CONTENTS

ARTICLES

The “Voluntary” Inpatient Treatment of Adults Under Guardianship ..................... Richard C. Boldt 1

Insuring Floods: The Most Common and Devastating Natural Catastrophes in America ......................... Christopher C. French 53

Taking Back Takings Claims: Why Congress Giving Just Compensation Jurisdiction to the Court of Federal Claims Is Unconstitutional ........ Michael P. Goodman, Ph.D. 83

Changing Course in the Anti-Domestic Violence Legal Movement: From Safety to Security .............. Margaret E. Johnson 145

Halliburton, Basic, and Fraud on the Market: The Need for a New Paradigm .................. Charles W. Murdock 203

The Villanova Law Review seeks to publish materials representing meritorious viewpoints on subjects of interest to the legal profession. Publication does not indicate that the views expressed are adopted by the Villanova Law Review.

As a matter of policy, the Villanova Law Review encourages and promotes the use of gender-neutral writing.

Except as otherwise expressly provided, the author of each article in this volume has granted permission for copies of that article to be made available for classroom use in a nationally accredited law school, provided that (1) copies are distributed at or below cost; (2) the author and the Villanova Law Review are identified; (3) proper notice of copyright is affixed to each copy; and (4) notice of the use is given to the Villanova Law Review.

The Villanova Law Review (ISSN-0042-6229) is published in February, March, June, August, and October by Villanova University School of Law at 299 North Spring Mill Road, Villanova, Pennsylvania 19085-1682. Periodicals postage paid at Villanova, Pennsylvania and at additional mailing offices. Annual Subscription $35.00 . . . . . . Single Issue of Current Volume $8.00 . . . . . . Special and Symposium Issues $10.00.

POSTMASTER: Send address changes to Villanova Law Review, Villanova University School of Law, 299 North Spring Mill Road, Villanova, Pennsylvania 19085-1682.

UNLESS NOTICE TO THE CONTRARY IS RECEIVED AT THE EDITORIAL OFFICE, IT IS ASSUMED THAT A RENEWAL OF THE SUBSCRIPTION IS DESIRED.

SUBSCRIPTIONS ARE BILLED ANNUALLY.

ONE MONTH'S NOTICE IS NECESSARY TO EFFECT A CHANGE OF ADDRESS.

Change of address notification, subscription requests, Article submissions, requests for expedited review, and all other correspondence should be sent to:

Villanova Law Review
Villanova University School of Law
299 North Spring Mill Road
Villanova, Pennsylvania 19085-1682

Or via e-mail:
Review@law.villanova.edu

Address requests for back issues to:
William S. Hein & Co., Inc.
1285 Main Street
Buffalo, New York 14209
VILLANOVA UNIVERSITY
SCHOOL OF LAW

The Villanova University School of Law is a graduate professional school providing a three-year course of instruction in the law and its history and philosophy. The School of Law is fully accredited by the American Bar Association and is a member of the Association of American Law Schools.

A J.D. degree is granted on completion of a three-year full-time course of study. An LL.M. degree in Taxation is granted to those who complete the Graduate Tax Program. An LL.M. degree in International Study can be earned through joint programs offered with universities in London, Edinburgh, Leiden, the Netherlands, Singapore, and Hong Kong. The School of Law offers a joint J.D./M.P.A. program with Villanova University College of Liberal Arts and Sciences. In addition, the School of Law offers a joint J.D./M.B.A. program in collaboration with Villanova’s top-ranked School of Business.

The School of Law is located on the Villanova campus on Philadelphia’s suburban Main Line, fifteen miles west of Philadelphia, Pennsylvania. The building—brand new in 2009—is devoted exclusively to the study of law. The innovative building offers state-of-the-art classrooms and facilities focused on student needs. It has received LEED Gold certification by the U.S. Green Building Council and includes five state-of-the-art courtrooms, including the Martin G. McGuinn Ceremonial Courtroom. The building also features computer labs, dining facilities, student lounges, and a 500,000 volume law library.

LAW REVIEW

The Villanova Law Review is a national legal periodical published five times per year by a board of student editors. The Law Review offers an opportunity for valuable experience in legal writing and research. In addition, the Law Review has created Tolle Lege, a dedicated online journal which showcases articles by members of the bench, bar, and academia on timely legal issues. Students are invited to join the Law Review on the basis of scholarship or performance in an open-writing competition.

ENVIRONMENTAL LAW JOURNAL

The Villanova Environmental Law Journal is a national legal publication dedicated to the scholarly review of environmentally related topics. The Environmental Law Journal is published two times per year in the winter and the summer by a staff of student editors and writers who have been invited to join the Environmental Law Journal based on their performance in an open-writing competition.
JEFFREY S. MOORAD SPORTS LAW JOURNAL
The Jeffrey S. Moorad Sports Law Journal is a national legal periodical published and edited two times per year by a staff of students chosen for membership based on their performance in an open-writing competition. The Sports Law Journal is committed to publishing scholarly articles on topics of importance in the sports law field.

JOURNAL OF CATHOLIC SOCIAL THOUGHT
The Journal of Catholic Social Thought is a national legal periodical published two times per year. It is a peer-reviewed interdisciplinary publication that brings together authors from the fields of law, theology, philosophy, and political theory. It is jointly published by Villanova University School of Law and the Office for Mission and Ministry at Villanova University.

