-and-pop operations are becoming relics of the past as large, multi-campus corporations now dominate the industry.157 In response to market demands, many of these institutions have also transcended their vocational moorings and now award degrees up to the doctoral level.158 Proprietary schools are now among the largest and most successful education providers in the country,159 validating their market-driven approach and exemplifying their superior adaptive ability.

III. THE RISE OF PROPRIETARY COLLEGES

Much has been written about the recent “arrival” of proprietary schools into the higher education market; however, “reemergence” might be a better descriptor. The history of proprietary schools in the United States is surprisingly long—pre-dating the signing of the Declaration of Independence. They were fixtures during Colonial times as alternatives to apprenticeships and the colleges of the day.160 The purposes of these early institutions evolved from teaching basic literacy to career training.161 These

157. See Foster, supra note 23, at 8 (referring to these schools as “super systems”).
158. See, e.g., Anctil, supra note 105, at 22.
159. The student enrollments of the five largest proprietary schools are as follows: Apollo Group (Parent Company of UOP, 420,700; Education Management Corporation, 112,700; Kaplan Higher Education, 103,300; Career Education Corporation; 93,100; and DeVry, 90,365. Erica R. Hendry, For-Profit Colleges See Larg Increases in Enrollment and Revenue, CHRON. HIGHER EDUCC., Aug. 23, 2009, http://chronicle.com/article/For-Profit-Colleges-See-Lar/40173/.
161. Ruch, supra note 110, at 52 (chronicling how student interest prompted early proprietary schools to expand their curricula to include courses that taught “skills that were in high demand by employers”).
institutions embodied the entrepreneurial spirit that would come to symbolize the founding of the United States and the spread of higher education. These institutions also embraced the ideal of educational access, an ethos that would hasten their reemergence in the early 1970s.

A. Title IV Expansion

The 1970s brought vast expansion of the proprietary school market. The 1972 Amendments to the Higher Education Act sought to broaden higher-education access by making proprietary institutions eligible to collect Title IV aid. Anxious to tap into Title IV, proprietary schools aggressively recruited students by touting programs that purported to provide job training; however, much of the promised training never materialized. The lack of effective oversight provided an environment in which sham schools and diploma mills operated with virtual impunity. As a result, student defaults on Title IV loans increased sharply. The 1990s brought closer scrutiny on the proprietary school sector, and between 1992 and 1997, almost 800 schools were shut down or stripped of their Title IV eligibility—which effectively shut down schools not closed outright. But while the amount of documented improprieties among proprietary schools has fallen, the sector’s disproportionate share of federal investigations shows that improprieties remain a problem.

B. Students

Proprietary schools serve students currently underserved by traditional institutions. They tend to enroll the “other 75 percent”—students “who

162. Id.
163. Id. at 57 (discussing how proprietary schools were among the first institutions to educate former slaves and Native Americans).
165. See e.g., GAO, OVERSUPPLIED OCCUPATIONS, supra note 22, at 7.
166. See, e.g., Foster, supra note 23, at 14.
167. Id. (“During the 1970s and 80s, institutions operated with little or no oversight and few constraints in recruiting and training students. A large number of institutions did not provide the training advertised, did not comply with fair consumer practices, and mismanaged finances.”).
168. See, e.g., id.
169. Linehan, supra note 7, at 760.
170. See, e.g., Anti-Fraud Hearings, supra note 10, at 2 (statement of Rep. John Boehner, Chairman, H. Comm. on Educ. and the Workforce) (“Proprietary schools . . . are playing a critical role in providing college access for some of our Nation’s most vulnerable students.
were not in the top 25 percent of their high school classes and who would be unlikely to enroll or be successful at other types of institutions.” Proprietary school students tend to be poorer and older than students at traditional schools. They are also more likely to be first in their families to go to college, be female, and to belong to a racial or ethnic minority group. Many proprietary schools award a disproportionately high percentage of degrees through the doctoral level to black and Hispanic students, and their graduates typically complete degree requirements faster than graduates of traditional schools. These statistics are laudable, and they confirm that proprietary schools help broaden access to higher education. However, this access is motivated more by profit than altruism. Moreover, this access comes at a high cost, due to high rates of attrition at these institutions and concomitant high rates of loan defaults among former students.

C. Outcomes

Proprietary schools suffer from poor student outcomes. Lower completion rates tend to create negative cascades that depress placement...
rates and increase loan default rates. 180 Less than 25% of proprietary school students graduate within six years of beginning their studies, compared to 55% and 64% at public and private nonprofit institutions respectively. 181 Regrettably, this statistic is probably inflated, given that it is based on self-reported data, and proprietary schools have many incentives to overstate graduation rates. 182 Placement rates also demonstrate inadequacies within the sector. Schools with the lowest completion rates tend to have the lowest placement rates. 183 This observation is particularly damaging because placement rates do not account for attrition; only students who complete the program are included in the calculations. The final links in the chain are student loan default rates. Proprietary school default rates exceed nonprofit schools at the two-, three-, and four-year intervals. 184 And once again, the validity of these numbers is questionable, as the method of calculation allows schools to understate their actual number of defaults. 185

There are many reasons for the lower completion rates at proprietary schools—and not all of them are nefarious. The primary reason concerns the negative association between reliance on Title IV aid and completion rates. 186 This association arises because poorer students persist towards college degrees at lower rates than wealthier students. 187 This phenomenon holds true across higher education, irrespective of sector, and is often used as a powerful justification for those seeking increased or better-targeted student aid. 188 And because proprietary schools enroll higher percentages of poor students, it makes sense that their completion and placement rates

---

180. GAO, POORER OUTCOMES, supra note 164, at 5.
181. KNAPP ET AL., supra note 6, at 12.
182. LOOON & DEVANTHERY, supra note 29, at 38 (“The reliability of the numbers in IPEDS is based solely on the reporting done by the institutions themselves. This is extremely problematic as it leaves nearly absolute discretion in the hands of schools that have every incentive to inflate the numbers.”).
183. See GAO, POORER OUTCOMES, supra note 164, at 9.
184. See GAO, STRONGER OVERSIGHT, supra note 13, at 14–15 (listing proprietary school default rates that are as much as 250% higher than the next highest rate—that of public schools).
185. See id. at 13–14 (“[T]he rate captures only a small portion of all student loan defaults at schools.”).
186. See GAO, POORER OUTCOMES, supra note 164, at 24–25 (demonstrating that increased reliance by schools on Title IV revenue leads to lower completion and graduation rates, and higher default rates).
187. See Long & Riley, supra note 142, at 40 (explaining that only 43% of students from families making less than $30,000 per year “immediately entered a post-secondary institution” compared to 75% of students whose families make more than $50,000).
188. See, e.g., id. at 38 (concluding that the Pell Grant will provide more access to higher education for low-income students).
are lower and their default rates higher. 189

However, not all factors contributing to these lower rates are benign. Proprietary schools enroll many students who clearly lack the ability to complete a postsecondary program of study. 190 In other words, they exploit their role as access providers—for profit. As discussed earlier, former admissions representatives allege that they were pressured to induce enrollment among unqualified students. And as 60 Minutes documented, proprietary schools engage in entrance exam improprieties in order to ensure that all prospective students attain satisfactory scores. Actions like these have created a perception that proprietary schools care less about their students and more about their students’ Title IV eligibility. 191

D. Industry

Today, the proprietary sector is dominated by five publicly-traded entities: Apollo Group [parent company of UOP], Education Management Corporation, Kaplan Higher Education, Career Education Corporation, and DeVry. Combined, these institutions enrolled more than 820,000 students in 2009. 192 All told, there are 2,900 Title IV-eligible proprietary schools providing both degree programs and vocational training. 193 The predominant niche of proprietary education remains career-focused education. 194 They have harnessed new technologies in delivering

189. See GAO, POORER OUTCOMES, supra note 164, at 20 (“We believe knowing a school’s completion rate helps predict its placement rate and knowing both completion and placement rates helps predict its default rate.”).

190. See KIRP, supra note 87, at 250 (“We accept students who, on paper, aren’t likely to make it . . . .”) (quotations omitted).

191. See Anti-Fraud Hearings, supra note 10, at 22 (statement of Rep. Maxine Waters) (“[T]he real motive behind wanting to enroll more minority and low income students is that they are the most profitable students since they qualify for the highest amounts of federal financial aid and the smallest expected family contribution, or none at all.”).

192. See Hendry, supra note 159.


194. See Foster, supra note 23, at 8 (“The primary purpose of for-profit postsecondary institutions is preparing graduates for jobs or career advancement. As a result, these institutions generally offer a small, focused range of programs limited to high-demand occupational or professional fields.”). But see generally GAO, OVERSUPPLIED OCCUPATIONS, supra note 22, at 4–5 (discussing the amount of federal financial aid used by students training for low-demand fields).

195. See KIRP, supra note 87, at 242 (listing the MBA program as an example of a predominant niche of proprietary education).
education programming, both online and in the classroom.196 Their programs are flexible and accelerated, thus appealing to older students who tend to be place-bound and limited in how much time they can spend attending classes.197 Given their adaptive skills, it should be no surprise that while traditional schools are dealing with enrollment and budgetary shortfalls, proprietary schools are experiencing vastly increased enrollment and revenue.198

E. Profitability

Many proprietary schools are highly profitable. In 2006, Apollo Group and ITT boasted returns on investment capital of 69% and 40% respectively, beating out companies such as Exxon Mobil and Microsoft.199 Since 1995, Apollo’s stock price has risen an unfathomable 7,000%, and ITT’s stock price has risen more than 3,500%.200 The stock prices for other publicly traded education providers have experienced precipitous

196. See Anti-Fraud Hearings, supra note 10, at 46 (statement of Nick Glakas, President, Career College Association) (“For-profit institutions are pioneering a wide array of innovative program delivery methodologies such as on-line, modular, and weekend programs to complement their traditional classroom offerings.”).

197. See id. (“Students choose to attend for-profit colleges because these delivery methods meet their time and geographical needs, allowing them to achieve their postsecondary education goals while continuing to meet the demands of their every day lives. On average, students attending career colleges earn their associates degree eleven months sooner than students at community colleges.”).

198. See Hendry, supra note 159 (“The recession has left nonprofit colleges and universities across the country struggling with budget cuts and uncertainties over enrollment, but many for-profit institutions are reporting record increases in student numbers and revenue—a sign that the recession is prompting more adults and nontraditional students to seek career training.”).


increases as well. A major component of proprietary schools’ profitability is operational efficiency.

F. Operational Efficiencies

Like any profit-seeking entity, proprietary schools “place a high value on running their operations efficiently and taking advantage of economies of scale.” They minimize inefficiencies in academic planning and teaching by designing curricula centrally and relying principally on untenured faculty to render instruction. Some proprietary schools even base instructor pay on the number of students in the instructor’s class. They use technology to minimize inefficiencies in administrative operations, such as in admissions. Some proprietary schools promote efficiency by calculating optimal facilities usage ratios. These calculations help institutions determine the size of their facilities and the types of leasing


202. Ruch, supra note 110, at 76 (citing “Scale Economies and Operating Efficiencies” as one of seven “Ingredients for Profitability” for proprietary schools).

203. Id. at 88.

204. See id. at 118–19 (stating that although the faculty “are the center of academic life,” the president is expected to maintain managerial control over decision making).

205. See id. at 119 (noting the apparent “lack of a tenure system at for-profit universities”).


208. Facility costs are the second largest expense incurred by proprietary schools, behind salaries. Blumenstyk, supra note 206.
agreements they enter.\textsuperscript{209} Enrollment management, however, is the primary method that proprietary schools use to promote efficiencies.\textsuperscript{210} Like all institutions, proprietary schools need students, but their need is more intense than that of most nonprofit schools.\textsuperscript{211} Most nonprofits are able to subsidize their expenses with non-tuition revenue, such as endowment income or public appropriations. Proprietary schools, however, do not benefit from such subsidies. As a result, each student represents a revenue stream that directly affects the company’s bottom line.\textsuperscript{212} As such, these institutions are under intense and constant pressure to increase enrollments. And given this pressure, it should be no surprise that improprieties in the sector almost always bear some relation to institutions’ enrollment management functions.

\textbf{G. Emphasis on Recruitment}

Admissions representatives at proprietary schools are essentially salespeople.\textsuperscript{213} At many proprietary schools, the number of admissions representatives is greater than full-time faculty.\textsuperscript{214} Proprietary schools also invest heavily in advertising;\textsuperscript{215} anyone who has ever watched daytime or late-night TV can attest to this. In fact, recruitment and marketing expenses typically dwarf the total salaries paid to faculty.\textsuperscript{216} Proprietary school advertisements often portray education as the path to a career and to financial security—and the advertising institution as the ideal provider of that education. Unfortunately, many of the electronic and human representatives of proprietary schools proffer misrepresentations as a means

\textsuperscript{209} This approach to managing space differs somewhat from that of most nonprofit schools, as these schools are often incentivized to own their facilities and expand such holdings. \textit{Id.}

\textsuperscript{210} See \textit{RUCH}, supra note 110, at 88 (discussing how proprietary schools track class enrollments closely and make adjustments accordingly).

\textsuperscript{211} See, \textit{e.g.}, \textit{KIRP}, supra note 87, at 242 (“Increasing enrollment has to be the paramount concern for any for-profit university, especially one whose stock is publicly traded.”).

\textsuperscript{212} See \textit{id.} at 242, 247 (claiming that for-profit universities, such as DeVry, have higher tuition and must use aggressive marketing and recruiting to bring in prospective students).

\textsuperscript{213} See \textit{id.} at 247 (explaining that for schools with multiple campuses, these individuals are often responsible for recruiting on behalf of all schools within a particular region).

\textsuperscript{214} \textit{Anti-Fraud Hearings}, supra note 10, at 8 (statement of Rep. Maxine Waters).

\textsuperscript{215} \textit{Id.}

\textsuperscript{216} \textit{Id.} (“The amount spent on advertising, lead creation, recruiting, and admissions representatives far exceeds the salaries paid to faculty.”).
of inducing enrollment.\textsuperscript{217} The primary targets of these misrepresentations are people who are most susceptible to being fooled by them and most likely to reap the negative effects of an unsuccessful educational experience. Therefore, a clear imperative to prevent schools from inducing enrollment using misrepresentations exists.

\section*{IV. The Need for Tighter Regulation of Higher Education Recruitment and Marketing}

The typical proprietary school student is undertaking education for very pragmatic reasons—most often, earning a degree that will soon result in a well-paying job.\textsuperscript{218} Cognizant of their niche, proprietary schools have done a convincing job of characterizing themselves as effective, if not obligatory, intermediaries between job seekers and the job market.\textsuperscript{219} And both through their advertisements and their recruitment practices, proprietary schools use the single-minded determination common among their students to their advantage. Take for example the following jingle, which is sung in very catchy fashion on a television commercial for a company that markets online programs for proprietary schools.\textsuperscript{220} The lyrics are written from the perspective of a prospective student:

\begin{quote}
I'm working for an hourly wage  
I went to high school—didn't do great  
Still I gotta make more cash—more education is what I'm looking at  
When I get a degree, I will make a bigger salary  
So now I've got to see—which college is right for me  
I went on the internet and found Education Connection  
I took some free tests to find out my direction  
I'm taking my classes online—getting my degree on my own time  
Education Connection matched me with the right college for free!\textsuperscript{221}
\end{quote}

The lyrics touch on all the common themes utilized by proprietary schools in pursuing their market, including monetary benefit and convenience. Moreover, the commercial is replete with graphical statements, such as “Make $25,000 More Each Year.”\textsuperscript{222}

\begin{flushright}
\textsuperscript{217} See \textit{Kirp}, \textit{ supra} note 87, at 242 (admitting that because recruiters are expected to meet enrollment quotas, that they sometimes oversell the school and “skirt[] the border of misrepresentation”).
\end{flushright}

\begin{flushright}
\textsuperscript{218} See, e.g., \textit{id.} at 245.
\end{flushright}

\begin{flushright}
\textsuperscript{219} See \textit{Linehan}, \textit{ supra} note 7, at 757–58 (“Students generally view proprietary schools as the gatekeepers to their trade of choice . . . .”).
\end{flushright}

\begin{flushright}
\end{flushright}

\begin{flushright}
\textsuperscript{221} \textit{id.}
\end{flushright}

\begin{flushright}
\textsuperscript{222} \textit{id.}
\end{flushright}
A. The Susceptibilities of the Market

Given the educational niche proprietary schools have carved out, the most common targets of proprietary school advertisements are poor, undere Educated, and older.\footnote{See, e.g., Linehan, supra note 7, at 757 (explaining the types of advertising that proprietary schools use when focusing on certain demographics).} Individuals who fit this profile are highly susceptible to being persuaded by misrepresentations due to their lack of insight about higher education.\footnote{See, e.g., id. at 757–58.} Unlike typical students at traditional colleges and universities, most proprietary school students are first-generation college students.\footnote{Anti-Fraud Hearings, supra note 10, at 46 (statement of Nick Glakas, President, Career College Association).} This lack of educational experience limits their ability to discern honest claims from deceptive ones. They are more likely to finance their education with student loans and eventually default on those loans.\footnote{See, e.g., GAO, STRONGER OVERSIGHT, supra note 13, at 28.} They are least likely to complain about unfair conduct to which they have been subjected,\footnote{See, e.g., Foster, supra note 23, at 24.} and even for those who do complain, current safeguards are inadequate in providing effective remedies.\footnote{See Linehan, supra note 7, at 754 (alleging that the current legal doctrine fails to protect students from the schools’ predatory practices).} Assurances that an educational program would quickly lead to a well-paying job are very compelling. Add a lack of higher education exposure and a burning desire to escape poverty to the equation, and it becomes clear why misrepresentations concerning graduation and placement rates are so dangerous.

B. The Costs of Failure

Proprietary schools are relatively expensive to attend. When compared to public and private institutions, proprietary schools charge by far the highest average tuition for non-degree and two-year degree programs.\footnote{Knapp et al., supra note 193, at 6.} Additionally, the average tuition for bachelor’s degree programs at proprietary schools is higher than the average public school tuition, though slightly lower than private school tuition.\footnote{Id.} These high tuition rates have encouraged high levels of borrowing among proprietary school students. Seventy-two percent of proprietary school students finance their education (at least in part) with Stafford loans; this is the highest percentage in higher education.\footnote{See Christina Chang Wei et al., NAT’L. CTR. FOR EDUC. STATISTICS, U.S. DEP’T
school students; once again, the highest in higher education.\textsuperscript{232} It is no wonder that while proprietary school undergraduates only account for \(8\%\) of students in higher education, they account for \(18\%\) of the loan volume.\textsuperscript{233}

The downside of these high borrowing rates is manifested mainly in the high default rates among former proprietary school students. Proprietary schools account for a disproportionate share of student loan defaults.\textsuperscript{234} In 2006, the sector's two-year cohort default rate was \(8.6\%\), the highest in higher education.\textsuperscript{235} Default rates among all borrowers increase over time, but the increase is much higher among proprietary school borrowers. Almost a quarter of proprietary school borrowers default on student loans within four years of entering repayment, greatly exceeding the public and private school sectors.\textsuperscript{236} The effects of student loan defaults are immense. Individuals who default acquire negative credit history that limits their ability to secure housing or other loans.\textsuperscript{237} They could also face income garnishments and restricted employment options.\textsuperscript{238} And to the disappointment of many, it is very difficult to discharge federal student loans in bankruptcy.\textsuperscript{239} For taxpayers, the costs of covering defaults are immense as well. Taxpayers cover virtually all the expenses associated with defaulted loans, including interest, and the price tag is in the billions of dollars.\textsuperscript{240}

\begin{footnotesize}
\begin{itemize}
  \item[232.] \textit{Id.} at 20 (listing average loan amounts for public two-year, public four-year, and private institutions in 2003–2004 as \(3,400\), \(4,900\), and \(5,100\) respectively).
  \item[233.] \textit{Id.} at 21.
  \item[234.] \textit{Id.} at 19.
  \item[235.] \textit{Id.} at 23.
  \item[236.] 
  \item[237.] See \textit{Id.} at 15 (providing a four-year default rate of \(23.3\%\) for proprietary school borrowers and \(9.5\%\) and \(6.5\%\) for public and private school borrowers, respectively).
  \item[238.] \textit{E.g.} \textit{Id.} at 12.
  \item[239.] \textit{Id.}
  \item[239.] See, \textit{e.g.}, U.S. Dep't of Educ., Federal Student Aid, Common Disputes Involving Defaulted Student Loans, \url{http://www.ed.gov/offices/OSFAP/DCS/disputes.html#Bankruptcy} (last visited Aug. 4, 2010) ("Whether a bankruptcy discharge relieves an individual of his or her obligation to repay a student loan or grant overpayment is now determined by whether a court has ruled that repayment would impose an undue hardship on the borrower and his or her dependents.").
\end{itemize}
\end{footnotesize}
Given the disadvantaged backgrounds from which proprietary school students often come, these students lack the social clout and political sophistication necessary to foster widespread dismay regarding their victimization. They are less likely to even complain about fraud perpetrated against them by proprietary schools.\textsuperscript{241} Even for those who complain, current legal and regulatory processes provide few options for redress.\textsuperscript{242}

V. INADEQUATE SAFEGUARDS

The prominence of higher education institutions makes them frequent targets of lawsuits. The diverse nature of this litigation represents a virtual microcosm of American jurisprudence. But in adjudicating disputes involving higher education institutions, courts have been rather consistent about their reluctance to intrude upon the inner workings of these institutions. This reluctance has been termed academic abstention.\textsuperscript{243} The concept has been applied in cases involving all types of educational institutions, and its fundamental premise is the judiciary’s belief that the professional judgment of educators should be protected from the unqualified assessments of judges or other fact finders. \textit{Paladino v. Adelphi University}\textsuperscript{244} illustrates this reasoning: The plaintiff alleged that the defendant institution failed to adequately educate his child, but the court dismissed the case in large part because adjudicating it would have required the “fact finder to enter the classroom and determine whether or not the judgments and conduct of professional educators were deficient.”\textsuperscript{245} The court was loath to evaluate the “complex educational determinations” made by the defendant.\textsuperscript{246} This type of judicial reluctance can greatly disadvantage plaintiffs by limiting the circumstances in which they can win, or even seek, recovery for damages.

Plaintiffs use tort law and contract law frequently as bases upon which to sue educational institutions, but both provide only narrow paths to recovery in cases where misrepresentation or fraud is alleged.

\textsuperscript{241} See Foster, \textit{supra} note 23, at 24 (“[L]egal aid attorneys believe that most students who have been misled by institutions do not complain; and, as a result, the number of complaints is not an adequate indication of the level of fraud and abuse perpetrated by some for-profit institutions.”).

\textsuperscript{242} Linehan, \textit{supra} note 7, at 764.

\textsuperscript{243} See id. (“[T]he doctrine of academic abstention reflects courts’ . . . reluctance to delve into the operation of educational institutions . . . .”).

\textsuperscript{244} 454 N.Y.S.2d 869 (App. Div. 1982).

\textsuperscript{245} \textit{Id.} at 873.

\textsuperscript{246} \textit{Id.} at 872.
A. Tort Law

In tort law, the most logical cause of action for victims of proprietary school fraud seems to be fraudulent misrepresentation. Generally, a target of a fraudulent misrepresentation may recover damages if the maker of the misrepresentation knew or should have known that it was false or baseless (scienter)\(^\text{247}\) and the target justifiably relied on the misrepresentation to his detriment.\(^\text{248}\) In order for the target’s reliance to be justified, the misrepresentation must be material—or in other words, the person making the misrepresentation must know or should know that the target will “attach importance to [the misrepresentation] in determining his choice of action in the transaction in question . . .”\(^\text{249}\) Additionally, liability can attach when the maker of a misrepresentation knows or should know that the target will rely on the misrepresentation, even if a reasonable person would not.\(^\text{250}\)

A typical scenario during which an admissions representative induces a student to enroll based on unjustifiably rosy future job prospects seems to comprise a textbook case of fraudulent misrepresentation. The representative knowingly makes a representation that is baseless, if not fraudulent, in order to induce enrollment, while the representative knows that the target will justifiably attach importance to the virtual promise of a well-paying job. However, courts are reluctant to award damages to plaintiffs in fraudulent misrepresentation cases against educational institutions; typically, only the most barefaced instances of fraud are successful.\(^\text{251}\) The primary difficulty plaintiffs face in these cases is proving scienter on the part of the defendant.\(^\text{252}\) In representing future job prospects, proprietary schools are able to hide behind the fact that much of what determines a graduate’s job prospects is outside of the school’s control.\(^\text{253}\) At least one court has characterized such representations as “no more than a prophecy,” in highlighting the limited power schools have in securing employment for their graduates.\(^\text{254}\) But this view allows schools to make baseless forward-looking claims with impunity by shielding them on the back end, without restricting their representations on the front end.\(^\text{255}\)

\(^\text{247}^\) Restatement (Second) of Torts § 526 (1977).

\(^\text{248}^\) Id. § 537.

\(^\text{249}^\) Id. § 538(2)(a).

\(^\text{250}^\) Id. § 538(2)(b).

\(^\text{251}^\) See, e.g., Linehan, supra note 7, at 770 (discussing the heightened pleading requirements for complaints alleging intentional fraud).

\(^\text{252}^\) Id.

\(^\text{253}^\) Id.


\(^\text{255}^\) See Linehan, supra note 7, at 768 (“By promising outcomes which in some way
TIGHTER REGULATION OF SCHOOL REPRESENTATIONS

It seems that academic abstention has fostered reluctance on the part of judges to critique what amounts to sales pitches, due to the tangential relationship of these pitches to the educational process. Therefore, fraudulent misrepresentation fails to provide a viable avenue of redress for most victims of proprietary school misrepresentations.

In addition, plaintiffs alleging negligence have been largely unsuccessful. For this discussion, the two most relevant negligence claims are negligent misrepresentation and educational malpractice. In order for a negligent misrepresentation claim to be successful, the plaintiff must show that the defendant, while acting in a business or professional capacity, supplied false information that was negligently obtained or communicated, upon which the plaintiff justifiably relied to his detriment. Educational malpractice is premised on the claim that the institution failed to provide the plaintiff an adequate education, thereby causing harm, such as failure to prepare the plaintiff to employment.

Both claims tend to fail because courts are reluctant to impose a duty of care upon educational institutions for their student outcomes. For instance, in Tolman v. CenCor Career Colleges, Inc., a group of former students of the defendant institution asserted various negligence claims relating to the quality of the education they received and the advertisements used by the defendant. In dismissing all of the negligence claims, the court cited the “collaborative and subjective process” through which education is undertaken and the “outside factors” that determine a student’s level of

---

256. Generally, negligent conduct can be found where (a) an act which the actor as a reasonable man should recognize as involving an unreasonable risk of causing an invasion of an interest of another, or (b) a failure to do an act which is necessary for the protection or assistance of another and which the actor is under a duty to do.

RESTATEMENT (SECOND) OF TORTS § 284 (1965). Negligence lawsuits against proprietary schools usually allege that the school failed to provide training or education it had a duty to provide.

257. See, e.g., Amaral v. Am. Sch. of Correspondence, 107 F. App’x. 497, 498–99 (6th Cir. 2004).


260. Specifically, the claims alleged negligence in informing plaintiffs of the type of education they would receive, negligence of a specialist in technical education, and negligence based on defendants’ failure to avoid false or misleading advertising. Id. at 204–05.
success. The court concluded that “there is no workable standard of care” that could be imposed upon schools. And of course, without a heightened standard of care or duty, no negligence claim can stand. Moreover, courts have almost universally rejected education malpractice as a recognized cause of action. They have cited various public policy considerations as reasons for this broad rejection. But the failure of both educational malpractice and negligent misrepresentation can be traced back to the judicial reluctance that characterizes the academic abstention doctrine.

B. Contract Law

Breach of contract suits stand a better chance of success than those asserting negligence; this is because it is generally settled that the relationship between a student and his or her educational institution is contractual in nature. Promises made by an institution or its representatives are binding. Catalogs and other materials made available to the student by his institution help define the contours of the contractual relationship. If certain promises are not kept by the institution, the student could have a claim of breach of contract. In Ross v. Creighton University, a former student-athlete brought a lawsuit claiming that the defendant institution failed to provide promised tutoring services and other academic accommodations. The court allowed the plaintiff’s contract

261. Id. at 205.
262. Id.
263. Ross, 937 F.2d at 412, 414, 416 (characterizing educational malpractice claims as “beloved of commentators, but not of courts,” and identifying Montana as the only state that allows educational malpractice claims to go forward).
264. See, e.g., id. at 414 (discussing reasons courts have rejected educational malpractice as a cause of action, including the lack of a standard of care, uncertainty about the cause and nature of damages, the potential for mass litigation, and the possibility that courts would be forced to oversee the operations of schools).
265. Linehan, supra note 7, at 771.
266. Wickstrom v. N. Idaho Coll., 725 P.2d 155, 157 (Idaho 1986) (“There seems to be almost no dissent from the proposition that the relationship [between institution and student] is contractual in nature.”).
268. Ross, 957 F.2d at 416.
269. Id. at 417.
270. Ross was a basketball player who entered Creighton with known academic deficiencies. In his complaint he averred that Creighton breached its promise to provide Ross “an opportunity to participate, in a meaningful way, in the academic program of the University despite his academic background” in return for his promise to play basketball. The breach arose from Creighton’s alleged failure to provide Ross with tutoring services, an opportunity and directive to take advantage of those services, an athletic redshirt that would
claims to proceed because they could “point to an identifiable contractual promise that the defendant failed to honor.”271 The Ross court’s emphasis on specific promises is central to how the propriety of these types of cases is assessed. The court articulated a standard for whether a contract claim of this sort can stand: Whether ruling on the issue would require the court to delve into the nuances of “educational processes and theories” or allow it to make “an objective assessment” of whether the institution failed to make good on promises.272 If the allegation requires the former, it cannot stand; if it allows the latter, it can be pursued on the merits.273 The former approach would, of course, run afoul of the academic abstention doctrine, while the latter would not. Unfortunately, the practical effect of this approach is similar to the effect on tort claims; institutions are allowed to make misrepresentations that are clear in their implications, but vague enough to evade legal obligation. As a result, when a student suffers damages arising from these legally vague but practically convincing misrepresentations, options for redress are limited.

C. Consumer Protection

State consumer protection statutes theoretically provide avenues for redress, as practically every state allows victims of fraud to sue for damages.274 Some states even have consumer protection statutes specifically addressing the operation of proprietary schools.275 However, the standards of proof required by these statutes often make winning damages difficult for victims.276 Some of these statutes require victims to prove scienter and proximate cause,277 creating the same difficulties

allow him to better focus on academics, and funds to attain a college degree. Id. at 416.

271. Id. at 417.
272. Id.
273. See id. (dismissing the plaintiff’s negligence claims while preserving his contract claims).
274. Sheila B. Scheuerman, The Consumer Fraud Class Action: Rein in Abuse by Requiring Plaintiffs to Allege Reliance as an Essential Element, 43 Harv. J. on Legis. 1, 23 (2006); see also Jon Mize, Comment, Fencing Off the Path of Least Resistance: Re-Examining the Role of Little FTC Act Actions in the Law of False Advertising, 72 Tenn. L. Rev. 653, 660 (2005) (listing “treble damages, punitive damages, statutory minimum damages, and attorney’s fees” as the most common damages allowed by state consumer fraud statutes).
276. Linehan, supra note 7, at 776. But see Mize, supra note 274, at 661 (characterizing state consumer fraud statutes as “the path of least resistance” in suits alleging false advertising).
277. See, e.g., Rizzo v. Michener, 584 A.2d 973, 980 (Pa. Super. Ct. 1990) (“Actual fraud has five elements which must coalesce. There must be (1) misrepresentation of a material
described earlier. Also, because some states limit attorney’s fees, it may be difficult for some victims to find lawyers willing to litigate cases in which only a few thousand dollars are at issue.

D. The “Triad”

Lastly, victims of proprietary school fraud are inadequately protected by the Title IV oversight mechanism—also known as the “triad.” The triad consists of the DOE, state regulatory bodies, and accrediting agencies. Its purpose is to ensure “that the ‘gate’ to student financial aid programs open only to those institutions that provide students with quality education or training worth the time, energy, and money they invest.” DOE’s primary functions within the triad are to verify institutional eligibility for Title IV funds and to certify accrediting agencies. States provide oversight in many ways, including through higher education regulatory agencies as well as through indirect means such as consumer protection and commerce laws. Accrediting agencies certify institutions as having met certain minimum standards of quality. Only institutions that are accredited by an agency certified by the DOE can receive Title IV aid.

None of the components of the triad provides much relief for victims of proprietary school misrepresentations. The DOE only provides a limited mechanism for victims to lodge complaints, and that mechanism does not

---

278 See, e.g., FLA. STAT. § 501.2105(1) (1998) (placing the awarding of attorney’s fees within the discretion of the trial judge).
279 GAO, ENSURING QUALITY, supra note 27, at 4.
280 An example of this function is DOE’s tracking of cohort default rates. Id. at 4.
281 See id. at 4–5 (noting that DOE “certifies that such agencies are reliable authorities as to what constitutes quality education or training provided by postsecondary institutions”).
282 See id. at 5 (stating that these agencies are often responsible for establishing standards for regulation of higher education institutions).
283 See id. ("Other state agencies define certain consumer protection measures, such as refund policies. In the normal course of regulating commerce, all states require postsecondary institutions to have a license to operate within their borders.").
284 See id. at 5–6 ("Accrediting agencies adopt criteria they consider to reflect the qualities of a sound educational program and develop procedures for evaluating institutions to determine whether they operate at basic levels of quality.").
285 See id. at 4–6 (describing the accreditation process).
286 See Anti-Fraud Hearings, supra note 10, at 23–24 (statement of Rep. Maxine Waters) ("[T]he Department does not investigate charges made by students regarding misrepresentations made to influence students to enroll . . . [;] these federal regulations have no private right of action, and can only be enforced by the Department, which does not do its job.").
include a private right to damages. Moreover, fines and other sanctions imposed upon schools by the DOE are often inadequate disincentives to unscrupulous behavior. The shortcomings of state oversight have already been discussed. And accrediting agencies often provide insufficient and conflicted oversight of the institutions they certify. Such lax oversight allows unscrupulous institutions to stay in operation and continue to victimize students.

VI. REGULATING HIGHER EDUCATION REPRESENTATIONS

The proposals presented in this section have a singular focus: to reduce, if not prevent, incidences of misrepresentations made by proprietary schools in order to induce enrollment. The proposals are not focused on deterrence per se, as penalties against offending schools are not presented. Similarly, avenues of redress for victims of misrepresentations are not directly proposed. Pragmatism is the motivation behind this narrow focus; the goal of this Part is to present solutions that account for the multifaceted nature of higher education oversight and the sensitive nature of commercial speech regulation without getting bogged down in their complexity. As such, the principle thrust of the proposals is to harness current regulatory frameworks in new ways.

The proposals are organized around two areas of focus: (1) proprietary school marketing, and (2) recruitment. While there is overlap between the two areas, there are certain distinctive hallmarks of each. For purposes of this Article, marketing pertains to the efforts of proprietary schools to promote their programs to prospective students via wide-reaching means. Advertising, whether on television, online, or in print, is the principle method of proprietary school marketing. Recruitment pertains to the representations made and methods used to enroll individual students. The tactics of admissions representatives are central to this area of focus.

287. See Linehan, supra note 7, at 788 (noting that “nothing is done to compensate the victim”).
288. See, e.g., Anti-Fraud Hearings, supra note 10, at 21 (statement of Rep. Maxine Waters) (“[T]he school doing the defrauding may be allowed to pay a few cents on the dollar to settle claims with the Department, or placed on reimbursement status so that they have to wait 45 days for payment of financial aid.”).
289. See id. at 17 (“[T]here is a built-in conflict of interest with respect to accrediting agencies, because they have no incentive to revoke accreditation since their income-stream is directly determined by the number of schools they accredit.”).
A. Proposals

To protect students and taxpayers from misrepresentations and impropriety from proprietary schools, tighter regulation of their marketing practices is necessary. First, proprietary schools should be required to place disclaimers on all advertisements making forward-looking claims. Second, the Federal Trade Commission should expand its regulations pertaining to proprietary school advertising. Third, the FTC and the proprietary school industry should encourage self-regulation of proprietary school advertising practice. In addition to tighter regulation in the marketing area, existing regulation of proprietary schools’ recruitment practices should be expanded. Specifically, proprietary schools should be required to make affirmative disclosures and provide relevant labor market information to students prior to enrollment.

B. Relevant Oversight Agencies

Five types of entities play significant roles in regulating higher education and commercial speech. These entities are the DOE, the Federal Trade Commission (FTC), state regulatory agencies, accrediting agencies, and self-regulatory bodies. Due to their integral role in the proposals, a brief overview of each entity’s oversight function is necessary.

1. The Department of Education

The primary role of the DOE within higher education is to certify institutional eligibility for Title IV financial aid funds. In addition to collecting various forms of data and providing oversight of accrediting agencies, the Department oversees entrance exam publishers290 and dictates education policy. The DOE also accepts complaints filed by persons “suspecting fraud, waste[,] or abuse involving [DOE] funds or programs.”291

290. See, e.g., GAO, STRONGER OVERSIGHT, supra note 13, at 9–11 (stating that the DOE is “responsible for overseeing test publishers” and setting standards).
2. The Federal Trade Commission

The FTC is responsible for regulating and enforcing advertising laws.292 The Commission is principally concerned with promoting fair competition through truth in advertising.293 The FTC promulgates both general and industry-specific advertising standards, and it also assists industries in developing their own standards and best practices.294 The FTC has devised guides that explain how its rules are applied to specific industries,295 including one that addresses advertising practices of private vocational and distance education schools.296 In pursuing sanctions against offending advertisers, the FTC can bring lawsuits and administrative actions.297 Through these actions, the FTC can seek various forms of relief, such as injunctions, corrective advertising, monetary penalties, and consumer redress.298

3. State Regulatory Agencies

Every state has a higher education regulatory body that oversees the operations of postsecondary institutions within its borders. Some states have agencies that specifically oversee proprietary institutions.299 The

292. See Patricia P. Bailey, Unfair Competition and Misleading Advertising: How Advertising is Regulated in the United States, 54 ANTITRUST L.J. 531, 532 (1985) (“Congress authorized the agency . . . to challenge ‘unfair or deceptive acts or practices’ to promote truth in advertising and fair merchandising practices. The Commission’s goal is . . . to ensure that consumers receive both the information necessary to make informed choices in the marketplace and the opportunity to use that information effectively . . . .”).

293. See id. (“[T]he Commission investigates commercial behavior that may be deceptive or unfair, including advertising . . . .”).

294. Over the last two decades, the Federal Trade Commission (FTC) has gradually shifted away from rulemaking, focusing instead on enforcement. The reason for this shift is that rulemaking is seen as labor-intensive and controversial. Rulemaking is now typically undertaken at the behest of Congress. Today, the FTC’s primary method of rulemaking is the public workshop conference, where industry stakeholders (e.g., business entities, consumer groups, other federal agencies, and state law enforcement officials) are brought together to discuss proposed rules changes. The input provided in these conferences help inform the rules promulgated by the FTC. See generally Lydia B. Parnes & Carol Jennings, Through the Looking Glass: A Perspective on Regulatory Reform at the Federal Trade Commission, 49 ADMIN. L. REV. 989, 998–99 (1997) (describing the role of workshop conferences).

295. Id. at 992 n.14 (noting that industry guides provide “an interpretation of the underlying statute, but [do not afford] an independent basis for enforcement action”).

296. Guides for Private Vocational and Distance Education Schools, 16 C.F.R. pt. 254 (2010).

297. See, e.g., Bailey, supra note 292.

298. Id. at 533.

299. For example, the North Carolina Office of Proprietary Schools is a division of the
extent of oversight provided by state higher education regulatory bodies can be broad, encompassing operational aspects of institutions as well as institutional marketing practices.300 Every state also has a consumer protection agency that provides oversight of various aspects of commerce, including advertising.301 Some of these agencies have divisions that specifically oversee proprietary institutions.302

4. Accrediting Agencies

Accrediting agencies assess and certify that institutions receiving Title IV funds are of sufficient quality. These agencies are non-governmental303 and are typically formed by peer institutions seeking to devise and promote certain educational standards.304 These standards, however, are rarely concrete, allowing individual schools to define their own missions.305 Federal law mandates this flexibility,306 though accrediting agencies are allowed to set standards that can trump institutional standards.307 Accreditation is voluntary; however, only accredited institutions can receive Title IV funds.308

---


302. See, e.g., Utah Div. of Consumer Prot., supra note 275 (providing instructions for postsecondary proprietary schools to complete registration with the Utah Division of Consumer Protection, as required by state law).

303. In most countries, accreditation is a governmental function. However, concerns about federalism have prompted Congress to place the responsibility of institutional quality assessment in the hands of these private entities. There are about 3,500 accredited institutions nationwide, and the vast majority of them are for-profit. Foster, supra note 23, at 18.

304. GAO, Ensuring Quality, supra note 27, at 5–6.

305. See, e.g., Foster, supra note 23, at 19.

306. 20 U.S.C. § 1099(b)(5)(A) (2006) (“[T]he standards for accreditation of the agency or association assess the institution’s success with respect to student achievement in relation to the institution’s mission . . . .” which may include different standards for different institutions or programs, as established by the institution.).


308. See, Foster, supra note 23, at 2, 4.
5. **Self-Regulatory Bodies**

Within the realm of advertising, various self-regulatory bodies promote good advertising practices. Generally, the purposes of self-regulation are twofold: to promote a set of industry norms and best practices, and to provide a means of applying and enforcing these norms. As it concerns advertising, self-regulation is also intended to protect consumers and foster fair competition—two goals that are highly compatible with free-market ideals. Like accreditation, participation in a self-regulatory scheme is voluntary. In the most developed arrangements, these bodies work directly with the FTC and state agencies in regulating advertising.

**C. Discussion**

Proprietary schools invest heavily in mass media advertising. They spend upwards of one billion dollars each year promoting their programs. Their commercials dominate non-prime-time television, and their online ads seem omnipresent. But when it comes to advertising, proprietary schools suffer from a problem common among all educational institutions: their product—education—is largely intangible. In most cases, the only tangible manifestation of education is the diploma that is received upon completion. In attempting to sell their product and differentiate themselves from competitors, nonprofit institutions often promote tangible ancillaries to the educational experience, such as

---

309. See generally Bailey, supra note 292.


311. See, e.g., Bailey, supra note 292, at 537 (“An often-stated goal of the self-regulatory apparatus is to protect consumers from deceptive advertising; there is no doubt in my mind, however, that another important goal served by it is to protect—if not necessarily to promote—fair competition.”).

312. See id. (describing how the National Advertising Division of the Council of Better Business Bureaus, Inc. and the National Advertising Review Board serve as valuable components to state and federal oversight of advertising).


315. See, e.g., ANCTH, supra note 105, at 31 (“Among the greatest challenges to successfully marketing higher education is the inherently intangible nature of the very thing that is being marketed.”).
attractive buildings and tangential student services. Proprietary schools, however, tend to take a different tack; they make more concerted efforts to sell their products by touting them to tangible end results, such as career advancement and financial stability. It is on these types of forward-looking ads that the FTC should require prominent disclaimers.

Free-market competition requires that commercial entities be allowed to communicate with consumers. Therefore, commercial speech is given many of the same First Amendment protections as regular speech. In the seminal case, *Virginia State Board of Pharmacy v. Virginia Citizens Consumer Council*, the Supreme Court reviewed the constitutionality of a Virginia state law banning pharmacies from advertising prices. The Supreme Court struck down the ban, reasoning that consumers in a free market must be empowered by the free flow of information. This case represented a departure from previous Court decisions suggesting that commercial speech fell outside the purview of First Amendment protection.

In *Virginia State Board of Pharmacy*, the Court insisted that First Amendment protection is available to such speech. The Court reasoned that most speech is commercial in nature when (1) it is a paid-for advertisement (2) that refers to a specific product, and (3) is published in the economic interest of the speaker; however, the Court indicated that not all three characteristics are necessary for speech to be considered commercial. Va. State Bd. of Pharm. v. Va. Citizens Consumer Council, 425 U.S. 740, 761 (1976) ("[S]peech does not lose its First Amendment protection because money is spent to project it, as in a paid advertisement of one form or another.").
Amendment protections only extend to truthful and non-misleading commercial speech, and in a later case, Central Hudson Gas & Electric Corp. v. Public Service Commission, it explained that such speech can only be restricted when the government proves it has a substantial interest and its restriction directly advances that interest in the least intrusive manner possible. In Central Hudson, the Court struck down a New York State ban on electric company advertising. The Court reasoned that even though the state had a substantial interest in energy conservation, the challenged restrictions were “more extensive than necessary”—and therefore could not stand. The holdings in both Virginia State Board of Pharmacy and Central Hudson affirm the Court’s view that the First Amendment “favor[s] the dissemination of truthful product information over government suppression of ideas.” So any proposed regulation of proprietary school advertising, including required disclaimers, must serve a compelling state interest and be narrowly tailored to serve that interest.

1. Requiring Disclaimers

The Supreme Court has weighed in on the issue of disclaimers. When the Court struck down a ban on attorney advertising in Bates v. State Bar of Arizona, the state bar association argued that attorney advertisements were inherently misleading due to the individualized nature of each potential client’s needs. The Court was unconvincing, however, reasoning that such a view “assumes that the public is not sophisticated enough to realize the limitations of advertising.” The Court further reasoned that correct but incomplete information was better remedied by more disclosure, not less. As such, disclaimers are preferred over broader restrictions on

323. Id. at 771 (“Untruthful speech, commercial or otherwise, has never been protected for its own sake.”); see also In re Int’l Harvester Co., 104 F.T.C. 949, 1056 (1984) (noting that deceptive advertising “is harmful to consumers, undermines the rational functioning of the marketplace, and . . . never offers increased efficiency or other countervailing benefits that must be considered”).
325. Id. at 566 (“For commercial speech to come within [First Amendment protection], it at least must concern lawful activity and not be misleading. Next, we ask whether the asserted governmental interest is substantial. If both inquiries yield positive answers, we must determine whether the regulation directly advances the governmental interest asserted, and whether it is not more extensive than is necessary to serve that interest.”).
326. Id. at 602.
329. Id. at 374–75.
330. Id. at 375.
speech, such as bans.

Disclaimers serve two basic purposes: to prevent deception and to prompt advertisers to weigh the benefits of making deceptive or incomplete claims in light of the costs of the disclaimer. 331 In arguing for required disclosures on forward-looking proprietary school ads, the author borrows language from the Bates Court and characterizes these ads as correct but incomplete. As the commercials assert, higher levels of education are positively associated with higher income levels. 332 In that sense, the information is correct. The incompleteness is in the suggestion that completion of the program assures higher income and that completion itself is assured—or even likely. The association between education and income is not absolute, and as discussed earlier, most proprietary school students fail to persist to degree. Required disclaimers would address the incomplete treatment of these realities. Regulation such as this would meet the test put forth in Central Hudson: the state has substantial interests in protecting its citizens from misrepresentations and reducing the public costs thereof, and disclaimers directly addressing incomplete information would be the least restrictive manner of serving these interests.

The FTC has required advertisers to use disclaimers when necessary to prevent deception. 333 Generally, disclaimers are required to be conspicuously placed and easy to understand. 334 These basic requirements make sense, because for a disclaimer to be effective, it must be noticed and understood by consumers. In its orders, the FTC is often very specific regarding the form, content, and placement of disclosures. For instance, in adjudicating In re La Salle Extension University, the Commission found that the respondent deceptively advertised its law degree program by not sufficiently disclosing its lack of accreditation. 335 As a result, it ordered the respondent to disclose the program’s limitations with disclaimers “in type the same size

331. Ippolito, supra note 318, at 950 (“[R]equiring [disclaimers] raises the firm’s cost of making the claims, and the ‘clutter’ of the added requirements may make the claims less effective as a marketing tool. If these effects are significant, they reduce firms’ incentives to make the claims at all.”).


333. See, e.g., In re La Salle Extension Univ., 78 F.T.C. 1272, 1284 (1971) (“Where . . . the mere offering of the product or service leads to deception . . . we believe that it is reasonable and necessary to demand that a disclosure required to dispel the deception be given equal prominence with the offer.”).


335. 78 F.T.C. at 1272–73.
and appearance as the advertising claims.” Further, the FTC dictated the placement and content of these disclaimers. Similar requirements could be placed on forward-looking proprietary school ads. In order to prevent deception, disclaimers relating to low completion rates and the relationship between education and earnings should be required components of these ads. The following disclaimers could be placed on ads:

Most students who begin academic or training programs at this institution do not complete them.

Completing the degree/training does not guarantee employment or a higher salary.

These disclaimers should appear conspicuously on ads, using the same font size and appearance as the advertised claims. In television ads, disclaimers could be displayed conspicuously on the screen or stated clearly by the narrator. These disclaimers would be particularly necessary for ads using consumer testimonials, a common marketing strategy for proprietary schools. The claims of consumer endorsers must be “typical” or a disclaimer is required. Given proprietary school completion rates, any consumer endorser touting the benefits of attaining a degree is arguably describing an atypical experience. Lastly, in determining whether an ad is deceptive, the FTC will consider the ad’s effect on a reasonable member of the targeted group. The previously discussed susceptibility of the targets

336. Id. at 1280–81.
337. Id. (requiring disclaimers to be placed “on the front page or cover and on each page of any promotional material or descriptive brochure wherein respondent’s law courses or law degrees are mentioned in type the same size and appearance as the advertising claims appearing thereon”).
338. Id. at 1281 (requiring disclaimers to state that “courses are not recognized or accepted as sufficient education or legal training to qualify the student to become a candidate for admission to the profession of law in any of the States in the United States or the District of Columbia”).
339. A recent Remington College commercial airing in the Nashville, Tennessee area displayed two disclaimers, including one stating, “Individual results may vary.” The disclaimers were displayed in very small font and only appeared for ten seconds of the sixty-second commercial. Vimeo.com, Remington College Commercials, http://www.vimeo.com/29015300 (last visited Aug. 4, 2010).
340. FTC Guides Concerning Use of Endorsements and Testimonials in Advertising, Consumer Endorsements, 16 C.F.R. § 255.2(b) (2010).
341. See John E. Villafranco & Andrew B. Lastigman, Regulation of Dietary Supplement Advertising: Current Claims of Interest to the Federal Trade Commission, Food and Drug Administration and National Advertising Division, 62 FOOD & DRUG L.J. 709, 723–24 (2007) (noting that “the degree of sophistication of the target audience is a significant factor in determining the reasonable message conveyed by the advertising” and explaining that the FTC has assessed both higher and lower standards of reasonableness).
of proprietary schools ads increases the need for disclaimers.

2. Expanding FTC Proprietary School Guides

In fostering good marketing practices, including the systematic use of disclaimers, the FTC should expand its Guides for Private Vocational and Distance Education Schools. The FTC promulgated the Guides in 1972 as a means of advising “proprietary businesses offering vocational training courses, either on the school’s premises or through distance education, how to avoid unfair or deceptive practices in connection with the advertising, promotion, marketing, or sale of their courses or programs.” As such, the Guides address prohibitions against various types of misrepresentations. However, the Guides only pertain to proprietary schools offering less than a two-year degree. This limited applicability does not reflect the current reality of proprietary school education. When the Guides were first enacted, very few proprietary schools were offering degree programs. Today, many of these schools offer degrees through the doctoral level. In fact, at some of the largest proprietary schools, most students are enrolled in degree-granting programs. But irrespective of their evolving programmatic focus, the marketing strategy used by these schools has remained rather consistent; they still tie their programs to labor market success. Thus, the dangers that the Guides were enacted to address have expanded beyond the scope of the Guides, necessitating a broadening of that scope.

343. Private Vocational and Distance Education Schools, Request for Public Comments, 74 Fed. Reg. 37,973, 37,975 (July 30, 2009).
344. These misrepresentations concern the description of the school, its accreditation, the transferability of credits, the content of ads and testimonials, teacher qualifications, courses offered, the availability of employment and financial aid, and enrollment qualifications. Id. at 37,973–74.
345. 16 C.F.R. § 254(a) (“These Guides do not apply to resident primary or secondary schools or institutions of higher education offering at least a 2-year program of accredited college level studies generally acceptable for credit toward a bachelor’s degree.”).
346. See GAO, STRONGER OVERSIGHT, supra note 13, at 1 (“In recent years, the scale and scope of proprietary schools have changed considerably. . . . Traditionally focused on certificate and associate programs ranging from cosmetology to medical assistance and business administration, proprietary institutions have expanded their offerings to include bachelors, masters, and doctoral level programs.”).
347. See KIRP, supra note 87, at 241 (describing a “new breed” of proprietary schools where the majority of students are enrolled “in degree programs for everything from the associate degree to the Ph.D.”).
348. In July, 2009, the FTC requested public comments on the Guides “as part of its systematic review of [agency] guides and regulations.” In the request, the FTC presented eighteen questions relating to how the Guides can be made more effective. None of the
3. **Encouraging Self-Regulation**

The proprietary school industry, with the encouragement of the FTC, should form a self-regulatory body to encourage good advertising practices within the sector. Industry self-regulation is an important component to FTC oversight and the overall prevention of fraudulent advertising. For example, the National Advertising Review Council (NARC), an umbrella self-regulatory agency, has set advertising guidelines for various industry-specific self-regulatory agencies, including those relating to electronic retailing and children’s advertising. Also, the FTC has incorporated self-regulatory agencies into its regulatory framework. The Children’s Advertising Review Unit (CARU) and the National Advertising Division (NAD) of the Council of Better Business Bureaus serve as initial reviewers of challenged advertisements. If an advertiser does not agree with a

questions directly related to expanding the scope of the Guides, though question three asks about possible modifications. 74 Fed. Reg. at 37,973–75.


350. See Better Bus. Bureaus, Electronic Retailing Self-Regulation Program, http://www.bbb.org/us/electronic-retailing-self-regulation-program/ (last visited Aug. 4, 2010) (“[The Electronic Retailing Self-Regulation Program]’s mission is to enhance consumer confidence in electronic retailing. ERSP provides a quick and effective mechanism for evaluating, investigating, analyzing[,] and resolving inquiries regarding the truthfulness and accuracy of the primary or core efficacy or performance claims that are communicated in national direct response advertising.”).


decision made by CARU or NAD, it may appeal to the National Advertising Review Board (NARB).353 In assessing challenged ads, the FTC gives great weight to precedent set by these quasi-judicial self-regulatory agencies.354

A proprietary school self-regulatory body could be chartered through an impartial agency such as the Better Business Bureaus.355 The body could serve as a clearinghouse for best practices in industry advertising, as well as a place where ad-related complaints could be brought by consumers and competitors alike. Like CARU and NAD, the body could serve as an initial arbiter of complaints, with appeals going to NARB. The proprietary school industry would benefit greatly from this type of self-regulation; it would improve the sector’s credibility with the public while encouraging healthy competition and possibly staving off closer governmental scrutiny of its advertising practices.

4. Requiring Affirmative Disclosures

The federal Student Right-to-Know and Campus Security Act requires all institutions receiving Title IV aid to make wide-ranging disclosures to prospective and enrolled students.356 The disclosures most pertinent to this discussion are graduation rates and placement rates. Under the Act, schools must make this information “readily available upon request” to prospective and enrolled students.357 Further, schools must “provide to all enrolled students a list of the information that is required to be provided . . . together with a statement of the procedures required to obtain such information.”358 By requiring schools to disclose this information, the Act is acknowledging the predominant motivation of students engaging in higher education; it is also making a powerful policy statement—one that places outcomes at the focal point of assessment.

Unfortunately, the manner in which the statute operationalizes the requirements lessens their effectiveness. The only documentation a school must provide is a list of information it is required to make available, and a process for obtaining that information. In effect, the Act places the onus on

354. Bailey, supra note 292, at 537 (noting that the advertising industry “has established two quasi-judicial regulatory bodies to review advertising,” and that these agencies are “a valuable complement to federal and state efforts to police against deceptive advertising”).
355. Cf. ICAP REPORTS, supra note 310, at 1 (“Impartiality is seen to be key to an effective [self-regulatory] code and public trust in it.”).
357. Id. § 1092(a)(1).
358. Id.
the student to not only request the information, but take the necessary steps to secure it. It stands to reason that these unnecessary steps limit the dissemination of this information; therefore, proprietary schools should be required to affirmatively disclose, at the very least, graduation rates and placement rates to students before enrollment and each academic year thereafter. Such a requirement would not be novel, as the Act already requires schools to disclose graduation rates and other data to athletes, their parents, and officials at their secondary schools.\textsuperscript{359} Also, individual states, like Utah, require proprietary schools to provide employment and graduation rate data prior to enrolling a student or accepting tuition payments.\textsuperscript{360} Such a requirement would also make it more difficult for schools to use bureaucratic inconveniences to discourage students from obtaining this information. Oversight of this requirement could be within the purview of the DOE, with assistance from state regulatory and accrediting agencies.

5. Expanding Disclosures

Disclosure requirements for proprietary schools should be expanded to include labor market data, specifically information relating to labor demand and salary.\textsuperscript{361} This expansion would be in direct response to proprietary school marketing and recruitment practices. Some of the occupational areas for which proprietary schools provide training have little to no demand.\textsuperscript{362} Additionally, salary data is often inflated by admissions representatives.\textsuperscript{363} Thus, providing this information to students prior to enrollment will better inform students, allowing them, as consumers, to make informed choices in the marketplace.\textsuperscript{364} Similar to a disclaimer, it

\textsuperscript{359} Id. § 1092(c)(2). \textit{But see id.} § 1092(e)(6) (waiving these requirements “for any institution of higher education that is a member of an athletic association or athletic conference that has voluntarily published completion or graduation rate data or has agreed to publish data that . . . is substantially comparable to the information required under this subsection”).

\textsuperscript{360} \textsc{Utah Code Ann.} § 13-34-108 (2009).

\textsuperscript{361} This proposal is based on a recommendation made by the GAO. See GAO, OVERSUPPLIED OCCUPATIONS, supra note 22, at 13–14.

\textsuperscript{362} \textit{See id.} at 8 (“The surplus of qualified job candidates, including proprietary school graduates, for some occupations occasionally reached dramatic proportions in some states, exceeding demand by ratios of 10 to 1 or more.”).

\textsuperscript{363} \textit{See Anti-Fraud Hearings, supra note 10, at 5 (statement of Honorable George Miller)} (“[C]ertain colleges . . . misrepresented graduation rates, promised inflated salaries to prospective enrollees, [and] enrolled students who did not have the ability to complete casework . . . .”).

\textsuperscript{364} GAO, OVERSUPPLIED OCCUPATIONS, supra note 22, at 5 (“Using labor market projections provides a rational basis for making training investment decisions . . . .”). \textit{But see}
would also prompt proprietary schools to consider the costs of making claims that may not be supported by the data. To ensure validity, the information should be compiled by a governmental agency or another entity certified by the DOE.

CONCLUSION

Proprietary schools play an important role in broadening access to higher education. They enroll a large number of students who are underserved by traditional, nonprofit institutions. These students tend to be poorer, less educated, and older than students at traditional schools, and they tend to undertake higher education for very practical reasons. These characteristics make them more susceptible to deceptive marketing and unfounded promises of higher education providers.

Proprietary schools invest heavily in marketing and recruitment. They appeal to the characteristics and motivations of their market niche by promoting tangible end results of educational study, such as career advancement and financial stability. Unfortunately, many of their ads and recruitment practices make representations that are incomplete, or worse, untrue. These behaviors contribute to low completion rates and high loan default rates among proprietary school students—outcomes that cost students and taxpayers billions of dollars.

To protect students and taxpayers from proprietary school misrepresentations and fraud, tighter regulation of their marketing and recruitment practices should be imposed. In the area of marketing, proprietary schools should be required to place disclaimers on forward-looking ads. Also, the FTC should expand its regulation of proprietary school marketing practices and encourage impartial self-regulation within the industry. In the area of recruitment, proprietary schools should be required to make affirmative and expanded disclosures. The goal of these reforms is to foster disincentives to misrepresentations and fraud. It must be noted that while the specific focus of this article is proprietary schools, the proposals could apply to any school that makes forward-looking representations in inducing enrollment. In the end, the message should be that while educational attainment can, and often does, yield benefits upon the possessor, these benefits are not assured—and because of this, “Your Results May Vary.”

id. (warning that labor market projections are “inherently imprecise”).
NEW GOVERNANCE, FINANCIAL REGULATION, AND CHALLENGES TO LEGITIMACY: THE EXAMPLE OF THE INTERNAL MODELS APPROACH TO CAPITAL ADEQUACY REGULATION

ROBERT F. WEBER*

TABLE OF CONTENTS

Introduction ............................................................................................................. 784
I. The Importance of Capital Adequacy Regulation of Financial Institutions ................................................................. 788
II. Banks and the Basel Committee .................................................................... 796
   A. Regulating Bank Capital Adequacy in the United States: A Brief Pre-Baselite History .................................................. 799
   B. The Basel Committee and Capital Adequacy: From Basic Risk Sensitivity to an Internal Models Approach .................. 800
      1. Birth of an International Policy Coordinator ......................... 800
      2. The Initial Basel Accord and the Introduction of Risk Sensitivity ................................................................. 802
      3. The Internal Models Approach: Delegating Assessments of Market Risks and Credit Risks to Regulated Banks ............ 805
         a. Banking in the 1990s: Competition, Globalization, Technology, Complexity, and Code ........................................ 805
            i. The New Financial Code .............................................. 807
            ii. Regulatory and Policymaker Reactions to Financial Complexity: In Their Own Words ........ 811

* Westerfield Fellow, Loyola University New Orleans College of Law. I thank Andrea Armstrong, Laura Beny, Carter Dillard, Trey Drury, Adam Feibelman, Cristie Ford, Erik Gerding, Liz Haynes, John Pottow, and JoAnne Sweeney for helpful comments on earlier drafts. I may be reached at rfw@loyno.edu.

783
(a) First Bank Regulatory Premise: Complexity Adds Risk, But Is Needed to Compete Effectively .................................................. 812
(b) Second Bank Regulatory Premise: The New Financial Code Provides an Effective Check on Risks to Safety and Soundness from New Bank Activities .................................................. 815
   c. Basel II Pillars 1 and 3: Internally Modeling Credit and Operational Risks and Public Disclosure ........ 826
      i. Pillar 1: Capital Requirements and Internal Models .................................................. 827
III. Broker–Dealers and the CSE Program ....................... 830
IV. Insurance Companies & Solvency II ......................... 835
V. Internal Capital Models Approach as New Governance 836
   A. What Is New Governance? ............................... 836
      1. The Benign Big Gun: Retaining a Public Administrative Presence ........................................ 839
      2. Decentering Governance: Increased Power-Sharing and Stakeholder Participation .................. 842
      3. Flexible and Dynamic Law: Overcoming Rule– Principle Polarity ........................................ 845
   B. Normative New Governance: Pro-Regulation, but Beware of Legitimacy Challenges ...................... 850
   C. Internal Capital Models as Flawed New Governance 855
      1. Enhanced Power-Sharing and Increased Participation, a Benign Big Gun, and Dynamic and Flexible Lawmaking .................................................. 855
      2. Internal Models and Normative Challenges to Legitimacy .................................................. 857
      3. Pillar 3’s Mandatory Disclosure as an Incomplete Tripartite Governance Response to Legitimacy Challenges .................................................. 863
Conclusion .................................................................................. 867

INTRODUCTION

This Article considers the application of “new governance” theory and scholarship to financial regulatory reform by examining the recent trend of incorporating proprietary internal risk models into capital adequacy
regulatory regimes. In the aftermath of the subprime credit crisis, recent academic and policy debate about the regulation and supervision of financial institutions has rightly focused on potential solutions to the manifold conflicts of interest and regulatory lacunae that exist in our current system. While most of these proposals concern the situs and scope of regulation, this Article contends that theoretical scrutiny of the methodologies and tools by which financial institutions are regulated—especially the “modes of interaction” between financial firms, their regulators, and other nonstate stakeholders—is relatively underemphasized in financial regulation legal scholarship. In recent years, political and socio-legal scholars have contributed to a rich new governance literature regarding the evolving methodologies and tools of governance. By focusing on these modes of interaction and tools as primary units of inquiry, this new governance scholarship offers important insights into the causes and potential remedies of regulatory dysfunction, as well as the dangers associated with increased involvement of regulatees and other nonstate actors in regulatory processes.

The central tenet of new governance literature posits that traditional command-and-control, top-down regulation has been supplanted, to varying degrees, by new forms of collaborative and polycentric governance, often involving dynamic cooperation between the public sector (formerly the “governors”) and the private sector (formerly the “governed”), and often characterized by an increased participation in governance by third-party nonstate actors. These new hybrid forms of managing events in social systems (including the financial system) have emerged in parallel with

1. Capital adequacy regimes are comprised of the legislative and regulatory rules requiring regulated financial firms to maintain levels of capital relative to assets (with appropriate adjustments) to foster the safety and soundness of institutions and the financial system.

2. For an explanation of the difference between “regulation” and “supervision” in the financial regulatory context, see generally R.M. Pecchio, PRUDENTIAL SUPERVISION IN BANKING (1987). To avoid confounding the terminology used in Parts I–V with the more general discussion of regulation in Part V, I have adopted regulation to refer to both supervision and regulation, recognizing that certain liberties were taken with these otherwise distinct terms.

3. See Miriam H. Baer, Governing Corporate Compliance, 50 B.C. L. REV. 949, 952-54 (2009), available at http://www.bc.edu/content/dam/files/schools/law/bclawreview/pdf/50_4/02_baer.pdf (discussing new governance in the context of deferred prosecution agreements); Saule Omarova & Adam Feibelman, Risks, Rules, and Institutions: A Process for Reforming Financial Regulation, 39 U. MEM. L. REV. 881, 920 (2009) (“[T]he ultimate goal is broader than formulating policy priorities or outlining the contours of substantive[] rules governing the conduct of private market participants. It is equally important to identify the most effective and efficient modes of interaction between the regulators and the regulated, which may vary across different segments of the financial services sector.”).
the attenuation of traditional state power and the increasing complexity of postmodern forms of life and social organization. New governance scholars have not attempted to apply their analysis to financial regulation, and scholars of financial regulation have not yet appreciated the rich insights that new governance theory offers to their field.

The internal models approach to capital adequacy regulation should be considered a new governance technique because, by incorporating regulated institutions’ internal capital models in the capital adequacy regime, it seeks to bridge intractable information asymmetries resulting from the complexity and dynamism of contemporary financial institutions. Despite its manifold advantages as a new governance tool in a highly complex and dynamic regulated field, the internal models approach falls into traps familiar to new governance reforms that render it susceptible to literal and softer forms of agency capture, thereby compromising its democratic legitimacy and effectiveness.

The example of the internal models approach is instructive: it represents a new governance initiative that, while credibly seeking to accommodate burgeoning complexity and overcome the ineffectiveness of traditional

---


5. Much of the debate about financial regulatory technologies has revolved around the juxtaposition of so-called “principles-based” and “rules-based” regulation. In Part V.A.3, this Article draws on recent scholarship arguing that the principles-rules binary distinction does little to advance our understanding of financial regulation, and argues instead in favor of a focus on new governance.
capital adequacy regulation, falls victim to complications that undermine its regulatory objective. Capital adequacy regulation aims to minimize social costs associated with financial institution insolvencies. It achieves that goal by fostering a loss buffer with reasonable confidence and imposing a risk tax in the form of a capital charge for riskier business lines. A premise of this Article is that since at least the early 1970s, when major powers jettisoned the Bretton Woods financial system, the global financial system has been positioned in a complexity paradigm in which innovation, competition, globalization, technology, and consolidation operate, sometimes interactively, to ratchet up the system’s complexity. This process accelerated in the 1990s as advances in theoretical finance and computer technologies combined to create a “new financial code” embodied in value-at-risk (VaR) and economic capital models. Since the effectiveness of capital adequacy regulation depends on the ability to secure a loss buffer and impose a risk tax, this complexity paradigm, which entails a dynamic and proliferating risk profile, presents a structural dilemma for regulators. In response to this dilemma, in recent years financial regulators have sought to incorporate the new financial code (in the form of internal models) into capital adequacy regimes as a means of bridging the informational chasm between the actual risk profile of regulated firms and the regulators charged with minimizing the social costs incurred if those risks materialize.

My hope is that this Article, by analyzing a financial regulatory technology as a response to complexity from a new governance perspective, will prompt a deeper appreciation for new governance theory within financial regulation scholarship. New governance theory offers notable insights into the regulation of social systems dominated by complexity and dynamism, as with the contemporary financial system. By analyzing the complex financial system according to a new governance framework, scholars and policymakers will likely be able to (1) improve their diagnosis of sources of regulatory dysfunction, including those in connection with the recent subprime credit crisis; (2) propose reforms that will better conduce the public goals of financial regulation; and (3) deepen their understanding of the normative challenges to democratic legitimacy that are implicated when regulators seek to govern complexity through increased involvement of nonstate actors in the regulatory process. Future new governance inquiries into financial regulatory reform must focus on this legitimacy challenge and safeguard against the possibility that in the rush to bridge information asymmetries inherent in complex markets, policymakers might

adopt a reform that creates more problems than it solves. This Article argues that this possibility materialized when financial regulators adopted the internal models approach. Sometimes, new governance theory will demonstrate how regulatory tools may be deployed to regulate complexity and dynamism more effectively; at other times, it will elucidate flaws with those tools. Because new governance theory focuses attention on these problems, it must remain open to a critique of complexity itself.

Part I explains the critical public regulatory role of capital adequacy regulation. Part II.A summarizes the genealogy of capital adequacy regulation for banks. Part II.B discusses the formation of the Basel Committee on Banking Regulations and Supervisory Practices (Basel Committee) as the international bank regulatory policy coordinator and its three central achievements: the initial Basel Accord of 1988, the substantial market risk amendments of 1996 (later amended in 2005), and Basel II. This Part also traces how the Basel Committee adopted internal models approaches as a means of more closely calibrating capital requirements to the actual risk profiles of regulated banks, which had become increasingly complex in the 1990s, in large part due to the development of a new financial code, including risk modeling technologies. Parts III and IV, briefly address how the ascendancy of this “Basel Brand” of regulation led to the Securities Exchange Commission’s (SEC’s) Consolidated Supervised Entities (CSE) Program for broker–dealers and the European Union’s (E.U.’s) proposed Solvency II regime for insurance companies.

Part V provides an overview of new governance theory and situates the use of internal models by financial regulators within the new governance literature. In particular, the discussion focuses on three attributes of new governance tools that are relevant to the internal models approach: increased participation by nonstate stakeholders, the retention by public regulators of a “benign big gun,” and dynamic and flexible lawmaking. In the process, the Article describes certain normative challenges posed by new governance initiatives. Finally, the Article (1) explains why the internal approach to capital adequacy regulation should be considered a new governance initiative; (2) shows how it poses normative problems typical of new governance initiatives that threaten its democratic legitimacy, expose it to risks of literal and soft capture, and ultimately undermine its effectiveness; and (3) evaluates its mandatory disclosure regime as a promising, but incomplete, solution to the central challenge to legitimacy.

I. THE IMPORTANCE OF CAPITAL ADEQUACY REGULATION OF FINANCIAL INSTITUTIONS

This Part explains the role of financial institutions, their inherent instability, and the role of capital adequacy regulation in achieving the
policy goals of consumer protection and systemic stability. Regulators require that regulated financial institutions maintain capital at specified levels. Regulators impose these capital requirements to ensure that firms have an available “cushion” or buffer to absorb unexpected losses without threatening the ability of the institution to satisfy claims of depositors (in the case of banks), insurance policyholders (in the case of insurers), or clients (in the case of broker–dealers), or the systemic integrity of the financial system. In its simplest form, the capital requirement consists of a minimum ratio of capital—defined to include equity (assets minus liabilities) and in some cases other add-ons thought to possess equity-like features, such as subordinated debt—to assets. Thus, as a firm’s equity declines, so does its capital ratio, which may result in noncompliance with the capital adequacy regime and a need to raise new capital.

Financial systems perform essential functions in a market economy, such as transforming savings into productive capital, providing information about users of capital, performing capital allocation functions, providing payment and funds transfer services, reducing agency costs by overseeing management, and efficiently spreading and pooling risk through derivatives and securitization. Finance is “the oil that lubricates the wheels of commerce.” These ends of finance can be frustrated by exogenous factors and even by the financial system’s internal operation. Financial institutions may also be controlled by rent-seeking constituencies, as evidenced recently when investment bank personnel “earned” tens of millions of dollars of income for their roles in securitizing collateral pools of subprime mortgages. While the optimal prevalence and complexity of the financial sector in a market economy is a matter very much up for debate,  

10. In the United States, finance has benefited from an increasing fraction of economic growth. See Martin Wolf, Why Is It So Hard to Keep the Financial Sector Caged?, FIN. TIMES, Feb. 5, 2008, http://www.ft.com/cms/s/0/9907c5c4-d41f-11dc-a8c6-0000779fd2ac.html (“The US itself looks almost like a giant hedge fund. The profits of financial companies jumped from below 5 per cent of total corporate profits, after tax, in 1982 to 41 per cent in 2007, even though their share of corporate value added only rose from 8 to 16 per cent.”). In order to determine whether the financial sector has outgrown its utility, policymakers should ask “what fraction of the economy’s total returns to productively invested capital is absorbed up front by the financial industry as the costs of allocating that capital.” Benjamin M. Friedman, The Failure of the Economy & the Economists, N.Y. REV. OF BOOKS, May 28, 2009, at
the critical role of a functioning financial sector is not.

The structural dilemma of all firms operating in this sector is the asset–liability mismatch: the basic conflict between guaranteeing a return of capital (e.g., insurance claims, interest on deposits) while also putting that capital at risk. This tension is omnipresent and engenders fragility in the financial system. Currently, capital adequacy regulation is the chief method regulators employ to foster the continuous healthy operation of the financial sector. Regulators impose minimum capital requirements to counteract the tendency for financial firm insolvencies to result in two categories of negative externalities: first, bailout costs to taxpayers resulting from the moral hazard effects of government safety nets and second, systemic losses to the broader economy from systemic risks that build up in the financial sector.

Bailout costs result when taxpayers foot the bill for public expenditures to fund government safety nets. When a nonfinancial firm maintains a capital buffer, it signals its financial health and ability to pay creditors as debts come due, and thereby reduces its cost of capital. Under such circumstances, creditors lacking recourse to assets or a third-party

42. Friedman correctly laments that this is a question “which no one seems interested in addressing.” Id. Professor Dirk Bezemer has recently made a similar point, acknowledging the critical role of finance, but noting the dangers of “financializing” the economy. See Dirk J. Bezemer, “No One Saw This Coming”: Understanding Financial Crisis Through Accounting Models 16 (Munich Pers. RePEc Archive, Working Paper No. 15892, 2009) (“Financial innovation . . . serves the real economy’s need, in that it boosts real-sector productivity and its ability to service its increased debts. But it also opens up the possibility of a sustained drain of liquidity from the real to the [financial] sector, so inflating asset prices—a credit bubble, or harmful ‘financialization’ of the economy.”). Under Bezemer’s “financialization” scenario, debt levels increase steadily, eventually using financial assets as collateral for new borrowings, and in the process the fraction of an economy’s total returns are absorbed up front by the financial sector. The real economy (government, households, and firms) will ultimately have to service the debt since debt backed by capital gains in financial assets cannot expand infinitely. Id. At a certain point in the cycle, the debt service burden increases to the point that the financial system becomes extractive to, rather than supportive of, the real economy. Id. In Bezemer’s terms, there emerged a “growing imbalance in the flow of funds between the real and financial sectors” and by analyzing the macro flows of funds from an accounting perspective reveals the “extent to which the economy had grown dependent on asset price gains.” Id. at 17.

11. This is not a necessary state of affairs. It is also possible to promote solvency through directly regulating the extension of credit. For instance, had U.S. federal banking regulators not removed loan-to-value (LTV) minimums on home mortgage loans extended by federally regulated banks, credit flow into the subprime housing sector would have been cut off at the spigot and the credit crisis would have been averted. Consider this context the experience of the comparatively salubrious Canadian banking sector during the 2008 credit crisis: Canada had a requirement that any mortgage loan with an 80% or greater LTV ratio had to be privately insured.
guarantee would be expected to price credit according to, among other factors such as loan indenture covenants, the amount of equity capital a debtor firm maintains; the greater the equity capitalization, the lower the costs of financial distress. This dynamic, which is routine in other industries, is inapposite in the financial context because of the taxpayer-funded safety net financial institutions receive in the form of de jure government safety nets (e.g., state insurance guarantee funds, Federal Deposit Insurance Corporation (FDIC) insurance) and de facto too-big-to-fail support (e.g., support for holding companies such as American International Group (AIG), Citigroup, Bear Stearns). Creditors, knowing there is at least a partial backstop guarantee from government, lack incentives to properly monitor risk and price credit accordingly. Since private parties lack incentives to monitor the use of the firm’s capital, levels of risk should outpace socially optimal levels, especially during times of stress. The rationale of capital adequacy regulation, then, is to ensure through public regulation that financial institutions enjoying government safety nets are subject to the discipline that capital markets and counterparties would normally impose but do not impose because of moral hazard. Safety and soundness are achieved through an enforced capital requirement serving as a proxy for market forces. In this respect, we can


13. See Richard Scott Carnell et al., The Law of Banking and Financial Institutions 300 (4th ed., 2009) (“By inhibiting market discipline of banks, deposit insurance creates incentives for excessive risk-taking and impels the government to protect the taxpayers through safety-and-soundness [i.e., capital] regulation.”); Jean-Charles Rochet, Why Are There So Many Banking Crises? The Politics and Policy of Bank Regulation 1 (2008) (“The two main components of this public intervention [into the banking market] are on the one hand the financial safety nets (composed essentially of deposit insurance systems and emergency liquidity assistance provided to commercial banks by the central bank) and on the other hand the prudential regulation systems, consisting mainly of capital adequacy (and liquidity) requirements, and exit rules, establishing what supervisory authorities should do when they close down a commercial bank.”).

14. Theoretically, the government could, as an effective guarantor of deposits and many debts, price its guarantee and secure covenant protections in a negotiation with each financial institution availing itself of a government safety net protection. Under normal circumstances, this is precisely what private creditors negotiate in credit documentation. However, because lender-of-last-resort facilities are provided in exigent circumstances when (1) the government, due to systemic risks or political realities, has little credible exit threat, and (2) no private market otherwise exists to perform a “market check” for the government credit, it is impractical to price the government’s extension of credit at the moment such pricing would be required. A far more sensible approach is embodied in the current state of affairs; it is better to set the rules of the game ex ante so as to minimize the likelihood of recourse to the safety net. Cf. Daniel K. Tarullo, Banking on Basel: The Future of International Financial Regulation 20 (2008) (noting that when extending discount
think of bank capital as a partial collateralization of the claims of bank creditors and, more importantly, future taxpayers that might be forced to bail out financial institutions.

The insolvency of a large financial institution might also pose systemic risks to the financial system due to the lack of adequate incentives for any particular market counterparty/creditor to mind the systemic risk that can build up. Knock-on effects from a systemically significant insolvency can threaten to paralyze the entire sector, which in turn can plunge an economy into a deep recession.15 Moreover, unlike other industries where insufficiently capitalized businesses can be broken up into their constituent assets and put to more efficient uses, a rash of insolvencies (or a series of defaults) in the financial sector can threaten to spur a systemic contagion that can freeze up credit formation and, as a result, economic growth. For instance, in the normal course of operation a bank obtains and processes extensive borrower-specific data and monitoring experience that is lost when the bank fails.16 The loss of this data and experience may result in a lack of availability of willing lenders to finance the borrower’s continued activities. Another systemic cost of financial insolvency is the threat to the integrity of the payments system, which can cause the cessation of inter-bank transfers and prevent the efficient flow of capital resources and international trade.17 More recently, we witnessed that the absence of capital behind AIG’s positions in the collateral default swap market exposed several systemically important counterparties to massive losses that were avoided only through government intervention. Again, with systemic risks as with moral hazards, the government steps in to provide in extremis capital at the expense of taxpayers.

Practically speaking, capital requirements foster solvency by (1) ensuring that regulated firms have a capital buffer to absorb unexpected losses and


2010] Internal Models Approach to Capital Adequacy Regulation

(2) imposing a de facto risk tax on financial institutions to provide a disincentive to engage in excessive risk-taking on all of its borrowed capital. From the viewpoint of regulated firms, in most cases minimum regulatory capital requirements constitute risk taxes because they restrict the bank’s ability to make profits. If a bank is required to hold capital against a risky asset portfolio, it is obviously not able to deploy that capital to seek returns. Its profit opportunity set is limited vis-à-vis its profit opportunity set absent regulatory capital requirements. Where firms perceive their capital requirements to be excessive, they will seek to engage in capital arbitrage by deploying capital in risky pursuits that are less burdened by the regulatory capital adequacy framework.

18. A minimum capital requirement will operate as a risk tax if the following assumptions obtain: (1) an imperfectly elastic supply of equity finance, and (2) temporal limitations for accumulating retained earnings as equity. In such a scenario, an institution will likely be required to reduce its risk-weighted assets to maintain acceptable capital ratios, foregoing returns. See, e.g., Sun Bae Kim & Ramon Moreno, Stock Prices and Bank Lending Behavior in Japan, 1 Econ. Rev. Fed. Res. Bank S.F. 31, 33–34 (1994).

19. Other policy justifications for capital regulation are of comparatively less importance, such as ensuring available capital for expansion into new activities, or are derivative of the enumerated factors above, such as protecting government deposit insurance funds and counteracting the moral hazard occasioned by these government safety nets. See Tabullo, supra note 14, at 16 n.3.

20. The “most cases” qualification is necessary to address cases when capital market expectations for capital reserves exceed minimum regulatory capital requirements. Capital markets might require that a firm hold capital in order to respond quickly to acquisition opportunities or market expansion, or to weather an expected downturn in asset values.

21. An example of this phenomenon is the proliferation of off-balance sheet liabilities such as asset-backed commercial paper conduits (ABCPCs) and structured investment vehicles (SIVs), sponsored by major international banks predominantly from 2004 to 2008. Banks established ABCPCs as off-balance sheet entities funded with minimal equity and large issues of commercial paper. In order to secure low funding costs, sponsoring banks would offer contingent recourse to their balance sheets in favor of the commercial paper investors in the form of liquidity or credit enhancement. The ABCPCs and SIVs would invest the proceeds from the commercial paper issuances in asset-backed securities, especially those backed by residential mortgages. The contingent commitments to the ABCPCs and SIVs carried a capital charge of 0.8%—one-tenth of the 8% capital charge required if the asset-backed paper were held on the banks’ own balance sheets. Eventually, the market value of the assets collapsed, the ABCPCs and SIVs were unable to roll over their commercial paper funding, and the investors looked to the contingent commitments of the banks. In September 2008, the Federal Reserve guaranteed investments in money market mutual funds, which were the principal funding entities of ABCPCs. Given the severity of the financial markets crisis, the Federal Reserve determined that it was necessary to, in effect, backstop with taxpayer funds a pure capital regulation arbitrage play! For a succinct summary of the rise and fall of the ABCPC industry, see Viral V. Acharya & Philipp Schnabl, How Banks Played the Leverage “Game,” in Restoring Financial Stability: How to Repair a Failed System 83, 83–100 (Viral V. Acharya & Matthew Richardson eds., 2009) [hereinafter Restoring Financial Stability].
Because capital promotes solvency and also limits profitability, regulators seek to set the capital requirement at a level that balances the need for prudent risk management with the need for firms to be competitive in the marketplace. Stated more technically, the tradeoff is between the marginal social benefit of reducing the risk of the negative externalities from financial institution failures (such as bailout costs and systemic risks) and the marginal social cost of reduced output and productivity due to the regulator-imposed reduction in the volume of finance provided to consumers and businesses. Of course, achieving the optimal balance between these countervailing considerations is only a theoretical possibility. Because the social benefits and costs are likely impossible to quantify and the periodicity of examinations has normally limited information gathering, regulators historically have resorted to uniform minimum capital levels to achieve this balance.

This uniform-rules-of-general-application approach, however, suffers from a lack of risk sensitivity, which compromises the balance and exacerbates the inherent arbitrage threat. Even if the capital adequacy rules are formulated precisely and at the optimal level of generality at time $T_1$, the limits of human foresight inhibit the extent to which those rules will achieve the regulatory objective at time $T_2$. The resulting mismatch between regulatory tools and regulatory objective is widest in highly dynamic industries such as financial services. Each successive period of financial innovation increases the complexity of the financial sector and, with it, the breadth and depth of risks facing regulated firms. As a result, regulator–regulatee information asymmetries widen and prevent regulators from exercising supervision over this new complexity. The major structural reforms in the capital adequacy regulatory regimes discussed below have occurred in response to periods of increasing complexity in the financial industry and a perceived dearth of effective risk sensitivity. In recent years financial innovation has overwhelmed capital adequacy regulatory systems and regulators have attempted to bridge the resulting informational gap by incorporating sophisticated, internally-generated, proprietary risk management and risk estimation models into the capital adequacy regulatory regime. Put another way, financial institutions have become too complicated for non-insiders to acquire the information necessary for effective regulation, and regulators have turned to insiders to mine the data.

---


23. See JULIA BLACK, RULES AND REGULATORS 7–11 (1997) (discussing how rules are inherently indeterminate and almost always over- or under-inclusive).
While this Article portrays capital adequacy reform as a reactive process, with regulators responding to the complexity paradigm, regulators need not assume such a posture regarding complexity. Complexity results from the dynamic interaction of technology, competition, and rent-seeking forces with regulation, which sets the rules under which the other factors operate. Regulation is both constitutive of, and responsive to, complexity. Such policy-created market conditions include, among countless others, globalized capital markets (e.g., removal of capital controls in late 1970s); cross-sectoral consolidation (e.g., repeal of Glass–Steagall encouraged formation of one-stop-shop financial bazaars like Citigroup and AIG); arbitrage (e.g., initial Basel Accord’s differential treatment of off-balance sheet exposures fueled securitization of mortgages); increased intersectoral competition (e.g., interest rate deregulation in early 1980s allowed banks to compete aggressively in deposit-like markets with nonbank institutions); credit expansion (e.g., abandonment of Federal Reserve margin requirements removed check on bubbles in housing and consumer finance); financial product deregulation (e.g., authorizing banks to offer adjustable rate mortgages enabled banks to offer complex loans that many homebuyers could not understand); and shifts to riskier lines of business (e.g., permitting banks to engage in the securitizations and derivatives business in the 1990s). Thus, to a large degree, the complexity paradigm qua regulatory dilemma is a creation of earlier decisions by the same regulators that now struggle to regulate it. Therefore, an alternative solution to regulating the complexity paradigm is to cut off the head of the monster: that is, to use state power to enforce simplicity. By reducing the complexity, the regulator could narrow the inherent mismatch between ex

24. See Philip Stephens, Cut the Banks (and Bonuses) Down to Size, FIN. TIMES, Sept. 1, 2009, at 11 (“Activity for the sake of it—a characteristic of bloated and ever more complex wholesale financial markets—serves only those who manage to extract large profits from the enterprise.”).

25. A neoclassical economic critique of enforced simplicity would emphasize that complexity develops when market participants harness technology to meet counterparty and customer expectations or to circumvent the effect of regulatory business restraints not demanded by counterparties (i.e., arbitrage). The former is benign, and the latter would only be exacerbated by enforced simplicity. Thus, according to this view, technology is instrumental to, rather than constitutive of, financial innovation and complexity. While on its own that principle is unobjectionable, a vulgar neoclassical position sometimes overstates the case and fails to appreciate how dynamic and unpredictable technology can expose financial markets to shocks. However, the neoclassical reminder of arbitrage serves as a useful reminder that any attempts to enforce simplicity are bound to struggle with the arbitrage problem. See, e.g., Ronald H. Coase, The Problem of Social Cost, 3 J.L. & ECON. 1, 6–8, 15–19 (1960) (theorizing that, absent transaction costs, market actors can “transfer and recombine” legal rights to achieve a desired ultimate result irrespective of the initial allocation of legal rights).
ante regulatory rules and regulatory objectives.\textsuperscript{26} A credible case may be made that such a “narrow banking” or “utility banking” approach is desirable for banks receiving the benefit of public insurance and lender of last resort support.\textsuperscript{27} These important complexity-related topics are beyond the scope of this Article, except to the extent that shedding light on inherent limitations of new governance techniques causes policymakers and regulators to question whether public goals of finance (e.g., credit formation, economic stability, monetary policy transmission, risk allocation) are better achieved by limiting complexity rather than accommodating it \textit{ex post} via new governance.

\section*{II. BANKS AND THE BASEL COMMITTEE}

Traditionally, the business of banking has consisted of performing the intermediary role of directing funds from entities with surplus (capital surplus savers) to entities in need of funds (capital deficit borrowers), pocketing a spread along the way. Fractional reserve banking permits banks to maintain only a fraction of their depositors’ funds as reserves on hand (i.e., cash in vault). They can then put the nonreserved funds to productive use by lending to businesses and consumers. Banks’ asset-liability mismatch results from their short-term deposit liabilities and their long-term loan assets that are vulnerable to interest rate risk, credit risk, call risk, operational risk, liquidity risk, foreign exchange risk, and macroeconomic risk.\textsuperscript{28} This fragility transforms into instability when a bank’s depositors begin to perceive, or even fear, that a bank is suffering losses in these longer-term assets. Thus, a positive feedback mechanism may be generated whereby investors withdrawing money on Monday cause more investors to withdraw money on Tuesday, and so forth. Often, due to the opacity of banks’ asset portfolios, even the most sophisticated counterparties have little more information on a bank’s solvency than the

\textsuperscript{26} See supra note 23 and accompanying text.


bank’s retail depositors.29 A run on deposits therefore can quickly transform into, and be stoked by, counterparty redemptions and collateral calls.30 While all financial institutions may pose systemic risks, bank systemic risk is particularly troubling because of the critical functions of banks in the payments system and the transmission of government monetary, credit, and exchange rate policies.31

In the case of banks, central banking and deposit insurance are designed to shore up confidence in the system and mitigate its inherent instability by, respectively, providing the option of emergency central bank liquidity (in the United States, via the discount window) to institutions suffering a sudden increase in deposit withdrawals and shifting the ultimate risk for bank failures from depositors to the entire industry (or taxpayers, if industry levies prove insufficient).32 In this context, it is important to recall that, as discussed above in Part I, the introduction of deposit insurance attenuates the force of market discipline by insulating banks from price and quantity reactions by depositors to alterations in a bank’s capital levels.33 Regulatory

29. See TARULLO, supra note 14, at 19 (noting that because banks’ assets are “notoriously difficult” to evaluate, the “asymmetry of information between corporate insiders and lenders that exists in any situation is compounded in the case of banks”).

30. The bank run contagion can also spread among the bank industry itself, as the massive interbank lending market is based largely on confidence and lacks robust documentation and monitoring. See Jack Guttentag & Richard Herring, Emergency Liquidity Assistance for International Banks, in THREATS TO INTERNATIONAL FINANCIAL STABILITY 150, 160 (Richard Portes & Alexander K. Swoboda eds., 1987).


32. Government safety nets include not only deposit insurance and access to central bank liquidity, but also unconditional FedWire payment guarantees, as well as the entire regulatory apparatus that watches over the banking sector. See, e.g., 12 C.F.R. § 210.28(b) (2010) (implying that banks participating in FedWire have access to overdraft facilities at the Federal Reserve Bank where such banks maintain their respective account); Allen N. Berger et al., The Role of Capital in Financial Institutions, 19 J. BANKING & FIN. 395, 400 (1995).

33. See, e.g., Donald O. Cook & Lewis J. Spellman, Repudiation Risk and Restitution Costs: Toward Understanding Premiums on Insured Deposits, 26 J. MONEY, CREDIT & BANKING 439, 440 (1994) (cautioning that the insulation of banks from market discipline is incomplete to the extent depositors believe that the insurer might be incapable of satisfying claims). The market discipline is transferred to the insuring institutions, which often levy fees (or premiums) for the insurance. It is theoretically possible to mirror the market discipline by accurately pricing the deposit insurance according to the likelihood a bank will require its support. In practice, deposit insurance premiums are largely thought to be dramatically underpriced, likely on account of a race-to-the-bottom approach to avoid handicapping domestic banks vis-à-vis their foreign competitors. See Viral V. Acharya et al., International Alignment of Financial Sector Regulation, in RESTORING FINANCIAL STABILITY 365, 366–67 (detailing the problems of arbitrage and the resulting lobbying by financial industries). The Federal Deposit Insurance Corporation Improvement Act, for the first time, pegged deposit insurance premiums to a bank’s risk-based capital, but only roughly.
capital requirements therefore aim to approximate the market discipline black hole left by the introduction of moral hazard associated with the government’s safety net. Recent bailout interventions in favor of stockholders and bondholders by central banks and governments in connection with the subprime credit crisis obviously deepen moral hazard problems.

Over time, and coincident with the increases in the complexity of the banking business and a gradual decline in capital ratios, bank regulators increasingly have focused on capital (and later, on so-called “risk-based capital”) as the primary means of ensuring the soundness of the banking sector. Generally speaking, minimum bank capital requirements consist of two quantities and a relation between them: (1) a definition of “capital” that comprises those claims that are first in line to absorb future losses (the numerator); (2) a measure of the exposure to risk that capital is intended to cover (the denominator); and (3) a required minimum ratio of capital to that risk exposure.

34. Former Chairman of the Board of Governors of the Federal Reserve System Alan Greenspan explained the moral hazard problem with U.S. banks’ safety net as follows:

The safety net—deposit insurance, as well as the discount window—has so lowered the risks perceived by depositors as to make them relatively indifferent to the soundness of the depository recipients of their funds, except in unusual circumstances.

With depositors exercising insufficient discipline through the cost of deposits, the incentive of some banks’ owners to control risk-taking has been dulled. Profits associated with risk-taking accrue to owners, while losses in excess of bank capital that would otherwise fall on depositors are absorbed by the FDIC.


35. See infra Part II.B.3.a.

36. See M.K LEWIS & K.T. DAVIS, DOMESTIC & INTERNATIONAL BANKING 149 tbl.5.1 (1987) (noting that the ratio of equity capital to total assets for U.S. banks decreased from 18.3% in 1914 to 6.9% in 1983 and that the same ratio for U.K. banks decreased from 12% in 1900 to 4.6% in 1985).


38. Ideally, regulatory capital (the numerator of the capital ratio) should consist of financial instruments that conduct to the threefold aim of (1) providing a loss buffer for the governmental providers of the bank safety net, (2) serving as “patient money” that will not be redeemed during a financial crisis, and (3) helping to discipline bank risk-taking behavior. As this Article focuses on the use of internal models in calculating risk-weighted assets (the denominator of the capital ratio), this element of the ratio and the required numerical ratio will not be the subject of further discussion.

39. Arturo Estrella, A Prolegomenon to Future Capital Requirements, 1 FED. RES. BANK OF
Prior to the agreement on the Basel Accord in 1988 (the initial Basel Accord), bank regulators measured the exposure to risk as a bank’s total assets, much in the same manner we calculate the still-required U.S. leverage ratio today. Since the initial Basel Accord mandated risk-sensitivity in calculating exposure to risk, the denominator of the capital ratio has measured the banks’ risk exposure by reference to the Basel Accord’s regulatory risk weighting system. Thus, instead of measuring regulatory capital against total assets in a risk-agnostic way, it is measured against assets to which risk weights have been applied to reflect assets’ perceived riskiness in a risk-sensitive way. As described in further detail below, for the first time Basel II now permits regulated banks to set the risk exposure component of their regulatory capital ratios by reference to their own internally-generated risk models.