EXPERIENTIAL LEARNING
Villanova School of Law provides students with a wealth of opportunities for practical, hands-on learning that goes beyond traditional legal theory and principles taught in the classroom. Students are encouraged to take part in a number of different programs: represent clients from the local community in one of our six legal Clinics; hone their litigation skills by participating in Moot Court, taking the Intensive or Advanced Trial Advocacy courses, or competing on the National Trial Team; put their legal writing and research skills to the test by joining the staff of one of our three published and two online journals; explore their interest in Sports Law by becoming involved with the Jeffrey S. Moorad Center for the Study of Sports Law; learn where the law meets business in Villanova’s new Center for Law, Entrepreneurship, and Innovation; go beyond the law school’s walls and complete an externship in a variety of legal settings from clerkships to small or large law firm employment; or explore the legal system of another nation through our Rome Summer Program or one of our semester abroad programs.

MOOT COURT
The Moot Court Board is the principal oral advocacy organization at Villanova University School of Law. The Board is chiefly responsible for the administration of the Theodore Reimel Moot Court Competition, an intra-school tournament in honor of the late Theodore L. Reimel, Judge for the Pennsylvania Court of Common Pleas from 1953 to 1973. The Board also competes in appellate advocacy competitions at law schools across the country. Finally, the board assists the legal writing faculty in the administration and judging of 1L oral arguments. Students are invited to join the board through a summer competition or by advancing to the Semi-Final round of the Reimel Moot Court Competition.

THE STUDENT BAR ASSOCIATION
The Student Bar Association of the School of Law is the official resource and liaison for the student body in matters involving the school administration, faculty, and students. Upon enrollment at the Villanova University School of Law, each student becomes a member of the organization. The SBA Board consists of nineteen positions that are elected by the student body and one appointed position. The Association conducts student government and extra-curricular activities at the School of Law and is affiliated with the Law Student Division of the American Bar Association.

THE ORDER OF THE COIF
The Order of the Coif, a national honor society with chapters in most leading law schools, selects members from among the School of Law’s graduating students who are in the top ten percent of the class and who have contributed substantially to the Law School.

ADMISSION
A bachelor’s degree, or equivalent, is required for admission. Applicants are required to take the Law School Admission Test. For further information, address the Admissions Office, Villanova University School of Law, 299 North Spring Mill Road, Villanova, Pennsylvania 19085-1682 or call 610-519-7010.
VILLANOVA UNIVERSITY
SCHOOL OF LAW

2014–2015
OFFICERS OF THE UNIVERSITY
THE REVEREND PETER M. DONOHUE, O.S.A., B.A., M.A., Ph.D., President
THE REVEREND KAIL C. ELLIS, O.S.A., Ph.D., Vice President for Academic Affairs

LAW SCHOOL ADMINISTRATION
JOHN Y. GOTANDA, B.B.A., J.D., Dean and Professor of Law
MICHAEL P. MORELAND, B.A., M.A., J.D., Ph.D., Vice Dean and Professor of Law
CHRISTINE G. MOONEY, B.A., B.S., J.D., Associate Dean for Academic Affairs, Director of Legal Writing Programs, Professor of Legal Writing, and Director of the J.D./M.B.A. and J.D./M.P.A. Programs
MICHELLE DEMPSEY, B.A., J.D., LL.M., D.Phil., Associate Dean for Faculty Research and Professor of Law
VICENC FELIU, B.A., M.S., J.D., LL.M., Associate Dean for Library Services and Associate Professor of Law
JAMES C. GARVEY, B.A., M.B.A., Ed.D., Associate Dean for Development and Alumni Relations
JENNIFER A. HENFEY, B.S., J.D., Interim Associate Dean for Finance and Administration, Executive Director of Career Strategy and Advancement

LAW FACULTY
TODD AAGAARD, B.A., J.D., M.S., Professor of Law
LESLE M. BOOK, B.A., J.D., LL.M., Professor of Law
RICHARD BOOKS, A.B., J.D., Professor of Law and Martin G. McGuinn Chair in Business Law
PATRICK MCKINLEY BRENNAN, B.A., M.A., J.D., Professor of Law and John F. Scarpa Chair in Catholic Legal Studies
DORIS DEL TOSTO BROGAN, B.A., J.D., Professor of Law
JOHN J. CANNON, A.B., J.D., Professor of Law
DAVID S. CAUDILL, B.A., J.D., Ph.D., Professor of Law and Goldberg Family Chair in Law
STEVEN L. CHANENSON, B.A., M.S., J.D., Professor of Law
JOSPEH W. DELAPENNA, B.B.A., J.D., LL.M., Professor of Law
MICHELLE DEMPSEY, B.A., J.D., LL.M., D.Phil., Associate Dean and Professor of Law
JOHN F. DOBBIN, A.B., J.D., LL.M., Professor of Law
VICENC FELIU, B.A., M.S., J.D., LL.M., Associate Dean for Library Services and Associate Professor of Law
T. KEITH FOGG, B.A., J.D., LL.M., Professor of Law
ABRAHAM J. GAFNI, B.A., B.H.L., J.D., Professor of Law
Ruth Gordon, B.A., J.D., LL.M., Professor of Law
John Y. Gotanda, B.B.A., J.D., Dean and Professor of Law
Ann Carey Juliano, B.A., J.D., Professor of Law
Catherine J. Lancot, B.A., J.D., Professor of Law
Beth Lyon, B.A., M.S., J.D., Professor of Law
James Edward Maule, B.S., J.D., LL.M., Professor of Law
Melanie McMenamin, B.S., B.A., J.D., Assistant Professor of Law and Director of the Clinic for Law and Entrepreneurship
Michael P. Moreland, B.A., M.A., J.D., Ph.D., Vice Dean and Professor of Law
Joy Sabino Mullane, B.A., J.D., LL.M., Professor of Law
Jennifer O’Hare, B.S., J.D., Professor of Law
Michele R. Pistone, B.S., J.D., LL.M., Professor of Law
Anne Bowen Poulin, B.A., J.D., LL.M., Professor of Law
Teresa E. Rayneill, B.A., J.D., Professor of Law
Michael Risch, A.B., J.D., Professor of Law
Chaim N. Saman, B.B.S., J.D., Professor of Law
Tuan Samahon, B.A., J.D., Professor of Law
Dviera Segal, B.A., M.A., J.D., Director of Clinical Programs and Professor of Law
Louis J. Sirico, Jr., B.A., J.D., Director of Legal Writing and Professor of Law
Walter John Taggart, A.B., J.D., Professor of Law
Ellen Wertheimer, B.A., J.D., Professor of Law

Visiting Faculty
Andrew Brandt, B.A., J.D., Executive Director of the Jeffrey S. Moorad Center for the Study of Sports Law and Visiting Professor
Michael Campbell, A.B., J.D., Visiting Assistant Professor
Jennifer H. Sperling, B.A., J.D., Visiting Assistant Professor

Distinguished Professor of Practice
J. Richard Harvey, Jr., B.S.B.A., MSc.

Emeritus Faculty
Gerald Abraham, A.B., J.D., Professor Emeritus of Law
Lewis Becker, B.S., LL.B., Professor Emeritus of Law
Donald N. Bersoff, B.S., M.A., Ph.D., J.D., Professor Emeritus of Law
Donald W. Dowd, A.B., J.D., Professor Emeritus of Law
John M. Hinson, B.A., M.A., LL.B., Professor Emeritus of Law
Don W. Llewellyn, B.A., J.D., LL.M., Professor Emeritus of Law
Howard R. Lurie, A.B., J.D., Professor Emeritus of Law
Michael Mulroney, B.S.C., J.D., Professor Emeritus of Law
John F. Murphy, B.A., LL.B., Professor Emeritus of Law
J. Willard O’Brien, B.S., J.D., Dean and Professor Emeritus of Law
Leonard Packel, B.S., J.D., Professor Emeritus of Law
Frederick P. Rothman, A.B., J.D., Professor Emeritus of Law

Legal Writing Faculty
Heather D. Baum, B.A., J.D., Professor of Legal Writing
Candace Mueller Centeno, B.A., J.D., Professor of Legal Writing
Diane Edelman, A.B., J.D., Director of International Programs and Professor of Legal Writing
Christine Mooney, B.A., B.S., J.D., Associate Dean, Director of Legal Writing Programs, Professor of Legal Writing, and Director of the J.D./M.B.A. and J.D./M.P.A. Programs
Mitch Nathanson, B.A., J.D., Professor of Legal Writing
Mary Ann Robinson, B.A., J.D., Associate Professor of Legal Writing
Jessica Webb, B.A., J.D., Assistant Professor of Legal Writing

Writing Specialist
Steven Schultz, B.A.

Legal Research Instructors
Benjamin Carlson, B.A., J.D., M.L.S., Emerging Technologies Librarian and Legal Research Instructor
Lori Strecker Corso, B.A., B.S., J.D., M.L.I.S., Electronic Services Librarian and Legal Research Instructor
Robert Hegadorn, B.A., J.D., M.L.S., Reference Librarian and Legal Research Instructor
MATTHEW McGOVERN, B.S., J.D., M.S.L.I.S., Reference Librarian and Legal Research Instructor
AMY SPARE, B.A., J.D., M.L.I.S., Faculty Services Librarian and Legal Research Instructor

ACADEMIC DIRECTORS

APRIL M. BARTON, B.S., J.D., Director of Academic Compliance and Distance Learning
MATTHEW CARLUZZO, A.B., J.D., Director of Academic Success and Bar Support Programs
DIANE EDELMAN, A.B., J.D., Director of International Programs and Professor of Legal Writing
RUFUS KING, A.B., B.Phil., J.D., Director of the Center for Law, Entrepreneurship and Innovation
CHRISTINE MOONEY, B.A., B.S., J.D., Associate Dean, Director of Legal Writing Programs and Professor of Legal Writing
PETER THOMPSON, B.A., M.Ed., J.D., LL.M., Director of Experiential Learning
LINDA LOVE VINES, B.S., J.D., LL.M., Instructor and Interim Director of the Graduate Tax Program

ADMINISTRATIVE DIRECTORS

ANTHONY CANCELLI, Director of Business Support Services
COLLEEN BELZ FRASCATORE, Executive Director of Development and Campaign
NICOLE GARAFANO, Director of Events
JENNIFER A. HENFEE, Associate Dean and Executive Director of Career Strategy and Advancement
JOSHUA LESS, Director of Annual Giving and Donor Relations
JOSEPH MARIANI, Director of Facilities
BRIAN McCABE, Campus Minister for Outreach and Student Formation
SUSAN K. MESICK, Registrar
SARAH N. PONZIO, Director of Alumni Relations
NANCY WHALEN, Director of Student Affairs
KIRSTINA WILHELM-NELSON, Associate Director of Financial Aid