A. Regulating Bank Capital Adequacy in the United States: A Brief Pre-Baselite History

The United States was the first jurisdiction to impose generally applicable minimum capital requirements on its banks. From the enactment of the National Banking Act of 1864 until the Great Depression-era legislation, U.S. regulators evaluated capital formally only when a national bank applied for a charter. Following the Great Depression, bank authorities turned more attention to capital and solvency of banks during examinations, but still lacked enforcement tools and, more critically, express statutory authority to mandate minimum capital levels. Capital regulation underwent significant formalization in the early 1980s, as bank regulators instituted systematized capital levels, in large part due to unprecedented exposures to riskier sovereign debt. In 1983, Congress

42. The reasons leading up to this exposure are manifold and outside the scope of this Article. For the moment, it is sufficient to note that the burgeoning Eurodollar markets
granted regulators express statutory authority to regulate capital on an ongoing basis. Moreover, in January 1986, U.S. banking regulators announced a plan to unilaterally implement a risk-based capital adequacy regime that took into account the various risks to which banks’ capital bases were subject. By the mid-1980s, Congress had transformed the regulation of capital from a matter of ad hoc enforcement actions to “an ongoing feature of basic bank supervisory policy.” 43 By moving first, Congress saddled U.S. banks with a competitive disadvantage vis-à-vis their European and Japanese competitors. U.S. banking regulators had in effect imposed a risk tax on U.S.-based banks—which now had to withhold capital as a reserve for each new extension of a loan—without any restrictions keeping capital from migrating offshore. To address each of these problems, the U.S. banking policy establishment pushed international policy coordination under the auspices of the Bank of International Settlements (BIS), eventually resulting in the initial Basel Accord.

B. The Basel Committee and Capital Adequacy: From Basic Risk Sensitivity to an Internal Models Approach

1. Birth of an International Policy Coordinator

In December 1974, with world financial markets reeling following the collapse of two major international banks, 44 the central bank governors of the G10 agreed to establish the Basel Committee, consisting of central bankers and bank supervisors under the auspices of the BIS, to address the immediate problems arising in connection with the 1974 international financial crisis. In the longer term, there was an understanding that the Basel Committee would work to promote some modicum of convergence of

(funded by U.S. balance of payments deficits) and the increasing internationalization and complexity of financial markets in the 1960s and 1970s, particularly subsequent to the collapse of the Bretton Woods system of managed exchange rates in 1971 and 1973, resulted in banks holding unprecedented levels of currency and interest risks. Banks not only were exposed to new types of international risks, but the amounts subject to these globalized risks increased as well. During the same period, large international banks made significant investments into Eurodollar markets and U.S. banks expanded into foreign markets via Edge Act and Agreement corporations, as well as joint ventures and foreign direct investments. Among U.S. banks, foreign-based assets increased from 3% in 1960 to 19% in 1975.


banking supervisory practices. The Basel Committee’s initial focus on formalizing a coordination framework pursuant to which insolvent institutions could be wound down gave way to increasing attention to the prudential concern of protecting the solvency of internationally active banks and the stability of the international financial system (i.e., the prevention, rather than the resolution, of insolvencies).

In a 1981 report, the Basel Committee indicated its concern with the erosion of bank capital ratios and advocated a greater approximation in the levels of capital employed by large banks. The decade of the 1970s witnessed a reduction of capital-to-total-assets ratios of all major industrialized nations except for Switzerland and Italy. While acknowledging that any formal attempt at harmonization of the widely disparate systems of bank capital regulation was impractical, the Basel Committee nevertheless envisaged a role for itself in achieving “greater convergence among its members with regard to national definitions of bank capital for supervisory purposes.” Against this background, a second goal of the convergence effort became central: if policy coordination was to be achieved, how could the Basel Committee do so on a competition-neutral basis? As the United States had been ratcheting up and formally “legalizing” its minimum bank capital levels in the early 1980s, bank capital ratios in other G10 countries, most notably Japan, continued to plummet. From 1980 to 1985, U.S. banks saw their share of the international banking business plunge from 30% to 23%, while Japan saw a corresponding increase for its banks from 20% to 26%. This trend continued throughout the 1980s and was coincident with Japanese banks operating with much lower capital ratios than U.S. banks.

These trenchant conflicts threatened to derail progress on the Basel Committee’s second goal of leveling the competitive playing field, until U.S.

45. See Walker, supra note 44, at 51 n.150.
46. See Bryce Quillen, International Financial Co-Operation: Political Economies of Compliance with the 1988 Basel Accord 14 (2008). The decline in the 1970s was a micro-event in a larger trend of declining capital ratios at U.S. banks from as high as 40% in the 1850s to 6–8% from the 1940s to the present. See Berger et al., supra note 32, at 402 fig.1, 403. The marked decline in the 1940s is attributed in large part to the establishment of deposit insurance at the FDIC and the related risk-taking incentives created thereby. See André Lucas, Evaluating the Basle Guidelines for Backtesting Banks’ Internal Risk Management Models, 33 J. Money, Credit & Banking 826, 827 (2001) (“[T]he creation of the [FDIC] in 1933 resulted in a decrease in capital ratios of about 50 percent over a ten-year period.”).
47. See Quillen, supra note 46, at 15.
49. See Quillen, supra note 46, at 14 fig.2.1.
and U.K. banking authorities reached an agreement in principle in 1987 on capital adequacy convergence that incorporated risk sensitivity for the first time. Importantly, the United States and United Kingdom announced their intention to apply these standards to foreign banks already present or wishing to enter their markets.\textsuperscript{50} Recalcitrant Committee members, including Japan, agreed to move forward on the Basel project.\textsuperscript{51}

2. The Initial Basel Accord and the Introduction of Risk Sensitivity

In July 1988, the Basel Committee published the initial Basel Accord, which set a minimum capital-to-risk-weighted assets ratio of 8\%. Like the 1987 U.S.–U.K. accord, the initial Basel Accord determined which risk-weighting methodologies to apply to bank assets. Under the Accord, assets are categorized into five risk-weighted buckets according to their perceived credit risk. The initial Basel Accord’s focus on credit risk reflected the then-current perception that credit risk posed the greatest risk to bank solvency.\textsuperscript{52} Each risk category carries a specified risk weighting factor (0\%, 10\%, 20\%, 50\%, or 100\%).\textsuperscript{53} Low risk assets such as U.S. Treasury securities receive a 0\% risk weighting (that is, they are not added into the denominator of the capital-to-assets ratio), and other assets representing incrementally higher credit risk receive higher risk weights. For example, claims against U.S. banks and securities issued by U.S. state and local governments receive a 20\% risk weight; revenue bonds issued by U.S. state and local governments and first mortgage loans on one-to-four family residences receive a 50\% risk weight; and (1) property, plant, and equipment, and (2) loans to households and commercial borrowers receive a 100\% risk weight. Amounts at risk pursuant to off-balance sheet items such as contingent liabilities and derivative instruments\textsuperscript{54} are then

\textsuperscript{50} Oatley & Nabors, supra note 37, at 49.


\textsuperscript{52} Recent events might have called this assumption into question, as the post-Gramm Leach–Bliley expansion of universal banking and the associated trading and derivatives activities built up significant market risks. As is discussed infra, by the mid-1990s the Basel Committee had incorporated market risks into the Basel Accord.

\textsuperscript{53} See CARNELL ET AL., supra note 13, at 259–60 (describing the four risk weight factor percentages that U.S. federal bank regulators adopted in their implementation of the initial Basel Accord); Hal S. Scott, International Finance: Transactions, Policy and Regulation 332 (14th cd., 2007).

\textsuperscript{54} The initial Basel Accord recognized that the total exposure of a bank with respect
multiplied by a “credit conversion factor” that reflects the Basel Committee’s judgment of the likelihood that the off-balance sheet item will give rise to a balance sheet liability. The normal risk weights then apply to the resulting “credit equivalent amounts” to yield the required capital to be held against the off-balance sheet item. However, a 50% reduction applies to derivative instruments.

For all its achievements in policy convergence, the initial Basel Accord only took into account one component of credit risk—counterparty default risk—and left other significant risks such as interest rate risk, foreign exchange risk, operational risk, legal risk, and market risk, as well as nondefault-related credit risks such as downgrade or migration risk, spread risk, settlement risk, and credit concentration risk outside of its risk-based framework. Conversely, banks received no credit from sound risk mitigation techniques such as requiring collateralization or guarantees of loans or hedging exposures. The risk-weighting categories (or “crude risk buckets” also were necessarily overbroad and failed to account for the

55. Id. at 12–13, 19–25.
56. Id. at 25 (asserting that the reduction in capital required to be held against such derivatives is in recognition that “most counterparties in these markets, particularly for long-term contracts, tend to be first-class names”). This provision of the accord shows the degree to which the Basel Committee, even before the internal models approach of Basel II, was operating under the assumption that banks themselves (especially those “first class names”) could ensure stability; however, recent experience confirms that first class names often house third-rate operations.
57. For a helpful summary of these sub-components of credit risk, see Arne Sandstrom, Solvency: Models, Assessment & Regulation 82–3 (2006). Technically, the initial Basel Accord permitted supervisors to develop capital adequacy requirements that accounted for other types of risk, or to supplement the risk-weighting methodology with “other methods of capital measurement” (e.g., the FDIC’s prompt corrective action regime and minimum leverage ratio requirements).
58. See Carnell et al., supra note 13, at 272; U.S. Gov’t Accountability Office, Risk-Based Capital: Bank Regulators Need to Improve Transparency and Overcome Impediments to Finalizing the Proposed Basel II Framework 3–4 (2007) (“Basel I’s simple risk weighting approach does not adequately differentiate between assets that have different risk levels, offers only a limited recognition of credit risk mitigation techniques, and does not explicitly address all risks faced by banking organizations.”).
different risk profiles of assets within the same category.\textsuperscript{60} Furthermore, the credit conversion factors often operated in a manner that obscured the true economic risks of bank positions.\textsuperscript{61} Banks exploited the resulting opportunities for arbitrage. These lacunae were all the more glaring in that it was foreign exchange risk and interest rate risk, and not credit risk, that most directly led to the 1974 bank failures leading to the formation of the Basel Committee and the subsequent sovereign debt crisis of the mid-1980s.

In 1991, Congress enacted the Federal Deposit Insurance Corporation Improvement Act (FDICIA), which implemented the initial Basel Accord but also required U.S. federal banking agencies to revise their risk-based capital standards for insured depository institutions to ensure that these standards take account of interest rate risk, concentration of credit risk, and the risks of “nontraditional activities.”\textsuperscript{62} National treatment of these other risks in non-U.S. jurisdictions was either nonexistent or significantly more complicated than credit risk. Banks were therefore able to engage in regulatory capital arbitrage across asset classes and jurisdictions.\textsuperscript{63} Nevertheless, and despite its manifold and evident unresolved problems, the initial Basel Accord represented a significant achievement in the realm of financial policy coordination and has been adopted by countries outside the Basel Committee and become an almost global standard.\textsuperscript{64}

\textsuperscript{60} As an example, debt securities issued by the Organisation for Economic Co-operation and Development (OECD) members, Mexico (which joined in 1994 and devalued its currency later that year in the face of a debt crisis) and the Republic of Korea (which joined in 1996 and experienced a massive currency depreciation in 1997 before receiving a $58 billion loan from the IMF), carried a 20% risk weight prior to their respective crises—reflecting an embedded assumption that sovereign credits from these countries were five times safer than a loan to IBM, which as a corporate exposure carried a 100% risk weight.\textsuperscript{61}

\textsuperscript{61} See Robert C. Merton, \textit{Financial Innovation and the Management and Regulation of Financial Institutions}, 19 J. Banking & Fin. 461, 468–69 (1995) for an illustration about how structuring a portfolio yielding the return on a group of mortgages through a swap (rather than holding the underlying mortgages themselves) resulted in an eightfold decrease in the bank’s capital charge.\textsuperscript{62}


\textsuperscript{64} See WALKER, supra note 44, at 569.
3. **The Internal Models Approach: Delegating Assessments of Market Risks and Credit Risks to Regulated Banks**

In the 1990s, financial regulators faced an unprecedented array of emergent risks to bank solvency, including most prominently the proliferation of derivatives and the expansion of banks’ trading, capital markets, and off-balance sheet activities. The U.S. bank regulatory regime which, as noted in Part II.A, was the first to impose capital adequacy requirements on its banks, was under significant competitive pressures from other jurisdictions with less onerous regulatory strictures. It is important to note in this context that during the half century following World War II, the United States and Japan were the only jurisdictions that mandated “narrow banking,” which prevented banks affiliating with insurance companies, securities firms, and hedge funds.65 Because the initial Basel Accord burdened banks from non-U.S. jurisdictions with a competitive disadvantage vis-à-vis their domestic nonbank competitors, those banks expanded their activities to search for returns not subject to regulatory capital taxes. In the United States, the Glass–Steagall Act’s (GSA) walling-off of the banking business from nonbanking activities constituted a serious competitive impediment for U.S. banks in a globalized financial market. In 1991, the Treasury Department unsuccessfully sought the repeal of GSA.66 Notwithstanding this initial congressional rebuff, the U.S. bank regulatory authorities had attenuated much of GSA’s effects and embraced the bank-as-financial-bazaar model by the time Congress abrogated GSA in 1999 with the Gramm–Leach–Bliley Act.

a. **Banking in the 1990s: Competition, Globalization, Technology, Complexity, and Code**

The late 1980s and early 1990s witnessed an accelerated shift of large, internationally active banks away from their traditional business of direct financial intermediation—that is, providing deposit-based financing to borrowers in exchange for loans to be held on the banks’ balance sheets—to (1) assuming a more indirect role of earning fee-based income as an underwriter-facilitator in capital markets activities; (2) acting as a dealer of, and deploying their own capital in, off-balance-sheet derivatives transactions; and (3) expanding their trading activities and their risk

---


tolerance for even traditional lending activities. Bankers and regulators referred to this sectoral dynamic as a shift to “noninterest” income to distinguish it from the traditional interest income earned on loans. These trends resulted from declining profitability of the traditional banking business due primarily to competitive and regulatory developments that undercut banks’ traditional advantages in their returns on assets (asset side of balance sheet) and their cost of capital (liabilities side of balance sheet). These funding pressures, deregulatory initiatives, globalization forces, and nonbank financial product innovations squeezed the competitive position of banks severely. Bank regulators permitted the expansion of banks into these new activities because, in the opinion of most bank regulators, “permitting wider activities is necessary to ensure that such organizations can remain competitive.” It was better to double down on risk than risk obsolescence. In the process, banking operations assumed increasing complexity that resulted in earnings and capital volatility, which in turn rendered the Basel Accord’s formulaic approach to credit risk an awkward fit for banks’ dynamic risk profiles.


69. While competitive pressures constituted, in my view, the proximate cause of the transformation of banking groups throughout the decade, other factors contributed to the scale and speed of the changes. For instance, as discussed below, technology fueled financial innovation. More indirectly, the ascendancy of finance theory in the late twentieth century (e.g., option pricing models, portfolio theory) focused on the varieties of finance as functions, with little regard for the institutions performing the functions. Accordingly, as finance theory was increasingly applied to practice, the traditional institutional demarcations between, on the one hand, banks and, on the other, insurance companies and securities firms, appeared (at least from a theoretical perspective) as anachronisms. Cf. Robert C. Merton, Future Possibilities in Finance Theory and Finance Practice 5–6, (Harv. Bus. Sch., Working Paper No. 01-030, 2000), available at http://www.signallake.com/innovation/FuturePossibilities01.030.pdf (explaining how a functional perspective on financial innovation is preferable to an institutional perspective because institutions often use finance theory to change their traditional product mix to better accommodate the functional demands of users of financial services).

The New Financial Code

During the 1990s and into the 2000s, new risk management technologies enhanced dramatically the perceived ability of large financial institutions to identify and quantify risks, leading to increased confidence in investment and hedging strategies that utilized data from those technologies. A new risk management industry hailed its ability to imbue advances in theoretical finance with mathematical precision through the use of new computer systems and software technologies that collect, organize, and analyze data in a systematized manner.71 A new cadre of risk management professionals emerged to administer these technologies, most of whom had scientific or mathematical training to complement their finance backgrounds. Risk managers aimed to measure, with technical precision, the “value at risk” (VaR) associated with certain products and events. Banks and securities firms generated proprietary intellectual property or purchased software or consultant services from third-party vendors to manage risk and promote new product offerings. This new financial code underlying the new banking business model developed outside the view of bank regulators, and formed part of a broader trend where information technology and information processing become an increasingly important source of competitive advantage and, ultimately, wealth.72

These new technologies facilitated three critical developments in late twentieth century finance: (1) VaR methodologies expanded from their initial role of gauging correlation of debt and equity market risks to measuring the impact of credit risks associated with longer-maturity assets such as loans; (2) credit assessment technologies were a contributing cause of banks’ loss of their traditional advantage in processing credit information, but banks were able to channel technologies to their advantage in developing new products that banks were well positioned to develop with their traditional clients, such as over-the-counter (OTC) derivatives; and (3) large banks employed computer technologies to develop sophisticated economic capital (as distinguished from purely regulatory capital) models that sought to ensure that firms had sufficient capital to meet the expectations of the capital market, and subjected these models to “stress testing” and “scenario analysis” to evaluate how a firm would be

71. For a brief discussion of the “systemic coding” of digital information by businesses, see Kenneth A. Bamberger, supra note 4, at 682–83.

affected by hypothetical sets of seriously adverse events.\textsuperscript{73}

First, the development of VaR models purported to enable banks to form a more complete understanding of potential loss associated with any given position. VaR attempts to measure the maximum possible loss of a portfolio over a given time horizon at a specified confidence level.\textsuperscript{74} In other words, VaR models yield a numerical assessment of the maximum loss a position might incur over a specified time period (the time horizon) and with a specified probability tolerance (the confidence interval). Because VaR is an expression of probability and maximum loss, it does not purport to assess literally the exposure at risk associated with a position, which in most circumstances would be 100% of the asset’s value. Moreover, the assumed distributions of VaR techniques are where the techniques become the most useful. Since there is no a priori reason for assuming a normal distribution in the context of financial outcomes, pioneers of the VaR techniques employed powerful new computer software to derive statistical model distributions based on which future returns could be predicted (within a given confidence interval).\textsuperscript{75} J.P. Morgan was largely responsible for the development of VaR technologies in 1994 and began to sell software including its proprietary VaR methods in 1996. The first wave of VaR techniques only quantified VaR due to market risks. It took little time, however, before risk managers appreciated the broader applicability of the putative ability of VaR techniques to quantify “maximum” possible loss amounts associated with particular risks, some of them non-market-related.\textsuperscript{76} By the late 1990s the most sophisticated banks had begun to use VaR to classify loan assets into credit risk classes and to use those classifications in allocating capital.\textsuperscript{77}

\textsuperscript{73} See James A. Fant, Anticipating the Unthinkable: The Adequacy of Risk Management in Finance and Environmental Studies, 44 WAKE FOREST L. REV. 731, 737–38 (2009).

\textsuperscript{74} See Philippe Jorion, Risk?: Measuring the Risk in Value at Risk, 52 FIN. ANALYSTS J. 47, 47 (1996); Lucas, supra note 46, at 829.

\textsuperscript{75} Typical modeling techniques include pure historical data approaches, which use historical data alone to determine loss distribution, and so-called “Monte Carlo” simulations, which use historical data to estimate loss but add the use of random sampling driven by advanced computing power.

\textsuperscript{76} As will be discussed in greater detail in Part V.C.2, some commentators view the VaR construct to be inherently suspect. See, e.g., Nassim Taleb, Against Value at Risk; Nassim Taleb Replies to Philippe Jorion, FOOLEDBYRANDOMNESS.COM (1997) (expressing skepticism in “[t]he act of reducing risk to one simple quantitative measure on grounds that ‘everyone can understand’ it”).

Second, the emergence of risk management and credit assessment computer and software technologies permitted nonbank lenders to intermediate in loan markets with increasing frequency. Those technologies also enabled banks to transition into new, often riskier, financial products that they began to offer more frequently in the 1990s. New code interacts dynamically with riskier business models; the former gives rise to the latter, which in turn encourages the assumption of additional risks on the grounds that those new risks are limited by the same risk management technologies. Furthermore, as risk management technology advances, it reduces costs for future innovations, propelling the innovation process.  

Banks also harnessed this new code to turbo-charge the fee-based securitization and derivatives business and migrate away from traditional forms of bank intermediation. Using new computer models to model cash flows, credit risk, and quality of credit support, originators could now transfer a pool of credit assets—such as credit card receivables, mortgages, and leases—without regard for the credit risk posed by any individual debtor. These new technologies transformed fundamentally the relationship between debtors and creditors by giving rating agencies, issuers, underwriters, and purchasers access to aggregated data predicting the likelihood of default on obligations constituting the securitized pool of assets. Rating agencies developed their own technologies to assess the credit quality of a securitized pool before stamping their credit rating imprimatur on the various tranches of a securitization transaction. Because a priori models predicted the performance of the pool itself, it became less important for securitization professionals, including bankers, to employ a banker’s traditional skill of assessing whether a specific debtor would be able to repay a loan. A debtor default would register only as a loss to the lowest tranche of securitized debt. The importance of computer technologies was not limited to VaR modeling techniques; originators also utilized marketing software, data mining programs, and advanced credit reporting to select creditworthy borrowers and set appropriate interest rates. For example, a credit card executive boasted that since the mid-1990s the sector has turned the analysis of consumers into a science rivaling the studies of DNA or the launching of the Discovery spaceship into orbit. The

---

78. See Merton, supra note 61, at 463 (“The total reduction in costs has the effect of reducing the threshold of benefit needed to cover the cost of a new innovation.”).
80. See Gerding, supra note 6, at 144.
mathematics of virtually everything consumers do is stored, updated, categorized, churned, scored, tested, valued, and compared from every possible angle in hundreds of the most powerful computers and by among the most creative minds anywhere.\textsuperscript{81}

In addition to growth in credit card and mortgage lending (and related securitizations), modeling advances contributed in large part to the burgeoning derivatives market throughout the decade.\textsuperscript{82} As a result of the reliance on computer modeling and aggregated data, an emergent preference for quantitative, standardized information over qualitative information—a “mathematization” of finance—took hold among financial institutions.\textsuperscript{83} A central premise of this Article is that this process of mathematization is better thought of as a concomitant to the complexity that took hold in banking business models during this period, rather than as an absolute enhancement in the accuracy of evaluating risk.

Third, large internationally active banks developed their own “economic capital models,” entirely separate from the regulatory capital framework, that computed the amount of capital needed to protect the bank from unexpected future losses within a given confidence level. Most typically, economic capital models are set up to ensure the firm will remain a going concern able to attract counterparties\textsuperscript{84} with a very high degree of probability. Economic capital modeling utilizes advanced market and credit VaR methodologies with long time horizons to assess how much capital a firm is putting at risk given its current portfolio and future business plans. Because the economic capital modeling techniques look forward to evaluate the probability of losses, risk managers submit the models to stress testing and scenario analysis that test how a bank would respond to future extreme adverse events, or sometimes a series of adverse events in so-called “stochastic” tests. These models allow bank management to identify


\textsuperscript{82} See Merton, \textit{ supra} note 61, at 463 (writing in 1995 that “[t]he rapid five-year growth in [OTC] derivatives [which are transacted away from a central market putting greater pressure on the underlying institution’s capability to price those derivatives and manage their risk] reflects a growing confidence in the issuing institutions’ modeling and evaluation skills”); \textit{cf.} Arthur E. Wilmarth, Jr., \textit{The Dark Side of Universal Banking: Financial Conglomerates and the Origins of the Subprime Financial Crisis}, 41 CONN. L. REV. 963, 1011 (2009) (highlighting internet-enabled mass marketing, computerized credit scoring programs, and automated loan processing as contributing to massive expansion in lending during 1990s).

\textsuperscript{83} \textit{See} Cannata & Quagliariello, \textit{ supra} note 63, at 12.

\textsuperscript{84} While minimum capital requirements are predominantly designed to limit the risk of a catastrophic insolvency resulting in public losses, they are agnostic as to a firm’s going concern value, which often depends on a firm maintaining capital ratios well in excess of minimum levels.
2010] Internal Models Approach to Capital Adequacy Regulation

811

centrations of risk and opportunities for diversification within a
systematized framework, and therefore represent a “more sophisticated, less
arbitrary alternative to traditional lending limit controls” that banks
otherwise use to limit risk. Economic capital could also focus on matters
of shareholder concern—such as market capitalization at risk and going
capital value—rather than simply focusing on book value as regulatory
capital requirements do. Moreover, banks use the economic capital
construct to ensure they maintain capital sufficient to maintain their debt
ratings as well as allocate capital across a company group.

ii. Regulatory and Policymaker Reactions to Financial Complexity: In Their
Own Words

The complexity paradigm and the ascendency of the new financial code
did not occur outside the supervision of bank regulators in the United
States and abroad who enabled and encouraged the sectoral shift into new
lines of business with higher risks and returns. Bank regulators authorized
the new business model for two related reasons. First, they believed that a
globalized market for capital and financial services made the traditional
banking model of holding long-term loan assets and short-term low-yielding
deposits untenable. In the process, external regulation by means of
“lay[ing] down, ex cathedra, common rules and ratios that all banks should
follow . . . [was] becoming both less effective and less feasible.”
Regulators perceived that the innovation of the 1990s was different in scale,
if not scope, than preceding bursts of financial innovation, but feared that
further regulation to curb risk would temper financial innovation and slow
economic growth. This concern was misplaced from the outset, but it is

85. See Michael B. Gordy, A Risk-Factor Model Foundation for Ratings-Based Bank Capital
86. Bank managers are responsible to a class of constituencies that overlaps, but is not
coincident, with the class of constituencies that bank regulators protect. Therefore, the
notion of VaR raises the question: whose value is at risk? For instance, the market value of a
banking group’s equity is expected to be much higher on the priority of a bank manager
than a bank regulator, who will be more concerned with the balance sheet liquidation value
that ensures safety and soundness of the banking system and the protection of a particular
class of creditors (that is, depositors). See Francesco Saita, Value at Risk and Bank
87. See id. at 20; Gordy, supra note 85, at 200. Decisions regarding distributions and
dividends are particularly sensitive in the context of operating subsidiaries with minority
shareholders.
89. In the words of former Federal Reserve Governor Laurence H. Meyer, “[t]he
growing scale and complexity of our largest banking organizations . . . raises as never before
the potential for systemic risk from a significant disruption in, let alone failure of, one of
evident in the remarks of bank regulators throughout the 1990s; it also evinces the extent to which the economy had come to depend on financial sector profitability.  

Second, they possessed a cautious, but untested, confidence in the predictive capability of banks’ proprietary risk management functions and internal risk and capital models, and the theoretical finance underlying those models. Faced with a banking sector in secular decline, they authorized banks to enter new markets and product lines with little regard for their effects on safety and soundness, and began to look to the risk management and modeling techniques that enabled much of the new banking industry to perform the task that regulators themselves had by the mid-1990s become incapable of performing: monitoring and understanding the safety and soundness of bank activities. In the capital adequacy context, regulators quickly perceived the poor fit between the initial Basel Accord’s formulaic “risk buckets” approach to measuring credit risk and the more dynamic risk profiles of actual banks in the 1990s. The regulators, overwhelmed by the complexity of a globalized, deregulated bank industry subjected to cross-sectoral competition and rapid technological developments, eventually decided to link regulatory capital to firms’ internal capital models.

(a) First Bank Regulatory Premise: Complexity Adds Risk, But Is Needed to Compete Effectively

Bank regulators recognized the risks inherent in new business models, but opted to step aside rather than risk “inefficient” regulation that could put banks at a competitive disadvantage. E. Gerald Corrigan, then the Chairman of the Basel Committee and President of the Federal Reserve Bank of New York, warned in 1992 that “[t]he speed, volume, value, and complexity of international banking transactions have introduced new linkages and interdependencies between markets and institutions that have the potential to transmit problems and disruptions from place to place and


90. For a discussion on the predominance of financial sector growth and profitability in developed economies, see supra note 10 and accompanying text.
institution to institution at almost breakneck speed.” In addition, Corrigan registered the Basel Committee’s concern that “banking groups by themselves, but especially in combination with insurance and securities firms, are becoming very complex organizations” and that “the world of banking and finance has become very complex and perhaps more risky as technology, competition, and deregulation irreversibly alter the framework within which financial institutions and their supervisors must function.”

The remarks also evince regulators’ growing fascination (which would become increasingly important throughout the decade) with the “management information systems that will provide the top management of financial institutions with the tools and the information to ensure that applied technology is being used in a safe, sound, and prudent manner.”

Governor Laurence H. Meyer stated in 2000 that “[f]or the past decade or so supervisors have recognized that snapshots of the balance sheets of complex banking organizations are not very helpful for supervisory evaluations. Positions just change too rapidly.”

U.S. lawmakers attempted several times to legislate fulsome derivatives regulation in the early 1990s. In 1994 and 1995, members of the U.S.

91. E. Gerald Corrigan, Chairman, Basel Comm., Remarks Before the Seventh International Conference of Banking Supervisors in Cannes, France: Challenges Facing the International Community of Bank Supervisors (Oct. 8, 1992), in FED. RES. Q. REV., Autumn 1992, at 6. In his initial remarks, Chairman Corrigan characterized the challenges of bank regulators in 1992 as “the most demanding and vexing in the post-World War II period.” Id. at 1. Within five years, the industry would be embroiled in the Mexican “peso crisis,” the Asian currency crises, and near collapses of institutions due to failure to monitor derivatives transactions (as with Barings Bank). Corrigan, then, was stating things too equivocally: in fact, the structural dilemma of post-Bretton Woods bank supervision seems to be that at any given point bank supervisors can state unequivocally that the current state of the banking industry is the most “demanding and vexing” from a systemic risk standpoint since World War II. For example, consider the remarks of then Secretary of the Treasury Lloyd Bentsen in testimony before the Senate Committee on Banking, Housing and Urban Affairs in 1994 following the collapse of the thrift industry: “Our country has just emerged from its worst financial crisis since The Great Depression. One of the lessons of that crisis is that our bank regulatory system is cumbersome and antiquated.” Banking Industry Regulatory Consolidation: Hearing Before the S. Comm. on Banking, Hous., & Urban Affairs, 103d Cong. 48 (1994) (statement of Lloyd Bentsen, Sec’y, Department of Treasury).

92. See Corrigan, supra note 91, at 4. Corrigan undid his qualification that the banking industry is “perhaps” more risky later in his remarks: “[W]hile we and others can engage in a lively debate about whether international banking in the nineties is likely to be more or less risky than it was in the past, I believe we would all be well served to operate on the assumption that systemic risk may be greater as we look ahead.” Id. at 6.

93. Id. at 9.


95. The FDIC Improvement Act of 1991 included provisions to improve the enforceability of netting contracts, which reduce the legal risks stemming from the failure of
Congress proposed several bills to introduce sweeping regulatory reform in the area of derivatives. These bills called for varying degrees of regulatory oversight of the derivatives activities of banks, ranging from disclosure requirements to full-blown capital requirements and an outright ban on derivatives transactions for FDIC-Insured depository institutions. Representative Henry Gonzalez of Texas, a co-sponsor of the bill that received the most attention—the Derivatives Safety and Soundness Supervision Act of 1994—warned that “growing bank involvement in derivative products is . . . like a tinderbox waiting to explode. In the case of many market innovations, regulation lags behind until the crisis comes, as it has happened in our case with [savings and loans] and banks.” By the time that bill was re-introduced, Representative Gonzalez lamented the “bank regulatory agencies[?] claims that legislation was not necessary.” Senator Byron Dorgan introduced a radical bill that aimed to achieve a simple result: prohibiting FDIC-Insured banks from engaging in derivatives activities. Senator Dorgan’s remarks when introducing the bill indicate his preference for a narrow banking industry:

What investors do with their own money is their own business. But what they do with money insured by the American taxpayers, is the business of Congress. The purpose of deposit insurance is to encourage saving. It is to promote a pool of capital that is available to build homes and businesses and jobs. Deposit insurance is not supposed to underwrite rampant speculation on Wall Street, and my bill will help prevent that from happening.

Notably, each of these legislative proposals emphasized to varying degrees the importance of management involvement in and oversight of banks’ derivatives activities. The Derivatives Supervision Act of 1994 went further, seeking to codify mandatory internal controls and procedures for derivatives programs. Congress never enacted any of these bills, in large part due to the Federal Reserve’s steadfast belief that “[w]e must not lose sight of the fact that risks in financial markets are regulated by private firms active in derivatives. That legislation also required regulators to increase capital standards for institutions with significant interest rate risk associated with derivatives or other instruments, and it required banks to limit their interbank credit exposures from derivatives and other sources.

parties.”  

(b) Second Bank Regulatory Premise: The New Financial Code Provides an Effective Check on Risks to Safety and Soundness from New Bank Activities

Federal bank regulators sought to harness banks’ internal risk models to promote their public regulatory goals of safety and soundness. For instance, 1995 interagency guidelines required banks to establish internal controls and information systems that provide for effective risk management and adequate procedures to safeguard assets. Regulators perceived risk management technologies as a tool to minimize the burgeoning special risks posed by derivatives. In 1993 the Office of the Comptroller of Currency (OCC) issued Circular No. 277, titled “Risk Management of Financial Derivatives.” Circular No. 277 set forth extensive risk management guidelines governing derivatives activities conducted by national banks, but limited its new capital adequacy discussion to a general admonition that banks’ boards of directors “should ensure that the bank maintains sufficient capital to support the risk exposures” from their banks’ derivatives activities. Similarly, the Federal Reserve issued trading- and derivatives-related risk management guidance in the mid-1990s, but without requiring additional capital or limiting the extent or magnitude of derivatives activities. In 1998 when the Commodity Futures Trading Commission (CFTC) published a request for


103. The remainder of that document’s discussion of capital adequacy was limited to a reminder that banks already had existing capital requirements in place arising under the Basel Accord’s treatment of off-balance sheet items. See id. at 24–25.

comments on the possible regulation of OTC derivatives by the CFTC—including supplementary capital requirements on OTC derivatives dealers—\textsuperscript{105} the reaction from the legislative and executive branches was swift. The Federal Reserve, the SEC, and the Clinton Treasury Department—which housed the OCC and the OTS—resisted, and Congress passed the Commodity Futures Modernization Act of 2000,\textsuperscript{106} which exempted nearly all OTC derivatives from CFTC and SEC oversight.\textsuperscript{107}

The regulatory focus on internal risk management techniques was hardly a U.S.-specific concern. From 1987 to 1992, bank regulators in the United Kingdom, France, Germany, Singapore, and Switzerland issued rules or guidance concerning risk management and internal control functions as a means of curbing derivatives risk at banks.\textsuperscript{108} The Basel Committee published a document titled “Risk Management Guidelines for Derivatives” in July 1994.\textsuperscript{109} These guidelines followed a 1993 Group of Thirty study\textsuperscript{110} addressing emergent risks attending mounting notional values of derivatives outstanding. The Basel Committee noted the increasing prominence derivatives had assumed to the overall risk profile and profitability of banks,\textsuperscript{111} and observed that the “growing complexity, diversity and volume of derivatives products, facilitated by rapid advances in technology and communications, pose increasing challenges to managing [credit, market, liquidity, operational and legal] risks.”\textsuperscript{112} The Basel Committee also welcomed the “increasing use of sophisticated models by major institutions as their principal means of measuring and managing risk.”\textsuperscript{113} The report anticipated the likelihood that the internal models would ultimately provide the most granular information concerning the risk

\begin{itemize}
\item \textsuperscript{105} See Over-the-Counter Derivatives, 63 Fed. Reg. 26,114, 26,114–27 (May 12, 1998).
\item \textsuperscript{107} See 144 Cong. Rec. E1505 (daily ed. July 31, 1998) (statement of Rep. James Leach) (stating, regarding the 1998 CFTC proposal, “three of the four government agencies which have responsibility for overseeing the derivatives market place—the Federal Reserve Board, the Treasury Department, the [SEC]—have come to the conclusion that the other principal regulator, the [CFTC], has embarked on a regulatory path at odds with the U.S. national interest”); David Barboza & Jeff Gerth, Who’s in Charge? Agency Infighting and Regulatory Uncertainty, N.Y. Times, Dec. 15, 1998, at C14.
\item \textsuperscript{108} U.S. GOV’T ACCOUNTABILITY OFFICE, FINANCIAL DERIVATIVES: ACTIONS NEEDED TO PROTECT THE FINANCIAL SYSTEM 114 (1994).
\item \textsuperscript{109} See BANK FOR INT’L SETTLEMENTS, supra note 104, at 1.
\item \textsuperscript{110} GROUP OF THIRTY, DERIVATIVES: PRACTICES AND PRINCIPLES (1993).
\item \textsuperscript{111} Id. at 2.
\item \textsuperscript{112} BANK FOR INT’L SETTLEMENTS, supra note 104, at 1.
\item \textsuperscript{113} Id. at 2.
\end{itemize}
of derivative products, and that it would therefore be important for supervisors “to assure that they (and external auditors) have staff with sufficient mathematical knowledge to understand the issues” associated with these models.114

The European Community’s 1993 Capital Adequacy Directive mandated that banks maintain capital to cover unexpected losses due to interest rate risks.115 In its recitals, the directive registered the importance of “the existence, in all institutions, of internal systems for monitoring and controlling interest-rate risks on all of their business” and noted that “such systems must be subject to overview by the competent authorities.”116 The directive also for the first time set minimum capital requirements in respect of market risk exposures relating to derivatives contracts.

Chairman Alan Greenspan’s 1994 testimony before the House of Representatives Subcommittee on Telecommunications and Finance concerning the 1994 GAO report on derivatives, and in particular its recommendations for increased regulation of banks’ derivatives and proprietary trading activities, provides a lens into the Federal Reserve’s intellectual framework during the 1990s and its concern over the inability of extant capital adequacy regulation to accommodate an increasingly complex and dynamic banking industry. Throughout his testimony, Greenspan remarked on the increasing complexity of financial markets and evidenced a profound confidence that enhanced risk management techniques and models developed by banks operated as effective checks on excessive risk-taking. His comments on market risk capital requirement proposals illustrated the connection between growing complexity and the intellectual reliance on internal capital models:

Although the market risks of many banking instruments, including many derivative contracts, can be accurately assessed using . . . simple models [such as the initial Basel Accord formulae to calculate credit risk capital requirements], a considerably more sophisticated approach is necessary to assess more complex instruments, especially those with options characteristics, and to aggregate different categories of market risk. The recognition of the need for a more sophisticated approach has led banking regulators in the United States and abroad to explore carefully the potential for allowing banks to use their own internal models to assess the need for capital to cover market risk.117

To summarize, complexity begets a need for a “more sophisticated

114. Id.
116. Id. at 2.
approach,” which “has led banking regulators” to explore the adoption of internal models.\textsuperscript{118} The expansion of derivatives markets was attributable to advances in risk management technology itself.\textsuperscript{119} Greenspan stated further that he expected that banks would achieve still further progress in risk management of derivatives activities and that regulators would need to incorporate consideration of those internally-generated, firm specific techniques and models into the supervisory process.\textsuperscript{120}

In other 1994 comments, Greenspan echoed his sentiments expressed throughout the early 1990s. He noted that the “evolving financial firm” was

becoming so complex that it not only challenges our ability to write laws and regulations, but—more important—is leading to overly complex rules and regulations that challenge the ability of managers to manage. At least part of the solution to the increasing complexity in bank risk positions may be to rely less on the writing of complicated and highly specific rules that apply to all banks, and to concentrate more on the development of common conceptual frameworks and flexible supervisory procedures that can accurately distinguish risks on a bank-by-bank basis.\textsuperscript{121}

While “traditional” capital adequacy rules have “served us well over the decades . . . as the complexity, if not the dimensions, of bank risk-taking has increased, the regulatory capital standards also have evolved and become more complex.”\textsuperscript{122} To continue ratcheting up the complexity of capital

\begin{itemize}
\item \textsuperscript{118} Id.
\item \textsuperscript{119} See id. at 7 (“It is important to recognize that significant advances in the management of market and credit risks, including improvements both in financial methodology and in the design of management information systems, lie behind the recent surge in derivatives activity. These advances have made independent, highly skilled risk management staffs and rigorous measurement and analysis of market and credit risks key elements of a sound risk management approach for trading activities, and more generally, for banking activities.”).
\item \textsuperscript{120} See id. at 7–8, 23; Alan Greenspan, Chairman, Bd. of Governors of the Fed, Reserve Sys., Remarks Before the 30th Annual Conference on Bank Structure and Competition, Chicago 11–12 (May 12, 1994), available at http://fraser.stlouisfed.org/historicaldocs/ag94/download/27980/Greenspan_19940512.pdf (“No matter how good we become at bank supervision, however, we should always keep in mind that the first line of supervisory defense must be the quality of the risk management systems used by banks themselves . . . . [W]e have recognized for some time that capital rules are often less meaningful than the sophisticated internal models used by some banks to test the sensitivity of their net worth to possible future changes in asset prices.”) [hereinafter Greenspan, Chicago Remarks].
\item \textsuperscript{122} Alan Greenspan, Chairman, Bd. of Governors of the Fed, Reserve Sys., Remarks
adequacy regulation in lockstep with bank risk-taking would “create[,] more problems than [it] would solve.” In fact, Greenspan continued, “[i]ntricate capital rules run the very real risk of causing inefficiencies resulting from complex bank strategies to avoid binding capital constraints and, at worst, may lead to less measurable and possibly greater bank exposure to losses beyond capital.” Instead, “[a]t least part of the solution to the problem of complexity in risk behavior is to rely less on the writing of rules, such as capital regulations, that apply uniformly to all banks . . .” While this position stopped short of expressly calling for calibrating regulatory capital to internal models, the upshot was clear: in light of the epistemic gap between what regulators could know and what they aimed to know, regulators needed to focus on institution-specific risk characteristics. In this context, Greenspan’s encouragement of the development of banks’ internal credit rating technologies and the rapid growth and use of the new financial code in broadening the scope of financial services is instructive. In short, “[t]hat is why the supervisory effort is increasingly focusing on the evaluation of risk management systems.”

At the 1995 G7 summit in Halifax, the G7 issued a communiqué highlighting the importance of policy coordination in addressing challenges

---

123. Id. at 9.
124. Id.; see also Greenspan, Chicago Remarks, supra note 120, at 9 (“Greatly increasing the complexity of capital regulations can only lead to inefficiency as I see it.”)
125. Greenspan, University of Utah Remarks, supra note 122, at 9 (emphasis added); Greenspan, September 1994 Senate Testimony, supra note 68, at 16 (“As we proceed through the 1990s, we should focus on enhancing supervisory practices, rather than on developing new laws and regulations.”).
126. See Greenspan, Boston College Remarks, supra note 121, at 1; Greenspan, University of Utah Remarks, supra note 122, at 3–6; Alan Greenspan, Chairman, Bd. of Governors of the Fed. Reserve Sys., Remarks Before the American Bankers Association 2–9 (Oct. 8, 1994), available at http://fraser.stlouisfed.org/historical/docs/ag94/download/27990/Greenspan_19941008.pdf [hereinafter Greenspan, ABA Remarks]; Greenspan, Chicago Remarks, supra note 120, at 4 (“A crucial difference between the banks of today and those of our traditions . . . is that risk information processing now lies more visibly closer to the core of banking business because of the blossoming of new financial products and services that rely so critically on fast and high quality risk information and risk analysis.”). Greenspan also cautioned that, in his opinion, regulators and Congress should not impede “technological change . . . because it seems so clear that bankers face significant ‘new horizons’ in the lending process, regulatory agencies must be especially careful not to place obstacles in the path of beneficial technological change.” Greenspan, ABA Remarks, supra, at 9.
relating to trade imbalances and currency and financial market instability. The Halifax communiqué is probably most noteworthy for its proposal to create an “Emergency Financing Mechanism” to serve as an enhanced in extremis financing facility under the auspices of the International Monetary Fund (IMF). However, the document also touched on the “growth and integration of global capital markets,” which had “created both enormous opportunities and new risks.” The group emphasized that the “accelerating pace of financial innovation,” combined with “the risks inherent in the growth of private capital flows [and] the increased integration of domestic capital markets,” posed significant systemic risks. To that end, the G7 envisaged a role for private industry participants to bridge the information gap between financial actors and their regulators: “We urge . . . strengthened policy advice from international financial institutions on the appropriate supervisory structures.”

In response to the G7’s 1995 Halifax Summit report, the Basel Committee and the International Organization of Securities Commissioners (IOSCO) issued a joint statement in 1996 noting that a tension existed in the “exponential rate of technological and financial innovation, including notably the increased use of derivative products.” On the one hand, innovation can increase systemic risk to the extent it is not captured in the ex ante regulatory-governance architecture; on the other hand, complexity often opens up possibilities for “significant enhancements to risk management procedures.” A year later at the G7 summit in Lyon, the G7 finance ministers emphasized the importance of “market-reinforced prudential supervision” and applauded the Basel Committee’s initiative to address market risk in the capital adequacy framework by reference to internal VaR models. In light of the increasing complexity of their business models, “primary responsibility for risk management rests

129. Id. at ¶ 14.
130. Id.
131. Id. at ¶ 22.
133. Id. Regarding this latter element, because derivatives facilitate the specific identification and management of risks, they offer the theoretical possibility of encouraging the safety and soundness of financial institutions, as well as the efficient allocation of risks. See Risk Management Guidelines, supra note 104, at 3.
with market participants.”

In 1996 the Group of Thirty\(^\text{136}\) levied a much bleaker assessment of the status of mid-1990s financial regulation, noting that “the global operations of major financial institutions and markets have outgrown the national accounting, legal and supervisory systems on which the safety and soundness of individual institutions and the financial system rely.”\(^\text{137}\) Prior to issuing its report, the Group of Thirty surveyed sixty-six large complex financial institutions, many of which expressed concerns about the quality of risk management knowledge systems.\(^\text{138}\) The Group of Thirty focused on the build-up of systemic risk that the Basel Committee and IOSCO addressed in their roughly contemporaneous joint report: “the increasing size, velocity, and complexity of international transactions, and the increasing concentration of trading activity in a relatively small number of institutions that play a leading role in multiple markets, suggest an increased potential for shocks as well as increasing difficulty in improvising effective crisis-management in the event of a shock.”\(^\text{139}\) The report warned that while a complete breakdown of the world financial system had never occurred, by the mid-1990s such a scenario was possible due to “the emergence of large integrated financial firms with corporate structures and finances of extreme complexity and global scope.”\(^\text{140}\) The former paradigm of capital adequacy regulation missed the mark:

[D]irect and indirect risk exposures within [such a] group are so complicated and opaque and change so rapidly that it is virtually impossible to monitor them in anything like real time. Accounting and disclosure practices have not begun to keep pace. Risk exposures can build up undetected by existing monitoring systems. In a crisis, both peer institutions and regulators may feel they have too little information about the condition of a faltering institution and insufficient time to assess this complex information to warrant taking action.\(^\text{141}\)

To remedy in part the ratcheting up of systemic risk, the Group of Thirty proposed “enhanced responsibility [for] financial institutions” in financial regulation, which “implies that supervisors will be readier [sic] to rely on the institutions that they supervise, and that the institutions

\(^{135}\) See Group of Seven, supra note 134, at C134a.

\(^{136}\) The Group of Thirty, or G30, is a group consisting of academics, former supervisors and regulators, and representatives of the financial industry.

\(^{137}\) Group of Thirty, Global Institutions, National Supervision and Systemic Risk v (1997).

\(^{138}\) See id. at ii.

\(^{139}\) Id. at 5–6.

\(^{140}\) Id. at 11.

\(^{141}\) Id. at 8.
themselves will accept the responsibility to improve the structure of, and discipline imposed by, their internal control functions."\footnote{142} Such a redesigned governance edifice would recognize that since the underlying causes of excess risk-taking are firm-specific, the top-down mode of governance will likely prove inadequate since regulators could not be expected "to evaluate the quality of traders or the current daily [VaR] in trading exotic derivative instruments."\footnote{143} Regulators would perennially be "behind the curve" and also unable to attract the talent required to monitor such complex systems with adequately high pay.