J.D. ADJUNCT FACULTY

THE HONORABLE PEGGY ABLEMAN, B.A., J.D.
STEPHEN ANDERER, B.A., J.D., Ph.D.
AARON AVILA, B.A.S., J.D.
DANIEL BACINE, B.S., J.D.
DAVID BAILEY, B.S., J.D.
JONI BERNER, B.A., J.D.
RON BROWN, B.A., J.D.
FRED BUCK, A.B., J.D.
APRIL BYRD, B.A., J.D.
JOHN CARLE, B.A., J.D.
BRIAN CARROLL, B.A., J.D.
PETER CLAUSS, A.B., LL.B., A. ROY DECARO, B.A., J.D.
DONALD DECARLO, Jr., B.A., J.D., LL.M., M.A.
JAMES DINNAGE, M.A., Lic. spéc. en Dr. Eur.
AMY DONNELLA, B.A., J.D.
JOHN T. DONOVAN, B.A., J.D.
MICHAEL DROSSNER, B.A., J.D., LL.M.
ISABEL LOPES DUFFY, B.A., J.D.
ROBERT DUNHAM, B.A., J.D.
CHARLES FORER, B.A., J.D.
BRYAN FRANZY, B.A., ISAT, J.D.
KEVIN FLANDREY, B.S., J.D.
DOUGLAS GANTON, B.A., J.D.
RICHARD J. GERING, M.A., Ph.D.
THE HONORABLE GARY S. GLAZER, B.A., J.D.
RICHARD GOEBBEL, B.A., LL.B.
JOHN C. GOODCHILD III, B.A., J.D.
JASON GOSSELIN, B.A., J.D.
CHRISTOPHER M. HARVEY, B.S., J.D., LL.M.
THE HONORABLE RENEE CALDWELL HUGHES, B.A., J.D.
JAMES A. HUNTER, Jr., A.B., J.D., LL.B.
ROSALIND T. KAPLAN, B.A., M.A., J.D.
AMY S. KLINE, B.A., J.D.

STEVEN KOFFS, B.S.M.E., M.S., J.D.
DALE LARRIMORE, B.A., J.D.
JORDAN A. LAVINE, B.A., J.D.
SHERY LEMONICK, B.S., J.D.
STUART LEV, B.A., J.D.
THE HONORABLE KATHRYN STREETER LEWIS, B.A., J.D.
ROBERT J. LICHTENSTEIN, B.S., J.D., LL.M.
BRYANT LIM, B.A., J.D.
JAMES McCARTNEY, B.A., M.A., M.S., Ph.D.
PATRICIA McCaULLAND, B.A., M.A., J.D.
RICHARD L. McMONIGLE, B.A., J.D.
COLEEN MEEHAN, B.S., J.D.
MARANNA MEEHAN, B.A., J.D.
JOHN MARSHALL, B.S., M.S., J.D.
THE REVEREND GERARD C. MUSE, B.A., M.S., J.D., J.D.
MARK MONJIAN, B.A., J.D.
CHERYL NICHOLSON, B.A., J.D.
WENDY NORMAN, B.S., J.D.
MARY DEBUCOS OCHIAI, B.A., J.D.
THOMAS O’KEEFE, B.A., J.D., M/Phil.
MICHAEL O’MARA, B.S., J.D.
MARK PRESSMAN, B.S., M.S., Ph.D.
THE HONORABLE SHELDON K. RENNIE, B.A., M.S., J.D.
THE HONORABLE EDUARDO ROBRENO, B.A., J.D.
LEE ROSENGARD, A.B., J.D.
DANIEL F. RYAN, III, B.A., J.D.
THE HONORABLE TODD SEELEG, B.S., J.D.
ROBERT SIEGER, B.A., J.D.
JUDY GOLSTEIN SMITH, B.A., J.D.
SHAUN SNITMAN, LL.B.                        SHEILA VANCE, B.A., J.D.
STEVEN D. SPENCER, B.S., J.D.              BRIAN VARGO, B.A., J.D.
GILBERT STEIN, B.A., J.D.                   BARBARA WALL, B.A., M.A., Ph.D.
TOMKA STEVENS, B.A., M.A., J.D.            MARGUERITE WEISE, B.A., J.D., LL.M.
STEVEN SUGARMAN, B.A., J.D.                R. SETH WILLIAMS, B.A., J.D.
JOSEPH TATE, A.B., LL.B.                    GEORGE WOLTERS, B.A., J.D.
JOSEPH A. TORREGROSSA, B.A., J.D.          SARA WOODS, B.A., J.D., M.P.A.
MICHAEL TRACTMAN, B.A., J.D.               ANDREW YOUNGMAN, B.A., J.D.
COLEEN VALLEN, B.S.                         MARC J. ZUCKER, B.A., J.D.

GRADUATE TAX PROGRAM
FULL-TIME FACULTY

SCHOOL OF LAW
LESLIE M. BOOK, B.A., J.D., LL.M., Professor of Law
T. KEITH FOGG, B.A., J.D., LL.M., Professor of Law
JOY SABINO MULLANE, B.A., J.D., LL.M., Associate Professor of Law
LINDA LOVE VINES, B.S., J.D., LL.M., Instructor and Interim Director of Graduate Tax Program

SCHOOL OF BUSINESS
SHELLY C. RHoades-Catanach, B.S., Ph.D., Associate Professor

EMERITUS FACULTY
MICHAEL MULRONEY, B.S.C., J.D., Professor Emeritus of Law

DISTINGUISHED PROFESSOR OF PRACTICE
J. RICHARD HARVEY, JR., BSBA, MSc.