To address the burgeoning trading book and derivatives activities of internationally active banks, the Basel Committee initiated a process in 1992 to establish rules that would require banks to set aside capital to protect against market risks.\footnote{144} The Basel Committee circulated a paper outlining a framework for measuring market risk and offering banks a menu of standardized computational methodologies. Banks reacted negatively to the proposals, on the grounds that their internally generated VaR models more accurately captured market risk.\footnote{145} The Basel Committee agreed with the banks, and in 1996 it amended the Basel Accord to include a capital requirement to cover market risks for assets held in the trading book, expressly permitting the use of internal VaR models in setting the capital requirement.\footnote{146}

\footnote{142}{Id. at 12.}
\footnote{143}{John Heumann & Lord Alexander Of Weedon, Global Institutions, National Supervision And Systemic Risk 87 (1997).}
\footnote{144}{See Nancy White Huckins & Anoop Rai, Market Risks for Foreign Currency Options: Basel’s Simplified Model, 28 Fin. Mgmt 99, 99 (1999).}
\footnote{145}{Id.}
The Amendment to the Capital Accord to Incorporate Market Risks (MRA) defined market risk to include interest rate risk and equity risk associated with a bank's trading book (including positions in derivatives and off-balance sheet instruments) and commodities risk and foreign exchange risk throughout a bank's asset portfolio. \(^\text{147}\) Importantly, and in distinction to the U.S. FDICIA initiative, \(^\text{148}\) the MRA did not require that banks submit to a single formula for computing market risk capital charges. Instead, regulated banks could choose one of two methodologies of computing the newly applicable capital charges: a "standardized" approach or a method "allow[ing] banks to use risk measures derived from their own internal risk management models." \(^\text{149}\) For those banks adopting the internal models approach, the MRA required them to determine their VaR over a one-day time horizon (that is, the maximum loss a bank would incur in a one-day period) at a 99% confidence interval. The VaR estimate would then be multiplied by a "safety factor" multiplier of three to set the regulatory capital requirement. \(^\text{150}\) This multiplier could be ramped up to four based on unfavorable model backtesting results showing the lack of reliability of an internal model.

As a condition precedent to the use of internal models for purposes of market risk capital, the MRA required the bank regulator to determine that the bank meets certain general prerequisites to ensure the conceptual

---


148. See supra note 62 and accompanying text.

149. 1996 MARKET RISK AMENDMENTS, supra note 146, at 3. The 1996 market risk capital charge exempts certain categories of covered assets (i.e., debt and equity securities in the trading book, all positions in commodities) from the operation of the base Basel Accord credit risk charge. Then, the market risk capital charge, whether calculated pursuant to the "standardized" method or a proprietary internal model, is added to the Basel Accord credit risk charge to yield a total capital requirement. See Thomas S.Y. Ho & Sang Bin Lee, THE OXFORD GUIDE TO FINANCIAL MODELING: APPLICATIONS FOR CAPITAL MARKETS, CORPORATE FINANCE, RISK MANAGEMENT, AND FINANCIAL INSTITUTIONS 602 (2004).

150. By requiring the application of the safety factor multiplier, the Basel Committee impliedly acknowledged the uncertainty of the VaR estimates. The selection of the multiplier was largely a political compromise, and lacked any putative scientific justification at the time of its selection. In 1997, Gerhard Stahl published a paper purporting to show that a multiplier of three was reasonable to incorporate consideration of model uncertainty with respect to a model's distribution of risk factors. See Gerhard Stahl, Three Cheers, 10 RISK 67, 67–69 (1997).
soundness of the models and the adequacy as well as day-to-day usage and integration of the risk management function.151

In recognition of the potential conflicts of interest inhering in an internal models approach, the MRA required that internal models be subject to backtesting and that unfavorable test results would result in a higher safety factor multiplier (the base multiplier was, as noted above, three).152 If the results of the backtesting indicate that the internal model’s measurement of VaR did not correspond with actual daily profit and loss data (or “trading outcomes,” as the MRA puts it), a penalty multiplier would apply to the VaR depending on the extent of the disjoint between VaR and actual loss. The MRA provides that the aim of the backtesting regime is to “incorporate suitable incentives” into the internal models approach, but the nature of the backtesting exercise is that it is retrospective and not anticipatory. The Basel Committee notably did not require that a model ensure capital adequacy in light of hypothetical future adverse scenarios or stress tests.

In 2005, the Basel Committee issued joint guidance with IOSCO to revise the MRA to address the burgeoning credit risk that was building up in banks’ trading books through an explosion in the trading of collateralized debt obligations (CDOs), credit default swaps (CDSs) and other structured and illiquid products.153 These amendments also aimed to reduce capital arbitrage opportunities. As bank trading books burgeoned and average

152. See Basel Comm. on Banking Supervision, Amendment to the Capital Accord to Incorporate Market Risks 41 (2005), http://www.bis.org/publ/bcbs19.pdf?noframes=1 [hereinafter 2005 MRA Amendments]; Basel Comm. on Banking Supervision, Supervisory Framework for the Use of “Backtesting” in Conjunction with the Internal Models Approach to Market Risk Capital Requirements 1–2, 11–12 (1996) [hereinafter Backtesting Framework], http://www.bis.org/publ/bcbs22.pdf?noframes=1. Notably absent from the Basel Committee reports, however, is a direct mention of conflicts of interest. Instead, the Basel Committee focused on “incorporating suitable incentives into the internal models approach[,]” though it appears from context that the Basel Committee was referring to “continual improvement” of the models rather than the obvious moral hazard involved from the delegation of responsibility. Backtesting Framework, supra, at 1. U.S. banking regulators adopted the backtesting methodology in nearly identical form to the Basel Committee paper.
asset holding periods shortened, opportunities for arbitrage abounded, chiefly by way of banks classifying as trading positions certain assets that properly belonged in the loan book (and accordingly carry a generally higher capital charge). Under the 2005 amendments to the MRA, most large banks would be required to measure and hold capital against the incremental default risk not captured in the bank’s VaR models. In particular, bank regulators considered the MRA’s specification of a ten-day trading holding period for a 99% confidence interval to be an unrealistic assumption for VaR in connection with illiquid credit default swaps. As was evident from the sudden collapse of AIG’s CDS business in 2008, a collapse in confidence with respect to an important CDS counterparty can lead to a rapid evaporation of liquidity. Although the Basel Committee insisted on an additional capital charge and ramped up the statistical “confidence” of the measurement, the banks were again instructed to utilize internal models to measure the relevant risks. Once again, we see in the 2005 amendments (1) a perception of a mismatch between extant capital

154. The Basel Accord only applies to “internationally active banks.” See Basel Comm. on Banking Supervision, Amendment to the Capital Accord to Incorporate Market Risks I (2005), http://www.bis.org/publ/bcbs119.pdf?noreferrer=1. In the United States, the proposed rules will apply to all banks with worldwide consolidated trading revenue equal to either (1) 10% of total assets or (2) $1 billion. See Risk-Based Capital Standards: Market Risk, 71 Fed. Reg. at 55,960.


158. See Risk-Based Capital Standards: Market Risk, 71 Fed. Reg. at 55,965 (“As under the current market risk capital rule, a bank would be required to use one or more internal models to calculate a daily VaR-based measure that reflects general market risk for all covered positions.”); Basel Comm. on Banking Supervision, The Application of Basel II to Trading Activities and the Treatment of Double Default Effects 67 (2005), http://www.bis.org/publ/bcbs116.pdf?noreferrer=1. Banks are also free to calculate their incremental credit risk capital charge by reference to a surcharge through an approach consistent with its approach for calculating credit risk in its Basel II risk-based capital internal model. If such a “surcharge” is applied, the bank can insulate itself from the backtesting requirement altogether. See 2005 MRA Amendments, supra note 152, at 47.
regulation and the risks to which banks were actually subjecting their asset portfolios and (2) a resolution to delegate the measurement of that risk to the regulated banks.

c. Basel II Pillars 1 and 3: Internally Modeling Credit and Operational Risks and Public Disclosure

Shortly after the Basel Committee had finalized the market risk-focused 1996 MRA, it commenced a reevaluation of the Basel Accord’s treatment of credit risk. Aside from a few amendments to tidy up unresolved issues in the Basel Accord, the credit risk regime had remained largely static throughout the 1990s.\textsuperscript{159} Bank supervisors were faced with an increasingly dynamic risk profile of banks and the need to anticipate new sources of vulnerability and uncertainty (of both the known and unknown variety), which “require[d] trying to understand how changing institutions, products, markets, and trading strategies create vulnerabilities to new kinds of shocks and new channels of contagion.”\textsuperscript{160} Among the measures that intrigued the Basel Committee was the incorporation of internal credit risk models, which had advanced considerably in the 1990s, into the capital adequacy regulatory context. The Basel Committee circulated a draft proposal to members and regulated banks in June 1999, and presented an agreed text five years later in June 2004 (Basel II).\textsuperscript{161} The Basel II reforms significantly retool bank capital adequacy regulation. Basel II comprises three “pillars”: Pillar 1 specifies the minimum capital requirements for credit risk and operational risk; Pillar 2 concerns the supervisory review process; and Pillar 3 sets forth new market disclosure requirements intended to enhance market discipline alongside regulation and supervision.

\textsuperscript{159} Amendments included (i) a 1991 amendment resolving a lacuna in the Basel Accord by setting of a 1.25\% ceiling (or 2\% in exceptional circumstances) for the extent to which banks could count general provisions or general loan loss reserves as Tier 2 capital; (ii) 1994 amendments facilitating the netting of obligations from a single counterparty; and (iii) 1994 amendments to address perceived risks with preferential risk weighting of OECD government-issued debt subsequent to the entry of Mexico, South Korea, and Turkey in the OECD. See Duncan Wood, Governing Global Banking 124–25 (2005).

\textsuperscript{160} Herring, supra note 59, at 396.

\textsuperscript{161} While increasing the risk sensitivity of the initial Basel Accord was the impetus behind the Basel II reform, I do not mean to give the impression that the members of the Basel Committee were pursuing this goal as enlightened philosopher kings. The agenda of the multi-year reform effort was driven in part by massive lobbying expenditures of large internationally active banks, which stood to gain from the internal models approach competitive advantages vis-à-vis their regional competitors. This story, while important for context, is outside the scope of this Article, which examines the internal models approach as it is in order to gauge its merits and demerits as a new governance technology.
i. **Pillar 1: Capital Requirements and Internal Models**

Pillar 1 allows banks to compute their regulatory capital requirements against credit risk in two ways: (1) a revised standardized approach based on the initial Basel Accord, or (2) one of two versions of an “internal ratings based” (IRB) approach whereby banks are permitted to develop and use their own internal risk ratings. The IRB approaches permit banks meeting certain qualitative and quantitative criteria to set their capital requirements by reference to inputs from their own internal VaR models rather than the Basel Accord’s multipliers. In the United States, bank regulators have mandated the use of the IRB for banks with either consolidated total assets of at least $250 billion or consolidated on-balance sheet exposure of at least $10 billion.

Specifically, there are two IRB approaches: the foundation approach (FA) and the advanced approach (AA). Both approaches require banks to categorize their assets according to five categories (sovereign, bank, corporate, retail, and equity). The IRB approaches are based on four key input parameters: (1) the probability of default (PD); (2) the loss given default (LGD); (3) the exposure at default (EAD); and (4) effective maturity (M). PD represents the “long-run average of one-year default rate” for a given borrower. LGD measures the anticipated loss, expressed as a percentage, of a total exposure upon the occurrence of a default. EAD measures the total exposure if a default occurred, expressed as an amount. Basel II sets forth elaborate asset class-specific computational formulae to be used in calculating the capital requirements based on whether a group is

162. The revised standardized approach further disaggregates the five risk-weighting categories of the Basel Accord, thereby heightening the risk sensitivity of the standardized capital adequacy requirements. Basel Comm. on Banking Supervision, International Convergence of Capital Measurement and Capital Standards—A Revised Framework (Comprehensive Version) 19–51 (2006), http://www.bis.org/publ/bcbs128.pdf?noframes=1 (hereinafter Basel II Document) (expanding risk weight categories to include, among others, a 35% risk weights for claims secured by residential real estate and 150% risk weights for claims on corporate rated below BB- by Standard & Poors). Basel II also includes provisions to take account of credit risk mitigation in the form of credit default swaps, financial guarantees and collateralization of claims. See id. at 51.  
163. See id. at 80–120.  
166. Id. at 99.  
167. A standardized definition of default is provided in paragraph 452. Id. at 100.
using AA or FA. Thus, the applicable formula differs between, e.g., an FA bank’s sovereign exposures and an AA bank’s sovereign exposures, or between an FA bank’s equity exposures and the same FA bank’s corporate exposures.168

The formulae include some combination of the input parameters. Generally, the FA formulae utilize internal models for PD estimates but not for LGD, EAD, and M estimates.169 The AA formulae utilize all internally modeled estimates.170 As a result, the derivation of the capital requirement for a given class of assets is dependant on internally generated estimates of the input parameters. These input parameters are the same as, or are very similar to, the inputs used in the VaR economic capital models that large groups have been using since the 1990s.171 It is important to point out that, unlike the capital models used in the MRA, banks’ internal credit risk economic models are not actually used in setting the regulatory capital level, though they share common inputs. Instead, the common inputs are inserted into the formulae set out in the Basel II documents. So the Basel Committee trusts banks about credit, but not as much as they trust banks about trading markets; embedded in the Basel formulae are particular assumptions about the underlying drivers of portfolio credit risk, including loss correlations.172

Basel II also includes for the first time a mandatory charge against “operational risk,” which was defined as the “risk of loss resulting from inadequate or failed internal processes, people and systems or from external events.”173 Under the advanced measurement approach (AMA), groups may “use their own risk metrics for calculating the operational risk capital requirement, including loss data, scenario analysis, and risk mitigation

168. There is an inertial aspect to the implementation of IRB to asset classes: any group using IRB for one or more asset classes (e.g., sovereign, bank) is required eventually to extend IRB treatment to all asset classes, and a group may not return to the standardized approach once initiating IRB. Id. at 61–62. Moreover, the enhanced flexibility to measure LGD and EAD provides a significant incentive for groups to migrate from FA to AA. See Til Schuermann, What Do We Know About Loss Given Default? 3 (Wharton Fin. Inst. Center, Working Paper Series No. 04-01, 2004), available at http://fic.wharton.upenn.edu/fic/papers/04/0401.pdf.


170. Id. at 59.


172. Id. at 69,292; TARULLO, supra note 14, at 155–59.

2010] INTERNAL MODELS APPROACH TO CAPITAL ADEQUACY REGULATION

829

measures.”174 Like the IRB approaches to credit risks and the MRA for market risks, AMA eligibility is conditioned on meeting certain criteria.175

ii. Pillar 3: “Market Discipline” Through (Some) Disclosure of Proprietary Information

With Pillar 3, the Basel Committee seeks to foster “market discipline” by requiring banks to make public, unaudited disclosures of certain qualitative and quantitative information about their regulatory capital positions, risk management infrastructure, and risk positions.176 In keeping with the Basel II theme, the disclosures are designed to “be consistent with how senior management and the board of directors assess and manage the risks of the bank.”177 Pursuant to Pillar 3, the top consolidated entity within a banking group must disclose information relating to its capital structure, regulatory capital requirements, risk exposures and risk management processes with respect to credit, market, interest, operational, and equity risks.178 All disclosures under Pillar 3 are subject to a materiality qualification.179

Groups using the IRB approach for credit risk must disclose quantitative details regarding the inputs used in the models (i.e., PD, LGD, EAD, M) and the backtested performance of the models. The time horizon for the required backtesting is unspecified, though the Basel Committee “expect[s] that banks would provide these disclosures for as long a run of data as possible.”180 The general characteristics of the internal models used by banks calculating the market risk AMA capital charge must similarly be described, and disclosures are to be made regarding the stress testing and backtesting that, as discussed above, must be applied to the portfolio.181 Groups using the AMA to set operational risk capital charges must provide a description of the AMA, including a “discussion of relevant internal and external factors considered in the bank’s measurement approach.”182

Additionally, qualitative disclosure regarding certain assumptions of the IRB credit risk internal models is required.183 With respect to these models, disclosure must be made of “[t]he definitions, methods and data for estimation and validation of PD, and . . . LGD and/or EAD, including

175. See BASEL II DOCUMENT, supra note 152, at 149–52.
176. See id. at 226.
177. Id.
178. Id. at 231–42.
179. Id. at 227.
180. Id. at 236 n.214.
181. Id. at 241.
182. Id.
183. Id. at 235.
assumptions employed in these variables.”

However, the disclosure “does not require a detailed description of the model in full—it should provide the reader with a broad overview of the model approach, describing definitions of the variables, and methods for estimating and validating” the model inputs.

III. BROKER–DEALERS AND THE CSE PROGRAM

In 2004, the SEC invited the largest U.S. broker–dealers to adopt an internal models approach to capital adequacy regulation in the style of Basel II under its Consolidated Supervised Entity (CSE) Program. From 1975 until 2004, all U.S. broker–dealers were subject to the so-called “uniform net capital rule” as Rule 15c3-1, which operated as a check on the proprietary trading activities of registered broker–dealers. The net capital rule gauges the adequacy of a broker–dealer’s capitalization by reference to availability of liquid assets to satisfy the obligations of its customers. Specifically, firms must elect either to “(a) maintain aggregate indebtedness at a level” not in excess of “fifteen times net capital” (the “basic test”) or “(b) maintain minimum net capital equal to not less than two percent of ‘aggregate debit items’” (the “alternative test”). When considering the regulatory purpose behind the net capital rule, most commentators have focused on the protection of customers and creditors from losses and delays that might arise when a broker–dealer fails, and the related protection of the SIPC insurance of customer accounts, though there is, as with banks, a systemic risk minimization rationale as well.

Both the basic and the alternative net capital ratio tests first require a calculation of net capital, which is really a regulatory assessment of the liquid capital available for prompt distribution in the event of liquidation.

184. Id.
185. Id. at n.207.
186. See generally Thomas Lee Hazen & David L. Ratner, Broker Dealer Regulation: Cases and Materials 521–33 (2003). Industry firms were subject to some form of capital regulation since enactment of the Securities Exchange Act of 1934. In 1942, the SEC promulgated the precursor to Rule 15c3-1, but exempted from its purview firms that were members of stock exchanges with similar capital adequacy regulation, such as the NYSE’s Rule 325. In response to the industry’s “paperwork crisis” in the late 1960s and a rash of insolvencies of U.S. broker–dealers, Congress and the SEC parried proposals and guidance for capital regulation that eventually resulted in Rule 15c3-1 in 1973.
187. John C. Coffee, Jr. & Hilary A. Sale, Redesigning the SEC: Does the Treasury Have a Better Idea?, 95 VA. L. REV. 707, 739 (2009). “Aggregate debit items” is a measure of assets that takes into account the “haircuts” discussed below to reflect illiquidity. See id. n.86.
188. Id. n.86.
To compute net capital, the broker–dealer starts with its GAAP assets and subtracts its GAAP liabilities to yield a net-worth-based capital number. Subordinated liabilities meeting certain criteria and deferred tax liabilities are added back into capital and illiquid assets (such as fixed property and exchange membership rights), intangible assets (such as goodwill), prepaid assets (such as insurance premiums and rent) and unsecured receivables are backed out to produce a “tentative net capital” number. Finally, a prescribed percentage of the market value of each broker–dealer proprietary position is subtracted from tentative net capital, based on the perceived market and asset liquidity risks associated with that particular position. These “haircuts” are essentially reserves that reflect the expectation that in a liquidation scenario, the proceeds to be obtained from liquidating securities are subject to adverse price movements. Proprietary assets that are less liquid or more volatile carry higher haircut percentages, which reduces the capital number more than, for example, Treasury bills (which carry a 0% haircut). These haircuts are aggregated and subtracted from the capital number to yield a final “net capital” amount that is compared to aggregate indebtedness (under the basic test) or aggregate debit items (under the alternative test). A firm with inadequate net capital may not open its doors for business until it corrects the capital shortfall. The haircut feature of the net capital rule is analogous to the initial Basel Accord’s risk-weighting regime.

It should be pointed out that both the basic and alternative variants of the net capital rule, like the IBA, incorporated only the most rudimentary risk sensitivities. In 1997 and 2002, the SEC and the CFTC, respectively, approved the use of private third-party statistical option-pricing models to set capital charges for certain options and futures contracts. In doing so, the SEC emphasized that statistical modeling techniques were capable of assessing risks and evaluating correlation of asset prices with greater detail and sensitivity than the rigid haircut regime that had formerly been in place. Under this new system, third-party vendors would be approved by the applicable self-regulatory organizations (e.g., NYSE or NASD) and would provide to the broker–dealers, for a fee, results of option pricing models that aimed to estimate potential loss on options. The highest amount of loss at any particular valuation point would be the charge to net capital.\footnote{U.S. GOV’T ACCOUNTABILITY OFFICE, RISK-BASED CAPITAL: REGULATORY AND INDUSTRY APPROACHES TO CAPITAL & RISK 144 (1998), available at http://frwebgate.access.gpo.gov/cgi-bin/getdoc.cgi?dbname=gao&docid=fagg98153.pdf.} The SEC therefore sought to leverage the expertise of industry to better synchronize risk and capital, much in the same way the SEC had pegged capital charge haircuts for nonconvertible debt securities to
“nationally recognized statistical rating organizations” in 1975.\textsuperscript{191} Other than the option pricing models, the net capital rule went largely undisturbed from 1982 until 2004.\textsuperscript{192} By October 2008, each of the five major independent broker–dealer institutions—Bear Stearns, Goldman Sachs, Lehman Brothers, Merrill Lynch, and Morgan Stanley—had undergone a massive and rapidly cascading liquidity crisis and been liquidated or, at the behest of the Federal Reserve, either converted to a financial holding company or been acquired in a shotgun wedding-style transaction. How had the liquidy-minded net capital rule failed to preserve the ability of these institutions to meet their obligations in due course? The answer is simple; since 2004, none of these institutions had been bound by the net capital rule and were instead subject to the CSE Program internal models regime that set no ceiling on leverage. In fact, all five institutions were operating by February 2008 with gross leverage ratios of approximately thirty times net capital.\textsuperscript{193}

Professors John Coffee and Hillary Sale have documented the connection between the E.U. Financial Conglomerates Directive (FCD)\textsuperscript{194} and the CSE Program,\textsuperscript{195} U.S. investment banks, desirous to avoid submitting to consolidated regulation in the E.U. (including requirements regarding internal controls, capital adequacy, intragroup transactions, and risk concentration) in addition to U.S. regulation, brought pressure on the SEC to regulate the broker–dealers in a substantially equivalent manner, so


\textsuperscript{194} Council Directive 2002/87, 2003 O.J. (L 35) 1 (EC). Major U.S. commercial banks were exempted from the operation of the FCD because they were subject to “equivalent” supervision in the United States in the form of group-level supervision by the Federal Reserve. See Coffee & Sale, supra note 107, at 738.

\textsuperscript{195} See Coffee & Sale, supra note 107, at 737–38.
as to fall within the FCD exemption for financial groups operating under equivalent supervisory standards. In 2004, the SEC offered to these broker–dealers the option of opting out of the net capital rule system and entering the new CSE regulatory regime of consolidated groups. Under the CSE Program, broker–dealers with at least $5 billion of capital would be permitted to avoid the net capital rule entirely, and instead subject themselves to an alternative net capital program resembling Basel II. A qualifying broker–dealer could become a CSE by applying for an exemption from the SEC standard net capital rule. CSE groups would have to “[c]alculate a group-wide capital adequacy measure consistent with the international standards adopted by the Basel Committee” and “maintain an overall . . . capital ratio of not less than the Federal Reserve’s 10 percent ‘well-capitalized’ standard for bank holding companies.” Firms were to maintain tentative net capital at a level above $1 billion and net capital above $500 million at all times.

As discussed in detail in Part II.B.3.c, the Basel II framework relies on the regulated firm to determine its own required capital levels based on inputs derived from internal models. Intriguingly, federal bank regulators proceeded very cautiously with the implementation of Basel II guidelines in the United States. The same banking regulators that participated in the creation of the guidelines were careful to ensure they were not applied too hastily domestically. The SEC, by contrast, pushed through the CSE Program in a matter of months. These risk models purported to gauge the market risks that by September 2008 had materialized and evaporated

196. See Alternative Net Capital Requirements for Broker–Dealers That Are Part of Consolidated Supervised Entities; Supervised Investment Bank Holding Companies, 69 Fed. Reg. 34,428 (June 21, 2004) (codified at 17 C.F.R. pt. 240). Technically, Citigroup and JPMorgan Chase were CSEs under the CSE Program, but the SEC did not exercise direct oversight over these institutions, which were separately overseen by the Board of Governors of the Federal Reserve. See CSFB–Bear Stearns Report, supra note 193, at v.

197. See Pickard, supra note 192, at 10 (lauding the track record of the net capital rule and lamenting the transition to the CSE Program regime).


any real capital cushion, requiring the provision of over $12 trillion in support in the form of financial guarantees, capital injections, asset purchases, and emergency liquidity assistance by the U.S. government and the Federal Reserve. And yet at the time of the Bear Stearns implosion, each of the major U.S. investment banks was sufficiently capitalized under the CSE Basel II program, according to its own VaR assessments. The capital market, however, was telling a different story; Bear Stearns lacked access even to secured funding.

The deregulatory effects of the CSE Program naturally raise the question of capture: was the SEC captured by the powerful investment banks that saw capital requirements as antiquated relics of a financial era before financial innovation had minimized costs of financial distress normally associated with leverage? It should go without saying that the investment banking industry and their well-placed alumni lobbied aggressively for abrogation of the net capital rule. The tight-knit and centralized group of the five largest investment banks might be expected to overcome the free-riding problem normally incident to legislative or regulatory change. The decentralized and dispersed groups of smaller broker–dealers (who would encounter a higher cost of capital vis-à-vis the largest investment


204. Cf. Fanto, supra note 73, at 742 (“[I]f financial institutions underestimate loss probability, they will not have adequate capital in extreme circumstances.”).

205. Id.


207. Cf. MANCUR OLSON, JR., THE LOGIC OF COLLECTIVE ACTION: PUBLIC GOODS AND THE THEORY OF GROUPS 1–3 (1965) (theorizing that small and concentrated interest groups may overcome free-riding problems that normally frustrate interest groups’ efforts to procure favorable legislation and regulatory policy).
banks) and the brokerage customers (who would lose the benefits of the inherent conservatism of higher capital charges applicable to their custodians) would arguably navigate the free-rider problem with greater difficulty. The five major investment banks constituted a group small enough that dropping out would be difficult, and the costs of abrogating the net capital rule could be spread among smaller competitors and industry customers.

While this basic Olsonian account is in some respects accurate, it is not a complete account of the CSE Program. Two other interrelated factors, familiar to the discussion above regarding Basel II, must be taken into account. Firstly, regulators perceived—not without good reason—that the increasing complexity of the investment banking and broker–dealer business rendered the formulaic approach of the net capital rule inadequately responsive to the risks underlying and affecting these regulated businesses. With each round of financial innovation, the correlation between the haircut formulae and the real risks underlying financial assets became increasingly arbitrary. Moreover, as risk disintermediation increases, the risk correlation among a firm’s positions is rendered more opaque. Secondly, the acculturation of financial regulators relying on leveraging private sector expertise in capital regulation—what Professor Geoffrey Miller has aptly termed, in a slightly different context, the “Basel Brand” —played a significant role in the SEC’s CSE initiative. The SEC borrowed the Federal Reserve’s Basel II method of calculating capital requirements for bank holding companies and applied them to the new consolidated supervised entities. In a sense, convergence of capital regulation among investment banking and commercial banking groups made eminently good sense; after all, in the years following passage of the Gramm–Leach–Bliley Act in 1999, global financial conglomerates came to dominate global lending and investment banking activities. If the two categories were collapsing into one, it made little sense to regulators to regulate them according to disparate logics.

IV. INSURANCE COMPANIES & SOLVENCY II

Insurance companies, like banks, must deal with an inherent asset–liability mismatch. In the case of insurance companies, though, the problem is the mirror image of the banking mismatch: an insurance

209. See CSE–BEAR STEARNS REPORT, supra note 193, at 10–11.
210. See Wilmarth, supra note 92, at 972–74 (discussing the Gramm–Leach–Bliley Act’s effect on the U.S. financial services industry both domestically and abroad).
company funds long-term future liabilities (i.e., policy claims) with a regular stream of premium income (i.e., short-term assets). Government-administered insurance guarantee funds support policyholders in the event a firm is insolvent. Insurance regulators therefore tightly regulate capital adequacy; it is, along with consumer protection, their chief regulatory objective. The European Union is poised to reform its insurance industry capital adequacy framework by instituting the so-called Solvency II framework. As proposed, Solvency II will allow firms to use internal market and credit risk models as the basis on which regulatory capital levels are established.

V. INTERNAL CAPITAL MODELS APPROACH AS NEW GOVERNANCE

For the most part, financial regulatory scholarship has not yet appreciated the important insights of new governance theory and its applicability to dynamic, complex financial markets implicating a wide range of public policy interests. Part V will (1) summarize new governance theory, providing illustrative examples where it is put to practical use and discussing its connections to and disjunctions from the familiar rules–principles debate; (2) describe how the internal models approach to capital regulation can be considered, in certain respects, a new governance regulatory technique; and (3) characterize the shortcomings of the internal models approach as falling within familiar traps that affect new governance reforms, and conclude that Pillar 3’s disclosure regime fails to redress these traps adequately.

A. WHAT IS NEW GOVERNANCE?

New governance refers to a wide range of administrative governance techniques and tools that differ in important ways from traditional top-down, command-and-control regulation. In particular, new governance scholarship highlights the increasing involvement of nonstate actors in the governance tools that shape and constitute public policy and regulation. The description of new governance techniques as tools is not accidental; their relation to policy objectives is instrumental rather than rival.  

211. See Lester M. Salamon, The Tools of Government: A Guide to the New Governance 1, 8–9 (Lester M. Salamon ed., 2002) (advocating a shift in the “unit of analysis” of public administrative study from public agencies to the tools and instruments of public administration); see also On Amir & Orly Lobel, Stumble, Predict, Nudge: How Behavioral Economics Informs Law and Policy, 108 COLUM. L. REV. 2098, 2131–32 (2009) (arguing that the ends and means of public policy are “inevitably intertwined,” and that new governance “unabashedly recognizes” that the choice of tools in public policymaking is therefore inherently normative); Orly Lobel, Setting the Agenda for New Governance Research, 89 MINN. L.
governance scholarship is ambitious in its scope, and it touches on a wide spectrum of the realms of public administration, from nursing home care, Medicare and Medicaid service delivery, workplace safety, and employment to crop insurance and endangered species conservation. Despite the ambition of new governance, comparatively little has been written to situate financial regulation reforms in the new governance context.

The units of inquiry according to the traditional conception of regulation are hierarchically-ordered administrative agencies, in possession of unique expertise, to which the legislature has granted discretion to pursue statutorily defined regulatory objectives. To the extent private involvement is contemplated, it is usually framed as a threat to proper administrative process. Public choice theorists have called into question administrative discretion itself, emphasizing the pervasiveness of strategic manipulation in public administration, and describing policy choices as a product of pressure on the part of discrete, well-organized private actors.


Still others focus on accountability: how to protect the integrity and expertise of the public administrative apparatus from narrowly self-interested, private-interested parties.\textsuperscript{214} 

Each of these accounts relies on a conceptual division between public policymakers and private nonstate actors. As Professor Jody Freeman has put it, they rely on “the illusion of a public realm”—that is, that there exists a purely public sphere of activity.\textsuperscript{215} Even public choice theory buys into this illusion, to the extent it advocates a normative deregulatory position; there is a cross-pollination of private and public realms, but only because the former is seeking to capture the latter. New governance theory posits that this traditional conception is outdated as a descriptive matter, and that, as a normative matter, in many contexts it is unlikely to conduce to favorable regulatory outcomes. Instead, it is better to evaluate the effectiveness of governance by reference to the tools through which governance is effectuated, rather than by reference to the \textit{dramatis personae} of the regulator–regulatee game. As such, new governance scholars tend to focus their analysis on the new modes of interaction between state and nonstate actors. New governance, as a dynamic toolset rather than an \textit{ex ante} ideal distribution of administrative power, is therefore analytically different than traditional accounts of administrative law discussed above. It can coexist with traditional administrative activity, either as a rival or a complement, and it can, though it need not, lead to wholesale transformation of old governance into new governance.\textsuperscript{216} New governance tools aim to respond to the continual changes of regulated society and knowledge itself, so “all solutions [to problems] should be regarded as provisional.”\textsuperscript{217} 

New governance programs and tools assume different forms, and often include clusters of the following characteristics: increased participation of and power sharing with nonstate actors; public adoption of rules negotiated by nonstate stakeholders; encouragement of experimentation; promotion of competition and diversity as a structural element of regulation; dynamic, flexible, and dialogic lawmaking process as a response to dynamism of regulated markets; multilevel functional integration and network-seeking among branches, departments, agencies of government, and among state

\textsuperscript{214} See Jody Freeman, \textit{Collaborative Governance in the Administrative State}, 45 UCLA L. REV. 1, 13 (1997) (describing this position as a “conceptual limitation” that inhibits more effective governance reform proposals).


\textsuperscript{217} \textit{Id.}
2010] *Internal Models Approach to Capital Adequacy Regulation*

839

and nonstate actors, and within regulatee firms; active pursuit of nonstate actor knowledge to supplement—and sometimes replace—public administrative expertise; promotion of subsidiarity; use of flexible, revisable rules and standards; use of broad framework agreements; use of benchmarks, indicators, and peer review to ensure accountability; and fostering of deliberation among stakeholders over the nature of problems and methods by which to solve them.

Of course, no new governance tool draws on all of these attributes. New governance measures are at times highly formalized, and at other times informal and consultative, consistent with “soft law” pronouncements.218 The common thread running throughout new governance scholarship is the deployment of novel techniques and tools in governance, usually involving increased nonstate involvement, to overcome emergent or intractable recurrent problems inhibiting the traditional command-and-control regulatory model from achieving its regulatory objectives. This Part discusses three of these characteristics in greater detail, since they are most relevant to regulatory reform in dynamic and complex financial markets, including Basel II’s internal models approach: (1) retention of a public role in lawmaking and enforcement; (2) active pursuit of nonstate actor knowledge to supplement, and sometimes replace, public administrative expertise; and (3) a dynamic, flexible, and dialogic lawmaking process.

1. *The Benign Big Gun: Retaining a Public Administrative Presence*

The dominant account of new governance contemplates retention of some formal public authority, even if it operates in the background. While broader definitions of “governance” include a wider range of measures impacting social and economic systems,219 for purposes of this discussion new governance must involve some element of public administrative ordering. This precondition can be thought of as a jurisdictional requirement; in other words, there is no ambiguity that the new governance regulatory event is within the purview of at least one, and perhaps multiple,

218. Not all new governance initiatives are characterized by “softness,” though softness is often an indicium of new governance. At times, however, new governance measures are characterized by a high degree of “hardness,” especially the procedural orderings designed to facilitate consensus-forming and power-sharing in the drafting of legislation and regulation. See Bradley C. Karkkainen, *Reply: “New Governance” in Legal Thought and in the World: Some Splitting as Antidote to Overzealous Lumping*, 89 MINN. L. REV. 471, 485–87 (2004).

219. See Scott Burris et al., *Changes in Governance: A Cross-Disciplinary Review of Current Scholarship*, 41 AKRON L. REV. 1, 7–9 (2008) (highlighting the risk that governance risks “becoming a point of false rhetorical convergence” and adopting an abstract definition of “governance” as “the management of the course of events in the social system”).
public administrative agencies. By locating the regulatory event within an administrative agency’s jurisdiction, the new governance measure enhances accountability and perhaps legitimacy as well. The central accountability challenge associated with new governance—namely, that such measures vest substantial policy discretion in nonstate actors that are not responsible for the results—is thus mitigated. In other words, the state is not dead; it remains a critical juncture of new governance networks, just not as an authoritative, directing regulator in a command-and-control system. By retaining a public element, new governance is also distinguished from deregulation. The deregulatory political mood of the 1970s and 1980s resulted in significant deregulation of the trucking and railroad industries, financial institutions, oil and gas prices, occupational safety, and environmental protection. Consider, for instance, the Garn–St. Germain Depository Institutions Act of 1982, which among other things removed restrictions on thrift banks that had previously prevented them from taking demand deposits and making commercial loans, lifted other restrictions applicable to all depository institutions, and scheduled the formal phase-out of all interest rate restrictions on demand deposits. With respect to each of these elements of the Act, the state dropped out of the picture, except for its general supervisory competence to ensure the “safety and soundness” of regulated banks and thrifts. The state decided

220. Cf. David A. Dana, The New “Contractarian” Paradigm in Environmental Regulation, 2000 U. ILL. L. REV. 35, 47 (2000) (“Thus, although it is true that contractarian regulation is a reform alternative to command-and-control regulation, it is also true that command-and-control regulation is a precondition for contractarian regulation.”).


222. See Adam Crawford, Networked Governance and the Post-Regulatory State? Steering, Roving and Anchoring the Provision of Policing and Security, 10 THEORETICAL CRIMINOLOGY 449, 459 (2006) (discussing the continued importance of the state even to radical governance reforms, seeing its importance “in its symbolic power and cultural authority; in its legitimacy claims and public perceptions of its legitimacy; as a distinctive (tactical) resource and source of information through which interests are pursued; [and] in its residual position as a back-up of last resort with regard to other forms of control’’); Louise G. Trubek, New Governance and Soft Law in Health Care Reform, 3 IND. HEALTH L. REV. 137, 159–60 (2006) (describing the state’s role in new governance initiatives as “disaggregated” but “necessary”).

223. The advances of public choice theory provided intellectual succor for the deregulatory politics of the Carter–Reagan–George H.W. Bush presidencies. For some public choice theorists, favorable regulatory outcomes were not achievable in the context of public politics, so the state’s proper role was a stage exit.

that achievement of the regulatory objective was no longer worth pursuing, and a private-market ordering of the thrift demand deposit markets materialized quickly. Again, for purposes of this Article, and most new governance scholarship, a purely private ordering will not be properly categorized as a new governance tool.

By retaining residual command-and-control powers, regulators wield what Ian Ayres and John Braithwaite have referred to as a benign big gun.225 The background threat of the benign big gun, in the form of rarely deployed but available severe sanctions, serves to incentivize regulated market actors to avoid defecting from the regulatory objective. In this respect, new governance excludes pure deregulatory initiatives and assumes that some payoff, whether through the avoidance of stringent regulation or a positive incentive to cooperate, exists for the regulated private-sector actors in exchange for their participation in the governance initiative. The new governance world, though, need not exclude self-regulatory regimes such as the Financial Industry Regulatory Authority (FINRA), provided that there remains public supervision of the self-regulatory organization (SRO).226 When the regulatory objective is best achieved through a collaborative relation between public regulator and private regulatee rather than a draconian set of sanctions accompanied by intrusive supervision, the background threat of a benign big gun will be preferable.227 Importantly, one need not adopt a rational actor model of behavior to accept the theoretical superiority of a benign big gun model; the presence of the public authority in the background can motivate socially responsible deliberation in the foreground, in effect leveraging off of regulatees’ pre-existing commitments to, e.g., professional integrity or law-abidingness.228

225. See generally IAN AYRES & JOHN BRAITHWAITE, RESPONSIVE REGULATION: TRANSCENDING THE DEREGULATION DEBATE 19–53 (1992) (describing a benign big gun approach to regulation that aims to appeal to the social responsibility of actors to obtain voluntary compliance, but also stands ready to deploy deterrent threat sanctions of increasing severity to motivate purely economically motivated “rational actors” and incapacitate chronic law violators).

226. In the case of FINRA, the SEC must approve any issuance of, or modification to, an SRO rule. See 15 U.S.C. § 78q(b)(1) & (2) (2006).

227. The concept of “penalty defaults” is analogous here. Contract law provides for penalty defaults, in the form of background sanctions that no party is likely to prefer, to induce contracting parties to engage in efficient contracting. See Ian Ayres & Robert Gertner, Filling Gaps in Incomplete Contracts: An Economic Theory of Default Rules, 99 YALE L.J. 87, 97–98 (1989) (characterizing the common law rule that courts will enforce only contracts with certain and definite terms as an incentive for contracting parties to make their expectations express in the contract). The penalty default thus encourages responsible ex ante contracting and deliberation rather than ex post litigation. See id.

228. AYRES & BRAITHWAITE, supra note 225, at 47.
2. Decentering Governance: Increased Power-Sharing and Stakeholder Participation

New governance measures often provide for actual power-sharing and increased participation of nonstate actors in multiple stages of the lawmaking process (legislation, rulemaking, implementation, enforcement). By including more stakeholders in the process, policymaking is decentralized and conceived “not as [a top-down ordering process] to be done by autonomous regulators but rather as a process of mutual problem-solving among stakeholders from government and the private sector.”

The expertise and knowledge of private actors can be “harness[ed] . . . to serve public goals,” and public lawmaking is oriented toward a collaborative, consensus-seeking form of governance. The actors might remain the same, but the modes of interaction are no longer tethered to the traditional administrative law paradigm. In this respect, new governance differs from pluralistic accounts of governance that acknowledge an important role for nonstate actor involvement, but never call into question the exclusively public dimension of policymaking power and authority.

In new governance, there are multiple public and private legal entities, but there is one public process.

Increased participation and power-sharing allow for structuring collaborative solutions to complex problems. Advantages to this

230. Freeman, supra note 215, at 549.
231. See Karkkainen, supra note 218, at 474 (referring to new governance as “consensus-oriented”).
232. See Freeman, supra note 215, at 559–60 (describing the pluralist “interest representation” model of administrative law in which interest groups strive to advance their perspectives in regulation and capture is checked by “democratizing” the regulatory process to include numerous groups). A prominent example of pluralist participation is the right of interested parties to provide comments to proposed rules under the APA. See 5 U.S.C. § 553(c) (2006).
233. See Lobel, supra note 221, at 375 (“[T]he governance model offers a framework that enables us to view the different sectors—state, market, and civil society—as part of one comprehensive, interlocking system. The focus is on government interactions with private actors in public action.”).
234. See, e.g., Dorf & Sabel, supra note 211, at 315–23 (theorizing an experimentalist governance regime of “directly deliberative polyarchy” in which “governance councils” collaborate with local citizens and “pool their experience to inform their separate decisions” to achieve “good government under conditions of volatility and diversity”); Michael A. Rebell & Robert L. Hughes, Schools, Communities, and the Courts: A Dialogic Approach to Education Reform, 14 YALE L. & POL’Y REV. 99, 114–56 (1996) (elaborating a “community engagement dialogic model” of school reform that seeks to build community consensus among teachers, administrators, parents, and students in order to achieve better outcomes); Charles F. Sabel & William H. Simon, Destabilization Rights: How Public Law Litigation Succeeds, 117 HARV. L.
approach include reduced likelihood that regulatees will defect from a mutually agreed upon policy solution because of its perceived legitimacy; after all, they helped write it, and they agreed to it. Therefore, provided appropriate stakeholders contribute to the process, enhanced participation and power-sharing bolster legitimacy. Of perhaps even greater importance, though, is the potential of negotiated regulatory deals to bridge knowledge gaps between regulators and regulatees in exceedingly complex regulated industries. As noted above in Part II.B.3.a.ii with respect to banks, complexity can overwhelm administrative agencies, and often regulatees are in possession of critical data and information. Since new governance posits a collaborative, nonadversarial relationship between regulator and regulatee, the latter will presumably be more inclined to disclosure and forthrightness than to preservation of its interests at all costs against public interference. A similar point may be made with respect to third-party public interest groups, whose inclusion in a tripartite lawmaking process will bring still further perspectives to the fore.

For example, the European Union’s Maastricht Treaty vests employers and labor representatives with co-lawmaking powers pursuant to the “social dialogue” lawmaking process. While traditional lawmaking in the European Union proceeds exclusively from public authority, under the “social dialogue,” employers and pre-certified labor representatives can negotiate generally applicable policy agreements governing labor relations, workplace safety, vocational training, and other areas of E.U. “social policy” that, upon Council approval, have binding force as directives under

---

235. According to socio-legal research, absent unusually high recourse to detection and prosecution command-and-control techniques, regulatees are more likely to comply with regulations that are perceived as legitimate. See Ayres & Braithwaite, supra note 225, at 113 (“[C]onsiderable evidence indicates that participation in a decision-making process increases the acceptance and improves the execution of the decisions reached.”); Donald C. Langevoort, The Social Construction of Sarbanes–Oxley, 105 Mich. L. Rev. 1817, 1818 (2007) (“[C]ompliance decisions are based at least as much on the perceived legitimacy of the law and prevailing norms in local context as any deliberate risk calculation.”); Orly Lobel, Interlocking Regulatory and Industrial Relations: The Governance of Workplace Safety, 57 Admin. L. Rev. 1071, 1089–91 (2005) (describing limited success of top-down, command-and-control OSHA regulation in inducing compliance from firms); Scott & Trubek, supra note 229, at 8 (discussing the efforts of new governance to secure legitimacy and the effects legitimacy concerns have on new governance design). See generally Tom R. Tyler, Why People Obey the Law 3–7, 170–78 (1990) (analyzing the link between legitimacy and the perception of governmental authority).

236. See infra Part V.B.
E.U. law.237 Thus, nonstate stakeholders wield substantial power over the direction of the lawmaking process, though public actors remain in the process (the Commission initiates the discussion, and the Council ratifies the negotiated agreement between the social partners). In the United States, regulatory negotiation (or “reg–neg”) has emerged as an alternative to traditional Administrative Procedure Act (APA) administrative rulemaking.238 In a reg–neg, either industry representatives or the agency proposes formation of a committee pursuant to the Federal Advisory Committee Act. If the agency agrees to constitute a committee for a reg–neg, stakeholders—including, unlike the E.U.’s social dialogue model, any additional interested nonstate third parties—may petition the agency for inclusion in the negotiations. The agency has discretion over which interested parties are included in the process. At the conclusion of the negotiation, if a consensus rule is achieved it is then channeled through the normal notice and comment rulemaking process under the APA. In the cases of reg–neg and E.U. social dialogue lawmaking, the public involvement remains substantial, but nonstate actors contribute as collaborative co-policymakers in reaching a negotiated governance solution.239

Another example of a new governance tool involving increased participation of nonstate actors is the use of Habitat Conservation Plans (HCP) to permit “incidental takings” of wildlife. Pursuant to a 1982 amendment to the Endangered Species Act (ESA), Congress permitted


239. Professor Kimberly Krawiec’s discussion of “negotiated governance” is illustrative of this collaborative, contract-based conception of new governance. See Kimberly D. Krawiec, Cosmetic Compliance and the Failure of Negotiated Governance, 81 WASH. U. L.Q. 487, 516–22 (2003). Krawiec is skeptical of negotiated governance initiatives inasmuch as they, like all agreements, leave certain terms undefined. These negotiated agreements are “incomplete contracts” that invite opportunistic behavior among parties—particularly lawyers and compliance professionals—during renegotiation, which, in the regulatory context, occurs during the implementation and enforcement stages. See id. at 521.
businesses that would otherwise run afoul of the strict prohibition of “takings” of any animal designated by U.S. regulators as an endangered species to engage in “incidental takings” of endangered species if they submitted a satisfactory HCP. An HCP must include commitments from the applicant to mitigate damage and not appreciably reduce the likelihood of survival and recovery of the species in the wild, but the plan design is still largely left to the discretion of the applicant. After an HCP enters into force, the supervision of the permit holder’s compliance with the HCP may, but need not necessarily, be delegated to a third-party intermediary with an interest in policing compliance, such as a nonprofit conservationist group. In place of the strict prohibition—which sets up a winner-take-all political conflict between pro-growth and pro-conservationist camps—the HCP alternative opens up the possibility that pro-growth and pro-conservationist groups can participate in a collaborative process.