GRADUATE TAX ADJUNCT FACULTY
RYAN M. Bornstein, B.S., J.D., LL.M.
TIMOTHY C. BOYLE, BSBA, M.B.A.
JOHN BAYO, B.S., J.D., LL.M.
DAVID A. BRENN, B.S., M.B.A., J.D.
RONALD S. COLLINS, B.A., J.D., LL.M.
THOMAS J. DONNELLY, B.S., J.D., LL.M.
JAMES C. FEE, JR., B.S., J.D., LL.M.
NEIL H. FEINSTEIN, B.S., J.D., LL.M.
JILL R. FOWLER, B.A., J.D.
JARED H. GORDON, B.A., J.D., LL.M.
MICHAEL P. HANEY, B.S., J.D., LL.M.
PETER J. HARDY, B.A., J.D.
CHRISTOPHER M. HARVEY, B.S., J.D., LL.M.
JOEL E. HOROWITZ, B.A., J.D., LL.M.
RICHARD S. HYMAN, B.S., M.T.
THOMAS J. IAnnozzi, B.A., J.D., LL.M.
BRIAN S. JONES, B.A., M.Ed., J.D., LL.M.
JASON KORATNICK, B.S., J.D., LL.M.
JONATHAN LANDER, B.A., J.D., LL.M.
DEBORAH M. LERNER, A.B., J.D., LL.M.
EDWARD A. LIVA, B.S., B.A., J.D., LL.M.
SCOTT MASATTS, B.A., J.D., LL.M.
MICHAEL J. McGoldrick, B.A., J.D., M.L.D.
LL.M. (Taxation), LL.M. (Corporate)
ERICA R. McREYNOLDS, B.A., J.D., LL.M.
DREW A. MORRIS, B.A., J.D., LL.M., M.B.A.
DANIEL C. MUNCE, B.A., J.D., LL.M.
THOMAS OSTRANDER, B.A., J.D., LL.M.
PHILLIP PILLAR, B.A., J.D., LL.M.
THOMAS M. RATH, B.B.A., J.D., LL.M.
NANCY B. ROMANO, B.A., J.D., LL.M.
JOSEPH E. RONAN, JR., B.A., J.D., LL.M.
DAVID SHAPIRO, B.A., J.D.
MICHAEL E. STEINSLATZ, B.S., J.D., LL.M.
CHRISTOPHER P. WOHRLE, B.A., J.D., LL.M.
DAVID B. ZELIKOFF, B.A., J.D.

LIBRARY
VICENC FELIU, B.A., M.S., J.D., LL.M., Associate Dean for Library Services and Associate Professor of Law
BENJAMIN CARLSON, B.A., J.D., M.L.S., Emerging Technologies Librarian and Legal Research Instructor
MARY CORNELIUS, B.A., M.L.S., Circulation Coordinator and Reference Librarian
LORI STRICKLER CORSO, B.A., B.S., J.D., M.L.I.S., Electronic Services Librarian and Legal Research Instructor
STEVEN ELKINS, B.A., M.A., M.L.S., Associate Director for Collection Management
KAREN GAUSE, B.A., M.A., M.L.S., Interlibrary Loan Librarian and Senior Reference Librarian
MARY JO HEACOCK, B.S., M.L.S., Catalog Librarian
ROBERT HEGADORN, B.A., J.D., M.L.S., Reference Librarian and Legal Research Instructor
MATTHEW McGOVERN, B.S., J.D., M.L.I.S., Reference Librarian and Legal Research Instructor
AMY SPARE, B.A., J.D., M.L.I.S., Associate Director for Law Library Research and Instructional Services
VILLANOVA LAW REVIEW

VOLUME LX

Editor-in-Chief
THOMAS P. REILLY

Executive Editor
MARK T. WILHELM

Managing Editors

Production
MICHAEL MONTALBANO

Research & Writing Operations
KELSEY HUGHES-BLAUM

Tolle Lege
AMANDA D. CRAWFORD

Manuscript Editors

Student Works Editors
NICHOLAS CARROLL
LISABEL CHEONG
NICOLE PEDI
MATTHEW SETZER

Articles Editors
STEPHANIE HAAS
NICHOLAS KARBWACKI
KYLE A. PINDER
ANTONIOS ROUSTOPOULOS
MATTHEW SKAROFF
BRETT A. ZAKEOSIAN

Online Editors
GABRIELLA PAGLIERI
DEAN J. SHAUGER

Associate Editors
LOGAN ANDERSON
EMILY BULLIS
JOCELYN COOPER
COURTNEY GROCHOWSKI
JASON KETELSEN

Online Editors
CAITLIN LAMBERT
CHRISTINE MARCOZZI
LAUREN MARVEL
KEVIN McGRATH
KIMBERLY MENAFRA
LISA ROTH
MEGAN ROUSH
ALEXANDRA SKARDA
ALEXANDRA SULLIVAN
ROSE MARIE WONG