3. Flexible and Dynamic Law: Overcoming Rule–Principle Polarity

The end governance output of each negotiated regulatory deal is generally an ex ante rule or principle. Moreover, these ex ante laws are often generally applicable to all market participants without regard to the specific circumstances of regulated firms. When regulating dynamic and complex market behaviors, such ex ante laws will be either over- or under-inclusive, and will often fail to achieve their objectives. Worse still, they might unnecessarily exacerbate market complexity by motivating a regulatory arbitrage game in which regulatees develop technologies to avoid the effects of the ex ante law. Unless enforced simplicity is the favored solution, the challenge for lawmakers is to construct a regulatory system that is flexible enough to keep up with the dynamism of these regulated activities.

Much of the scholarship and popular discourse concerning the need for flexibility and dynamism in financial regulatory reform has occurred in the context of the rules–principles debate. While the rules–principles divide has already proven its substantial marketing value in jurisdictional


241. See Lee P. Breckenridge, Nonprofit Environmental Organizations and the Restructuring of Institutions for Ecosystem Management, 25 ECOLOGY L.Q. 692, 697 (1999) (noting a trend that nonprofit organizations, such as the Nature Conservancy, are taking an increasingly active role in monitoring HCPs).

242. This qualification is not trivial in the case of financial regulation. See supra notes 25–27 and accompanying text.
competition, its meaning is notoriously inexact and it is never an accurate description of an entire system. In most cases, it is more sensible to analyze regulatory reforms in dynamic financial markets according to a new governance framework and ask if lawmaking technologies, whether “based” in principles or rules, are being deployed effectively and in a flexible manner.

Professor Larry Cunningham has described three ways in which the rules–principles distinction is commonly understood. First, the analytical distinction refers to a temporal division between rules, the content of which is set out ex ante, and principles, the content of which is filled in ex post. Second, the conceptual distinction refers to a distinction between particularized, concrete rules and general, abstract principles. Third, the functional distinction refers to principles, which repose discretion in designated actors, and rules, which do not. All three ways of understanding the divide resonate with new governance scholarship. In particular, the analytical and conceptual rubrics shed light on the new governance attribute of flexible and dynamic lawmaking.

Professor Louis Kaplow has touched on the analytical and conceptual rubrics, theorizing that precise rules are costly ex ante, as their content is the result of an extensive deliberative or negotiated process. Applying that precept to the reg–neg and E.U. social dialogue processes discussed above in Part V.A.2, it is unobjectionable. On the other hand, general, abstract principles are costly ex post, when the content of the law must be implemented by regulators, interpreted by practitioners, and enforced by adjudicatory authorities. According to Kaplow’s formulation, it then


244. See Cristie L. Ford, Principles-Based Securities Regulation in the Wake of the Global Financial Crisis, 55 MCGILL L.J. (forthcoming 2010) (manuscript at 9) (available at Social Science Research Network) (“Rules and principles are also best understood as points on a continuum rather than discrete concepts, and there is a good deal of overlap and convergence among them.”).

245. See Cunningham, supra note 243, at 1417, 1420. Cunningham ultimately eschews the rules–principles distinction due to its unavoidable imprecision, and recommends that legal scholars and policymakers “retire” its usage.


247. There is some evidence that less time is required to fully deliberate on, and negotiate, rules under the reg–neg process than traditional APA rulemaking. See Cornelius M. Kerwin & Scott R. Furlong, Time and Rulemaking: An Empirical Test of Theory, 2 J. PUB. ADMIN. RES. & THEORY 113, 124, 134 (1992). That said, it is certainly true that the time and expense of producing an agreed-upon rule pursuant to reg–neg or a similar negotiated lawmaking process exceeds the amounts expended to draft an open-ended principle.
follows that, given economies of scale, “the greater the frequency with which a legal command will apply, the more desirable rules tend to be relative to standards.” In other words, when a regulatory scenario is likely to be frequently occurring (as with the “repeat player” regulator– regulatee encounters in the financial industry), it is preferable to invest in precise rulemaking upfront, rather than incur expenditures in frequent and shifting reinterpretations of principles, the content of which is filled in ex post.

Professor Cristie Ford elaborates on Kaplow, noting that the law drafter might lack adequate information to draft ex ante rules governing frequently applicable occurring transactions or events. Under such circumstances, it is therefore preferable that the content of laws—or, more broadly, the means and tools by which regulatory objectives are to be achieved—be flexible and dynamic. Therefore, synthesizing Kaplow and Ford, a highly complex and high-frequency regulatory scenario characterized by regulator–regulatee information gaps is susceptible to neither rules (because of information asymmetries) nor principles (because of the frequently recurring nature of the regulatory encounter), but more likely some mixture of the two. From a new governance perspective, though, this theoretic dilemma is easily remedied by focusing on the ways in which new tools, whether they consist of rules or principles, will achieve regulatory objectives in a flexible and dynamic market.

From a new governance perspective, it is more sensible to refer to many so-called “principles-based” regulatory systems as flexible systems that are open to diverse forms of articulation. One way to retool governance in these markets is to equip public administrative processes with structural mechanisms to engage in continuous learning and revision, monitoring and error detection, benchmarking, and peer review. An “experimentalist” strain of new governance theorists has emphasized that solutions in new governance regimes are provisional and should remain subject to constant

---

248. Kaplow, supra note 246, at 577.
250. See Ford, supra note 244, at 9.
251. See Lohel, supra note 221, at 391. Julia Black proposes a “conversational model of regulation” to remedy the inherent uncertainty, indeterminacy, and over- or under-inclusiveness of regulatory rules and principles. See Black, supra note 23, at 37–44. Professor Black notes that “[i]n conversation, the problems of generalizations and to an extent of open texture can be, and are, resolved by explanation and latitude in interpretation and understanding on the part of those participating.” Id. at 38. A conversational model of regulation raises accountability concerns that could be addressed by widening the conversational constituency to include other affected parties. See id. at 42–43; cf. Part V.B. The participants in the conversation will then constitute the “interpretive community” that collectively develop the content of regulation. See Black, supra note 23, at 37.
revision in light of observed experience. The experimentalist scholars apply insights from management theory and practice to public governance dilemmas. For example, the Lamfalussy rulemaking process governing E.U. financial regulation rulemaking provides for mandatory four-year sunset reviews of framework laws, which guarantees that the E.U.’s lawmakers revisit the assumptions of earlier laws. Moreover, the Lamfalussy process requires the E.U. lawmakers to consult repeatedly with so-called Level 3 committees comprising financial regulators from the member states. The Level 3 committees ensure consistent implementation, supervision, and enforcement of the E.U.-level rules by the member state regulators. One important new governance aspect of the Level 3 committee role is a mandatory peer review of member states’ implementation efforts, which forces the committees to evaluate implementation, share information, and even create ad hoc groups to address specific shortcomings. In their composite, these peer review interactions permit the identification, on a rolling basis, of best practices and create a forum to exert moral suasion on underperforming member states.

---

252. See Dorf & Sabel, supra note 211, at 286 (“The immediate instigation of our desig[n] for democracy is a series of innovations by private firms that suggest institutional devices for applying the basic principles of pragmatism to the master problem of organizing decentralized, collaborative design and development under conditions of volatility and diversity. The innovations, inspired by organizational breakthroughs in Japan, but no longer limited to Japanese firms or those in close association with them, are a response to markets that have become so differentiated and fast changing that prices can serve as only a general framework and limit on decisionmaking.”); Simon, supra note 4, at 6–11 (contrasting the “static optimization” managerial perspective with a “reliability” perspective of managers operating in dynamic industries that is more responsive to uncertainty and, therefore, a better model for financial regulation).


255. See Posner, supra note 253, at 47.

256. See id. at 55.

257. Cf. Dorf & Sabel, supra note 211, at 350–53 ( theorizing a rolling best practice
While Europe has been more creative in making use of experimentalist best practice and continuous learning regimes in the financial regulatory arena, U.S. regulators have recently demonstrated a greater openness to these alternatives. For example, in October 2009 the Federal Reserve announced a proposed supervisory initiative to conduct a horizontal review of compensation practices of twenty-eight “large, complex banking organizations” (LCBOs). The announcement listed several principles that the Federal Reserve expects will guide LCBO compensation practices. The policies implemented by LCBOs in response to the listed guiding principles will become subject to the ongoing safety and soundness review by Federal Reserve personnel. The Federal Reserve set forth a broad set of principles and left the articulation of the specific policies to the discretion of the LCBOs. As LCBO compensation practice evolves, the Federal Reserve will be able to identify best practices and establish them as benchmarks, and then press the lagging LCBOs during supervisor reviews to update compensation in light of their more innovative competitors. Here we see an oft-overlooked attribute of transparency: not only does it enhance democratic legitimacy, it is also a learning device. As firms and regulators learn more on a rolling basis, the regulatory landscape can and should adjust.

Another recent example of open articulation of law in the financial arena is the U.S. insurance regulators’ move to a “principles-based reserving” regime according to which life insurers will be able to estimate their liabilities with respect to policy and annuity reserves based in part on company specific experience and the results of internal stochastic models.

environmental regime that would require firms to ramp up environmental standards to meet the industry best practices, and explaining how such a regime would result in significant innovation as market participants would aim to out-innovate competitors, which would be saddled with additional expenses to ensure compliance with constantly evolving best practice standards; id. at 345 (“The chief purposes of administrative agencies in democratic experimentalism are . . . to set—again by a variation of benchmarking—regulatory standards for market actors; and to undertake such changes in their own activities and organization as cumulative self-scrutiny indicates will further these purposes.”).

259. Id. at 55,232–38.
260. Id. at 55,238.
261. See id. (“The review is designed to . . . [among other things, i]dentify emerging best practices through comparison of practices across organizations and business lines.”). Additionally, the guidance itself is open-ended: the Federal Reserve commits to “update this guidance as appropriate to incorporate best practices as they develop over time.” Id.
262. See Sabel & Simon, supra note 234, at 1072 (“Transparency is both an accountability norm and a learning device.”).
263. See Elizabeth K. Brill et al., Modernization of U.S. Insurance Regulation: Principles-Based
The Model Standard Valuation Law, some form of which is in effect in nearly all U.S. states, is based on a legislative model that has not changed for over 150 years. The “principles-based reserving” reform will allow the valuation of liabilities to evolve with advances in mortality risk modeling technology and in consideration of a firm’s experience when reserving for its own products.

B. Normative New Governance: Pro-Regulation, but Beware of Legitimacy Challenges

There is a normative pro-regulatory implication of new governance, and its techniques are instrumental tools to be deployed to achieve legitimate policy objectives. However, the tools of new governance pose significant normative challenges for governance stakeholders. Lester Salomon has referred to the “legitimacy challenge”\(^\text{264}\) posed by new governance. According to the traditional conception of regulation, there is a trade-off between flexibility and legitimacy.\(^\text{265}\) Increased private involvement in governance risks blocking the channels through which democratic legitimacy flows. Professor Orly Lobel describes the legitimacy challenge as follows: “A . . . challenge posed by the shift to a [new] governance model is striking a balance between the value of direct participation and the need for a high-quality representative democracy.”\(^\text{266}\) We have seen above in Part V.A.1 how retention of a benign big gun presence reinforces legitimacy, but what bulwarks protect the state from excessive influence by nonstate actors? It is here in the discussion of legitimacy that the danger of capture arises: if

\(^{264}\) See Salomon, supra note 211, at 38. Salomon also describes the “accountability challenge” and the “management challenge.” I read Salomon’s legitimacy challenge as, at bottom, a more radical version of the accountability challenge, as it goes to the compatibility of the new governance regime with our current system of government. Cf. Edward Rubin, The Myth of Accountability and the Anti-Administrative Impulse, 103 Mich. L. Rev. 2073, 2073, 2121 n.138 (2005) (distinguishing between accountability—“the ability of one actor to demand an explanation or justification of another actor for its actions and to reward or punish that second actor on the basis of its performance or its explanation”—and legitimacy, which “refers to the acceptability of a political regime in its entirety”); Jerry L. Mashaw, Accountability and Institutional Design: Some Thoughts on the Grammar of Governance, in PUBLIC ACCOUNTABILITY: DESIGNS, DILEMMAS AND EXPERIENCES 115, 120–22 (Michael W. Dowdle ed., 2006) (referring to legitimacy concerns, under Rubin’s definition, as questions of “legal accountability”).

\(^{265}\) Michael C. Dorf & Charles F. Sabel, Drug Treatment Courts and Emerging Experimentalist Government, 53 Vand. L. Rev. 831, 837 (2000) (“[E]mergent experimentalist bodies . . . point the way beyond a parallel dilemma that has long been taken to be a defining feature and limit of bureaucratic administration, the conflict between accountability and the flexibility required for effectiveness.”).

\(^{266}\) Lobel, supra note 221, at 453.
the administrative state can be captured under the traditional regulatory model, should we be all the more cautious that regulatees will use new governance as a Trojan Horse through which to achieve further agency capture?

Regulators need to be concerned not only with direct capture, but also with subtler, but related, threats to legitimacy. Professor Lobel describes one such danger as the risk that a new governance measure, because of its emphasis on ground-level information gathering from a larger array of stakeholders, will internalize the “illusion of information and transparency—that the information age, through its own mechanisms, can solve all problems.”\textsuperscript{267} Lobel warns in particular against the dominion of economic expertise, which, due in part to its putative quantifiability and quasi-scientific qualities, might come to dominate a new governance lawmaking process. The nonstate actors with the best access to, and the greatest ability to process and present, information might be expected to exert disproportionate influence over the information-based policymaking process.\textsuperscript{268} In reality, power, in both direct and hegemonic manifestations, is inevitably enmeshed in decisions about which information to credit, and how to organize information. The challenge, then, is to ensure that substantive policy deliberation still occurs (thus ensuring a proper democratic exercise of public power), and that the regulatory exercise is not reduced to the stamping of a state imprimatur on policy prescriptions based solely on data presented by particular interest groups. A related danger arises when a financial market adopts an idea, method or innovation purporting to streamline a process, or improve a matter of regulatory concern, that becomes “branded” such that its recurrent use is considered by stakeholders to be an authentic improvement. In financial market regulation, this “branding” phenomenon is likely to be most parous during periods of rising asset prices and corporate profitability, when, to borrow Robert Shiller’s phrase, “social contagion” attenuates counterparty market discipline\textsuperscript{269} and regulators are eager to avoid conflict. There may even be

\textsuperscript{267} \textit{Id.} at 455. Professor Lobel describes the “illusion of information and transparency” as possessing two components: “[f]irst, it elides the tension between the desire of a society to radically disperse decisionmaking” and the existence of qualified experts capable of making meaningful decisions; and second, it adopts the misguided notion that the more information that is disclosed and circulated, the higher the likelihood that stakeholders will converge on a single position.

\textsuperscript{268} Similarly, Professors Dorf and Sabel warn against the danger that underperforming entities might propose performance measures that “conceal more than they reveal,” in order to obstruct the proper functioning of an experimentalist benchmarking system. \textit{See} Dorf & Sabel, supra note 211, at 348.

\textsuperscript{269} \textit{See} ROBERT J. SHILLER, THE SUBPRIME SOLUTION: HOW TODAY'S GLOBAL FINANCIAL CRISIS HAPPENED, AND WHAT TO DO ABOUT IT 51–55 (2008) (discussing social
psychological tendencies to develop unconscious biases that privilege knowledge systems—such as those embedded in the new financial code—that purport to normalize and assimilate events and data that are in fact unexpected and unpredictable.\textsuperscript{270} If the human mind abhors an authentic encounter with uncertainty, we should be skeptical of crediting knowledge systems that tend to minimize uncertainty. The legitimacy challenge deepens further as financial regulators, believing markets to be self-correcting or relying on limited information, allow product innovation to proceed until a problem, along with the institutions involved in its production, becomes “too big to fail.”\textsuperscript{271} While not forms of capture in a strict sense, these risks can be thought of as a sort of soft, hegemonic capture whereby agencies behave as if captured without any direct expenditures by regulatees.

Nevertheless, as noted above, the New Deal-era formula of prescriptive rules drafted \textit{ex ante} and issued \textit{ex cathedra} will likely be ineffective to achieve stated regulatory objectives in a dynamic and complex market environment. This is certainly true with respect to the complexity paradigm of contemporary financial markets. We have seen above in Part II.B how complex financial innovation resulting from technology, globalization, and increased competition rendered the crude risk buckets of the initial Basel Accord arbitrary nearly from the outset. Certain activities have become too complex to be regulated or supervised by goal-pursuing regulators, no matter how expert; they will almost always be behind the curve.\textsuperscript{272} Moreover, even if regulations could be drafted with requisite precision, bureaucratic slippage in an environment with little margin for error (because in complex systems, small differences often make worlds of difference) would frustrate regulatory objectives.\textsuperscript{273} As a result, it is to some degree inevitable that those regulators charged with supervision of complex contagion phenomenon in financial markets).

\textsuperscript{270} Simon, \textit{supra} note 4, at 4–5.

\textsuperscript{271} See Kenneth C. Kettering, \textit{Securitization and Its Discontents: The Dynamics of Financial Product Development}, 29 Cardozo L. Rev. 1553, 1670 (2008) (“[A] financial regulator dealing with a product ‘too big to fail’ will tend to behave in much the same way as it would if it had been captured by the firms invested in the product . . . .”).

\textsuperscript{272} See Trubek & Trubek, \textit{supra} note 216, at 542 (“[A]s society becomes more complex and problems harder to solve, there is a need for more experimentation. Because stakeholders often have the requisite knowledge [to solve problems], increased participation becomes not only desirable, but also necessary.”).

\textsuperscript{273} See, e.g., William R. Freudenburg & Robert Gramling, \textit{Bureaucratic Slippage and Failures of Agency Vigilance: The Case of the Environmental Studies Program}, 41 Soc. Prob. 214, 214 (1994) (defining “bureaucratic slippage” as “the tendency for broad policies to be altered through successive reinterpretation, such that the ultimate implementation may bear little resemblance to legislated or other broad statements of policy intent.”).
regulatees operating in complex markets will turn to nonstate actors who are actually “on the ground” to supplement their understanding of the regulated market.274

The tripartite model of new governance can erect structural bulwarks against the risk that a regulator is captured by a regulatee that has been invited into a new governance process. The “social dialogue,” HCP, and reg–neg examples discussed above contemplate a tripartite model involving some participation from third-party interest group organizations in addition to regulators and regulatees.

Wisconsin’s Green Tier environmental program provides another example of a tripartite new governance structure. The Green Tier allows qualifying firms to “opt out” of much of the command-and-control framework, including the permitting process.275 A firm is eligible if it (1) constructs and commits to an “environmental management system” (EMS) self-regulatory regime that, in the estimation of state clean air and water regulators, is functionally equivalent to an ISO-certified system276 and (2) presents a plan that will ensure “superior environmental performance.”277 Green Tier firms are permitted to tailor their own solutions, embodied in the EMS, to regulatory objectives, and enter into a “participation contract” with the state environmental regulator that may specify certain derogations, on a case-by-case basis, from the command-and-control regulatory regime.278 The Green Tier program contemplates including third-party interest group organizations directly in the regulatory contract negotiation process (with appropriate rights of action against the regulatee)279 and the formation of the Green Tier advisory committee comprised of academics, municipal government officials, and representatives from industry and environmental groups.280

---

274. While the focus of this Article is on contemporary financial complexity, complexity is hardly a uniquely financial phenomenon. The factors impacting the complexity of financial markets (e.g., globalization, competition, technology) underlie analogous processes in other industries and cultural contexts.

275. See Trubek & Trubek, supra note 216, at 558.

276. See Wis. Stat. § 299.83(1)[g] (2009).

277. Id. § 299.83(3)[d]. “Superior environmental performance” is defined as “performance that results in measurable or discernible improvement in the quality of the air, water, land, or natural resources, or in the protection of the environment, beyond that which is achieved under [extant] environmental requirements.” Id. § 299.83(1)[g].

278. Wis. Stat. § 299.83(6)[j] (“The department shall ensure that the incentives provided under a participation contract are proportional to the environmental benefits that will be provided by the participant under the participation contract.”).

279. Wis. Stat. § 299.83(1)[e].

280. Wisconsin Department of Natural Resources, Green Tier Advisors, http://dnr.wi.gov/org/caer/cea/environmental/advisors/index.htm (last visited July 4,
The involvement of third-party stakeholders such as public interest groups to supervise and contribute to regulatory compliance enhances the legitimacy and effectiveness of a broad participatory regime and minimizes the risks of capture.\textsuperscript{281} Provided they are adequately resourced and deputized, these groups may improve governance in many other respects, including (1) contributing more constitutively in a dialogic governance process, rather than in a confrontational \textit{ex post} litigation capacity;\textsuperscript{282} (2) increasing the knowledge base on which policy is made by bridging information gaps and providing additional perspectives to be considered; (3) eliminating the risk of soft capture, whereby administrative agencies are captured without any lobbying expenditures on the part of regulatees;\textsuperscript{283} and (4) embedding in governance participation a counterweight to well-organized constituencies that more effectively mobilize political support, thereby enhancing policy from a republican standpoint as well. These groups become part of an “interpretive community” that collectively develop the content of regulation.\textsuperscript{284}

Such open participation of, and communication with, third-party interest groups not only polices the regulator–regulatee relationship, it also provides incentives for those groups’ representatives to remain faithful to the stated mission. In a regime characterized by confrontation and litigation, these groups often resort to pursuing symbolic, rather than tangible, rewards, especially when their constituency is diffuse and tangible rewards are only obtained with great difficulty.\textsuperscript{285} When the stakes and results of the regulatory process are open, it becomes more difficult to justify symbolic victories to represented constituencies. It goes without saying that not every group organized as “Citizens for X” is characterized by unbending fealty to X. It is therefore important that the public agency acts as a gatekeeper, inviting to the governance game only regulatees and

\textsuperscript{281} See generally AYRES & BRAITHWAITE, \textit{ supra} note 225, at 54–100 (elaborating a theory of tripartite governance involving regulators, regulatees, and third-party public interest groups).

\textsuperscript{282} See id. at 91–92 (“[F]ace-to-face negotiation will often transform confrontational disputes into accommodative encounters where the concerns of the other are internalized.”); cf. Dorf & Sabel, \textit{ supra} note 211, at 349 (“[T]here are some first signs that advocacy groups are in fact realizing that they have more to gain by participating in decentralized problem solving than by using strong-arm techniques to set limits on centralized decisions.”).

\textsuperscript{283} See AYRES & BRAITHWAITE, \textit{ supra} note 225, at 79–80, 90.

\textsuperscript{284} BLACK, \textit{ supra} note 23, at 30–37.

\textsuperscript{285} See Murray Jacob Edelman, \textit{The Symbolic Uses of Politics} 4, 22–43 (1985) (explaining how unorganized political groups can provide “symbolic reassurance” to their constituencies more reliably than tangible benefits for when aggregate political promises diverge from actual possible allocations).
authentically representative third-party interest group organizations.286

C. Internal Capital Models as Flawed New Governance

By studying the causes of regulatory reforms, we can understand their objectives and assess their effectiveness as regulatory tools. In the case of Basel II and the CSE Program, internal models were included in the capital adequacy framework as a response to growing complexity that amplified the mismatch between asset portfolios and the ex ante system of risk weights embodied in the initial Basel Accord. Part II.B.3 described the integration of internal market risk and credit risk models into the Basel II capital adequacy regime as an attempted response to increased complexity in banking institutions. In Part III, the adoption of the CSE Program for large U.S. investment banks was explained in part as a similar response to complexity. The internal models approaches to capital adequacy possess attributes of new governance initiatives, especially power-sharing, enhanced stakeholder participation, and flexible and dynamic lawmaking. They also retain a benign big gun in the form of residual command-and-control authority. However, as currently formulated, the utilization of internal models in capital adequacy regulation fails to address adequately the legitimacy challenges of new governance programs.

1. Enhanced Power-Sharing and Increased Participation, a Benign Big Gun, and Dynamic and Flexible Lawmaking

Each internal models approach involves increased participation of and power-sharing with regulatees. The capital charges for risks incurred are set by reference to criteria generated by the firms themselves. In the case of Basel II, the regulator provides the formula and the regulated banks furnish their own internal estimates of the credit risk inputs, and permits banks to calculate market risk and organizational risk on their own. Similarly, the CSE Program, which measures market risk, permits banks to utilize their own market risk calculations. Both Basel II and the CSE Program share power and enhance governance participation vertically by harnessing the information to which bank groups, but not regulators, have traditionally had access in order to address a complex regulatory objective.287

286. This tripartite mode of governance is surely more natural to those familiar with the European socio-legal order, with its rich history of social dialogue among institutions representing wide spectra of workers and employers, which antedates the development of the E.U. social dialogue governance discussed above by several decades.

287. See Cannata & Quagliarello, supra note 63, at 11 (“[S]ince the identification of the adequate combination of capital and risk is not an easy task and financial markets and products are increasingly complex, it is reasonable that any assessment starts up with the
“Basel Brand” of regulation is also characterized by an emphasis on the cross-learning that occurs as a result of enhanced regulator understanding of how regulatees actually operate. In this respect, Basel II invites comparison with cross-learning resulting from stock investments from capital-rich developing jurisdictions in sophisticated and complex financial institutions in capital-importing jurisdictions.\footnote{288}

In addition, the internal models approach introduces significant flexibility into the capital adequacy regime by aligning the capital charges in a more risk-sensitive manner with the dynamic asset portfolios of contemporary financial institutions. The models are plastic and may be adjusted to reflect market experience or results of model backtesting. The SEC hailed the CSE Program for its ability to “monitor for and act quickly in response to financial or operational weakness in a CSE holding company or its unregulated affiliates that might place regulated entities, including U.S. and foreign registered banks and broker dealers or the broader financial system at risk.”\footnote{289} For a firm operating under Basel II’s IRB approach, lead risk managers at the firm will be able to adjust the model assumptions directly through channels of corporate authority, rather than wait for a cumbersome notice–and–comment rulemaking or similar proceeding at the administrative level. Under Basel II, regulators are more concerned with process of model updating than the substantive details of any particular model. For instance, banking groups using the IRB approach must ensure that their models’ inputs are used in the business operating units, and that results of backtesting models are incorporated into the models. As a process-oriented reform, the internal models approach recalls Lester Salamon’s observation that new governance initiatives reconceptualize regulators as procedural arrangers seeking to achieve regulatory objectives.

Each reform also adopts, to varying degrees, a benign big gun approach. The Basel Committee only permits firms to utilize internal models to set their capital adequacy requirements if they meet, both initially and on an ongoing basis, certain qualitative and quantitative criteria. Moreover, bank

\begin{footnotesize}
\end{footnotesize}
regulators provide Basel II’s credit risk formulae, which neatly circumscribe
the authority of regulatee banks to set their own credit risk capital charges. Required backtesting for market risk VaR models with step-ups in capital charges for underperforming models similarly ensures a continuing regulator presence. In certain respects, the CSE Program, while increasing regulatee participation and power-sharing, also ramped up regulator involvement in U.S. investment banks by, for example, submitting them to consolidated supervision and requiring consent to SEC examination of all affiliates, regular reporting requirements to SEC, and mandatory provision of reports filed by affiliates subject to inspection by other regulators. Overall, though, the SEC’s authority and resources to police the new governance aspects of the CSE Program proved disastrously inadequate.\textsuperscript{290} Due to its abdication of effective enforcement, the CSE Program straddles the line between new governance and deregulation.

In summary, the internal models approach in theory holds potential for finally bridging the information gaps that impede regulatory understanding of firms’ risk profiles. It also might foster a more collaborative relation between regulators and regulatee firms, and minimize the tendency for firms to resort to capital arbitrage.\textsuperscript{291}

\section*{2. Internal Models and Normative Challenges to Legitimacy}

The limitations of using statistical models from the new financial code as tools to measure risk are manifold and well-documented,\textsuperscript{292} but their main inherent limitation is their inability to locate an \textit{ex ante} data set from which reliable conclusions may be drawn as to the probability and impact of future events, particularly large impact events.\textsuperscript{293} In other words, a predictive statistical model is necessarily self-referential, so it is incapable of supporting inferences with respect to future “fat-tail” phenomena, certain

\begin{footnotesize}
\begin{itemize}
\item \textsuperscript{290} See Coffee \& Sale, \textit{supra} note 187, at 741–44 (pointing out that the SEC assigned only three staff to each CSE firm and that the SEC technically lacked the basic authority to order a firm to increase its debt-to-equity ratio).
\item \textsuperscript{291} See \textit{supra} notes 20 and 23 and accompanying text.
\item \textsuperscript{292} Note that VaR-type models pose additional challenges unrelated to their accuracy as risk measurement tools, such as the so-called VaR negative feedback loop, which occurs when declines in an asset’s price cause firms utilizing modeling technologies to sell off that asset, exacerbating the downward pressure on the asset price as all firms attempt to sell.
\item \textsuperscript{293} See, e.g., Nassim N. Taleb, \textit{Common Errors in Interpreting the Ideas of The Black Swan and Associated Papers}, N.Y.U. POLY INSTITUTE (2009), available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1490769 (bemoaning what Taleb calls the “ludic fallacy”: namely, that statistical models assume a probability structure similar to closed games with a priori known probability when they often lack sufficient data from which to construct a probability distribution in the first place).
\end{itemize}
\end{footnotesize}
of which may lack any historical precedent data at all.\textsuperscript{294} The importation of statistical “certainty” into complex phenomenon like contemporary financial markets is especially problematic. Some have suggested that induction, and even cause-and-effect relationships, are unintelligible in certain complex environments characterized by circularity, self-referentiality, and unpredictability.\textsuperscript{295} Asset and credit markets, in short, are not like coin flips.\textsuperscript{296}

Moreover, all VaR models have mechanical limitations. Consider as an example the MRA internal models approach, which permits banks to use VaR models based on whatever simulations they prefer.\textsuperscript{297} Banks may choose among so-called historical simulation, Monte Carlo simulation, and variance-covariance techniques to measure VaR.\textsuperscript{298} A key assumption of the variance-covariance technique is a normal distribution of financial market returns, which is not only empirically false,\textsuperscript{299} but false with respect to arguably the primary justificatory purpose of the capital charge: namely, to ensure that banks had a capital buffer to withstand low-probability adverse developments (including, most importantly, non-normally distributed fat tail events). Though historical simulations have the advantage of avoiding the assumption of a normal distribution of returns, two significant problems exist with their use: first, accuracy is a function of the sampling period; and second, limiting inputs to historical data necessarily assumes that future extreme return distributions will not exceed

\footnotesize
\begin{itemize}
\item \textsuperscript{294} Cf. Riccardo Rebonato, \textit{Plight of the Fortune Tellers: Why We Need to Manage Financial Risk Differently} 252 (2007) (questioning utility of a 99.9 percentile confidence interval for one-year loss because of a lack of sufficient data points for 1000-year events from which patterns might be drawn); Ford, \textit{New Governance in the Teeth of Human Folly}, supra note 4, at 49 (discussing “Knightian uncertainty [as] a breeding ground for pathologies in decision making and human conduct”).
\item \textsuperscript{295} Taleb, supra note 293, at 5–6. A similar point is made with respect to the impossibility of actuarially pricing risks for purposes of setting risk-based deposit insurance premiums: “The risks of an ‘extreme’ event, in the form of a banking crisis resulting in massive losses to a deposit insurance fund, defy the sort of probabilistic quantification based on experience that insurers conduct to anticipate losses from insured events. In the face of this uncertainty, an insurer cannot calculate the resources it may need and thus cannot price efficiently.” Tarullo, supra note 14, at 26.
\item \textsuperscript{296} See Emanuel Derman, \textit{Models}, 65 Fin. Analysts J. 28, 32 (2009).
\item \textsuperscript{297} 1996 \textit{MARKET RISK AMENDMENTS}, supra note 146, at 44.
\item \textsuperscript{298} Id.
\item \textsuperscript{299} See Goodhart et al., supra note 15, at 80 (“[I]t is now well known that financial market returns are only imperfectly described by the normal distribution. In particular, the empirically observed frequency distributions have fat tails, i.e. large market movements occur more often than predicted by the normal [distribution].”); Shahin Shojai & George Feiger, \textit{Economists’ Hubris—The Case of Risk Management}, 28 J. Fin. TRA\textit{NSFORMATION} 27, 32 (2010), available at http://www.ssrn.com/abstract=1550622.
\end{itemize}
historical extremes. 300

Using internal models for regulatory capital adequacy purposes poses incentive-related problems as well. In the years leading up to the subprime credit crisis, most firms’ internal models had not addressed the possibility that house prices might decline. 301 This methodological error might not be casual; it is precisely the method one might expect a bank, cognizant of the safety net it enjoys, to take. 302 The bank’s model allows the bank to participate fully in the rising tide of financial profits. This recalls Professor Steven Schwarz’s discussion of systemic risk as a tragedy of the commons problem: because the benefits of exploiting finite capital allocations accrue disproportionately to financiers performing the allocation and the costs fall onto the broader class of users of finance, financiers lack sufficient incentives to internalize the externalities and their misjudgment of risk might cost the real economy in the long run. 303 To adapt the infamous phrase of former Citigroup CEO Chuck Prince, “If the music is still playing, the capital model will be made to dance.” Finally, this point on the incentive structure of banks applies on an intra-firm basis as well: profit center units have little interest in accurate risk profiling of activities that are likely to yield greater compensation. 304

300. Goodhart et al., supra note 15, at 82. On the other hand, adopting a longer-term observation period would include more data observation points, but would be unavailable for certain standard risk factors (if the risk factor has developed recently) and increase the likelihood that a return distribution has changed in recent periods, which could lead to biased VaR estimates.  See id. at 79.


302. Note that one need not maintain that banks were engaging consciously in fraudulent, “heads I win, tails you lose” behavior vis-à-vis taxpayers to ascribe to this theory. There may be a psychological bias in favor of unrealistic modeling assumptions. See supra note 270 and accompanying text. Credit rating agencies played an important role in the subprime credit crisis by providing a veneer of credibility to the models’ predictions. Even the Federal Reserve, as late as 2006, claimed that, notwithstanding historical price trends, the housing bubble was not worrisome because housing prices always tracked income growth (despite the fact that mortgage loan growth, which inevitably flowed through to housing prices, doubled that of GDP growth). See Testimony Before the Financial Crisis Inquiry Commission Hearing on the Financial Crisis, slide 3 (2010) (statement of Michael Mayo, Managing Director, Caylon Securities), available at http://www.cspan.org/Watch/Media/2010/01/13/HP/A/28382/Financial+Crisis+Inquiry+Commission+++Day+One.aspx (documenting 8%-21% loan growth in all real estate sub-sectors and 4.5% GDP growth, in each case during the 2000s).


304. Cf. Tarullo, supra note 14, at 101 (“The various bank divisions had little interest in promoting clear and well-developed risk profiles of their activities, since this might mean
For present purposes, it suffices to note that the inherent limitations of modeling technologies are present in all their uses, and the high-impact events against which we should guard most cautiously will almost never be susceptible to financial modeling. It is therefore imperative that any regulatory use of internal models is undertaken with full knowledge of these inherent limitations and a healthy dose of skepticism.

Notwithstanding the positive new governance attributes discussed above in Part V.C.1, the extant regulatory structures within which internal models are put to use present normative challenges typical of new governance reforms that accentuate the models’ inherent limitations. The retention of a benign big gun attenuates the immediate danger of outsourcing an important policy objective to largely self-interested regulators. However, relying on effective supervision of internal models’ adequacy opens up possibilities of literal or soft capture, especially with respect to bank regulators (like those in the United States) that are funded by assessments and fees from their regulators. The adoption of a new governance technique in response to financial complexity raises challenges to legitimacy and effectiveness, particularly with the forms of soft capture discussed above in Part V.C. As discussed above in Part II.B.3.i, that very complexity was facilitated by sophisticated technological code that is treated in Basel II and the CSE Program as the solution to the complexity problem.

Regulators resorted to internal models because there was no competing method to bridge the information asymmetry between banks and regulators. Banks tightly guard their propriety risk models, which constitute a significant store of value for risk-intermediating financial institutions. When a single group of stakeholders offers to bridge a wide information deficit through introduction of a new governance technology relying on proprietary information or methods, the challenge to effective policymaking and democratic legitimacy is acute. Recalling Lobel’s warning against the “illusion of information” (i.e., that privileged access to

more constraints on the very activities that—at least in the short term—were most likely to yields the highest profits,”); Steven L. Schwarz, Regulating Complexity in Financial Markets, 87 WASH. U. L. REV. 211, 224–25 (2009) (describing how risk management personnel had incentives to allocate assets to achieve low VaR estimates, notwithstanding the gradual build-up of massive fat tail risk that was not taken into account in the VaR model). In this context, it is also noteworthy that though most large financial groups had risk committees on their boards of directors, outside audits of risk management function (as opposed to the financial statements and internal controls functions) is not required. See Fanto, supra note 73, at 744–45.

and processing of information, especially economic information, can compromise new governance initiatives), the resort to internal risk models raises this unique danger of regulator-controlled economic information in the form of the new financial code.\textsuperscript{306} The danger is that this information will foster an unjustified perception of safety and soundness due to its putative quantifiability. When David Li, a J.P. Morgan researcher, published a paper in 2000 using a statistical technique known as the Gaussian copula model to measure default correlation among corporate bonds underlying a CDO, he dramatically enhanced bankers’ confidence in their ability to quantify the risks associated with each CDO tranche.\textsuperscript{307} As bankers relied more and more on Li’s method, they unwittingly created a new systemic correlation risk: if the model contained a faulty assumption, the effects of the error would be amplified considerably because all market participants would suffer losses. Li himself said of his model: “The most dangerous part is when people believe everything coming out of it.”\textsuperscript{308}

Moreover, the SEC’s uncritical adoption of the Basel Brand in the CSE Program further compromises the legitimacy of resorting to internal models to calculate capital requirements, inasmuch as regulators proved to suffer from the same social contagion as market participants, perhaps stoked by biases that encouraged the use of internal models that purported to minimize uncertainty.\textsuperscript{309} Even where inherently skeptical regulators attempt further inquiry into a model’s assumptions, a desire to avoid conflict will frequently inhibit effective supervision when the political cost of dissenting from the Basel Brand is high, as occurs during periods of rising asset prices, returns, and risks. The discussion in Part II.B.3.i.b regarding regulators’ largely uncritical embrace of banks’ ability to understand risk exemplifies the social contagion problem.

Finally, as noted above in Part V.C, financial technologies are vulnerable to “too big to fail” dynamics where regulators behave as though they have been captured. First, the rapid proliferation of internal models created a path dependency in that they became too important to bank competitiveness and to look to any other risk gauge could constitute a disruption of bank business that could pose systemic risks. The models themselves had become “too big to fail” and no other method of bridging the information asymmetries was feasible. Second, on an ongoing basis, the procyclicality of using VaR models to compute capital requirements

\textsuperscript{306} See Lobel, supra note 221, at 455.


\textsuperscript{308} Id. at 122.

\textsuperscript{309} See supra note 270 and accompanying text.
heightens the regulatory stakes: if regulators or other stakeholders do not recognize mismatches between modeled risk and actual risk at an early stage, regulators might be tempted to “forbear” prompt corrective action until the next administration so as to forestall a painful contraction of credit (in the process exacerbating its inevitable effects). 310

While no comprehensive empirical study has been conducted to date, it is safe to say these problems are not merely theoretical. In the years leading up to the recent subprime credit crunch, large international commercial banking groups had, for the most part, not transitioned fully to an internal model approach for capital adequacy purposes, though U.S.-based investment banks were subject to the CSE Program. U.S. commercial bank regulators adopted the IRB approaches in December 2007, but the final rule provided for: (1) a minimum sixty-day period of review of a bank’s implementation plan; (2) a four-quarter “parallel run” period, during which a bank was required to comply with all of the IRB prerequisite criteria; and (3) a minimum of three four-quarter “transitional floor” periods before which a bank could transition to a “stand alone” IRB approach. 311 The earliest a U.S. banking group can fully transition to the IRB, then, is April 2012. U.S. commercial banks—but not investment banks—were, however, operating under 3%–4% leverage ratio constraints under FIDICIA. European banks were subject to Basel II effective January 2008, well after the credit crisis was a foregone conclusion. Since Basel II was not in effect in the U.S. commercial banking sector, it can hardly be blamed for the capital collapse during the subprime credit crunch. However, the losses suffered by the industry lay some blame on the modeling technologies on which Basel II is based, inasmuch as they failed to understand the risks involved. 312

In 2006, Tim Geithner, then-president of the Federal Reserve Bank of New York, ordered a review of how well large banks measured their ability to withstand a severe market downturn; the results were not encouraging, as firms were failing to account for worst-case scenarios in their models. 313

310. Forbearance by regulators during the late 1980s contributed to the thrift debacle of the early 1990s. See William K. Black, Why Is Geithner Continuing Paulson’s Policy of Violating the Law?, HUFFINGTON POST, Feb. 23, 2009, http://www.huffingtonpost.com/william-k-black/why-is-geithner-continuing_b_169234.html (“[FIDICIA]’s premise was that regulatory discretion led to cover-ups of failed banks and excessive losses to the taxpayers.”).


312. At present, the models are not publicly disclosed, so it is impossible to individuate with precision the models’ effects.

In 2005, Swiss bank UBS had already transitioned to Basel II when it decided to pursue aggressively the mortgage securitization business. UBS’s risk management department utilized a model that assumed that the so-called "super senior" risk that UBS retained on its balance sheet in securitization transactions could only lose 2% of its value. Its Basel II capital requirements plunged and UBS piled on its super senior exposure. By 2007, it had $50 billion in exposure; not coincidentally, within two years UBS had booked about $50 billion in asset write downs and was shored up only by a series of private and public capital infusions. As Rodge Cohen of Sullivan & Cromwell put it, describing his eleventh-hour sessions with the Federal Reserve and bank executives during the week AIG was nationalized and Lehman Brothers filed for Chapter 11 protection: “If there is a single factor which is the principal source of what has happened, it is the absence of knowledge of how much risk is in the system, and where it was. I think those who were optimistic simply did not realize how much risk was there.” This conclusion implicitly supports the prescription that if regulators are to continue to use internal models in capital adequacy regulation, the current regulatory approaches require modification.

3. Pillar 3’s Mandatory Disclosure as an Incomplete Tripartite Governance Response to Legitimacy Challenges

Pillar 3 of Basel II attempts to address soft and direct capture concerns by fostering “market discipline” through disclosure requirements. Public disclosure transmits information concerning capital positions, risk exposures, and internal models to capital markets participants, other counterparties, and rating agencies, and therefore affects the access to and price of debt and equity capital as well as deposits/premiums. If recipients of the information possess the requisite sophistication to process and understand it, the capital markets can impose discipline on disclosing banks by increasing the cost of capital for, or diverting business from, firms that are perceived to be engaging in riskier activities. In this way, market

---

314. TETT, supra note 307, at 161–63.
315. Id. at 162.
316. Id. at 242.
319. The CSE Program did not require any public disclosure relating to internal models and therefore did not address the tension at all.
320. Cf. BANKING CRISIS, supra note 13, at 11–13 (explaining how market participants can create market information sources that complement banking supervision by increasing
responses to banks’ uses of internal models would contribute to what Robert Shiller has referred to as a “new information infrastructure” aspiring to “alter[] the social contagion and information cascades that underlie the formation of speculative bubbles.” 321 In particular, mandatory disclosure concerning the assumptions, mechanics, and backtesting results of internal models would subject the models themselves to outside scrutiny. 322 Provided the disclosure is meaningful, third-party organizations representing stakeholder constituencies affected by allocations of risk capital (e.g., depositors, capital markets investors, consumer credit customers, mortgage loan recipients) could subject the internal models to their own scenarios and tests, altering assumptions where necessary, in order to gauge institution-level risks. Moreover, because these interest groups represent wide arrays of actors, they are well positioned to examine the aggregate assumptions of bank industry capital modeling practices to test for correlations that could pose systemic risks. That is, to the extent that banks and regulators as a group are failing to focus on a particular distributive contingency—because, for example, banks enjoy limited liability and regulators have been captured—an unaffiliated interest group particularly affected by the contingency could test and publish a model assuming an adverse course of events. Such an approach should foster both legitimacy and also the quality of the models. 323 Third-party surveillance of the internal models represents an attractive tripartite governance solution in which regulators, regulatees, and affected third-party stakeholders collaborate.

A tripartite regime based on mandatory disclosure leaves unresolved the problem of what to do with the disclosed information. Interest groups can run and publish their own tests, modifying assumptions and data sets, but unless they have standing to object to the internal models used by banks, such a regime will rely on regulator suasion. If nothing else, a bevy of ever-skeptical third-party interest groups will raise the stake (the “price”) of regulatory capture or forbearance: it will be harder to capture or deceive

information flows to regulators as well as the costs of forbearance, and could even serve as triggers for regulatory intervention).

321. Shiller, supra note 269, at 121.
322. A mandatory disclosure regime should also disrupt intra-firm politics by shifting power away from senior executives and towards risk management personnel, the sole firm constituency with the operational expertise to fulfill the firm function now subject to disclosure. Cf. Sabel & Simon, supra note 234, at 1078–79 (explaining how a liability determination in a case against a public agency, by mandating enhanced transparency, tends to decentralize power within the agency and empower “lower tier workers” who receive “increased discretion to cope with contingencies with which they are most familiar”).
323. See supra note 262 and accompanying text.
bank regulators that, in possession of alternative scenario tests, are adequately informed and subject to scrutiny. In most cases, though, the represented third-party interest will be widely dispersed, so it is not reasonable to expect that such interest groups will marshal their resources as effectively as regulated financial institutions to engage regulators in a debate about internal models. Interest groups could be formally deputized with some form of administrative standing to challenge a model’s assumptions, or institutionally resourced with public financial support. Credit rating agencies, which have extensive modeling experience and have even begun to incorporate consideration of internal capital models into their financial strength ratings, could be enlisted to review capital models.\textsuperscript{324} Provided the banking groups themselves are not funding the rating agencies, the familiar rating agency conflicts of interest would not be expected to hamstring such a system. Alternatively, interest group participation could be formalized by a thirty-day comment period during which groups could submit comments or modifications to a bank’s internal models.\textsuperscript{325} The bank’s regulator would then be required to explain publicly why it did or did not adopt the proposed revisions to the bank’s regulatory capital requirement. In this way, capital adequacy regulation would borrow tripartite new governance features from, e.g., reg-neg and the Green Tier Program.

Though Pillar 3’s emphasis on disclosure of internal models is a welcome addition to the bank regulatory toolkit, its requirements are at present too limited to foster an authentic tripartism that can credibly protect the capital adequacy regime against legitimacy challenges. As mentioned above, under Pillar 3 banks provide general descriptions of the market risk and operational risk models, and more particularized information concerning the “definitions, methods, data, and assumptions” used in the credit risk

\textsuperscript{324} Large internationally active bank competitors domiciled in foreign jurisdictions (and their regulators), too, would have incentives to monitor the use of internal models to ensure their local industries are not competitively disadvantaged by lax enforcement in other jurisdictions. While monitoring from other jurisdictions might improve the quality of global banking regulation, it would not necessarily conduce to the public objectives of a particular jurisdiction’s bank regulatory apparatus, which is concerned with the safety and soundness of the domestic institutions. In recognition of these cross-jurisdictional conflicts of interest, the Basel Committee has established Accord Implementation Group and a Capital Interpretation Group to foster uniformity in the implementation and interpretation of Basel II. Notwithstanding the domestic focus of regulation, foreign banks and regulators would be expected to contribute to the discussion, and raise the international political costs of domestic regulators endorsing flawed models.

\textsuperscript{325} Cf. Dorf & Sabel, supra note 211, at 349 (proposing that “authorizing legislation would confer on aggrieved citizen users a statutory right to participate” in administrative processes resulting in benchmarking and standard setting).
models, subject to the qualification that Pillar 3 “does not require a detailed description of the model in full—it should provide the reader with a broad overview of the model approach, describing definitions of the variables, and methods for estimating and validating” the model inputs.326 This exception threatens to swallow Pillar 3’s general disclosure rules, as it provides banks with a justification to avoid disclosure of sensitive proprietary information. Pillar 3’s fatal flaw as a tripartite new governance tool is its overemphasis on the proprietary aspect of the models and an underemphasis on their informational aspect.

The current disclosure requirements did not come without protestations from banks, which lamented the required disclosure of confidential and proprietary information.327 The Basel Committee recognized that any disclosure regime must balance the benefits of increased disclosure against the utility created by protecting proprietary information. According to the Committee, this information, if shared with competitors, “would render a bank’s investment in these products/systems less valuable, and hence would undermine its competitive position.”328 We have seen earlier how banks’ VaR and other modeling technologies did not merely facilitate the business of banking; they had become part and parcel of the business of banking, and comprised a key competitive asset and repository of value for banking groups.329 The same is true with the internal models used in setting regulatory capital requirements, which borrowed technology directly from firms’ VaR models. In particular, mandatory disclosure of proprietary models estimating credit risks strikes at the basic value proposition of banks: their effective intermediation between savings and investment. The introduction of widespread statistical modeling (and, as a result, securitization) in the banking business has shifted the focus of credit intermediation from qualitative on-the-ground assessment of debtors’ likelihood-of-repayment to behavioral and economic predictions based on

326. Basel II Document, supra note 162, at 235 n.207; cf. Agency Information Collection Activities; Submission for OMB Review; Joint Comment Request, 73 Fed. Reg. 4222 (Jan. 24, 2008) (stating that “certain summary information would be made available to the public for reporting periods after a bank has qualified to use the advanced approaches for regulatory capital to provide a sufficient degree of public disclosure to market participants”) (emphasis added).


329. See supra note 72 and accompanying text.
aggregated pools of debtor data. Statistical models are, in short, *what banks do anymore*. It might be possible to accommodate bank concerns about proprietary interests by embedding modeling software into interactive disclosure interfaces, which would permit other parties to adjust assumptions and parameters to determine the outputs a model would yield. Regulators could require the use of eXtensible Business Reporting Language (XBRL) to present interactive data, much like the SEC currently does for some reporting companies, mutual funds, and rating agencies. The challenge would be to design a compromise XBRL-type disclosure system that neither permits third-parties to deduce proprietary models nor lacks information required to assess a model’s performance.

It is at least partially true that mandatory disclosure of risk models will erode competitive position of disclosing banks vis-à-vis two sets of competitors. First, while large U.S. banks are required to adopt Basel II’s IRB approach, other jurisdictions offer IRB as an option. As such, mandatory internal model disclosure could impose a regulatory tax on U.S.-based banks. Second, the IRB approach is designed, in all jurisdictions, to apply to large internationally active banks. Accordingly, a free-rider problem might arise as smaller banks adopt modeling technologies developed by larger competitors subject to IRB. A different but related danger arises if large internationally active banks from non-U.S. jurisdictions eschew a disclosure-enhanced Basel II altogether, resulting in social loss through frustration of IRB’s new governance attributes.

Notwithstanding these potential costs of enhanced disclosure, the legitimacy challenge, discussed above, poses its own substantial economic costs and democratic deficits. I propose a new emphasis on tripartism through fuller disclosure, on the grounds that third-party interest groups representing other constituencies will (1) discipline banks by subjecting their internal models to enhanced scrutiny, thereby attenuating risks of literal and soft capture, and (2) experiment with new modeling assumptions, techniques, data sources and scenarios to improve the accuracy of the models.

**CONCLUSION**

New governance theory provides an analytical framework to assess financial regulatory reform initiatives designed to regulate and manage the complex and dynamic networks of risks in which financial institutions put our surplus capital to work. New governance is always oriented around the problem of complexity; it sees in it a source of opportunity, but also danger. It shines light on normative challenges to the legitimacy of efforts to involve regulatees and third-parties more directly in public regulation of complexity, such as heightened risks of literal and soft capture. This Article
poses as a complexity problem the structural dilemma of capital adequacy regulation resulting from the disjoint between its traditional prescriptive risk-weighting approach and the actual risks financial institutions face. Capital adequacy is not unique in this respect; the complexity paradigm of contemporary finance is characterized by widening information asymmetries between regulators and regulatees generally. In the face of dizzying complexity, the tools and methodologies of traditional regulation begin to appear arbitrary. I argue that a new governance perspective addresses this structural dilemma of financial regulation as a problem of regulatory technology: that is, how to harness new modes of interaction between public and nonstate actors without calling into question the legitimacy of public administrative goals?

Like most technologies, new governance tools often have unintended consequences. Basel II’s internal models approach, which privileges proprietary interests in new financial code (i.e., the internal models) over an open discussion over the models as tools of public governance, is a case study in the tendency of soft capture to compromise a new governance initiative. In the rush to harness the socially productive financial code, the internal models approach embraces the illusion of information efficiency embodied in the Basel Brand. Future new governance inquiries into financial regulatory reform must focus on soft capture and safeguard against the possibility that, like the internal models approach, in the rush to bridge information asymmetries we adopt a reform that creates more problems than it solves.

Because new governance theory sees complexity as both a source of opportunities and dangers, new governance financial regulatory scholarship should always remain critical of complexity itself. When regulating complex financial systems, the alternative is never between arbitrary regulation on the one hand and deregulation, “voluntary self-regulation,” or flawed new governance on the other. Regulators and policymakers need to consider a third alternative: using command-and-control techniques (e.g., maximum LTV ratios, bans on trading activities, central clearing and collateral posting for derivatives) to shape the regulated market dynamics so that regulation is not rendered arbitrary from the outset by rapid and volatile market changes. Depending on the nature of the market’s complexity and the soft capture threats posed by new governance techniques, in many cases public objectives (e.g., credit formation with an acceptable risk tolerance) will be more effectively achieved by ramping back complexity itself. In this respect, new governance—as applied to financial regulation—resonates with skeptics of neoclassical economic explanations

---

330. See supra notes 25–27 and accompanying text.
of complexity as an instrumental phenomenon in pursuit of capital arbitrage or customer demands. In the real world, complexity entails significant costs, particularly for the “greater fool,” whether an investor or a regulator, on the wrong end of the information asymmetry.331

* * *
COMMENT

IN NAME ONLY: EMPLOYEE PARTICIPATION PROGRAMS AND DELEGATED MANAGERIAL AUTHORITY AFTER CROWN CORK & SEAL

JOSEPH D. RICHARDSON*

Table of Contents

Introduction .............................................................................................................. 872
   I. What it Means to Deal With Under § 2(5) of the NLRA 874
      A. The Statutory Framework of Prohibited Employer Conduct Vis-à-Vis Employee Participation ......................... 876
      B. Dealing With Encompasses More than Bargaining. ............. 878
      C. The Managerial Authority Exception to Dealing With .......... 879
      D. The Board Defines Dealing With ........................................... 884
      E. Crown Cork & Seal: Redrawing the Lines of the Managerial Exception ................................................................. 887
   II. The Delegated Managerial Authority Exception Since Crown Cork & Seal .............................................................................. 891
      III. The Consequences of Crown Cork & Seal ......................... 893
         A. The Crown Cork & Seal Managerial Authority Exception is Imprecise ................................................................. 894
         B. Unbalanced Expansion of Employee Involvement
            Undermines § 8(a)(2) of the NLRA ........................................ 895
      IV. A New Test for Delegated Managerial Authority......... 902

Conclusion ............................................................................................................. 904

* J.D. Candidate, 2011, American University Washington College of Law; B.A., 1999, St. John’s College. I am indebted to Harold Datz for giving me the benefit of his insight and experience. Special thanks are due to my editor, Caitlin Mandel, and to senior editor Lauren Clark for their hard work. I thank my wife, Jessica, for her steadfast love and support. Most importantly, I’d like to thank my father, Charley Richardson, without whom I simply could not have written this, and whose commitment to justice for working men and women will always humble and inspire me.
INTRODUCTION

Seventy-five years ago, Congress passed the National Labor Relations Act (NLRA or the Act),\(^1\) regulating the often violent and economically disruptive relationship between workers and management, characterized by strikes, lockouts, and similar manifestations of industrial strife.\(^2\) Through the Act, Congress sought in part to prohibit employer-dominated labor organizations,\(^3\) known as “company unions,” which were seen as unfairly impeding the effectiveness of organized labor.\(^4\) The prohibition extended to many—but not all—of the less formal organizations through which employees communicate with employers and participate in workplace governance. Such organizations, sometimes referred to as employee committees or employee involvement programs, have undergone a resurgence in recent years as employers try to find new ways to leverage the intellectual capital of their employees in a global economy.\(^5\)

Proponents of the NLRA believed that cooperative relationships between employers and workers could only be maintained by protecting the equality of bargaining power between labor and management and the freedom of both parties from restraint by the other.\(^6\) One way the Act protects that equality is by reserving a broad range of representative conduct exclusively for independent unions and prohibiting competition from employer-dominated representation organizations. The Act accomplishes this by encompassing within the term “labor organization,” under § 2(5),\(^7\) any entity that engages in a process of dealing with

---

2. See, e.g., Labor Disputes Act, S. 2926, 73d Cong. § 2 (1934), reprinted in 1 NLRB, LEGISLATIVE HISTORY OF THE NATIONAL LABOR RELATIONS ACT, 1935, at 1 (1959) [hereinafter LEG. HIST.] (“Inadequate recognition of the right of employees to bargain collectively through representatives of their own choosing has been one of the causes of strikes, lockouts, and similar manifestations of economic strife . . . .”).
4. See 78 CONG. REC. 4229–30 (1934) (article by Sen. Wagner), reprinted in 1 LEG. HIST., supra note 2, at 22–23 (“At the present time genuine collective bargaining is being thwarted immeasurably by the proliferation of company unions.”).
6. 78 CONG. REC. 4229 (1934) (article by Sen. Wagner), reprinted in 1 LEG. HIST., supra note 2, at 22.
management about statutorily delimited workplace issues. The Act does not define dealing with, but the National Labor Relations Board (NLRB or the Board) has construed it to encompass a range of employer–employee interactions, the scope of which has evolved over the years.

Some have criticized the apparent inconsistency of the Board’s evolving interpretation of dealing with. Undeniably, the scope of permissible employer conduct related to employee participation has taken several evolutionary leaps, each marked by key decisions from the Board and the courts. However, the flexibility of dealing with under the Act—and thus the scope of permissible employer conduct relative to organized employee participation—is limited, and the Board can change its construction of the term only so far before it ceases to be a reasonable interpretation of the statute.

*Crown Cork & Seal,* marked the most recent leap forward in the Board’s dealing with doctrine, and arguably opened the door for significant expansion of employer-influenced employee participation programs.

---

8. Section 2(5) of the National Labor Relations Act (NLRA) provides: “The term ‘labor organization’ means any organization of any kind, or any agency or employee representation committee or plan, in which employees participate and which exists for the purpose, in whole or in part, of dealing with employers concerning grievances, labor disputes, wages, rates of pay, hours of employment, or conditions of work.” *Id.*

9. The National Labor Relations Board (NLRB) is composed of five members, appointed by the President to five-year terms, except that an appointee chosen to fill a vacancy will only serve the remainder of the unexpired term of the member whom he shall succeed. *Id.* § 153. The NLRB is charged, *inter alia,* with determining bargaining units, conducting and certifying elections, and preventing unfair labor practices as defined under the Act. *Id.* §§ 159(b), 159(c), 160(a).


Comment contextualizes the Crown Cork & Seal managerial authority exception to dealing with, concluding that it is, analytically, a radical departure from prior Board decisions that potentially disrupts the equality of bargaining power provided for in the Act. Part I examines the meaning of dealing with under § 2(5), giving a brief history of the somewhat convoluted line between unlawful employer–employee communication mechanisms and permissible delegation of managerial authority to employee committees. Part II examines employee participation cases since Crown Cork & Seal, analyzing the effect of that decision on the landscape of the dealing with doctrine. Part III presents two criticisms of the Crown Cork & Seal delegated managerial authority test: that it is not sufficiently clear to provide meaningful guidance to the agency, courts, or employers; and that it upsets the balance reflected in § 8(a)(2) of the NLRA. Part IV proposes a test for dealing with that would clarify the doctrine, enabling both management and labor to structure employee committees with greater confidence in their legality under the Act. The Conclusion urges the Board to adopt the test proposed in Part IV, imposing meaningful limitations on the managerial authority exception and restoring the important balance reflected in § 8(a)(2).

I. WHAT IT MEANS TO DEAL WITH UNDER § 2(5) OF THE NLRA

The concept of employee participation encompasses a broad range of employer–employee interactions, many of which fall outside the ambit of dealing with under § 2(5). For example, employer–employee interactions that occur between a single employee acting alone and the employer do not come under § 2(5). Before examining the statutory and historical frameworks of dealing with, it is helpful to have a general understanding of how groups of employees can interact with their employers around workplace decisionmaking. One way to do so is to divide the range of potential employer–employee interactions into three classifications or models: final decisionmaking, unilateral information flow, and bilateral exchange.

First, an employer can delegate to an employee committee the power to make final decisions on issues raised before the committee by individual employees. An example of this is an employee committee that issues final rulings on employee disciplinary actions. This type of committee does not

16. See § 152(5) (encompassing an agency, committee, organization, or plan but not individual employees); Gen. Foods Corp., 231 N.L.R.B. 1232, 1233 (1977) (finding no dealing with where employees in a team meeting raised issues with management representatives because the employees acted as individuals and were not represented by the team).
generally deal with management because in the exercise of its
decisionmaking authority it need not seek approval from—and thus interact
with—management at all.\textsuperscript{17}

The second class of employer–employee interactions is a unilateral,
bottom-up information flow, where employees make suggestions and
management takes that information without making a response to it,
although it may choose to incorporate those suggestions into its operations.
A common example of this is an employee suggestion box. This practice is
not dealing with, because there is no exchange between the employer and
employees; in this model the information flows in only one direction.\textsuperscript{18}

The third class of employer–employee interactions involves a bilateral
exchange between the employer and the employee organization.\textsuperscript{19} An
element of this is an employee committee that recommends disciplinary
action be taken against an employee, or that the employer change
workplace safety policies, but requires management review and approval
before it can implement its recommendation. This process of exchange
distinguishes bilateral exchanges from second-class upwards information
flow because here management demonstrates responsiveness to the
employee committee. This class comes within the ambit of dealing with
because the employer and the employee committee are engaged in a
process that, although perhaps lacking the formality of collective
bargaining, begins to closely resemble—and potentially compete with—the
representative function of a union.\textsuperscript{20}

Over the years, the Board’s construction of dealing with under § 2(5) has
traced a somewhat meandering line between the permissible employer
conduct of the first two classes and the employer conduct of the third class,
which risks running afloat of the Act’s prohibition of employer-dominated

\textsuperscript{17} See 231 N.L.R.B. at 1235 (finding no dealing with where employee committees issued
final decisions, exercising “flatly delegated” managerial authority).

\textsuperscript{18} Cf. E.L. du Pont de Nemours & Co., 311 N.L.R.B. 893, 894 (1993) (noting that
where an employer gathers information from an employee and simply does what it wishes
with that information, the committee is not a labor organization under § 2(5)).

\textsuperscript{19} See id. (describing “dealing” as a “bilateral mechanism” that “entails a pattern or
practice in which a group of employees, over time, makes proposals to management,
management responds to these proposals by acceptance or rejection by word or deed, and
compromise is not required”).

\textsuperscript{20} Particularly troubling to some management lawyers, the definition of a bilateral
mechanism in Electromation, discussed infra, left open the potential that the Board could find
dealing with where management implements suggestions that come from a committee, since
such an interaction could reasonably be characterized as acceptance or rejection by word or
deed. See Pauling & McGuire, supra note 10, at 226 (noting that Electromation left employers
confused by the “subtle and often elusive” distinction between lawful information sharing
and unlawful proposal-and-response).
labor organizations. The Sections that follow will review the statutory framework under which dealing with is prohibited and examine shifts in the Board’s construction of dealing with since the enactment of the NLRA, contextualizing current trends in the Board’s approach to employee involvement.

A. The Statutory Framework of Prohibited Employer Conduct Vis-à-Vis Employee Participation

Under § 8(a)(2), an employer may not “dominate or interfere with the formation or administration of any labor organization or contribute financial or other support to it.”21 In determining whether an employer violates § 8(a)(2), the Board’s inquiry is two-fold.22 First, the Board inquires whether the entity involved is a “labor organization” as defined in § 2(5).23 If it is, then the Board inquires whether the employer’s conduct vis-à-vis that labor organization constitutes domination or interference with the organization’s formation or administration, or unlawful support of the organization.24

Section 8(a)(2) preserves a balance between traditionally less-powerful employees and more-powerful employers, proscribing employer interference with employees’ right to organize in representative organizations independent from the employer.25 Even before the enactment of the NLRA, the Supreme Court recognized that employer-dominated unions “obstruct[] self-organization of employees and their choice of their own representatives for the purpose of collective bargaining.”26 While the term company-dominated union—describing employer-controlled entities that usurped the role of independent unions—

23. See id. (identifying the elements of proof for a § 2(5) labor organization as: (1) employee participation; (2) a purpose to deal with employers; (3) concerning itself with conditions of employment or other statutory subjects; and (4) if an employee representation plan or committee is involved, evidence that the committee is in some way representing the employees).
24. Id. This Comment focuses on the Board’s analysis in the first step: whether the employee committee is a labor organization under § 2(5).
25. See 78 CONG. REC. 4230 (1934) (article by Sen. Wagner), reprinted in 1 LEG. HIS., supra note 2, at 22 (declaring that a fair balance between industry and labor “can be accomplished only by cooperation between employers and employees, which rests upon equality of bargaining power and the freedom of either party from restraints imposed by the other”).
suggests a formal labor organization, in fact the most prevalent form of company-dominated union was not a formal labor organization, but rather a “loose organization, if you can call it an organization, that has no members, no dues, that [was] merely a method of electing representatives.”

Congress intended the term labor organization to encompass not only formal unions, but also a wide range of employee committees and plans. By doing so, it ensured that § 8(a)(2) would reach a broad range of processes through which management exercises control over groups that purported to represent the interests of employees. Congress intended the Act to prohibit such organizations, known as “representation committees” or “plans,” indicating that even the simplest representation processes are covered under the term labor organization. The Board, backed by the Supreme Court, indicated early on that the Act would be construed broadly to prohibit employers from establishing or supporting representative organizations that could interfere with employees’ right to self-organize.

By using the term dealing with, instead of the narrower term bargain collectively, Congress indicated that it intended to cover a broad range of employer activity. Over the years, rulings by the Board and the courts...
have varied the effective scope of the term *dealing with* in § 2(5), and thus the range of organizations protected from employer domination by § 8(a)(2). The Sections that follow will examine some of those decisions, framing the context of recent developments in the *dealing with* doctrine.

**B. Dealing With Encompasses More than Bargaining**

As early as 1959, the Supreme Court indicated that the term *dealing with* applied to a broad range of representational conduct. In *NLRB v. Cabot Carbon*, the Court declined to exempt certain employee committees from § 8(a)(2) and rejected the argument that under the Labor Management Relations Act of 1947 (LMRA), *dealing with* encompassed only bargaining with employers in “the usual concept of collective bargaining.” The Supreme Court found no support in the LMRA for narrowing the

between “*dealing with*” and “*bargaining collectively*”) (citing *Comparison of S. 2926, 73d Cong. (1933) and S. 1958, 74th Cong. (1934)*, reprinted in 1 LEG. H. SR., supra note 2, at 1319, 1347).

34. *Id.* In *Cabot Carbon*, the employer established an employee committee at each of its plants to “provide a procedure for considering employees’ ideas and problems of mutual interest . . . .” *Id.* at 205–06. The committees handled employee grievances, submitting findings and recommendations to a department head, who issued a decision on the recommendation, appealable up to the General Manager. *Id.* at 206 n.3. The committees also met monthly with management to consider issues including “safety; increased efficiency and production; conservation of supplies, materials, and equipment; encouragement of ingenuity and initiative;” and employee grievances. *Id.* at 205 n.2. The committees made—and discussed with plant officials—proposals with respect to many “aspects of the employee relationship, including seniority, job classification, job bidding,” timecards, overtime, “a merit system, wage corrections, working schedules, holidays, vacations, sick leave, and improvement of working facilities and conditions.” *Id.* at 207. While the committees’ jurisdiction clearly included wages, hours, and conditions of work, none of the committees attempted to negotiate a collective bargaining agreement. *Id.* at 209.


36. *Cabot Carbon Co. v. NLRB*, 256 F.2d 281, 285 (5th Cir. 1958), re’id, 360 U.S. 203 (1939). The Fifth Circuit broke with the traditionally broad reading of § 2(5) under the NLRA, relied on the legislative history of the Taft–Hartley Act amendments to § 9(a), which clarified that rights of individual employees or groups of employees to present grievances to their employer and to have those grievances adjusted so long as the adjustment is consistent with the terms of a collective bargaining agreement then in effect, and construed § 2(5) to apply only to organizations within the collective bargaining process. *Id.*
traditional broad construction of dealing with.\textsuperscript{37} The Court also rejected the employer’s argument that there was no dealing with when management exercised final decisionmaking authority on what amounted to recommendations from the committees, holding that the retention of final decisionmaking authority is “true of all such ‘dealing,’ whether with an independent or company-dominated ‘labor organization.’”\textsuperscript{38}

\textit{Cabot Carbon} established that the term dealing with encompasses more than the back-and-forth negotiation of collective bargaining.\textsuperscript{39} However, \textit{Cabot Carbon} did not specify which other forms of employer–employee interaction the term would cover. In a series of decisions, the Board addressed this question, defining and carving out exceptions to dealing with under § 2(5).

\textbf{C. The Managerial Authority Exception to Dealing With}

In the 1970s, the Board carved out an exception to the broad construction of dealing with upheld by the Court in \textit{Cabot Carbon}, protecting management’s prerogative to delegate its managerial decisionmaking authority to employee committees.\textsuperscript{40} In this line of cases, the Board clarified two aspects of the delegated managerial authority exception: first,

\begin{flushright}
\footnotesize
\textsuperscript{37} See \textit{Cabot Carbon}, 360 U.S. at 211–12 (noting that when considering the NLRA Congress declined to adopt the Secretary of Labor’s proposal that “bargaining collectively” replace the “dealing with” language in § 2(5) and that the Taft–Hartley Act re-enacted § 2(5) without changes). A proposed amendment to the Taft–Hartley Act would have permitted employers in nonunion workplaces to form employee committees for the purpose of discussing subjects covered by § 2(5), however this amendment was not included in the final Act. \textit{Compare H.R. 3020, 80th Cong.} (as passed by House, Mar. 24, 1947), \textit{reprinted in 1} \textsc{Leg. Hist. of the LMRA}, supra note 35, at 158, 183, \textit{with H.R. 3020, 80th Cong.} (as amended by Senate, May 13, 1947), \textit{reprinted in 1} \textsc{Leg. Hist. of the LMRA}, supra note 35, , at 226, 244 (1948).

\textsuperscript{38} 360 U.S. at 214 & n.15 (quoting \textit{NLRB v. Jas H. Matthews & Co.}, 156 F.2d 706, 708 (3d Cir. 1946) (“Final decision is always with management, although when a claim is made by a well organized, good sized union, management is doubtful more strongly influenced in its decision . . . .”)).

\textsuperscript{39} See \textit{id.} at 214 (noting that employer–employee interactions need not constitute collective bargaining to “establish that the Committees were ‘dealing with’ [the employer] . . . within the meaning of § 2(5)”).

\textsuperscript{40} See generally \textit{Gen. Foods Corp.}, 231 N.L.R.B. 1232 (1977) (adopting without comment the administrative law judge’s conclusion that no § 8(a)(2) violation occurred where employee teams were empowered to make job assignments, assign job rotations, schedule overtime, and occasionally interview applicants for vacant positions); \textit{Mercy-Mem’l Hosp. Corp.}, 231 N.L.R.B. 1108 (1977) (declining to find a § 8(a)(2) violation where an employee committee, composed of four elected employee representatives and one management representative, heard employee grievances and issued decisions which could be appealed to the company’s personnel board); \textit{Sparks Nugget, Inc.}, 230 N.L.R.B. 275 (1977) (distinguishing an employee committee that adjudicated employee grievances from those that deal with management under § 2(5)).
\end{flushright}
the substantive question whether the exception included employee committees that adjudicated employee grievances; and second, the procedural question whether the Board would scrutinize interactions between employer- and employee-members of an employee committee for evidence of dealing with under § 2(5). The Board concluded affirmatively on the first issue but left unanswered the procedural question of how closely it would scrutinize intracommittee interactions. This Section outlines briefly the decisions that developed the delegated managerial authority exception.

In *Sparks Nugget*, the Board held for the first time that an employee committee is not a “labor organization” if it performs a purely adjudicatory function and does not interact with management for any purpose or in any manner other than to convey a final decision on the grievance. The Board reasoned that committees resolving employee grievances, as opposed to representing employees by bringing grievances to management, perform a “function for management,” distinguishing the adjudicatory function of the Council in *Sparks Nugget* from the advocacy or representative function of the committee in *Cabot Carbon*. In reaching this decision, the Board focused on the relationship between the committee as a whole and management, ignoring intracommittee interaction between management representatives and employees. Thus, *Sparks Nugget* carved out a new substantive exception to dealing with under § 2(5) for employee committees that adjudicate employee grievances.

42. Id. at 276 (describing the resolution of employee grievances as a function of management). The employer created an “Employees’ Council”—consisting of one elected employee representative and two management representatives—that reviewed employee grievances and issued binding decisions. Id. at 275–76.
43. Id. at 276.
44. Id. at 276 n.9 (noting that in *Cabot Carbon* the employee committee “discussed with management, inter alia, such topics as safety, efficiency and production, and grievances”).
45. See *Sparks Nugget*, 230 N.L.R.B. at 276 (“[T]he Employees’ Council... does not interact with management for any purpose or in any manner other than to render a final decision on the grievance.”). The Employees’ Council was comprised of two management representatives and one employee representative, but the Board did not scrutinize the Council’s internal decisionmaking process for dealing with. Id. at 275. In contrast, both the Chairman of the Board and the Administrative Law Judge (ALJ) scrutinized the internal relationship between management and employee representatives on the Council, finding implicitly that dealing with occurred inside the Council, where management and employee representatives jointly addressed employee grievances. See id. at 277 (Fanning, Chairman, dissenting in part) (“[A] committee or plan in which employees participate and which exists for the purpose, in whole or in part, of dealing with employers concerning grievances is a labor organization.”); id. at 282 (Gilbert, A.L.J.) (“The Council meets the criteria of a labor organization... in that employees participate through the election of representatives... and the Council does address itself to the solution of employee grievances...”).
Just a few months later, in *Mercy-Memorial Hospital*, a still-divided Board continued to develop the grievance adjudication exception. In that case, the employer instituted a grievance procedure that provided for an investigation and decision by a committee comprised of four representatives elected from among the employees and one selected from management. During the resolution of one employee grievance, the employer’s personnel director called a meeting with the committee and solicited from it a policy recommendation.

Applying the grievance adjudication exception from *Sparks Nugget*, the Administrative Law Judge (ALJ), whose decision the Board adopted, characterized the committee’s decisions as “final” and found that the one occasion on which management discussed policy with the committee did not suffice to make the committee a statutory labor organization. Chairman Fanning dissented, distinguishing this case from *Sparks Nugget* because here the committee’s written procedures called for making recommendations to management, unlike the final decisions issued by the *Sparks Nugget* committee. The majority chose not to address this point; thus, *Mercy-Memorial Hospital* marked a tacit expansion of the exception created by the Board only months before in *Sparks Nugget*, characterizing as final a decisionmaking process that, at least in its written policies, purported to make recommendations to management.

---

47. Id. at 1119–20. The employer also charged the committee with making recommendations on changes to rules, regulations, and standards in effect at the hospital. Id. at 1120. The *Mercy-Memorial Hospital* grievance committee resulted from revisions to an earlier grievance procedure implemented during a protracted labor dispute arising out of an organizing campaign by the Service Employees International Union and the employer’s subsequent refusal to bargain. Id. at 1110.
48. Id. at 1121. The grievance pertained to the different methods used to determine eligibility for a service longevity pin. Id. The personnel director called a meeting of the committee, attended by his successor, who solicited the committee’s recommendation for a uniform policy, which was later adopted by management without further discussion. Id.
49. Id. at 1108.
50. Id. at 1121. The ALJ reached this conclusion despite the fact that a grievant “could appeal the Grievance Committee’s decision to . . . the personnel committee of the [employer’s] board of directors.” Id.
51. Id.
52. Id. at 1109 (Fanning, Chairman, dissenting) (“[T]he Committee has ‘the right and the obligation to recommend . . . any change in the rules, regulations, and standards. These recommendations will then be discussed and acted upon by the administrative head and the committee members . . . .’” (quoting the employer’s August 1975 policy statement)).
53. *Mercy-Memorial Hospital* suggests that the written procedures of employee committees would not be subject to serious Board scrutiny. Unlike Chairman Fanning, the Board declined to address the employer’s written procedures, which specified that while the committee was not empowered to change rules or policies, it did have “the right and
In General Foods, a unanimous Board held that an employee committee that exercised flatly delegated managerial authority was not a statutory labor organization. In that case, the employer divided its entire non-supervisory workforce into four “teams,” which used consensus decisionmaking to designate job assignments, job rotations, and allocation of overtime. The ALJ, whose decision the Board adopted essentially without comment, determined that in as much as the committees made decisions about job assignments, job rotations, and overtime, they exercised flatly delegated managerial authority, and thus did not deal with the employer. The ALJ also concluded that the team did not deal with the employer, even though in some instances employees raised issues or concerns during team meetings, because the employees were speaking on their own behalf without the team intervening as a representative.

To understand the lasting impact of these cases, it is helpful to contrast Chairman Fanning’s dissent in Sparks Nugget and Mercy-Memorial Hospital with the unanimous decision in General Foods. Chairman Fanning argued that an exception for a committee that adjudicates employee grievances ignores the plain language of the statute. Whatever the merits of such an obligation to recommend to the director of personnel . . . any changes in rules, regulations, and standards.” 231 N.L.R.B. at 1120.


55. Id. at 1235 (finding no § 2(5) labor organization where the employer divided its workforce into four teams each of which made job assignments, assigned job rotations, and scheduled overtime). Ad hoc committees composed of members designated by the plant director also conducted interviews, performed safety inspections, and prepared job descriptions. Id. at 1234.

56. Id. at 1232–33. Employees also attended meetings facilitated by an outside consultant intended to improve internal communication and build trust levels through team building exercises. Id. at 1233.

57. See id. at 1232 (“The Board has considered the record and the attached Decision in light of the exceptions and briefs and has decided to affirm the rulings, findings, and conclusions of the Administrative Law Judge . . . .”).

58. Id. at 1235.

59. See id. at 1235 (“Thus, [an employee] complained at a team meeting about supervisory favoritism [sic] toward female employees . . . [t]his was [the employee] speaking individually to the [employer]. [The employee] did not call upon, did not receive, and in the nature of things could not receive the intercession of the team on his behalf in pressing these matters.”).


61. See Sparks Nugget, 230 N.L.R.B. at 277 (Fanning, Chairman, dissenting) (“[A] committee or plan in which employees participate, and which exists for the purpose, in whole or in part, of dealing with employers concerning grievances is a labor organization.”); Mercy-Mem’l Hosp. Corp., 231 N.L.R.B. at 1108 n.2 (Fanning, Chairman, dissenting)
argument, today grievance adjudication is a settled subset of the managerial authority doctrine. However, Chairman Fanning also scrutinized employee committees more carefully than the majorities in both *Sparks Nugget* and *Mercy-Memorial Hospital*. Since neither case established clear limits for the Board’s factual analysis of written procedures or intracommittee interactions, both left open the possibility that a future Board might follow Chairman Fanning’s example, narrowing the adjudicatory exception through closer examination of the record.

By contrast, in *General Foods* a unanimous Board found no labor organization where the employer completely delegated decisionmaking concerning job schedules, rotations, and overtime to rank-and-file employee teams. Looked at together, *General Foods*, *Mercy-Memorial Hospital*, and *Sparks Nugget* stand for the proposition that an employee committee does not deal with the employer when the committee exercises independent discretion to make recommendations even if such recommendations are “potentially disagreeable to the employer.”

(disagreeing that *Sparks Nugget* is valid precedent and distinguishing the committee in *Mercy-Memorial Hospital* on the grounds that it did not render final decisions).

62. See, e.g., Electromation, Inc., 309 N.L.R.B. 990, 995 (1992), enforced, 35 F.3d 1148 (7th Cir. 1994) (“[A]n organization . . . limited to performing essentially a managerial or adjudicative function is not a labor organization under Section 2(5).”).

63. Compare *Sparks Nugget*, 230 N.L.R.B. at 276–77 (noting that the employee committee as a whole does not advocate on behalf of a grievant, finding it unnecessary to know the precise process of decisionmaking inside the committee, and affirming the Administrative Law Judge’s finding of a § 8(a)(2) violation), with *Mercy-Mem’l Hosp. Corp.*, 231 N.L.R.B. at 1120–21 (describing written procedure for an employee committee to make recommendations to management, but ignoring the same procedure in addressing potential dealing between the committee and management), and id. at 1109 (identifying the policy statements setting out the employee committee’s “right and . . . obligation” to make recommendations to management as a plain indication of dealing with (quotation omitted)).

64. 231 N.L.R.B. at 1235 (emphasizing that management delegated its power unilaterally, and could unilaterally withdraw the delegated authority).

65. See supra text accompanying notes 18–20.

66. The use of the term *exception* to describe the delegated managerial authority doctrine is inapposite in some respects. The delegated managerial authority doctrine is arguably not a true exception to § 2(5) at all, but rather an explicit formulation of what is implicit in *Cabot Carbon*: that employees must always deal with management when management retains final decisionmaking authority. Cf. *NLRB v. Cabot Carbon Co.*, 360 U.S. 203, 214 (1959) (“[T]hat final decision remained with [the employer] . . . is true of all such ‘dealing,’ whether with an independent or a company-dominated ‘labor organization.’ The principal distinction lies in the unfettered power of the former to insist on its requests.”). The converse is also true—when management cedes final decisionmaking authority to an
D. The Board Defines Dealing With

In the early 1990s, the Board clarified the scope of dealing with under § 2(5), at the same time signaling the beginning of a period of closer factual scrutiny of employee participation programs. Whereas Cabot Carbon merely established that dealing with included more than collective bargaining, Electromation and E.I. du Pont de Nemours & Co. expanded on Cabot Carbon, developing an affirmative test for dealing with under § 2(5).

In Electromation, the Board focused on whether the purpose of an organization, “shown by what the organization actually does,” is to engage in bilateral communication with an employer concerning conditions of employment. There, in response to employee dissatisfaction with unilateral changes in bonus and wage policies, a non-union employer created five joint worker-management “action committees,” in which management would “sit down and work with” employees. The action committees operated for only a few weeks before the company withdrew its management representatives in response to a union demand for recognition.

Although it reaffirmed the managerial exception carved out in earlier cases, the Board abrogated its prior practice of turning a blind eye to employee committee there is no residual need for communication or authorization by management that could qualify as dealing with.

67. See supra text accompanying note 42.
68. 309 N.L.R.B. 990 (1992), enforced, 35 F.3d 1148 (7th Cir. 1994).
70. See 309 N.L.R.B. at 996 (declining to follow the Sixth Circuit requirement that the employees believe their organization to be a labor union).
71. Id. at 991 (quoting testimony of employer’s president, John Howard). The “action committees” each consisted of up to six employees and two management representatives; the company’s Employee Benefits Manager sat on all five committees. Id. Each team addressed one of five subjects designated by management: “(1) Absenteeism/Infractions, (2) No Smoking Policy, (3) Communication Network, (4) Pay Progression for Premium Positions, and (5) Attendance Bonus Program.” Id. The employer’s action committee coordinator testified that management intended the employees on the committees to “kind of talk back and forth’ with” the other employees in the plant about the work of the committee. Id. (quoting testimony of Loretta Dickey). In the Attendance Bonus committee, for example, a management representative—the company’s comptroller—declared that the committee’s proposal on bonuses was too costly, prompting the committee to devise a second proposal, which the management representative deemed acceptable. Id. at 991–92.
72. Id. at 991–92.
73. See id. at 995 (“Notwithstanding that ‘dealing with’ is broadly defined under Cabot Carbon, it is also true that an organization whose purpose is limited to performing essentially a managerial or adjudicative function is not a labor organization . . . .” (citing Gen. Foods Corp., 231 N.L.R.B. 1292 (1977); Mercy-Meri'l Hosp. Corp., 231 N.L.R.B. 1108 (1977); Sparks Nugget, 230 N.L.R.B. 275 (1977)).
intracommittee interactions between employers and employees. The Board also announced an affirmative definition for dealing with: a bilateral mechanism involving proposals from the employee committee, coupled with real or apparent consideration of those proposals by management. Finally, the Board adopted a realist approach to its inquiry into the purpose of an employee committee, noting that an organization’s purpose can be shown not only by what it was set up to do, but also by what it actually does.

Although many in business and labor anticipated that the Seventh Circuit would take the opportunity to make a broad ruling on the legality of employee participation programs, it declined to do so. Treating the question of what is a labor organization under § 2(5) as a factual one, the Court concluded that there was substantial evidence to support the Board’s finding. Thus, the Seventh Circuit implicitly approved the Board’s discretionary application of a more intensive review of intracommittee employer–employee interactions, a level of scrutiny by the Board noticeably absent in prior managerial exception cases such as Mercy-Memorial Hospital and Sparks Nugget.

Six months after deciding Electromation, the Board clarified the “bilateral mechanism” test for dealing with under § 2(5) in E.I. du Pont de Nemours & Co. The Board explained that a bilateral mechanism ordinarily entails a

---

74. See id. at 997 (noting that bilateral communication between management representatives and employees concerning conditions of employment occurred inside of the Attendance Bonus Committee, generating a proposal whose formal acceptance by the company president was disrupted by the onset of a union campaign). The Board also took the opportunity presented by this case to clarify that the employees’ subjective view of whether a committee is a labor organization is irrelevant under § 2(5), abrogating a line of cases developed by the Court of Appeals for the Sixth Circuit. See id. at 996 (declining to follow Airstream, Inc. v. NLRB, 877 F.2d 1291 (6th Cir. 1989); NLRB v. Scott & Fetzer Co., 691 F.2d 288 (6th Cir. 1982)).

75. See id. at 995 n.21 (comparing the bilateral mechanism of dealing with to unilateral communications such as a suggestion box or brainstorming groups).

76. See id. at 996 (declining to read into § 2(5) a requirement that the employer created the employee committee to avoid unionization, or that employees view the committee as the equivalent of a union).

77. See Electromation, Inc. v. NLRB, 35 F.3d 1148, 1151 (7th Cir. 1994) (“[T]his case potentially raised the rather novel and important issue whether modern ‘employee involvement’ or ‘employee participation’ organizations are unlawful under section 8(a)(2) and (1) of the Act . . . We find it unnecessary to address this much broader issue.”). The Seventh Circuit observed that—despite the serious policy arguments in favor of revising the scope of § 8(a)(2)—making such a change was more properly the role of Congress. Id. at 1157.

78. Id. at 1157–59.

79. See du Pont, 311 N.L.R.B. 893, 895 (1993) (finding a § 2(5) labor organization where management implemented several employee committees, each with one management
pattern or practice in which a group of employees, over time, make proposals to management; management responds to these proposals by acceptance or rejection by word or deed; and compromise is not required. Moreover, stating explicitly what it implied in Electromation, the Board signaled it might find dealing with even if the management representative with whom employees communicated bilaterally was a member of the committee. Taking a functional approach to dealing with, the Board focused on whether management representatives could decisively control the committee’s decisionmaking process—for example, by constituting the majority of a committee that decided questions based on majority vote. The Board also distinguished between bilateral dealing with—where proposals are made to and rejected by management—and permissible unilateral mechanisms such as brainstorming, where management solicits employee ideas.

In a line of cases following Electromation and du Pont, the Board continued to take a restrictive view of the adjudicatory and managerial authority exceptions. Notably, in Keeler Brass the Board found that an employee grievance committee was a labor organization because, although management’s written policy stated that the committee’s decisions were “final,” the evidence showed that management consistently did not consider

---

80. Id. at 894. The Board noted, however, that dealing is not present if there are only isolated instances in which the group makes ad hoc proposals to management. See id.

81. Id. at 895 (abrogating the Board’s implied policy of examining only the relationship between the committee as a whole and management representatives outside of the committee). Cf. Sparks Nugget, 230 N.L.R.B. 275 (1977) (finding no § 8(a)(2) violation where an employee committee consisting of one employee representative and two management representatives heard and issued final decisions on employee grievances).

82. See du Pont, 311 N.L.R.B. at 895 (“[T]here would be no ‘dealing with’ management if the committee were governed by majority decision-making, management representatives were in the minority, and the committee had the power to decide matters for itself . . . .”).

83. Id. at 894. For example, the Board noted that the employer did not deal with employees at its quarterly safety conferences, where workers were encouraged to share suggestions and concerns on workplace safety. Id. at 897 (“Nothing in the Act prevents an employer from encouraging its employees to express their ideas and to become more aware of the safety problems in their work.”).

84. See Polaroid Corp., 329 N.L.R.B. 424 (1999) (finding a statutory labor organization where management polled members of an employee committee to determine the majority view, to which management responded); Keeler Brass Co., 317 N.L.R.B. 1110 (1995) (finding dealing with where employee grievance committee recommended reinstatement of a grievant, then reversed itself after management responded to the recommendation).

85. 317 N.L.R.B. 1110.
the committee’s decisions to be final, but instead treated them as recommendations it was free to accept or reject.86

*Electromation* and *du Pont* signaled that going forward the Board would scrutinize processes of employee involvement for *dealing with* irrespective of whether management exercised its decisionmaking authority from outside the committee. Importantly, the Board left intact the managerial and adjudicatory function exception from *General Foods, Sparks Nugget* and *Mercy-Memorial Hospital*, observing that it is compatible with *Cabot Carbon*.87 However, the Board’s close scrutiny of intracommittee interactions between employees and management signaled a restrictive construction of those exceptions, and a return to the emphasis on determining the locus of decisionmaking authority seen in *Cabot Carbon*.88

*Electromation* and its progeny clarified the Supreme Court’s limited explanation of *dealing with*,89 but they left unanswered two important questions. First, employers were left to guess about when lawful brainstorming or information sharing would cross the line into unlawful acceptance or rejection of a proposal.90 Second, although employers now knew that employee committees could make “ad hoc proposals” in “isolated instances” without committing an unlawful pattern or practice,91 the Board gave no guidance on how frequently an employee committee could make such proposals before it crossed the line into impermissible *dealing with* the employer.

E. Crown Cork & Seal: Redrawing the Lines of the Managerial Exception

In 2001, the Board announced a potentially radical shift of its delegated

---

86. *Id.* at 1114 n.16. The Board distinguished the *Keeler Brass* committee from those in *Mercy-Memorial Hospital* and *Sparks Nugget*, which “could definitively resolve grievances without further recourse to the employer.” *Id.* at 1114.
88. *Cf. NLRB v. Cabot Carbon* Co., 360 U.S. 203, 214 n.15 (1959) (“Final decision is always with management, although when a claim is made by a well organized, good sized union, management is doubtless more strongly influenced in its decision . . . .” (quoting *NLRB v. Jas. H. Matthews & Co.*, 156 F.2d 706, 708 (3d Cir. 1946))).
89. *Compare E.I. du Pont de Nemours & Co.*, 311 N.L.R.B. 893, 894 (1993) (defining dealing with as a pattern or practice in which a group of employers, over time, make proposals to management, management responds to these proposals by acceptance or rejection by word or deed, and compromise is not required); with *Cabot Carbon*, 360 U.S. at 210–11 (holding that dealing with encompasses more than bargaining collectively).
90. *See du Pont*, 311 N.L.R.B. at 897 (“Nothing in the Act prevents an employer from encouraging its employees to express their ideas . . . .”)
91. *Id.* at 894.
managerial authority analysis in *Crown Cork & Seal*. Crown Cork & Seal Co. used the Socio-Tech system, an industrial organization model wherein all employees in the plant were members of one of four production teams with the power to “decide and do” on a variety of workplace issues, although some higher-tier committees did not exercise complete discretion, but rather made recommendations to management. In a surprising departure from its traditional delegated managerial authority analysis, the Board rejected the General Counsel’s argument that the employee committees in this case dealt with management because they issued recommendations rather than final decisions, reasoning instead that the employee committees were analogous to a frontline supervisor, making management recommendations that are reviewed by higher management. The Board distinguished Crown Cork & Seal Co.’s employee committees from the committee in *Keeler Brass*, not on the more predictable grounds that the employer had never overturned the committees’ recommendations, making the committees’ decisions functionally final, but rather on the grounds that the committee here had “management” status, and thus could not deal with management.

92. *Crown Cork & Seal*, 334 N.L.R.B. 699, 701–02 (2001) (comparing employee committees that make recommendations to front-line supervisors and distinguishing the exercise of such delegated managerial authority from dealing with under § 8(a)(2)).

93. *Id.* at 699. *Cf.* Gen. Foods Corp., 231 N.L.R.B. 1232 (1977) (finding delegated managerial authority where employee teams made job assignments, assigned rotations, scheduled overtime, and occasionally interviewed applicants for positions). Two employees from each production team sat on three teams one administrative tier above the production teams (for a total of eight employee representatives in each higher-tier team), along with four management representatives. *Crown Cork & Seal*, 334 N.L.R.B. at 699. All teams used a form of consensus decisionmaking, under which team members who did not agree with the group on an issue would abstain. *Cf.* du Pont, 311 N.L.R.B. at 895 (finding dealing with where a committee in which management participated operated under a pure consensus model, giving management a veto on all committee proposals). The three higher-tier teams monitored policy among the four production teams, issuing decisions “in the form of recommendations forwarded to the Management Team or the plant manager.” *Crown Cork & Seal*, 334 N.L.R.B. at 700. In practice, management never overruled the committees’ recommendations. *Id.*

94. *Crown Cork & Seal*, 334 N.L.R.B. at 701. The Board’s rationale is difficult to square with its previous rule that recommendations from employee committees, rejected or accepted by management, constitute a bilateral mechanism of “dealing with” under § 2(5). *Cf.* du Pont, 311 N.L.R.B. at 894 (“[T]he concept of ‘dealing’ does not require that the two sides seek to compromise their differences. It involves only . . . a pattern or practice in which a group of employees, over time, makes proposals to management, [and] management responds to these proposals by acceptance or rejection by word or deed . . . .”).

95. *Crown Cork & Seal*, 334 N.L.R.B. at 701 (“Indeed, it is the fact that the interaction is occurring between two management bodies that distinguishes this case from cases such as *Keeler Brass* and persuades us that the statutory element of dealing is absent.”).
Crown Cork & Seal signaled, albeit impliedly, two changes in the Board’s approach to the managerial authority exception: a narrowing of the concept of bilateral communication developed in du Pont and Keeler; and, more significantly, an expansion of the delegated managerial authority safe haven. First, quoting Keeler Brass, the Board hinted that dealing with might require a process whereby the company and the employees “went back and forth explaining themselves until an acceptable result was achieved.” The “back and forth” analysis was not strictly necessary to its decision, it suggests that the Board adopted a narrower view of dealing with under § 2(5) than it had in du Pont. Second, Crown Cork & Seal announced a new test for delegated managerial authority under which a pattern or practice of bilateral communication—one that would clearly constitute dealing with if it occurred between an employee group and a management body or representative—would not constitute dealing with if it occurred between management and an employee committee acting as management. This marked an important change from the historical managerial authority safe haven, which applied only when that authority was “flatly delegated,” that is, when the committee exercised final decisionmaking authority. Under the Crown Cork & Seal analysis, the Board found no dealing with even though the employee committee issued recommendations that were at least formally subject to management

96. Id. at 700 (quoting Keeler Brass Co., 317 N.L.R.B. 1110, 1114 (1995)). The comparison with the “back and forth” language here is dicta—the Board ultimately rests its decision on the idea that the committee is management, distinguished from Keeler Brass on that basis rather than on the basis of whether there was a bilateral mechanism of communication. Id. at 701.