Staff Writers
KRISTEN ASHE
ELIZABETH GIORDANO
LOGAN C. MILLER
SARA MOYER
ELIZABETH MULLIN
ALEXANDRIA MURPHY
MICHAEL NACHAJSKI
NICHOLAS C. NEEDLE
MELISSA RUTH
BRENNAN SCHREIBMAN
MICHAEL STEGNER
DONALD VAN BUREN
KATHLEEN VERMILION
JESSICA WATKINS

Faculty Advisor
MICHAEL V. RISCH

THE HARRY S. CHERKEN, JR. LAW REVIEW FUND
PROVIDES FINANCIAL SUPPORT FOR THE LAW REVIEW
(x)
The Editorial Board of Volume 60 of the Villanova Law Review would like to recognize and thank Professor Walter John Taggart for his dedicated service to the Law Review. After over four decades of counseling the student editors of forty-three volumes, Professor Taggart has decided to retire from his position as Faculty Advisor for the Law Review. Professor Taggart, thank you for your longtime support, advocacy, and mentorship of the Law Review and its members.

Editorial Board, Volume 60
Articles

THE “VOLUNTARY” INPATIENT TREATMENT OF ADULTS UNDER GUARDIANSHIP

RICHARD C. BOLDT*

I. INTRODUCTION

SOME people who have severe mental illnesses or other severe mental disabilities derive benefit from the care and treatment provided in psychiatric hospitals and other inpatient settings.1 Frequently, the very disabilities that call for inpatient treatment also disrupt an individual patient’s capacity to participate fully in the decision-making process by which hospital admission is elected.2 All jurisdictions within the United States maintain statutory schemes that allocate decision-making authority with respect to psychiatric hospital admissions, often by creating separate procedures and standards for “voluntary” admissions and involuntary commitments.3 A number of states have adopted a preference for voluntary hospitalization over involuntary civil commitment, in part based on the theory that voluntary psychiatric patients are more likely to form a therapeutic alliance with clinicians and are therefore better able to benefit from treat-

* Professor of Law, University of Maryland Francis King Carey School of Law. For help in planning and developing this Article, I thank Eileen Canfield, Cameron Connah, Susan McCarty, and Aaron Parker. This project was supported by a research grant from the University of Maryland Francis King Carey School of Law.

1. See THOMAS G. GUTHEIL & PAUL S. APPELBAUM, CLINICAL HANDBOOK OF PSYCHIATRY AND THE LAW 73 (1982) (stating that inpatient care “provide[s] structure, a supportive milieu, protection, intensive care, closely supervised pharmacotherapy, electroconvulsive therapy, or other forms of treatment”). As used throughout this Article, the phrase “mental disability” includes mental illnesses; significant developmental disabilities (particularly intellectual disabilities); and the impairments associated with geriatric dementia, Alzheimer’s disease, and the like.


ment. Not all states, however, provide clear statutory or judicial guidelines on the degree of capacity required for a patient to consent to admission.

Determining whether (and how) patients with impaired decision-making capacity may consent to inpatient treatment is a complex question that often turns on an interpretation of state laws read against the backdrop of constitutional doctrine governing the due process rights of individual patients. In addition, in some cases, impaired patients have court-appointed guardians who are authorized to make financial decisions, treatment decisions, or both, for their wards. In instances in which guardians have received either general authorization to make decisions for their wards or have been granted specific authority to make healthcare decisions, difficult legal questions may arise as to the power of the guardian to consent to the ward’s admission for inpatient psychiatric treatment. In a number of states, the law is clear that a guardian may not consent to the ward’s admission to a psychiatric hospital, thus requiring the use of the state’s involuntary civil commitment process in order to obtain inpatient treatment. In other states, it is clear that a guardian may provide the necessary consent for voluntary admission. In a third group of states, the guardian may consent so long as the ward also consents or, in some jurisdictions, does not object. In a fourth group of states, the


5. See Gutheil & Appelbaum, supra note 1, at 50.


8. See Brakel et al., supra note 3, at 377; Gutheil & Appelbaum, supra note 1, at 36.

9. See, e.g., Alaska Stat. § 13.26.150(e) (2012) (prohibiting guardians from placing wards in psychiatric facilities: "A guardian may not [ ] place the ward in a facility or institution for the mentally ill other than through a formal commitment proceeding under [Alaska Stat. §] 47.30 in which the ward has a separate guardian ad litem").

10. See, e.g., GA. CODE ANN. § 37-3-20 (2012) (authorizing guardians to make applications for wards’ voluntary commitment).

11. See, e.g., Mich. COMP. LAWS ANN. § 330.1415 (West 2014) ("[A]n individual 18 years of age or over may be hospitalized as a formal voluntary patient if the
guardian’s authority to arrange for voluntary inpatient care depends on the guardian obtaining specific court authorization, although the substantive standards and procedural requirements for securing such an order may differ from those that govern involuntary commitment. 12 Finally, a range of additional variations exists in other states, often created by the interplay between the laws regulating mental hospital admission and other provisions governing the powers and responsibilities of guardians. 13

Several rationales have been provided for the statutory or case law rules operating in a number of states to prohibit guardians from consenting to a ward’s psychiatric hospitalization, or withholding that authority when the ward objects. One widely cited state court has reasoned:

If we were not to require at least substantial compliance with the [involuntary commitment] law to fully protect the rights of incompetents it would be possible for an unscrupulous person to have himself appointed as guardian and then lock his ward in a mental institution and proceed to waste the ward’s estate. 14

It is worth comparing this skepticism about the good faith of an appointed guardian of an adult with the presumed good faith of parents who seek the “voluntary” hospitalization of a child against the child’s wishes, which was central to the United States Supreme Court’s decision in Parham v. J.R. 15 In Parham, Chief Justice Burger, writing for the Court’s majority, made plain that an independent medical assessment by hospital individual executes an application for hospitalization as a formal voluntary patient or the individual assents and the full guardian of the individual, the limited guardian with authority to admit, or a patient advocate authorized by the individual to make mental health treatment decisions under the estates and protected individuals code, 1998 [Mich. Pub. Acts] 386, [Mich. Comp. Laws Ann. §§] 700.1101 to 700.8102, executes an application for hospitalization and if the hospital director considers the individual to be clinically suitable for that form of hospitalization.”). 12

12. See, e.g., ARIZ. REV. STAT. ANN. § 14-5312.01 (2012) (providing that guardians must obtain court order and comply with other notice requirements and procedural provisions before placing wards in inpatient mental health settings).

13. See, e.g., N.M. STAT. ANN. § 45-5-312(B)(3) (LexisNexis 2013) (“[A] guardian may consent or withhold consent that may be necessary to enable the incapacitated person to receive or refuse medical or other professional care, counsel, treatment or service.”). However, guardians may only present their wards for evaluation for inpatient mental health treatment:

A guardian appointed under the Uniform Probate Code, an agent or surrogate under the Uniform Health-Care Decisions Act or an agent under the Mental Health Care Treatment Decisions Act shall not consent to the admission of an individual to a mental health care facility. If a guardian has full power or limited power that includes medical or mental health treatment or, if the individual’s written advance health-care directive or advance directive for mental health treatment expressly permits treatment in a mental health care facility, the guardian, agent or surrogate may present the person to a facility only for evaluation for admission . . . .

Id. § 43-1-14.


officials is a sufficient check against possible abuse by parents whose motives for seeking inpatient care may be at odds with their child’s best interests.\textsuperscript{16} Some writers have suggested that a similar process of medical review should suffice to ameliorate the concerns about abuse by the guardians of adult wards that animate the restrictive rules operating in a number of states.\textsuperscript{17} Other writers have suggested that state law limitations on the authority of guardians to approve voluntary hospitalization for incompetent adults stem from a concern that permitting such authority would provide a built-in end-run around these jurisdictions’ involuntary commitment criteria and procedures, including their rigorous standard of proof.\textsuperscript{18}

Since the 1970s, the laws governing involuntary civil commitment in virtually every state have required a finding that an individual subject to commitment must not only be mentally disabled but also dangerous to himself or herself or others.\textsuperscript{19} Voluntary hospitalization, by contrast, generally is based on the patient’s need for and amenability to treatment.\textsuperscript{20} Significantly, the development of more demanding civil commitment standards in the 1970s represented a move away from the mid-century approach of permitting involuntary hospitalizations on a medical-needs basis, in which the decision maker often was a board of physicians rather than a judge or jury.\textsuperscript{21} If a guardian is not permitted to consent to voluntary admission, however, the more restrictive involuntary commitment standard—with its dangerousness criterion in particular—may make it difficult or impossible to arrange for inpatient care for some patients with severe mental disabilities who do not present a risk of harm to themselves or others, but who would benefit from such treatment.\textsuperscript{22}

There are other important consequences as well that may flow from the requirement that a patient under guardianship be civilly committed rather than voluntarily hospitalized, including possible differences in the rules governing the provision of psychotropic medications, discharge, discharge, discharge.

\begin{itemize}
\item 16. See id. at 607. While Chief Justice Burger’s rationale—that the law generally presumes the good faith of parents because of their natural love and affection for their children—would appear to be a sufficient basis to distinguish the court-appointed guardians of adult wards, the \textit{Parham} decision also included minor wards hospitalized by a state department of social services. \textit{See id.} at 587–88.
\item 17. See Halverson, \textit{supra} note 4, at 167.
\item 18. See Slobogin, Rai \& Reisner, \textit{supra} note 3, at 864; Stone, \textit{supra} note 4, at 42; \textit{see also In re} Gardner, 459 N.E.2d 17, 19 (Ill. App. Ct. 1984) (holding that placement by guardian under voluntary admission statute without ward’s consent is precluded because “grant of such power would contravene the involuntary commitment provisions” of state law).
\item 19. See Gutheil \& Appelbaum, \textit{supra} note 1, at 40; Slobogin, Rai \& Reisner, \textit{supra} note 5, at 705.
\item 20. See Stone, \textit{supra} note 4, at 30–32.
\item 21. See Gutheil \& Appelbaum, \textit{supra} note 1, at 40; Slobogin, Rai \& Reisner, \textit{supra} note 5, at 703.
\item 22. See Stone, \textit{supra} note 4, at 37.
\end{itemize}
length of stay, and the like. The question, then, is whether a set of legal standards and practices either exists or can be developed that would provide for the treatment needs of those patients who have an impaired capacity to consent to voluntary hospitalization but who do not meet the criteria for involuntary civil commitment. Such an approach should be attentive to (and seek to minimize) the potential for coercion or abuse these patients may face and should be designed to maximize the self-determination of which they are capable. Some patients whose disabilities are chronic but episodic may be able to provide adequate consent through the use of advance directives, although the use of such instruments—especially by persons with mental disabilities—is relatively rare. Others may be able to participate in the decision-making process if the criteria by which their competence is assessed are defined to require only minimal functionality. But patients who are assisted by third-party decision makers—particularly those with court-appointed guardians—should be able to gain access to inpatient psychiatric treatment without running the gauntlet of involuntary civil commitment, if the substantive standards and procedural requirements put in place by state law can be made adequate to insure that the third-party decision makers are acting with respect for the values held by these patients and, to the extent possible, are seeking to serve their best interests.