97. Cf. du Pont, 311 N.L.R.B. at 894 (“[T]he concept of ‘dealing’ does not require that the two sides seek to compromise their differences. It involves only . . . a pattern or practice in which a group of employees, over time, makes proposals to management [and] management responds to these proposals by acceptance or rejection by word or deed . . . .”); NLRB v. Cabot Carbon Co., 360 U.S. 203, 211 (1959) (“Certainly nothing in [section 2(5)] indicates that the broad term ‘dealing with’ is to be read as synonymous with the more limited term ‘bargaining with.’”).

98. 334 N.L.R.B. at 701. The Board’s analysis rests on the idea that the committee exercised group authority “that, in the traditional plant setting would be considered supervisory.” Id.

99. See Gen. Foods Corp., 231 N.L.R.B. 1232, 1235 (1977) (“The employees, acting within the team, were authorized within certain narrow limits to set their own starting and quitting times . . . . These are managerial functions being flatly delegated to employees and do not involve dealing with the employer . . . .”).

100. See Keeler Brass, 317 N.L.R.B. at 1114 (applying Mercy-Mem’l Hosp. Corp., 231 N.L.R.B. 1100 (1977), and Sparks Nugget Inc., 230 N.L.R.B. 275 (1977), in ruling that employee committees exercise managerial authority where they can definitively resolve grievances without further recourse to the employer).
The Board’s surprising analytic shift in *Crown Cork & Seal* reflects the difficulty of applying traditional labor regulation to nontraditional management structures like Socio-Tech. Both the ALJ and the Board recognized that the Socio-Tech system represented “a significant variation on the traditional plant organizational structure.” 102 Analyzed under the traditional interaction-based models, the industrial organizing model used by Crown Cork & Seal Co. skirts the edge of the bilateral exchange identified above as indicative of a potentially unlawful third class of employer–employee interaction. 103 A careful reading of *Crown Cork & Seal* reveals that, while the higher-tier committees issued recommendations to management rather than decisions, the committees often implemented changes without waiting for management approval. 104 The plant manager could remember no instance in which he overturned such a recommendation. 105 In the only recorded instance in which management overruled one of the higher-tier teams, the management overruled the team not because it disagreed with the substantive decision, but because the team could not make the decision without exceeding its delegated authority. 106

Because management never overruled the employee teams, and the teams implemented changes without waiting for management approval, employer–employee interaction in *Crown Cork & Seal* arguably lacked the element of exchange necessary to push it into the third class, despite the characterization of the team decisions as “recommendations.” 107 However, the Board did not rely on these narrow factual grounds to decide that the employee committees in *Crown Cork & Seal* were not a labor organization under § 2(5), relying instead on the “comparable to a front line supervisor” test to determine that the employee committees were themselves managerial and thus not labor organizations under § 2(5). 108

The employee-committee-as-management analysis was a significant analytical departure from previous *dealing with doctrine*. Before *Crown Cork & Seal*, the Board analyzed the management exception as falling under the

101. *Crown Cork & Seal*, 334 N.L.R.B. at 701. Instead of issuing final decisions, the employee committees issued recommendations to management, which were “rarely, if ever, overruled,” or, said another way, were routinely accepted. *Id.*

102. *Id.*

103. See supra text accompanying notes 18–20.

104. 334 N.L.R.B. at 700.

105. *Id.*

106. The team recommended a layoff procedure based on seniority, after which management informed the team that there was no seniority in the plant. *Id.*

107. See supra notes 102–06 and accompanying text.

108. See supra notes 90–102 and accompanying text.
first class identified above—a category of employer–employee interactions wherein the power to act is flatly delegated to the employee committee—and therefore outside the ambit of dealing with under § 2(5). In Crown Cork & Seal, the Board sidestepped this analysis completely. Thus, the Board created a loophole in the dealing with analysis whereby any committee deemed to be management is arguably immune from further scrutiny, expanding significantly the scope of permissible employee participation programs under §§ 2(5) and 8(a)(2).

While Crown Cork & Seal arguably expanded the scope of the delegated managerial authority exception, it left unanswered two key questions. First, the Board declined to scrutinize employer–employee interactions inside the Socio-Tech committees, leaving open the procedural question of whether Crown Cork & Seal signaled the end of the Board’s policy of scrutinizing those interactions for dealing with under the framework made explicit in du Pont. Second, the Board indicated that the exercise of managerial authority “comparable to that of the frontline supervisor in the traditional plant setting” suffices to make an employee committee management for the purposes of § 2(5), but gave no detailed guidance as to the minimum exercise of discretion necessary to meet the managerial authority “floor” above which a committee would be considered management.

II. THE DELEGATED MANAGERIAL AUTHORITY EXCEPTION SINCE CROWN CORK & SEAL

Although Crown Cork & Seal can be read to expand significantly the

109. See Keeler Brass Co., 317 N.L.R.B. 1110, 1114 (1995) (applying the rule from Mercy-Mem’l Hosp. Corp., 231 N.L.R.B. 1108 (1977), and Sparks Nugget, Inc., 230 N.L.R.B. 275 (1977), that employee committees exercise managerial authority where they can definitively resolve grievances without further recourse to the employer); Gen. Foods Corp., 231 N.L.R.B. 1232, 1235 (1977) (“The employees, acting within the team, were authorized within certain narrow limits to set their own starting and quitting times . . . . These are managerial functions being flatly delegated to employees and not involve dealing with the employer . . . .”). Cf. NLRB v. Cabot Carbon Co., 360 U.S. 203, 214 (1959) (observing that the retention of final decisionmaking authority is “true of all such ‘dealing,’ whether with an independent or company-dominated ‘labor organization’”).

110. See Crown Cork & Seal, 334 N.L.R.B. at 701 (“[T]he seven committees in issue do not deal with management within the meaning of Section 2(5). Rather . . . the seven committees are management.”); id. (observing that a pattern of recommendations and either acceptance or rejection is not dealing with if it occurs between levels of management).

111. Compare id. at 701 n.2 (“The General Counsel does not contend that ‘dealing’ occurs between the management and nonmanagement members within the committee.”), with E.I. du Pont de Nemours & Co., 311 N.L.R.B. 893, 895 (1993) (“[I]f management representatives can reject employee proposals, it makes no real difference whether they do so from inside or outside the committee.”).

delegated managerial authority exception, some commentators have argued that it did little to alter the law of employee committees. The Board has decided several § 8(a)(2) cases related to employee committees since Crown Cork & Seal. However, few of these cases relied on the delegated managerial authority exception to dealing with, and in none of those cases did the Board find that the employee committees exercised non-final, yet managerial, authority. Apparently, while the Board has adopted the language of the managerial authority exception from Crown Cork & Seal, it remains unclear how broadly the Board will read the “comparable to a frontline supervisor” test.

One case that provides some indication of the potential breadth of the new managerial authority test is Dow Chemical. There, applying the managerial authority test from Crown Cork & Seal, an ALJ found that two

113. See Pauling & McGuire, supra note 10, at 231 (noting that while Crown Cork & Seal arguably clarified and expanded the managerial exception, it can also be argued that it does little to alter the existing landscape of § 8(a)(2)).

114. See Syracuse Univ., 350 N.L.R.B. 755, 758 (2007) (finding an employee grievance committee subject to limited managerial review did not deal with management, and therefore was not a statutory labor organization, because the employee committee was not required to give any weight to management’s recommendations); Dow Chem., 349 N.L.R.B. 104 (2007) (granting motion to sever and remand for withdrawal of union charge that employer’s human resource teams (HRT) were labor organizations under § 2(5), where the ALJ found that the HRTs were management under Crown Cork & Seal); Ead Motors E. Air Devices, Inc., 346 N.L.R.B. 1060, 1076–77 (2006) (finding that “Have Your Say” committee unlawfully dealt with employer because it discussed with management, and made recommendations relating to, wages, hours, and conditions of work); Miller Indus., 342 N.L.R.B. 1074, 1083, 1089–90 (2004) (finding dealing with where committee of management-selected employee representatives recommended changes to working conditions and employer policies, upon which employer acted); Ga. Power Co., 342 N.L.R.B. 192, 200 (2004) (applying the Crown Cork & Seal managerial exception test to employer’s grievance process and finding no dealing with because “process resulted in management level decisions at the end of each procedure”); Music Express E., Inc., 340 N.L.R.B. 1063, 1077 (2005) (finding that an employee committee was a labor organization where employer’s memorandum described committee members as representatives, announced that the committee would meet with management to “go over problems,” and “try to reach a reasonable conclusion,” and the committee sent a letter to the employer demanding wage and benefits increases); Baptist Med. Ctr/Health Midwest, 338 N.L.R.B. 346, 371 (2002) (finding that Nurses Practice Committee, which the employer stated was intended in part to discuss wages, salaries, and benefits, was a labor organization under § 2(5)).

115. Cf. Dow Chemical, 349 N.L.R.B. at 104 (declining to address the merits of the ALJ’s finding that the HRT were management under Crown Cork & Seal); Ga. Power, 342 N.L.R.B. at 200 (finding employee committees issued final decisions because the process “resulted in management level decisions at the end of each procedure”).

116. See 342 N.L.R.B. at 200 (“[The] process resulted in management level decisions at the end of each procedure.”).

employee committees did not deal with the employer under § 2(5) because they “engaged in activity which management had previously done,” including posting jobs, interviewing, making hiring recommendations, and setting work schedules. The committees passed their decisions on to a management representative who could overrule them. This pattern of bilateral communication with the employer about covered subjects, where the employee committee made recommendations that were rejected or accepted by management, clearly constitutes dealing with under Electromation and its progeny. However, applying Crown Cork & Seal, the ALJ declined to reach the issue of dealing with once he determined that the employee committees were management.

The Board remanded the case without deciding the merits of the § 8(a)(2) charge. Thus, Dow Chemical hints at how the Board could apply Crown Cork & Seal going forward, but ultimately does little to clarify the scope of the new managerial authority exception. As one Board Member said in Dow Chemical, the permissible scope of employee committees after Crown Cork & Seal is “a nettlesome area of Board law,” one in which a decision on the merits could provide needed guidance to the agency, employers, and employees.

III. THE CONSEQUENCES OF CROWN CORK & SEAL

Crown Cork & Seal left the managerial authority exception—and by extension, the scope of dealing with—in a state of uncertainty. The traditional managerial exception to § 2(5) is not really an exception at all, but Crown Cork & Seal introduced an analysis of managerial authority under which the exception could easily swallow the rule. By removing the requirement that committees falling under the managerial authority exception make final decisions, Crown Cork & Seal can be criticized for introducing ambiguity into a settled area of law and moving beyond the rational underpinnings of the managerial authority exception, potentially upsetting the balance between employers and employees reflected in

118. Id. at 110.
119. Id.
120. See supra text accompanying notes 79–80.
121. See 349 N.L.R.B. at 110 (“The [employee committees] acted as part of management and its recommendations and decisions could be overruled by higher management. They were not labor organizations within the meaning of Section 2(5) of the Act.”). In making this determination, the judge relied in part on the fact that management had previously made hiring and work schedule decisions. Id.
122. Id. at 105.
123. Id. at 107 (Schumber, Member, dissenting in part).
124. See supra note 68 and accompanying text.
§ 8(a)(2) and prior case law. This Section addresses both criticisms in turn.

A. The Crown Cork & Seal Managerial Authority Exception is Imprecise

Crown Cork & Seal expanded the scope of permissible employee participation programs, but the cost of that expansion was the loss of a clear test for the managerial authority exception and a corresponding risk increase for employers designing programs to fall under it. The comparable to a front line supervisor test\textsuperscript{125} is flawed because it is potentially all-encompassing; one can hardly imagine a level of reviewable discretion so small that it would be inconsistent with that exercised by a frontline supervisor in a traditional plant setting.\textsuperscript{126} Such a test provides little meaningful guidance to the Board, the courts, or employers about the permissible scope of employee participation programs, and obscures the line drawn by earlier § 2(5) cases.\textsuperscript{127} The ALJ’s decision in Dow Chemical shows that some adjudicators will read Crown Cork & Seal to preclude further dealing with analysis of employee committees deemed to be management, even in workplaces not organized under the Socio-Tech model or a similarly comprehensive system of employee participation.\textsuperscript{128}

Thus, under Crown Cork & Seal an employer could arguably avoid § 8(a)(2) liability by delegating to an employee committee minimal discretion to make reviewable decisions, provided the discretion exercised met the ambiguous comparable to a frontline supervisor test. Although this would certainly give employers an incentive to grant employee committees “managerial” status, under such a broad definition that status would no longer be meaningful.

If future Board decisions follow the current trend, Crown Cork & Seal may result in little more than an exchange of one managerial authority test for another, without a significant expansion of the scope of permissible participation.

\textsuperscript{125} Under Crown Cork & Seal, an employee committee is management if it exercises authority “comparable to that of the frontline supervisor in the traditional plant setting.” Crown Cork & Seal Co., 334 N.L.R.B. 699, 701 (2001).

\textsuperscript{126} The Board’s assertion that managers in traditional plants do not exercise final decisionmaking authority is no doubt correct. See id. (“Few, if any, supervisors in a conventional plant possess authority that is final and absolute.”). However, such an observation misses the point of the traditional managerial authority exception, which is based not on a talismanic approach to management authority, but rather on the lack of employer–employee interaction when the employee committee wields final decisionmaking authority. See supra note 66 and accompanying text.

\textsuperscript{127} See E.I. du Pont de Nemours & Co., 311 N.L.R.B. 893, 895 (1993) (finding dealing with where management representatives discussed proposals with employee committees and had the power to accept or reject any proposal).

\textsuperscript{128} See supra text accompanying notes 118–21.
employee involvement programs. However, because the Crown Cork & Seal managerial authority test lacks precision, the Board could expand broadly the scope of permissible employee involvement programs. The next Section explains how such a broad expansion could significantly alter the balance of power between employers and employees and argues that this alteration would be contrary to the clear purpose of § 8(a)(2).

B. Unbalanced Expansion of Employee Involvement Undermines § 8(a)(2) of the NLRA

The Crown Cork & Seal decision highlights how evolutionary changes in the workplace are pushing against the constraints on employee participation imposed by § 8(a)(2). Policymakers have long recognized that employee involvement committees, in order to reach their full potential, will need to address subjects traditionally considered out of bounds under the NLRA. Such an expansion, however, could lead to unjust and potentially disruptive results if it does not preserve the historical balance of power between employers and employees reflected in the traditional construction of § 8(a)(2). This Section briefly describes the policy rationale for increased employee involvement, the risks associated with relaxing employee involvement regulations, and how the Crown Cork & Seal managerial authority test unbalances § 8(a)(2).

The drive for greater employee participation reflects real changes in American workplaces. Today, jobs demand higher skills, employees are increasingly more educated, and global competition is pushing employers

---

129. Cf. supra note 114.

130. See S. REP. NO. 103-12, at 22 (1997) (“Some of the matters of mutual interest which employee involvement structures address will unavoidably include discussions of conditions of work. The processes by which a company ‘produces’ its product are inextricably linked to the terms and conditions of individual’s employment in those processes.”); see also id. (“What is productivity? It’s who does what, it’s whether ‘A’ works certain hours, whether ‘B’ gets relief . . . . And to say that you can abstract productivity from working conditions is something that I have a great deal of difficulty with.”) (quoting Transcript of Proceedings Before the National Labor Relations Board of Electronation, Inc. (Case No. 25–CA–19818) 61–62 (Sept. 5, 1991) (statement of Lawrence Gold, counsel to the AFL-CIO); Electronation, Inc. v. NLRB, 35 F.3d 1148, 1157 (7th Cir. 1994) (“Today’s evolving industrial environment may require reconsideration of Section 8(a)(2) . . . . or its interpretation and application to certain modern employee organizations.”).

131. Cf. Bruce E. Kaufman, Does the NLRA Constrain Employee Involvement and Participation Programs in Nonunion Companies?: A Reassessment, 17 YALE L. & POL’Y REV. 729, 803 (1998) (“Greater competition . . . . is a ‘plus’ for the public interest. But for competition to serve the public interest it must take place on a level playing field . . . .”). Kaufman observed that unbalanced changes to the scope of permissible employee involvement would impede fair competition. Id.
to maximize productivity by working cooperatively with their employees to leverage their insights into how to increase the efficiency of production processes—unique intellectual capital that represents a relatively untapped value center for employers.\textsuperscript{132} According to a survey of workers and management,\textsuperscript{133} employees want a greater voice in the workplace, both as individuals and as part of a group.\textsuperscript{134} Employees prefer cooperation with management to the institutionalized labor–management conflict of unions and collective bargaining,\textsuperscript{135} and employee participation programs have been linked to increases in productivity and increased employee satisfaction.\textsuperscript{136} Employers also view expanded employee participation as an important mechanism for leveraging the intellectual capital of their employees.\textsuperscript{137}

However, cooperation does not mean ceding control to employers. Workers prefer independence—backed up by government protections—in their dealings with management.\textsuperscript{138} In contrast, employers broadly oppose programs that would cede final decisionmaking authority to employees.\textsuperscript{139}

\textsuperscript{132} See U.S. COMM’N ON THE FUTURE OF WORKER–MANAGEMENT RELATIONS (DUNLOP COMMISSION), FINAL REPORT (1994), at 18 (“[T]he best way to compete in the marketplace and to secure both profits for the firm and good jobs for workers is through cooperative worker–management relations.”); S. REP. NO. 105-12, at 2 (“As this Nation enters the 21st century, the committee believes it important that U.S. workplace policies reflect a new era of labor–management relations—one that fosters cooperation, not confrontation.”); cf. Michael H. LeRoy, “Dealing With” Employee Involvement in Nonunion Workplaces: Empirical Research Implications for the TEAM Act and Electromotive, 73 NOTRE DAME L. REV. 31, 63 (1997) (describing the NLRA as “a 1935 law intended to outlaw rampant sham-unionism . . . now applied in a different economic and organizational-management context.”). For a revealing investigative report about the initial success and eventual failure of an employee participation program under the “Toyota production system” in a U.S. automobile factory, see This American Life: Episode 403: NUMMI (WBEZ Chicago Public Radio broadcast Mar. 26, 2010), available at http://www.thisamericanlife.org/radio-archives/episode/403/nummi.

\textsuperscript{133} See RICHARD B. FREEMAN & JOEL ROGERS, WHAT WORKERS WANT 32–34 (1999) (discussing the Worker Representation Participation Survey methodology).

\textsuperscript{134} Id. at 4.

\textsuperscript{135} Id. at 5.

\textsuperscript{136} S. REP. NO. 105-12, at 4.

\textsuperscript{137} See, e.g., TEAMwork for Employees and Managers (TEAM) Act of 1997: Hearing on S. 295 Before the S. Comm. on Labor & Human Resources, 105th Cong. 59 (1997) (testimony of Charles Cohen) (testifying that American companies are under pressure from foreign competition to increase quality, productivity, and efficiency, which is achieved in many instances “through employee participation committees—which facilitate employer–employee communication and allows managers to tap into their most valuable resource: their employees.”).

\textsuperscript{138} FREEMAN & ROGERS, supra note 133, at 5.

\textsuperscript{139} See id. at 7 (“Managers broadly confirm worker assessment of their unwillingness to share power . . . . Many oppose programs that would keep them from making the final decisions about workplace governance.”).
In this picture of the evolving workplace, employers and employees both want increased employee participation and more cooperation between employers and employees but differ significantly on how power should be distributed in the participation process. Arguably, the drive for increased employee participation also reflects a common perception that employer–employee relations no longer fit the confrontational labor–management paradigm that existed when Congress enacted the NLRA.

Some scholars argue that employee involvement programs, or employee committees, are unlike the company unions that Congress outlawed in § 8(a)(2), suggesting that the law has not kept up with changes in the workplace. While the Board’s handling of employee involvement in Crown Cork & Seal to some extent validates this criticism of existing labor regulations, unbalanced liberalization of employee participation regulations would open up real potential for employer abuse, frustrating in part the purposes of the Act. As this Comment observes, at the time the NLRA was enacted, the most prevalent form of company union was not a formal labor organization, but rather a relatively unstructured representation plan. These representation plans usurp the role of an independent labor organization by purporting to represent employees’ interests. Often employers formed these organizations in response to union organizing activity. Clearly, one reason employers formed company unions was to avoid the economic costs of dealing with an independent union by creating a mechanism that canalized worker dissatisfaction. Today employers often form employee committees during union organizing campaigns.

140. See LeRoy, supra note 132, at 80 (“[T]his study finds no evidence that [employee committees] are like company unions . . . .”). But cf. Kaufman, supra note 131, at 772–73 (observing that two of six companies studied by the author had employee committees that “are in a number of respects closely akin to the 1920s-era employee representation plans”).

141. See supra notes 27–31 and accompanying text.

142. See BUREAU OF LABOR STATISTICS, U.S. DEP’T OF LABOR, BULL. NO. 634, CHARACTERISTICS OF COMPANY UNIONS, at 4 (1937) (noting that all of the company unions studied “were, more or less consciously and explicitly, offered to the workers as an alternative either to dealing with the employer individually or through a trade-union”).

143. Id. at 79 (“In more than half of the cases of company unions formed during the N.R.A. period, recently established trade-union locals contended for the right to represent the workers.”). In those cases, the trade unions either collapsed or were ignored by the company in favor of the company union. Id. at 79–80.

144. See Kaufman, supra note 131, at 780 (“[T]here is no pretense that [employee involvement programs are] not a part of a union avoidance strategy.”).

this suggests union avoidance remains a driving force behind employee participation.

Of course, avoiding unionization can be beneficial to employers and employees. Union avoidance can create better working conditions for employees by spurring employers to offer better wages and benefits, including providing employees with a voice in conditions of work—a union substitution approach.\textsuperscript{146} This approach looks at employee representation from a market perspective; employees will choose employee involvement programs over unions because unions offer less benefits or higher transaction costs.\textsuperscript{147} On its face, using competition to increase labor-management cooperation and improve work conditions makes sense. However, union substitution cannot accomplish these goals unless it results in fair competition where employees make informed choices whether to unionize.\textsuperscript{148}

Employee committees or involvement programs can be used deceptively by less scrupulous employers who may implement programs that give the appearance of representation without offering employees a meaningful voice.\textsuperscript{149} What was true in 1935 remains true today—employee independence is a predicate of meaningful workplace participation.\textsuperscript{150}

Certainly, employees represented by nonunion employee committees can experience a degree of independence.\textsuperscript{151} However, a workplace

\begin{flushleft}
employer committees set up during organizing campaign; Garney Morris, Inc., 313 N.L.R.B. 101 (1993) (upholding a bargaining order where the employer instituted employee committees during union organizing drive).

\textsuperscript{146} Kaufman, supra note 131, at 780-81 (distinguishing between “high-road employers”—who avoid unionization by offering above-market wages, job security, formal dispute resolution, and a culture of fair dealing—and “low-road employers”—who suppress union activity through coercion).

\textsuperscript{147} Id.

\textsuperscript{148} See id. at 803 (“[F]or competition to serve the public interest it must take place on a level playing field . . . .”).

\textsuperscript{149} See id. at 787 (observing that establishment of an employee committee “can, for example, buy the employer more time by deceptively convincing the employees that the company is seriously interested in resolving their complaints and promoting improved relations”).

\textsuperscript{150} See 78 CONG. REC. 4230 (1934) (article by Sen. Wagner), reprinted in 1 LEG. HIS., supra note 2, at 22 (“[O]hlen representatives who are not subservient to the employer with whom they deal can act freely in the interest of employees.”); cf. BUREAU OF LABOR STATISTICS, U.S. DEP’T OF LABOR, BULL. NO. 634, CHARACTERISTICS OF COMPANY UNIONS, at 4 (1937) (“The ability of employee representatives to secure favorable action on their recommendations through joint committees was affected in some instances by fear of discrimination . . . . Employee representatives expressed their fear of bringing up grievances under such circumstances.”).

\textsuperscript{151} Cf. Gen. Foods Corp., 231 N.L.R.B. 1232 (1977) (permitting delegation of managerial functions such as safety inspections, setting hours within established limits, and
committee’s power, as expressed in its ability to have management agree to its proposals, is only as great as the committee’s bargaining power. Employee bargaining power in turn relies on management’s interest in avoiding the costs of replacing its workforce. Thus, employees who must rely on market forces to guarantee meaningful employee representation can be disadvantaged during periods of economic slowdown or in other situations where employers have access to relatively large applicant pools.

In contrast, the risk of unionization is an independent check on employers, who must guard against worker dissatisfaction to avoid organizing. That check becomes less effective when the employer can usurp the representational role of the union through an unrestricted employee committee that competes unfairly with independent unions. Unfair competition in turn erodes unions’ effectiveness and thus their market impact as a check on employer conduct, disturbing the balance reflected in § 8(a)(2).

The potential impact of Crown Cork & Seal on this balance is perhaps best illustrated by a comparison to congressional efforts to expand the scope of permissible employee involvement under § 8(a)(2). In the mid-1990s Congress considered legislation—the Teamwork for Employees and Managers (TEAM) Act—that would have broadened the scope of permissible employee participation programs without corresponding interviewing applicants).


153. The economic cost associated with the loss of skilled workers is an important source of leverage for employees. See Kaufman, supra note 131, at 794 (“One key constraint on companies is employees’ capacity to quit one firm and find alternative employment at another.”).

154. Cf. id. at 797 (noting that during the Great Depression, due to high unemployment, workers who felt unjustly treated by company unions were powerless to obtain alternative representation).

155. See id. at 795 (“[M]anagement is strongly motivated to operate nonunion employee representation committees in a responsible, above-board manner to the extent that there is a viable threat from a union.”).

156. Cf. id. at 787 (“Establishment of an employee committee . . . can, for example, buy the employer more time by deceptively convincing the employees that the company is seriously interested in resolving their complaints and promoting improved relations.”).

increases in protection for employees’ bargaining power.\textsuperscript{158} The TEAM Act did not become law due in part to strong opposition from organized labor,\textsuperscript{159} and in part because opponents feared that it would upset the “balance between employee protections and the legitimate business needs of the employer.”\textsuperscript{160}

The TEAM Act would have amended § 8(a)(2) to permit employers in nonunion workplaces to establish a broad range of employee participation organizations, empowered to address matters of mutual interest, including statutory subjects of bargaining under the NLRA, such as hours, wages, and conditions of employment.\textsuperscript{161} This expansion of permissible employee participation programs was proposed without a concomitant increase in protections for employee organizing such as expedited elections or stronger penalties for employer unfair labor practice.\textsuperscript{162} Thus, the TEAM Act would have permitted employers to form and support organizations to compete with unions by providing an alternative, non-collective-bargaining mechanism for dealing with statutory subjects of bargaining,\textsuperscript{163} an expansion that, in the absence of stronger protections for union organizing, would have changed substantially the balance of power between employers and employees reflected in § 8(a)(2).\textsuperscript{164}

\textsuperscript{158} See S. 295 § 3 (carving out an exception to § 8(a)(2) for employee involvement programs in nonunion workplaces that “address matters of mutual interest, including but not limited to, issues of quality, productivity, efficiency, and safety and health . . . .”) (emphasis added); S. REP. NO. 105-12, at 49 (1997) (“[T]he TEAM Act bolsters employer prerogatives without a commensurate enhancement of employee rights under the NLRA.”).

\textsuperscript{159} See Adam Clymer, Clinton and Dole Clash on Delay of Wage and Gas-Tax Measures, N.Y. TIMES, May 9, 1996, at B12 (reporting Democratic and union opposition to the TEAM Act).

\textsuperscript{160} S. REP. NO. 105-12, at 31.

\textsuperscript{161} S. 295 § 3; S. REP. NO. 105-12, at 21–22. Under the TEAM Act, employee participation organizations would have been prohibited from seeking exclusive bargaining representative status, or from entering into collective bargaining agreements. Id. at 22.

\textsuperscript{162} See Kaufman, supra note 131, at 803 (noting that the NLRA contains “shortcomings and weaknesses that allow employers to exercise undue influence and coercion on employees’ choice of a representational agent” which the TEAM Act does not remedy).

Congressional opponents of the TEAM Act viewed it largely as an end-run around employee protections in the collective bargaining process, granting employers broad power to “engage in unlimited bargaining” with employees, without statutory good faith protections, so long as those negotiations did not result in a written collective bargaining agreement. S. REP. NO. 105-12, at 34.

\textsuperscript{163} See Kaufman, supra note 131, at 802–03 (noting that the TEAM Act allowed employers to “set up any and all forms of nonunion employee representation plans and thus promote[ ] greater competition between union and nonunion forms of representation”).

\textsuperscript{164} See S. REP. NO. 105-12, at 53 (letter from Acting Secretary of Labor Metzler) (“Rather than promoting genuine teamwork [the TEAM Act] would undermine the delicate system of checks and balances between employer and employee rights and obligations that
Like the changes proposed in the TEAM Act, *Crown Cork & Seal* arguably expanded the permissible scope of employee involvement programs far beyond the boundaries developed by the Board in previous § 8(a)(2) case law.\(^{165}\) The TEAM Act purported to do so by creating a statutory exception for employee participation programs in nonunion workplaces that could talk about, *inter alia*, wages, hours, and conditions of employment, whereas in *Crown Cork & Seal* the Board sidestepped the boundaries of § 8(a)(2) by expanding the existing managerial authority exception to include employee committees that issued recommendations to management on issues including pay increases, employee terminations for cause, layoff procedures, and overtime schedules.\(^{166}\) Where the TEAM Act would have explicitly precluded finding an unfair labor practice where employers “address” issues with employee organizations,\(^{167}\) in *Crown Cork & Seal* the Board avoided the *dealing with* analysis by construing the managerial exception to include bilateral communication between the employer and employee committees.\(^{168}\) Clearly, *Crown Cork & Seal* opened the door for some of what the TEAM Act purported to do, permitting a back and forth exchange between employers and employee committees and effectively expanding the range of permissible subjects of communication between employers and employer-dominated employee committees to include some of the statutory subjects under § 2(5).\(^{169}\) Thus, like the TEAM Act, *Crown Cork & Seal* arguably changed the balance between employers and employees reflected in § 8(a)(2), expanding the scope of permissible employee involvement without providing concomitant increased protections for union organizing.

---

165. *See* Pauling & McGuire, supra note 10, at 231 (“*Crown Cork & Seal*... potentially is the first board decision representing a dramatic shift away from... *Electromation and E.I. du Pont de Nemours.*”).

166. *Compare* S. 295 § 3 (“[T]o address matters of mutual interest, including but not limited to, issues of quality, productivity, efficiency, and safety and health...”), and S. REP. No. 105-12, at 34 (statement of Sen. Reed) (“[I]n fact there are no limits to what they could talk about, so effectively they could talk about... wages, hours [, and] conditions....”) (emphasis omitted), with *Crown Cork & Seal Co.*, 334 N.L.R.B. 699, 700 (2001) (identifying workplace issues on which employee committees would make recommendations to management).

167. S. REP. No. 105-12, at 55. Under the TEAM Act, the term *address* would have supplanted *dealing with* as the primary descriptor for employer–employee interaction. *See* id. at 34 (“In place of the term ‘dealing with,’ which has a well established meaning, the authors of the TEAM bill use the term ‘address’...”).

168. *See* 334 N.L.R.B. at 701–02 (finding that the employee committees did not deal with management because they were themselves management).

IV. A NEW TEST FOR DELEGATED MANAGERIAL AUTHORITY

The scope of permissible employee participation under § 8(a)(2) is due for congressional attention. However, the recent history of legislative action in this area of law indicates that no attention is likely forthcoming. Congress’s failure to pass the TEAM Act in the 1990s and the stalling of current legislative efforts to increase protections for union organizing both illustrate the difficulty of legislating in this area.¹⁷⁰

Until Congress acts, it is up to the Board to continue to define the scope of permissible employee participation within the bounds of § 2(5). In doing so, the Board should avoid disturbing the balance reflected in § 8(a)(2) while leaving room for legitimate employee participation in workplace decisionmaking. One way to accomplish this is to construe Crown Cork & Seal narrowly, applying the employee-committee-as-management analysis only in those rare instances when an employer has instituted the Socio-Tech system or a similarly comprehensive program of employee participation.¹⁷¹ For workplaces without such systems, the Board should return to analyzing functionally the decisionmaking authority of employee committees the way it did in Electromanation and du Pont. Narrowing Crown Cork & Seal’s application would reduce the likelihood that employers will abuse the managerial exception to unfairly impede union organizing activity; after all, implementation of a comprehensive system of employee involvement equivalent to the Socio-Tech system would require considerable time and expense on the employer’s part and thus would be less attractive as a mere pretext to temporarily disrupt union activity. However, should it choose to apply Crown Cork & Seal broadly, the Board need not look far afield for a settled definition of managerial with which to clarify the managerial test for employee committees.

While the NLRA does not define the term managerial,¹⁷² there is a well-settled, judicially developed exclusion for “managerial employees” which


¹⁷¹. Cf. Crown Cork & Seal Co., 334 N.L.R.B. 699, 701 (2001) (“[T]he Socio-Tech System represents a significant variation on the traditional plant organizational structure . . . .”); Pauling & McGuire, supra note 10, at 231 (“It is unlikely, however, that most employers would be inclined to adopt a structure as elaborate as socio-tech even if those employers did place a high value on sharing production authority.”).

applies to employees who “formulate and effectuate management policies by expressing and making operative the decisions of their employer.”\textsuperscript{173} Under this definition, managerial employees are “much higher in the managerial structure” than mere supervisors, and exercise discretion within or even independent of established employer policy.\textsuperscript{174} Normally, an employee is categorized as “managerial only if he represents management interests by taking or recommending discretionary actions that effectively control or implement employer policy.”\textsuperscript{175} The purpose of the managerial employee exception is to avoid conflicts of interest that could arise for employees from whom management requires loyal exercise of discretion, which could be complicated by their participation in a union.\textsuperscript{176}

The managerial employee test has already been successfully applied outside of the traditional industrial management context. In \textit{NLRB v. Yeshiva University},\textsuperscript{177} the Supreme Court applied the managerial exclusion to the faculty of a university, finding that while such employees are not completely analogous to an industrial manager, in the context of the university they were “managerial.”\textsuperscript{178} Similarly, the discretionary authority of an employee committee, as a whole, can be analyzed under the managerial employee test. Application of the well-settled managerial employee standard to the \textit{Crown Cork \& Seal} managerial authority exception would harmonize that decision with the existing managerial exclusion framework for nontraditional management structures reflected in \textit{Yeshiva} and avoid confusing multiplicity in the meaning of managerial under the NLRA.

However, that harmony would not come without costs. Incorporating the \textit{Yeshiva} managerial employee standard into the \textit{Crown Cork \& Seal} managerial authority exception leads to the anomalous result that employers would be permitted to deal with an employee committee on issues at the highest levels of workplace decisionmaking without running afoul of § 8(a)(2) but might find themselves at risk of committing unfair labor practices if the same committee engaged in lower-level workplace

\textsuperscript{174} \textit{Id.} at 682-83 (quoting \textit{Bell Aerospace}, 416 U.S. at 283, 286-87).
\textsuperscript{175} \textit{Id.} at 683.
\textsuperscript{176} \textit{Bell Aerospace}, 416 U.S. at 290 n.20. On remand, the Board addressed the question of whether buyers in \textit{Bell Aerospace} were managerial employees, and determined that they were not, despite their discretion to make purchases up to $5,000, because the buyers were guided by numerous manuals and instructions, there was one supervisor for every three buyers, and the buyers’ salaries and working conditions were comparable to those of non-managerial employees. 219 N.L.R.B. 384, 386 (1975).
\textsuperscript{177} 444 U.S. 672 (1980).
\textsuperscript{178} \textit{Id.} at 689.
decisionmaking subject to managerial review. The intellectual capital of employees—the untapped added value that they can bring to an employer—pertains as a rule to the workplace systems with which they interact daily, not to the control or implementation of employer policy performed in a traditional workplace by high-level managerial employees. Integration of the managerial employee standard into the delegated managerial authority exception to \textit{dealing with} under \textit{Crown Cork \& Seal} thus would promote consistency in Board law but ultimately do little to promote beneficial, effective employee participation.

The most useful alternative would be for the Board to abrogate the \textit{Crown Cork \& Seal} managerial authority exception and return to the functional analysis of \textit{dealing with} applied in \textit{Electromation} and \textit{du Pont}. From this starting point, the Board could proceed with a more nuanced development of the delegated managerial authority exception, an exception based on identifying the locus of decisionmaking authority in employer–employee interactions, that acknowledges the increasingly complex and permeable boundaries between employers and employees in workplace decisionmaking processes and that protects workers from employers for whom employee involvement is simply a method of avoiding legitimate employee representation.

\textbf{CONCLUSION}

\textit{Crown Cork \& Seal} arose in the context of an ongoing debate about the benefits and risks of increased employee participation in workplace decisionmaking. While purporting to rely on settled Board doctrine, the analysis of managerial authority in \textit{Crown Cork \& Seal} represents a radical departure from the traditional understanding of the managerial exception, one that extends the exception well beyond its rational underpinnings. Such an extension may well be warranted by the substantial changes in employer–employee relations since the 1935 enactment of the National Labor Relations Act.\textsuperscript{179} However, the Board should leave to Congress the task of managing the expansion of permissible employee involvement and return to a pre-\textit{Crown Cork \& Seal} functional analysis of employer–employee relationships as the basis for expanded employee participation while preserving the careful balance of bargaining power between employers and employees expressed in § 8(a)(2).

If the Board is unwilling to overturn \textit{Crown Cork \& Seal} outright, it can still limit the risk posed by it in one of two ways. First, it can apply \textit{Crown Cork \& Seal} narrowly, limiting the managerial exception for employee

\begin{footnotesize}
\begin{itemize}
\end{itemize}
\end{footnotesize}
committees that issue nonfinal decisions to only those workplaces where the employer has instituted the Socio-Tech system or a similarly comprehensive employee involvement program. Second, it can reject the “comparable to a front line supervisor” test as fatally imprecise, and apply the well-settled “managerial” test applied to another nontraditional management paradigm in Yeshiva University, gaining internal consistency in Board law at the expense of a wholly effective solution. In any case, the Board should act at the earliest opportunity to shore up the dealing with doctrine in the wake of Crown Cork & Seal, returning balance and clarity to this area of the law.
***
RECENT DEVELOPMENTS

CONGRESS AND THE COURTS CLOSE THEIR EYES: THE CONTINUING ABDICATION OF THE DUTY TO REVIEW AGENCIES’ NONCOMPLIANCE WITH THE CONGRESSIONAL REVIEW ACT

SEAN D. CROSTON

TABLE OF CONTENTS

I. Does Anyone Comply with the Congressional Review Act? 908
II. Where Is Congress?............................................................................................................ 909
   III. Even More Importantly, Where Are the Courts?........... 912
      A. What the D.C. Circuit Got Wrong and the Indiana District Court Got Right........................................... 915
      B. What the Indiana District Court Missed........................................... 917
      IV. So What Can the Courts Do?........................................... 917

F Sean Croston is an Attorney at the United States Nuclear Regulatory Commission. The views expressed in this article are his alone and do not represent those of the United States Nuclear Regulatory Commission. He received his J.D. in May 2007 from the William & Mary School of Law and wishes to thank Steve Crockett for his thoughtful ideas and support.
“The clear message to the watching agencies is that they need not be concerned . . . So what if there is congressional review?”¹

I. DOES ANYONE COMPLY WITH THE CONGRESSIONAL REVIEW ACT?

The very first sentence of the Congressional Review Act (CRA or Act)² states that, “Before a rule can take effect, the Federal agency promulgating such rule shall submit to each House of the Congress and to the Comptroller General” several items.³ These items include a copy of the rule, a concise statement explaining whether it is a “major” rule under the CRA, the proposed effective date of the rule, and any regulatory analyses required by law.⁴ There is nothing particularly mysterious or complicated about this mandate.

But in the final week of 2009, the Congressional Research Service (CRS) issued a report on federal agencies’ compliance with the CRA.⁵ The report revealed that between 1998 and 2008, the agencies had failed to submit over 1,000 substantive final rules to the Comptroller General,⁶ as required by the Act.⁷ The agencies did not submit 101 rules during Fiscal Year 2008, indicating that CRA compliance has not improved over time.⁸ And this new report may actually understate the problem of CRA noncompliance.

In 2006, a congressional report indicated that “perhaps thousands of covered rules have not been submitted for review.”⁹ An earlier investigation revealed that three agencies alone had failed to submit over 7,500 covered rules for CRA review between 1996 and 1999.¹⁰

So why aren’t the agencies complying with this law? The requirement that agencies submit these items for inspection “[b]efore a rule can take

³. § 801(a)(1)(A), (A)(i)–[B(iv)] (emphasis added).
⁴. § 801(a)(1)(A)(i)–(B(iv).
⁶. Id. at 9–10.
⁷. § 801(a)(1)(A).
⁸. COPELAND, supra note 5, at 10.
¹⁰. Id. at 88 n.261 (citing COMM. ON GOV’T REFORM, NON-BINDING LEGAL EFFECT OF AGENCY GUIDANCE DOCUMENTS, H.R. REP. NO. 106-1009, at 539 (2000)).
effect” seems to indicate a substantial incentive to comply with the CRA: If the agencies do not make the necessary submissions, their rules will not take effect.\textsuperscript{11} Regulated entities could then ignore these rules, quickly frustrating the agencies and spurring them to follow the process set forth in the CRA.

At least, that’s how it should work. But alas, it’s not that simple.

II. WHERE IS CONGRESS?

The most obvious agent to enforce the CRA is Congress whose acts created the various agencies and gave them their statutory duties. Moreover, Congress enacted the CRA for the very purpose of making sure that agencies were not deviating from Congress’s script. In a post-enactment statement, the CRA’s bipartisan sponsors\textsuperscript{12} explained:

As more and more of Congress’[s] legislative functions have been delegated to federal regulatory agencies, many have complained that Congress has effectively abdicated its constitutional role as the national legislature in allowing federal agencies so much latitude in implementing and interpreting congressional enactments.

In many cases, this criticism is well founded. Our constitutional scheme creates a delicate balance between the appropriate roles of the Congress in enacting laws, and the Executive Branch in implementing those laws. This legislation will help to redress the balance.\textsuperscript{13}

The CRS report agreed that “the CRA was enacted in an attempt to reclaim a measure of congressional control”\textsuperscript{14} by requiring a period of congressional review before rules could go into effect.

But there are some problems with congressional enforcement of the CRA. For one, it is difficult for Congress to “review” agencies’ actions if the agencies never submit those actions for review in the first place, as required by the statute. If the agencies ignore the CRA’s mandate and simply issue their rules, then an already quite busy Congress must carefully read through the voluminous daily editions of the Federal Register to see what it can catch. Congress would need some chart of rules that had been properly submitted for CRA review in order to compare and catch those that had slipped through its grasp. Moreover, this approach assumes that the agencies publish all rules covered by the CRA in the Federal Register. But “such covered documents are rarely if ever published in the Federal

\textsuperscript{11} § 801(a)(1)(A).

\textsuperscript{12} The sponsors included Senators Reid (Democrat), Nickles (Republican), and Stevens (Republican). See 142 Cong. Rec. S3683 (daily ed. Apr. 18, 1996) (statement of Sen. Nickles submitting bill on behalf of himself and Sens. Reid and Stevens).

\textsuperscript{13} \textit{Id.} (joint statement of House and Senate sponsors).

\textsuperscript{14} \textit{COPeland, supra note 5, at 21.}
Register and thus will come to the attention of committees or Members only serendipitously.”

Then again, as former Solicitor General, former Dean of Harvard Law School, and recent Supreme Court appointee Elena Kagan commented, Congress could always conduct its own “independent oversight activity. With or without a significant presidential role, Congress can hold the same hearings, engage in the same harassment, and threaten the same sanctions in order to influence administrative action.” But there are structural problems inherent in any model of congressional enforcement.

First, because the agencies are in the Executive Branch and at least nominally under the President’s control, “Congress rarely is held accountable for agency decisions.” If regulated entities are upset because agencies are passing secretive, burdensome rules without complying with the CRA, they will probably not take out their anger on Congress. They will blame the agencies, which are naturally at fault, and perhaps complain to the Office of Information and Regulatory Affairs (OIRA) or other executive actors. The general result is a congressional “lack of interest” in CRA enforcement.

In addition, as Kagan noted, the “partisan and constituency interests of individual members of Congress usually prevent them from acting collectively to preserve congressional power—or, what is almost the same thing, to deny authority to the other branches of government.” Certainly, some members of Congress who might be in the minority but who agree with the agencies’ policy decisions will be pleased that the agencies are circumventing Congressional review. Or at least they will not object. Nor will other members of Congress, who might not be thrilled with the substantive decisions reflected in the agencies’ actions, but have bigger fish to fry—they will be more concerned with immediate constituent concerns and the weightier policy issues of the day. And perhaps “many in Congress prefer business as usual and don’t want to risk political capital on attacking regulations that the public seems to support,” even “highly controversial, costly rules.” The director of regulatory studies at the libertarian Cato Institute suggested that many members of Congress did not “want to be perceived as being nasty” in opposing certain rules.

Moreover, “because Congress’s most potent tools of oversight require

15. INTERIM REPORT, supra note 9, at 88.
17. Id.
21. Id.
collective action (and presidential agreement), its capacity to control agency discretion is restricted.” If agencies’ noncompliance with the CRA pleases the congressional majority, perhaps because it allows quick implementation of favored policies, then there are not many things a disturbed congressional minority can do to crack down on the agencies and enforce the law.

Like any other law, the biggest weapon in Congress’s CRA arsenal—a successful disapproval resolution nullifying the agency rule—requires passage in both houses of Congress along with the President’s signature (or two-thirds of Congress in the event of a presidential veto). In fact, in over fifteen years of the CRA, only one rule, the Clinton ergonomics regulation, has been undone. Congress simply cannot do much to threaten agencies daring to defy the CRA. And “the ready realization by agencies over time that passage of a disapproval resolution is highly unlikely . . . substantially reduce[s] the efficacy of such a threat.” Thus, agencies probably “will not factor in [the possibility of] Congressional disapproval as part of the rule development process.”

But Congress has its agents. For example, the Government Accountability Office (GAO), located in the Legislative Branch, has a statutory role in the CRA process, along with OIRA in the Executive Branch. “GAO has voluntarily taken on the task of determining which Federal Register rules it has not received, and has periodically notified OIRA of these missing rules.” But GAO would be hard-pressed to find rules issued but not published in the Federal Register, and OIRA often ignores GAO’s letters, which GAO does not forward to Congress.

Even if GAO wanted to change its practices and more aggressively consult Congress regarding agencies’ noncompliance with the CRA, it cannot necessarily fix the situation. GAO reported to the CRS that “it follows up with the agencies regarding any major rules that are missing,” but “currently ha[s] limited resources to take on additional responsibilities for CRA compliance enforcement.” So GAO tries its best, but is inherently limited by resources and does not manage to make much of an impact in enforcing the CRA submission requirements.

In sum, Congress cannot, or at least does not, do much to force agencies

24. INTERIM REPORT, infra note 9, at 84.
25. Id. (citations omitted).
27. Id.
28. Id.
to submit their rules in compliance with the CRA. But that still leaves another branch of government.

III. EVEN MORE IMPORTANTLY, WHERE ARE THE COURTS?

So why don’t regulated entities simply file lawsuits challenging agency enforcement of rules that were not submitted in compliance with the CRA? After all, the statute says that these submissions must occur “[b]efore a rule can take effect.” Presumably, a challenge could assert that an unreviewed rule has not taken effect, and therefore cannot be enforced.

The key impediment to this obvious approach to increased CRA compliance is a provision in the statute itself—§ 805—which contains one short sentence: “No determination, finding, action, or omission under this chapter shall be subject to judicial review.” As envisioned by Congress, this would rule out judicial challenges to, “for example, a determination by the OIRA Administrator that a rule is major or not, a Presidential determination that a rule should become effective immediately, an agency determination that ‘good cause’ requires a rule to go into effect at once,” or a challenge regarding “the adequacy of a Comptroller General’s assessment of an agency’s report.”

But the CRA’s legislative history indicates that lawsuits asking a court to consider whether a rule had gone into effect were not barred. “A key sponsor of the legislation, Representative McIntosh, commented during a floor debate that ‘[u]nder section 801(a), covered rules . . . may not go into effect until the relevant agency submits a copy of the rule and an accompanying report to both Houses of Congress.’ And the clearest indication appears in a joint statement issued shortly after the CRA’s enactment, in which its bipartisan sponsors emphasized that “[t]he limitation on judicial review in no way prohibits a court from determining whether a rule is in effect. For example, the authors expect that a court might recognize that a rule has no legal effect due to the operation of subsections 801(a)(1)(A) or 801(a)(3).”

The CRA sponsors’ joint statement makes it clear that § 805 was never meant to preclude review of whether rules that had not been submitted were in fact effective. And the subsequent congressional report simply

30. § 805.
31. INTERIM REPORT, supra note 9, at 79.
stated, “An unreported rule [can] not be enforced.”

As that report indicated, there was no committee hearing and little debate on the floor, so this joint explanatory statement “represent[s] the most authoritative contemporary understanding of the provisions of the law.” Of course, it is “post-enactment legislative history” and therefore “does not carry the weight that committee report explanations and floor debates provide.” For example, the Supreme Court has held that “post hoc statements of a congressional Committee are not entitled to much weight,” and “even the contemporaneous remarks of a single legislator who sponsors a bill are not controlling in analyzing legislative history.”

But in this case, several legislative sponsors, not one, issued the statement shortly after the Act passed. And the Supreme Court has favorably cited post-enactment legislative history on several occasions. For instance, the Court cited a drafter’s explanatory statement made eleven years after a bill’s passage, as well as a bill summary placed in the Congressional Record by its sponsor after passage, along with explanatory remarks made two years later. The joint explanatory statement for the CRA was also a bill summary placed in the Congressional Record by its sponsors after the bill’s passage.

In any event, taking the legislative history and the structure of the CRA into account, the intent of § 805 appears clear: “[T]his preclusion of judicial review would not apply to a court challenge to a failure of an agency to report a rule.”

But shortly after the CRA’s enactment, the Executive Branch began trying to limit review of federal agency action through § 805. In 1997, the Department of Justice (DOJ) stated that the language of § 805 “precluding

34. INTERIM REPORT, supra note 9, at 72.
35. Id. at 86 n.253.
36. Id.
39. See Pac. Gas & Elec. Co. v. State Energy Res. Conservation & Dev. Comm’n, 461 U.S. 190, 211 n.23 (1983) (noting that even though expressions of a subsequent Congress are not “particularly useful” in determining congressional intent, the post hoc statements in this case were made by the senator who was “an important figure” in drafting the Act at issue in the case).
41. INTERIM REPORT, supra note 9, at 79.
judicial review is unusually sweeping.” Then DOJ began arguing that § 805 effectively blocked all judicial enforcement of the CRA. In an early test case, a Texas federal district judge agreed with this view, shrugging that the “plain English” of the statute led the judge to conclude: “Apparently, Congress seeks to enforce the [CRA] without the able assistance of the courts.” But this opinion avoided all discussion of the CRA scheme and legislative history. A few years later, another federal district court in Ohio agreed that “the language of § 805 is plain” and precluded judicial review.

More recently, the District of Columbia (D.C.) District Court and then the Court of Appeals for the D.C. Circuit chimed in, supporting the earlier district court opinions in Texas and Ohio precluding judicial review of rules’ effectiveness under the CRA. In a 2008 opinion, the D.C. District Court considered an allegation that the Forest Service failed to submit rules to Congress as required by the CRA, and that the rules were therefore void. The court relied upon the “limited case law” above and the “unambiguous[]” language of § 805 and held that it was “statutorily barred from reviewing” the matter. Again, the court did not discuss the CRA’s overall structure or legislative history.

In 2009, the D.C. Circuit addressed this issue on review for the first time. In Montanans for Multiple Use v. Barboletos, the court considered the effectiveness of language in § 801 of the CRA along with the judicial review provision in § 805. Without considering the legislative history, the Court determined that § 805 “denies courts the power to void rules on the basis of agency noncompliance with the Act. The language of § 805 is unequivocal and precludes review of this claim.”

But like the three district court opinions noted above, the D.C. Circuit panel was wrong. The D.C. federal courts, in particular, failed to consider the rationale in an interesting but unpublished 2002 federal court decision from the Southern District of Indiana, United States v. Southern Indiana Gas and


46. Id. at 20–21.

47. 568 F.3d 225 (D.C. Cir. 2009).

48. Id. at 229.
Electric Co.\textsuperscript{49}

\textbf{A. What the D.C. Circuit Got Wrong and the Indiana District Court Got Right}

\textit{Southern Indiana Gas and Electric} involved the same basic legal question covered in the other cases: whether a federal court had jurisdiction to determine the effectiveness of an agency rule that should have been reported to Congress under the CRA.\textsuperscript{50} The particular case involved an allegation that the Environmental Protection Agency changed its policy regarding the applicability of its regulations to existing utilities without reporting this new interpretation to Congress under the CRA.\textsuperscript{51} The government argued that the court lacked jurisdiction to review the agency’s failure to submit its policy to Congress under § 805 of the CRA, relying upon the interpretations in earlier district court decisions.\textsuperscript{52}

But Chief District Judge Larry McKinney was not persuaded by the earlier cases.\textsuperscript{53} Rather than finding the judicial review language in § 805 “plain,” Judge McKinney perceptively determined that it was susceptible to two plausible meanings: (1) . . . Congress did not intend for courts to have any judicial review of an agency’s compliance with the CRA; or (2) Congress only intended to preclude judicial review of Congress’[s] own determinations, findings, actions, or omissions made under the CRA after a rule has been submitted to it for review.\textsuperscript{54}

Judge McKinney found the first interpretation, adopted by the district courts and since then by the D.C. Circuit, especially troubling. As he noted, it would mean that agencies could evade the strictures of the CRA by simply not reporting new rules, and courts would be barred from reviewing their lack of compliance. This result would be at odds with the purpose of the CRA, which was to provide a check on administrative agencies’ power to set policies and essentially legislate without Congressional oversight.\textsuperscript{55}

Because the CRA lacks an internal enforcement mechanism, Judge McKinney noted that this interpretation “would render the statute

\textsuperscript{50} See id. at *10–18 (discussing the court’s opinion that it has jurisdiction to review whether an agency rule is in effect even though the agency did not report it to Congress under the CRA).
\textsuperscript{51} Id. at *2–3.
\textsuperscript{53} Id. at *13.
\textsuperscript{54} Id.
\textsuperscript{55} Id. at *13–14.
ineffectual,” which could not have been the intent of Congress.

In support of the second interpretation, Judge McKinney pointed out that § 805 “precludes judicial review of a determination, finding, action, or omission under this chapter.” He stated that “[a]gencies do not make findings and determinations under this chapter,” but Congress does. Thus, he found it “logical” to hold that § 805 precludes judicial review of congressional findings made “after a rule is submitted by an agency” but allows review of “whether or not an agency rule is in effect that should have been reported to Congress.”

In addition, Judge McKinney relied on the CRA’s legislative history, which in his opinion supported his interpretation of the ambiguous judicial review provision in § 805.

He determined that the legislative history supported his view, citing the joint statement by the CRA’s sponsors that:

Section 805 provides that a court may not review any congressional or administrative “determination, finding, action, or omission under this chapter.” Thus, the major rule determinations made by the Administrator of the Office of Information and Regulatory Affairs of the Office of Management and Budget are not subject to judicial review. Nor may a court review whether Congress complied with the congressional review procedures in this chapter.

Moreover, Judge McKinney relied upon the conclusion of that joint statement, which noted that “the limitation on judicial review in no way prohibits a court from determining whether a rule is in effect.” That statement clearly and directly conflicts with the other district court decisions and the D.C. Circuit’s opinion in Barboletos, which held there could be no judicial review of a rule’s effectiveness under the CRA.

Finally, Judge McKinney concluded that “[i]f Congress wanted to bar [all] judicial review . . . concerning the applicability of any of the provisions of the CRA, it would have clearly done so.” But Congress did not do that. Instead, “Congress limited its judicial review preclusion by referring

56. Id. at *14
57. Id. (quoting 5 U.S.C. § 805 (2000)).
58. Id. at *14.
59. Id.
60. Id. at *14–15 (citing Bd. of Trade of Chicago v. SEC, 187 F.3d 713, 720 (7th Cir. 1999) (for the use of legislative history to interpret genuinely ambiguous statutes)).
2010] Noncompliance with the Congressional Review Act

917

toward subs under requires effective as repor th stated determinations. Congress’s So
review judicial Congress and to 2 6 CROSTO 01 0

n e
S d determinations, [70. 69. 65. Judge review 0

it is q rangely, later qu §

§

eter

N
r

E

er

O

M

N

OMB, was referring back to those duties when it enacted the CRA judicial review provision.”65

B. What the Indiana District Court Missed

Strangely, Judge McKinney held that § 805 simply “preclude[d] judicial review of Congress’[s] own determinations,”66 but he cited legislative history showing that it was also meant to preclude review of determinations made by OMB, an executive actor. He then stated that § 805 precluded only review of the “determinations, findings, and actions that the OMB and Congressional committees would be required to make under the CRA.”67 So it is not clear whether he thought that § 805 precluded review only of Congress’s actions, or whether it also precluded review of OMB’s determinations.

A later congressional report came to a different conclusion, although it stated that it was “certainly arguable that the Southern Indiana court’s view of the limited preclusiveness of section 805 is plausible and persuasive.”68 The report concluded that § 805 precluded judicial review of both of the congressional and OMB determinations noted by Judge McKinney, as well as review of “a Presidential determination that a rule should become effective immediately” or “an agency determination that ‘good cause’ requires a rule to go into effect at once.”69 As the report found, these actions are also clearly “determinations” made under the CRA, and should not be judicially reviewed under § 805. But that still leaves one thing.

IV. So What Can the Courts Do?

Judge McKinney’s opinion persuasively noted that any interpretation of § 805 implicitly allowing agencies to forego submitting rules to Congress under § 801(a)(1) would defeat the entire purpose of the statute. As the subsequent congressional report concluded, “the statutory scheme is geared toward Congressional review of all covered rules . . . and a reading of the statute that allows for easy avoidance defeats that purpose.”70 The report

65. Id. at *18.
66. Id. at *13.
67. Id. at *18.
68. INTERIM REPORT, supra note 9, at 92.
69. Id. at 79.
70. Id. at 93.
continued, arguing that “[i]nterpreting the judicial review preclusion provision to prevent court scrutiny of the validity of administrative enforcement of covered but non-submitted rules appears to be neither a natural nor warranted reading of the provision.”

Finally, the report predicted that “without the potential of court invalidation of enforcement actions based on the failure to submit covered rules, agencies are not likely to comply with submission requirements.”

Given agencies’ unfortunate and continuing history of ignoring the CRA’s submission requirements by failing to submit thousands of covered rules, and in light of Barbouletos and the other court decisions that condoned this behavior, the report’s prediction has unfortunately come true.

The CRA sponsors’ joint statement makes it clear that § 805 was never meant to preclude judicial review of whether rules that had not been submitted to Congress were in fact valid. Instead, as Judge McKinney and at least one academic reviewer have concluded, “The limitation on judicial review in no way prohibits a court from determining whether a rule is in effect.” And once the courts begin to step in, agencies’ compliance with the CRA should drastically improve.

71.  Id.
72.  Id.
73.  See, e.g., COPELAND, supra note 5, at 11 tbl.1 (detailing the number of substantive final rules not received by the GAO for Fiscal Year 2006).
74.  Rosenberg, supra note 32, at 1073.
THE ECONOMICS OF RAILROAD
“CAPTIVE SHIPPER” LEGISLATION

RUSSELL PITTMAN*

TABLE OF CONTENTS

Background ....... 920
I. The Problem: Recovering the Cost of the Network .............. 924
II. Proposed Legislation .................................................. 926
   A. Paper Barriers ...................................................... 926
   B. Refusals to Deal .................................................. 930
   C. Mergers .......................................................... 933
Conclusion ...................................................................... 934

Recent rate increases by freight railroads have refocused attention on regulation, deregulation, and regulatory reforms in the railroad industry. Specifically, some shippers have complained that a lack of competition among railroads adversely affects their shipping options and makes them “captive” to the high rates charged by the railroad companies serving them. While such complaints have a long history in the railroad industry, significant rate increases since the early 2000s, coupled with increased profits by the large railroads, have given the complaints greater weight. Depending on the index used and components included, average rail freight charges declined in the 1980s, may have leveled off or even increased slightly in the 1990s, but seem to have increased significantly since 2000. Similarly, railroad industry profitability increased slowly and

* Director of Economic Research, Economic Analysis Group, Antitrust Division, U.S. Department of Justice, and Visiting Professor, New Economic School, Moscow. The author thanks Ken Heyer, Donna Kooperstein, Henry Posner III, Bob Potter, Oliver Richard, Chris Sagers, Carl Shapiro, and Lou Thompson for comments on an earlier draft.

The Antitrust Division encourages independent research by its economists. The views expressed herein are entirely those of the author and are not purported to reflect those of the U.S. Department of Justice.
gradually from 1998 through 2004 and then began a more rapid increase.\(^1\) The current economy-wide recession has halted this upward trend for now, though the rail industry seems to be suffering less than others.\(^2\)

In January 2009, legislation was introduced in the U.S. House of Representatives and Senate to address these issues. The Railroad Antitrust Enforcement Act of 2009 (H.R. 233 and S. 146) would amend existing federal statutes to render railroad mergers, acquisitions, collective ratemaking, coordination, and other “anticompetitive conduct” subject to the jurisdiction of the antitrust statutes, with potential enforcement both by the Department of Justice Antitrust Division (Antitrust Division or Division) and the Federal Trade Commission (FTC), and through lawsuits brought by state attorneys general or private parties.\(^3\) Similar legislation has been introduced regularly in previous years. In this Article, I consider the economic issues raised by such legislation and seek to evaluate its likely effects on competition and welfare.

**BACKGROUND**

The types of railroad behavior addressed by legislation to protect “captive shippers”—mergers, as well as some categories of coordinated and unilateral actions—have in the past enjoyed a broad antitrust exemption, subject to enforcement by the Surface Transportation Board (STB) and formerly the Interstate Commerce Commission (ICC) rather than the courts, and judged under a “public interest” standard rather than the more targeted competition standards established by the Sherman and Clayton Acts and jurisprudence. Remedial legislation introduced in previous Congresses would have gone further than H.R. 233 and S. 146 in the direction of command-and-control regulation, not only removing antitrust exemptions but also, for example, requiring railroads serving captive shippers either to: (a) permit trains from competing railroads to operate over the track of the serving railroad in order to provide competitive pick-up or delivery to the shipper (i.e., compulsory “trackage rights agreements,”

---


with the access charge regulated); or (b) offer to carry freight to or from the nearest connection to a competing railroad and to interchange it there, rather than insisting on providing the “long haul” themselves (so-called “reciprocal switching agreements,” with the “short haul” tariff regulated).4

Captive shipper is an STB term of art describing a goods shipper lacking economic alternatives to the single railroad serving it, those alternatives being either intramodal—competition from other railroads—or intermodal—competition from carriers using other modes, such as road or water.5 A former STB chairman has estimated that 15–20% of all rail movements involve captive shippers;6 he has also noted that, under the STB’s distinction between smaller captive shippers, who may bring cases under the streamlined “small rate case procedures,” and larger captive shippers, who must use the “large rate case procedures,” the latter class consists entirely of about seventy-five coal shippers.7 Recent rough STB calculations suggest that around two-thirds of captive traffic (as measured by revenues) consists of coal and chemical shipments.8

Under existing statutes and STB jurisprudence, the STB may intervene in the rate set by a railroad to a particular shipper if three conditions are met: (1) the rate exceeds 180% of the variable cost of carrying the traffic; (2) a “qualitative” STB assessment determines that there is no feasible, economic transportation alternative for the traffic involved; and (3) the rate

---


5. Whether an alternative is an “economic” alternative is often not a close call; water transport may be either available or not, and truck transport is generally uneconomic for bulk commodities such as coal.

6. See JOHN FRETTELLI, CONG. RESEARCH SERV., REPORT NO. RL34117, RAILROAD ACCESS AND COMPETITION ISSUES 1 n.2 (2006) [hereinafter CRS REPORT] (An estimate by the former chairman of the Surface Transportation Board (STB) is that about 80% of rail customers are served by only one railroad, but that because most of these customers can also ship by other modes, only about 15% to 20% of all rail movements would be judged captive by the STB.” (citation omitted)). Another extremely rough estimate is the calculation of the Government Accountability Office (GAO) that in 2004, 12% of rail industry revenue and 10% of industry tonnage traveled on origin–destination routes with access to only a single Class I railroad in both origin and destination Bureau of Economic Analysis economic areas.


8. Calculated from Table 5 in Simplified Standards for Rail Rate Cases, STB Ex Parte No. 646 (Sub-No. 1) (Sept. 4, 2007), 2007 WL 2495309, at *27.
is found to cross-subsidize other traffic on the railroad.9 The variable cost measure is defined precisely in STB jurisprudence, as is the cross-subsidization test—the latter using a complex regulatory construct called the “stand alone cost” test.10

The background to the broader issue is that the U.S. freight railroad industry has become much more concentrated in the decades since the substantial deregulation implemented by the Staggers Act of 1980, with only two Class I railroads now serving most of the western United States (the Union Pacific (UP) and the Burlington Northern Santa Fe (BNSF)), two Class I’s serving most of the east (the Norfolk Southern (NS) and the CSX), one running north-south in the heartland (the Kansas City Southern), and two Canadian Class I railroads with some U.S. operations and major U.S. connections (the Canadian National and the Canadian Pacific).11 The Antitrust Division argued at the STB in favor of more stringent conditions than were imposed in the merger that formed the BNSF and against the merger that formed the UP, but under existing law it did not have jurisdiction to challenge the mergers directly under the Clayton Act.12 The Division has more recently argued against those claiming efficiencies and benefits from the UP merger ex post.13

In recent years, the STB has apparently determined that the merger era has gone far enough—first imposing a moratorium on major consolidations and then lifting the moratorium but announcing that significantly more stringent conditions than in the past must be met for approval of further major consolidations:

11. The STB divides railroads into Class I, Class II, and Class III based on annual revenues. 49 C.F.R. § 1201.1-1(a)–(b)(1) (2009). Currently a Class I railroad is one with annual revenues exceeding $401 million; a Class II has revenues between $32 million and $401 million; and a Class III has revenues less than $32 million. See Surface Transportation Board, FAQs, http://www.stb.dot.gov/stb/faqq.html (last visited Aug. 2, 2010) (explaining that Class I, II, and III figures are updated annually for inflation using the Railroad Freight Index Price).
Because of the small number of remaining Class I railroads, the fact that rail mergers are no longer needed to address significant excess capacity in the rail industry, and the transitional service problems that have accompanied recent rail mergers, we believe that future merger applicants should bear a heavier burden to show that a major rail combination is consistent with the public interest. 14

A separate and relevant recent development is that since around 2000 the Class I railroads have increased their earnings to the point where earnings are arguably—though not uncontroversially—close to the railroad industry’s cost of capital. According to the recent Christensen Associates study for the STB, in 2005 the profitability of the BNSF rose above the industry’s cost of capital, whether measured by the STB’s old (discounted cash flow (DCF)) or new (capital asset pricing model (CAPM)) methodology; NS profitability rose above the cost of capital according to the new methodology in 2004 and 2005; CSX profitability did the same in 2005; and only the UP earned less than both measures of the cost of capital in 2005—though UP profitability had exceeded the newer measure in 2002 and 2003. 15

Finally, among shippers who are not “captive” there remains great variation in the degree to which they have economic options to substitute away from a particular railroad to ship their goods, whether to other railroads (intramodal competition) or to shippers using modes such as road and water (intermodal competition). In particular, the standard assumption and practice in goods transport is that bulk goods (for example, coal) and goods traveling long distances (for example, over 500 miles) travel more economically by rail (or, where feasible, water), while high-value goods (for example, many manufactured goods other than bulk chemicals) and goods traveling shorter distances travel more economically by truck. 16 A complication to this standard practice is that railroads have recaptured some of the higher value traffic lost over the years to trucks by utilizing containers rather than boxcars or other specialized rolling stock—a method called “intermodal shipping.” 17

---


15. See LAURITS CHRISTENSEN ASSOC., supra note 1, at 8-58 to 8-60.


17. Note that this definition of “intermodal shipping” coexists uneasily with the definition of “intermodal competition.” See infra Part I.
I. **The Problem: Recovering the Cost of the Network**

Arguably, the principal economic issue here concerns the means used by network owners to recover the common costs of constructing and operating the network. Like other network industries, railroads have large common, fixed, and sunk costs that are by definition not uniquely allocable to particular lines of business or operations. In the case of railroads, a widely accepted working figure is that fixed costs of infrastructure constitute about 25% of total costs.  

There are a number of ways for an industry with high fixed costs to set prices for their recovery, and each has its own advantages and disadvantages. At the high price end, an unregulated profit-maximizing seller would set a monopoly price, with society suffering from the associated reduction in output and deadweight loss in welfare. At the low price end, a regulated seller could be ordered to satisfy traditional “first best” conditions by charging a price equal to marginal cost, with society benefiting from the associated high level of output but forced to make up the fixed costs in some other way—for example, through subsidies paid out of taxation, with the corresponding deadweight losses associated with taxation as well as the negative incentive effects of subsidies. The traditional “happy medium” chosen between those two in U.S. regulatory practice has been “fully allocated cost” pricing with a deadweight welfare loss remaining but smaller than in the monopoly case—basically prices set at average costs, with a portion of fixed costs added to marginal costs according to some unavoidably arbitrary formula associated with, for example, shipment weight, distance, value, or a combination of any of the above.

With the Staggers Rail Act of 1980 and then the *Coal Rate Guidelines* of 1985 and the *Non-Coal Rate Guidelines* of 1996, the ICC and STB adopted the economists’ preferred “second best” alternative to these three options: price discrimination. In particular, the agency decisions recognized that a particular type of third degree price discrimination—“Ramsey pricing”—constitutes a methodology for an enterprise to recoup a given amount of fixed costs with a minimum loss of social welfare. Under

---

18. Of course the precise figure depends on time frames and the definition of “fixed” costs. In discussing the current economic downturn, one analyst noted that “[r]ails can usually drop operating expenses about 40%+ of the volume decline.” E-mail from Anthony Hatch, ABH Consulting, to author (July 7, 2009; 16:38 EST) [on file with author].


Ramsey pricing, rates are set in inverse proportion to the elasticity of demand of each customer in order to minimize the reduction in output stemming from the necessity of charging prices exceeding marginal cost.\(^\text{24}\) In the case of freight railways, the elasticity of demand of an individual shipper is most closely related to that shipper’s ability to switch from one rail carrier to other rail carriers or to carriers using other modes, though—as this elasticity is one of derived demand—it also reflects to a generally smaller degree the elasticity of demand for the product being shipped, the products for which that product is an input, and so on down the vertical chain. Although the setting of prices by a railroad according to individual shipper demand elasticity sounds complex, Laffont and Tirole show that this requires no more information than that used by any non-regulated enterprise to set discriminatory prices in a profit-maximizing way.\(^\text{25}\)

Unfortunately, however, Ramsey pricing, being a form of price discrimination, is discriminatory: different shippers, and especially different classes of shippers, pay rates for rail shipping that vary a great deal, both in absolute terms and in relation to the railroad’s cost of serving them, and those who pay the higher rates question the fairness of the system. Economists argue that price discrimination is not necessarily welfare-harming and may in fact be welfare-improving, especially when used as here to cover large fixed costs. For example, if too high a percentage of common costs is charged to customers with more elastic shipping demand, those customers may switch to road or water carriers instead—thereby depriving the railroad of what coverage of fixed costs those customers were providing, and increasing the amount that must be covered by remaining customers.\(^\text{26}\) And the discrimination suffered by captive shippers under STB jurisdiction is limited: the stand-alone-cost test cited above provides a


\(^{25}\) See Laffont & Tirole, supra note 24, at 60–66.

\(^{26}\) As one industry analyst notes, advocates of the proposed legislation are in effect attacking “other rail customers” instead of railroads. “Railroading is a capital intensive industry that needs a certain amount of revenue to pay for day-to-day operations and support capital spending programs that shippers demand. . . . If customers must provide the money for rail capital programs, the battle really is about how cost responsibility is apportioned.” Larry Kaufman, *Shipper vs Carrier or Shipper vs Shipper? Argus Rail Bus.*, Sept. 28, 2009, at 3.
ceiling beyond which shippers may not be forced to pay for the fixed costs of the system. Nevertheless, it is true, as shippers argue, that the stand-alone-cost test is complex and expensive to utilize and adjudicate—and it certainly does not, and is not intended to, prevent discrimination.27

II. PROPOSED LEGISLATION

The primary thrust of recent legislative proposals responding to complaints that rail freight rates are both high and discriminatory has been to eliminate the exemption from the antitrust laws afforded to much railroad behavior under current statutes. This is also the primary thrust of the endorsement of the 2007 legislation by the Antitrust Section of the American Bar Association—almost half of the content of its “Comments” discusses the importance of limiting exemptions of any kind to the antitrust laws.28 While the antitrust exemption does not protect all anticompetitive activity,29 eliminating the antitrust exemption may indeed both improve the situations facing some captive shippers and increase overall economic welfare.30

There are two classes of railroad behavior which those supporting such legislation may hope will come under increased scrutiny and attack if jurisdiction is granted to the antitrust enforcement agencies: “paper barriers” and “refusals to deal.”31

A. Paper Barriers

The first issue, paper barriers, is a seemingly obscure railroad issue that

29. Letter from STB to S. Comm. on Commerce, Sci., and Transp. 2 (Sept. 13, 2007) (footnote omitted) (on file with author) (explaining that railroads already “face civil and criminal liability for [cartel] violations of the Sherman Act (e.g., price-fixing, market allocation, bid rigging), and have been successfully sued for violating that Act”); see also Daniel Machalaba & John R. Wilke, Railroads Face Probe Over Prices for Shipping Coal, WALL ST. J., Feb. 17, 2003, at A2.
may have significant competitive implications. A common phenomenon since the passage of the Staggers Act—and indeed a generally acknowledged success story for the industry—has been the frequent spin-offs (or, alternatively, leasing) of small, lightly used rail lines by Class I railroads either to groups of local shippers to form a new Class III railroad or to an existing operator of a Class II or Class III railroad. In some cases these lines may have been abandoned by the Class I if they could not have been sold or leased. One reason for the survival and success of the spin-offs is arguably that the newly independent local lines are less constrained than the Class I railroads by union rules and wages; indeed, the president of the new line may also be its locomotive operator. Another reason is that these spin-offs constitute a reallocation of assets from large firms efficient at the “wholesale” level to smaller but more cost- and service-focused firms efficient at the “retail” level.32 The American Short Line and Regional Railroad Association estimates that “short lines now originate or terminate one out of every four rail cars moved by the domestic railroad industry.”33

The issue to be addressed by many legislative proposals is that the spin-offs creating or extending these short lines often come with a contract restriction: the newly independent local line must hand its interline cargo over exclusively to its former parent, rather than to other Class I’s to which it may connect. Alternatively, there may be strong financial incentives to favor the former parent. These contractual prohibitions to dealing with competitors of the former parent line are termed paper barriers by shippers—the rail industry prefers the more neutral term “interchange commitments”—and reformers wish to prohibit them.34 No one knows how many spin-offs have been accompanied by these restrictions; the STB imposed a reporting requirement only recently and did not include a requirement to account for existing agreements.35 However, the widespread nature of shipper complaints suggests that the number may be large.

32. See Rail Operator: Re-Regulation Would Hurt Shortlines, ARGUS RAIL BUS., Nov. 12, 2007, at 3; see also Tom Murray, A Different Way to Run a Railroad: Regional Versus Network Carriers, 71 J. TRANSP. L. LOGISTICS & POL’Y 295, 296–99 (2004); CRS REPORT, supra note 6, at 7 (“Especially in agricultural states, short-line railroads perform a gathering function, linking mostly rural shippers to high-volume Class I main lines.”).


34. A good case for elimination of paper barriers is provided by former Department of Justice Antitrust Division attorney Salvatore Massa in A Tale of Two Monopolies: Why Removing Paper Barriers Is a Good Idea, 41 TRANSP. J. 47, 50–56 (2001).

35. See generally Disclosure of Rail Interchange Commitments, STB Ex Parte No. 575 (Sub-No. 1) (May 21, 2000), 2000 WL 2216060, at *2.
Limiting or even getting rid of paper barriers may well be a good idea, but the potential costs should not be ignored. First, in some cases the efficiency gains that would be the subject of bargaining between the buyer and seller railroads may be smaller than the monopoly profits lost to the seller by the requirement to introduce competition. Second, in some cases potential short-line operators may lack access to the capital necessary for the purchase or lease of these lines at fully remunerative prices, so that transactions at lower prices accompanied by long-term traffic agreements may act as financing arrangements for these transactions. For both reasons, possible results of a prohibition on paper barriers would be more abandonments and fewer spin-offs. These costs of a rules change may not outweigh the benefits of increased competition for shippers on those lines that are spun off despite the rules change, but they should not be ignored. On the other hand, one well-known analyst who is generally sympathetic to the rail industry view of regulation and legislation regards this issue as a “no-brainer” that does not “appear worth fighting over,” as it would not “cost railroads much revenue.”

More broadly, it seems likely that paper barriers would be treated less sympathetically by antitrust enforcers and courts than they have been by the STB (though the two recent STB rulings in this area may suggest a tightening of STB restrictions as well). In general, contracts for the sale of assets that impose restrictions on the future competitive conduct of the buyer—for example, agreements not to compete with the seller in the future—are treated under the rule of reason in antitrust jurisprudence: they are not per se illegal, but they are frowned upon if they impose restrictions.


37. Note, however, that the abandonment option is not a simple one. The Staggers Act significantly reduced the barriers to a railroad abandoning a line, but abandonment remains regulated and constrained. STB permission must be applied for, and the STB will consider petitions from other potential users of the line, including Class II and III railroads. See, e.g., Siew Hoon Lim & C.A. Knox Lovell, Profit and Productivity of US Class I Railroads, 30 MANG. & DECISION ECON. 423, 425 (2009).

38. The American Association of American Railroads argues that interchange commitments are “core requirements without which [spin-off] transactions would not and could not take place.” Review of Rail Access and Competition Issues, supra note 33, 2007 WL 3170981, at *3 (citation omitted).


greater than necessary to achieve the efficiencies of the transaction itself.\textsuperscript{41} In particular, courts have ordered that such restrictions be limited in scope “with respect to duration, territory, and type of product,” and permanent agreements not to compete are frowned upon.\textsuperscript{42}

So the likely outcome of the imposition of antitrust jurisdiction to the creation of paper barriers would be not only some minor reduction in the number of sales of small, local lines to local shippers and Class II and III railroads, but also a reduction in the time frame of paper barriers for those sales that do remain, and thus an increase in competition for shippers at the end of the new time frame—perhaps three to five years. To the degree that paper barriers are a contractual response to financial constraints, and to the degree that barriers in the three to five year range are sufficient to address those constraints, the discouragement effect on spin-offs may be small. As argued by one shippers’ group in its petition for the STB to address the issue, “even if paper barriers may help preserve some trackage for continued use, it does not necessarily follow that paper barriers imposed as a condition of track sale or lease should be continued in perpetuity.”\textsuperscript{43}

Note, by the way, that an agreement among the Class I and smaller railroads, endorsed by the STB, has already limited the ability of a Class I to impose paper barriers for future business attracted by a new local spin-off.\textsuperscript{44} Note also that an ex post removal of paper barriers on lines spun off in the past would constitute the forced rewriting of one provision of complex contracts: the Class I railroads losing the contractual protection would generally have insisted on a higher price for spinning off the line had they known that they would subsequently lose its business. As argued by Henry Posner III, chairman of the Railroad Development Corporation and operator of the Class II Iowa Interstate Railroad:

If I were a Class I, I at least would certainly want a much higher purchase price from a buyer if I didn’t have the prospect of holding on to traffic. Many Class III railroads have been formed from pieces of Class I’s—at a discount—because of the Class I’s ability to implement paper barriers. . . .


\textsuperscript{42} 1 A.B.A. SECTION OF ANTITRUST LAW, ANTITRUST LAW DEVELOPMENTS 130 (6th ed. 2007) (citations omitted). Regarding the application of the tying provisions of antitrust law to paper barriers, see Massa, supra note 34, at 50–51, \textit{The Intersection of Competition Policy and Surface Transportation Regulatory Policy: An Examination of S. 772, the Railroad Antitrust Enforcement Act: Hearing Before the S. Subcomm. on Antitrust, Competition Policy, and Consumer Rights of the S. Comm. on the Judiciary, 110th Cong. 94–95 (2007) (statement of Darren Bush, Assoc. Professor of Law, Univ. of Houston Law Center).


\textsuperscript{44} See Massa, supra note 34, at 48.
On the other hand, since future revenue streams are discounted when deals such as these are made and priced, this objection loses its force the longer past the contract date the policy change is imposed.

B. Refusals to Deal

The second way in which the types of legislative proposals under discussion here might be expected to help captive shippers may be in their supplementation of STB jurisdiction with that of the Antitrust Division and the courts (including private enforcement) for “refusals to deal” by the railroad serving the shipper. In this case the “refusal” is that of the railroad serving the captive shipper to either: (a) allow the trains of a competing railroad to serve the shipper over the monopoly railroad’s tracks, or (b) offer to carry the shipper’s goods only to the nearest interchange with a competing railroad, rather than insisting on hauling the goods for the entire route itself.

These two services are termed in the industry “trackage rights” and “reciprocal switching,” respectively. Some legislative proposals for mandatory reciprocal switching or quotation of a rate to a nearby connecting carrier would order or permit the STB to impose the same outcomes in a regulatory fashion that the current legislation would aim for under the Sherman Act.

An important question that has apparently not been addressed about such a requirement is how it would work and what would be its ground-level implications in the real world—especially, how far a captive shipper typically is located from the nearest competing railroad, and thus how intrusive a requirement to deal would be on the railroad serving the captive shipper directly. The Christensen Associates report argues from a set of stylized facts that if policies such as reciprocal switching are required over short distances only, the resulting loss of system economies will be relatively small. The cost of creating and operating a trackage rights regime almost certainly increases with the distance of track involved, but the direct cost of a reciprocal switching regime does not seem obviously to increase with the distance from the shipper to the switch. Professor Curtis Grimm has suggested that the majority of captive shippers may be within perhaps 100 kilometers of a competing railroad, with the most important exceptions probably electric utilities located outside metro areas and some rural grain

45. Rail Operator, supra note 32, at 3.
46. LAURITS CHRISTENSEN ASSOC., supra note 1, at 22-7.
shippers, while rail expert Louis Thompson believes that something like 100–150 kilometers might be necessary to account for the majority of captive shippers.\footnote{E-mail from Curtis Grinn, Dean’s Professor of Supply Chain and Strategy, Robert H. Smith School of Business, to author (Feb. 10, 2009) (on file with author); e-mail from Louis Thompson, Principal, Thompson, Galenson & Associates, to author (Feb. 17, 2009) (on file with author).} One hundred kilometers would be a fairly long distance for frequent regulatory or judicial imposition of trackage rights.\footnote{For a broader discussion of the workability of shared access regimes, see José A. Gómez-Ibáñez, When Open Access Works: Lessons from North America’s Railroads (Sept. 2009) (unpublished manuscript, on file with author).} Lawsuits brought by the Antitrust Division, by state attorneys general, or by shippers charging illegal refusals to deal, might indeed force concessions by the railroads to captive shippers, either in the forms just suggested or in lower rates charged to settle the complaints. There are, however, two factors that could limit the amount of relief that this change in the law would actually bring to captive shippers.

First, a refusal to deal is a relatively uncommon antitrust offense. The Supreme Court has stated that, in general, a seller “has a right to deal, or refuse to deal, with whomever it likes, as long as it does so independently.”\footnote{Monsanto Co. v. Spray-Rite Serv. Corp., 465 U.S. 752, 761 (1984) (citations omitted).} At the same time, the Supreme Court has recently stated: “Under certain circumstances, a refusal to cooperate with rivals can constitute anticompetitive conduct and violate § 2.”\footnote{See generally A.B.A. SECTION OF ANTITRUST LAW, supra note 42, at 160–72.}

Refusals to deal have generally been found unlawful only when they have been part of a clear scheme to reduce competition—for example, to enforce an illegal tying arrangement—rather than simply because the “deal” refused would have been more advantageous to the buyer than the one offered.\footnote{Verizon Commc’ns, Inc. v. Law Offices of Curtis V. Trinko, LLP, 540 U.S. 398, 408 (2004) (citing Sherman Act, 15 U.S.C. § 2 (2000)).} In this respect, a refusal by one railroad company to allow trackage rights to a competing railroad or to hand over traffic at a nearby interchange point with a competing railroad may be at least somewhat comparable to its simply charging a high price: unpleasant for the customer, but not anticompetitive in the sense of the Sherman and Clayton Acts, and so not easily subject to antitrust challenge.

The possibility of court orders to allow switching would also raise an enforcement issue that is at the heart of judicial reluctance to impose dealing requirements generally: the necessity of monitoring and regulating the price and quality of the “deal” supplied under duress.\footnote{See Verizon Commc’ns, 540 U.S. at 408 (“We have been very cautious in recognizing...”}. On the other
hand, Professor Grimm points out that the Canadian regime of mandatory reciprocal switching is administered with a short and simple rate schedule.\textsuperscript{53} A further complication is that these would be unusual refusal-to-deal cases, as a complaining shipper would be seeking not only compulsory “deals” between itself and the defendant but also between the defendant and its competitor.

Second, recall that, thanks to the string of large mergers, U.S. railroading is now a very concentrated industry. The experience with mandatory switching or short hauls to hand-offs in Canada and Mexico has been somewhat disappointing, at least in part for a reason that is likely to restrict benefits in the United States as well: a competing duopolist fears that if it takes advantage of the opportunity to serve a shipper captive to its rival, the rival will in turn take advantage of the opportunity to serve its own captive shippers, and competition will break out throughout the system.\textsuperscript{54} This argument has some echoes of the facts alleged in \textit{Bell Atlantic Corp. v. Twombly}, in which each of the Regional Bell Operating Companies chose not to seek to enter the geographic markets of the other, despite legislated encouragement to do so.\textsuperscript{55} Even if antitrust enforcers were to begin bringing Sherman Act or Clayton Act cases against refusals to deal by railroads serving captive shippers, and even if those cases received favorable receptions in court, strong relief would require the cooperation of the competing duopolist.\textsuperscript{56}

such exceptions, because of the uncertain virtue of forced sharing and the difficulty of identifying and remedying anticompetitive conduct by a single firm.”).

53. E-mail from Curtis Grimm, Dean’s Professor of Supply Chain and Strategy, Robert H. Smith School of Business, to author (Mar. 4, 2009) (on file with author).


56. “Let’s assume that legislation reverses the bottleneck doctrine that says a railroad cannot be required to quote or file a rate for part of a move that it already can handle in its entirety, thereby retaining an effective monopoly on the traffic in question. BNSF, for example, might set a bottleneck rate that would require UP to compensate it fully for the use of its line and the opportunity cost of its investment in the line. In the limited capacity environment, UP might choose not to bid for the traffic, choosing to allocate its capacity to business where it has more attractive pricing power. How do you force two normally competitive railroads—the same concept applies to NS and CSX in the East—to act contrary to their self-interest?” Larry Kaufman, \textit{What happens if Shippers Prevail on Legislation?}, \textit{ARGUS RAIL BUS.}, Dec. 7, 2009, at 4.
2010] THE ECONOMICS OF RAILROAD “CAPTIVE SHIPPER” LEGISLATION 933

In fact, there is something of a natural experiment whose outcome supports this concern. Under the “bottleneck” rulings of the STB, a railroad that serves a captive shipper and interconnects near the shipper with a competing railroad is required to offer the shipper the alternative of a short haul to its competitor at a (potentially) regulated rate if the shipper can present an already signed contract with the competitor for its portion of the route. Consistent with the “concentrated industry” hypothesis, in practice the competing railroads have not signed such contracts with shippers.57

The best outcome for captive shippers might be that the threat of such litigation outcomes would force the railroads to negotiate lower rates rather than risk going to court.

C. Mergers

One reason for the current market power enjoyed by U.S. Class I railroads is the past mergers that have already been allowed by the STB—some of which, as noted above, were either opposed by the Antitrust Division or recommended only with more stringent conditions than were imposed by the STB. The result of these mergers has been two mammoth regional duopolies in which neither duopolist aggressively seeks to poach business from the other. Thus, had antitrust jurisdiction rested with the Antitrust Division at the time these mergers were proposed, the industry likely would be more competitive today.

As noted above, it is some consolation that the STB, having permitted these mergers in the past, has apparently decided that enough is enough: it has issued new regulations requiring a much stronger showing of public benefits to future major mergers—not simply a showing of a lack of competitive harm—than in the past. These requirements may take one or more forms, depending on the circumstances: (1) requiring future merger applicants to suggest conditions that would enhance, not merely preserve, competition; (2) requiring specific plans for how congestion problems will be addressed should they appear; (3) strengthening post-merger monitoring, and requiring compensation or compensatory acts if promised benefits do not appear; or (4) requiring forecasts as to how the proposed merger will cause subsequent merger scenarios to play out.

Future proposals for major U.S. railroad consolidations will thus arguably face a much more skeptical STB than have past proposals, and it is probably no accident that no such proposals have been put forward since

57. CRS REPORT, supra note 6, at 5 (“In practice, however, the non-bottleneck railroad generally has not entered into a contract with a shipper under these circumstances.”).

At this point in the evolution of the U.S. railroad industry, it seems much more likely that any merger of Class I railroads would be of the end-to-end rather than the parallel variety; neither the STB nor (if they are granted jurisdiction) the Antitrust Division nor the courts seem likely to approve mergers between either the two remaining western Class Is or the two remaining eastern Class Is. Whether further end-to-end mergers—perhaps resulting in one or two U.S. transcontinental railroads—would be a good thing or a bad thing is a separate matter. No one disagrees that end-to-end rail mergers have the potential to increase efficiency as they create single line service from former interline moves; indeed some of the Class Is currently devote significant management attention to reducing the inefficiencies inherent in interconnection.58

On the other hand, the econometric literature suggests that the U.S. Class Is have already exhausted potential economies of system size and are at or near the point of exhausting potential economies of density.59 And reducing the number of major U.S. railroads any further runs the risk of forfeiting other aspects of competition that have value for the economy—in technological and service improvements, for example, or in competition to provide incentives for new shippers to locate on a particular line. Given the merger record of the STB, granting merger authority to the Antitrust Division seems sensible to protect existing competition into the future.

CONCLUSION

Captive shippers are currently subject to some regulatory protection. Whether they are subject to enough protection would seem to be a matter of equity and fairness more than economic efficiency, since economic efficiency dictates that precisely those shippers who have the fewest economic

options—those who are captive—should contribute the most to the fixed costs of the railroad network. Moreover, railroading is a capital-intensive industry that requires continual investment for both maintenance and expansion in order to support economic growth; one cannot expect that investment if the investors are not permitted to earn a good return on it. A direct result of reduced investment in rail infrastructure would be higher costs, congestion, higher rates, and thus more freight traffic switching from railroads to trucks.\(^{60}\) Having said all this, however, it does appear that the Class I railroad companies are finally earning something like normal returns on capital, and perhaps even something in excess of normal returns, so that measures to restrict or even reduce the amount contributed by captive shippers to railroad returns would be appropriate.

The most direct and obvious path for this would be tighter STB regulation of rates charged to captive shippers. It is not the purpose of this paper to analyze possible revisions to STB regulations. However, the three factors listed above as predicates to STB action—P/VC exceeding a floor of 180\%, a qualitative finding of no economic options, and the stand-alone-cost test—may provide a framework for considering possible changes. For example, the stand-alone-cost test is an expensive, complex, and time-consuming regulatory requirement,\(^ {61}\) and as the journal article on which it is based makes clear, its primary claim to welfare enhancement is as a condition for the elimination of incentives for hit-and-run entry to serve a subset of customers—a factor that the serving railroad has every incentive to take account of on its own.\(^ {62}\) A possible route for simultaneously strengthening and streamlining the regulatory process would be to eliminate this portion of the exercise and replace it with a P/VC ceiling.

---

60. For a discussion of railroad investment requirements going forward, see CAMBRIDGE SYSTEMATICS, NATIONAL RAIL FREIGHT INFRASTRUCTURE CAPACITY AND INVESTMENT STUDY, 7-1 to 7-6 (2007).

61. As summarized on the website of the shippers’ organization Consumers United for Rail Equity,

“‘Stand alone cost’ refers to the expenses associated with a captive rail customer building and operating its own hypothetical efficient railroad. To develop this hypothetical railroad, the captive rail customer must retain lawyers, accountants, railroad economists and other such experts in this multi-million dollar exercise. The STB uses ‘stand alone cost’ to determine if a captive rail customer rate is ‘unreasonably high’ . . . . The only successful rate cases under this process have involved those coal movements that are moved over densely traveled rail lines.” Consumers United for Rail Equity, Unfair Federal Policies, http://www.railcure.org/issue/issue_unfair.asp (last visited July 6, 2010).

However, this is the topic of a different paper.  

Proposed legislative changes that would place more railroad behavior under the jurisdiction of the antitrust laws—in particular, restrictions on the creation of paper barriers in conjunction with the spin-off of local and regional rail lines, and restrictions on the ability of railroads to refuse to deal with competing railroads regarding the traffic of captive shippers—may act to limit the ability of railroads with market power to exploit that power at the expense of captive shippers. In particular, it seems likely that antitrust jurisdiction would impose time limits on the “exclusive dealing” requirements or incentives that often accompany the spin-off of a line from a Class I railroad to a Class II or III, and this would increase the level of competition enjoyed by shippers in the longer term. Other types of policies, such as mandatory switching—whether achieved through regulation or through antitrust challenge—face greater hurdles owing to the already highly concentrated structure of the U.S. railroad industry, but may nevertheless have some effectiveness in protecting captive shippers as well.

63. See generally Pittman, supra note 27.
Subscription and Order Form

The following are the prices for subscriptions and individual issues. Please contact us at 312-988-5522 with inquiries and orders.

1. Individual eligible for membership in the ABA (includes membership in Section of Administrative Law and Regulatory Practice) ........................................... $40.00

2. Organization or individual located in the United States or its possessions not eligible for membership in the ABA ................................................................. $40.00

3. Organization or individual located outside of the United States and its possession not eligible for membership in the ABA ....................................................... $45.00

4. Law students attending an ABA accredited law school (includes membership in Section of Administrative Law and Regulatory Practice) ...................... $10.00

5. Back issue(s) less than two years old (For issues published more than two years ago contact William S. Hein & Co., Inc. at 800-828-7571) ......................... $10.00 each

6. Quantity discount for five (5) or more copies of the same back issue ...... $8.00 each

<table>
<thead>
<tr>
<th>Volume/Number</th>
<th>Quantity</th>
<th>Cost/Issue</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Handling charge when ordering back issues $3.95.

Subscription (if student indicate name of school, if attorney give jurisdiction where admitted):

Total ..............................................................................................................

Please print
Name: ............................................................................................................
Title: ..............................................................................................................
Company/Organization: ...................................................................................
Address: .........................................................................................................
City/State/Zip: ............................................................................................... 
Telephone: .................................................................................................
Facsimile: ....................................................................................................
E-Mail: .........................................................................................................
Would you like to be placed on our e-mail mailing list G Yes G No

Method of payment
G Check enclosed (make check payable to The American Bar Association)
G Charge my credit: G MasterCard G Visa - Card Number: .............................

Expiration Date: ................................ Signature: ..............................................

Send this form via facsimile to 312-988-5568 or mail to Member Services, American Bar Association, 750 North Lake Shore Drive, Chicago, IL 60611-4497. Orders also accepted via telephone at 312-988-5522.

Please allow two to three weeks for delivery.

THANK YOU FOR YOUR ORDER