Part II of this Article recounts the long history of the development of involuntary civil commitment law and practice in the United States and the more recent development of voluntary inpatient admissions. It also explores the role that informed consent plays in voluntary hospitalizations and the due process considerations regarding consent introduced into

---

23. See, e.g., MD. CODE ANN., HEALTH—GEN. § 10-708(b)(2) (LexisNexis 2013) (providing right to refuse medication); id. § 10-803 (providing for release of voluntary patient).


this area by the United States Supreme Court’s decision in Zinermon v. Burch. Part III takes up the topic of guardianships and the related question of how incompetency—the threshold requirement for the appointment of a guardian—is defined. Part IV reviews the variety of approaches adopted by states for regulating the authority of guardians to provide consent for inpatient treatment. Part IV then concludes with an analysis of the interests in tension and offers a framework for an effective statutory approach to the area.

II. VOLUNTARY INPATIENT ADMISSION, INVOLUNTARY CIVIL COMMITMENT, AND THE PROBLEM OF CONSENT

The practice of permitting the voluntary psychiatric hospital admission of persons with mental disabilities did not become widespread in the United States until the last quarter of the twentieth century. While the first state law governing voluntary admissions was enacted in Massachusetts in 1881, and while over half of the states had statutes regulating the practice by 1924, few patients with mental disabilities were admitted as inpatients on their own initiative. Even as late as the decades following World War II, only about 10% of psychiatric inpatients were voluntarily admitted. In the 1970s, however, the practice became much more common, accounting for a majority of mental hospital admissions in many jurisdictions.

Today, voluntary admission laws are in force in virtually every state, and voluntary patients make up more than 50% of the population in specialized mental hospitals and about 85% of the population in psychiatric units of general hospitals. The law in a number of states sets out a preference for voluntary hospitalization over involuntary civil commitment, and “almost all mental health professionals, and probably a majority of the mental health bar, favor the retention of voluntary admissions as the most frequently used means of ingress to a psychiatric hospital.”

27. See Gutheil & Appelbaum, supra note 1, at 48.
28. See Brakel et al., supra note 3, at 177–78.
29. See Gutheil & Appelbaum, supra note 1, at 48.
30. See id.
32. See Brakel et al., supra note 3, at 178; Stone, supra note 4, at 29 (describing statutory preference for voluntary admission in Minnesota, New York, Louisiana, and Florida).
33. Gutheil & Appelbaum, supra note 1, at 49. The modern trend has been for states to enact parallel statutes governing the voluntary admission of developmentally disabled persons and individuals with substance use disorders as well as patients with mental illness. In some states, “the ‘voluntary’ admission procedures for developmentally disabled persons are not so much identical as analogous to those for the mentally ill. These procedures are sometimes labeled ‘administrative,’” permitting the developmentally disabled person to apply for himself if he is of age and competent but providing for application by the parent or guardian.
This increase in the rate of voluntary admissions for persons with mental disabilities has taken place within the context of a public mental health system in the United States that has been focused, for most of the past two centuries, primarily on involuntary commitment. In order to frame further consideration of voluntary admissions for persons under guardianship, therefore, attention to the course of development of involuntary commitment law and practice in the United States is due.

A. The Development of Civil Commitment Law and Practice

The first civil commitment statutes were enacted in American states in the late eighteenth century. By the middle of the nineteenth century, the practice of civil commitment had become somewhat common, in part due to a growing perception—fostered by the founding of the American Psychiatric Association (APA) and the advocacy of Benjamin Rush and others—that effective “scientific” techniques for treating mentally ill persons were available.

The ensuing path of this history of involuntary treatment for persons with mental disabilities is anything but linear. A number of commentators have described the historical record as “cyclical” or marked by the swinging of a “pendulum” between the poles of paternalist concern for those with mental illness and other mental disabilities on the one side, and an opposing libertarian pole animated by fears of government (and physician) overreaching and centered on the values of individual autonomy and self-determination on the other. The libertarian impulse was first expressed in the period immediately following the Civil War, when patients’ rights advocates raised concerns that some idiosyncratic individuals were being unnecessarily targeted for involuntary commitment by others who found their religious beliefs or political opinions to be excessively “novel.” Particularly troubling stories also emerged about malevolent husbands who arranged to have their wives committed for reasons of simple convenience. During the 1860s and early 1870s, these crusaders persuaded a number of states to enact procedural protections in the civil commitment context that included a right to counsel and trial by jury.

when he is not, in which case the procedure is—by our definition—not voluntary.

Brakel et al., supra note 3, at 179.

34. See Slobogin, Rai & Reissner, supra note 3, at 701–09.
35. See Brakel et al., supra note 3, at 14.
36. See Slobogin, Rai & Reissner, supra note 3, at 702.
37. See id. at 701–04. See generally Albert Deutsch, The Mentally Ill in America: A History of Their Care and Treatment from Colonial Times (2d ed. 1949).
38. See Slobogin, Rai & Reissner, supra note 3, at 703.
39. See, e.g., Brakel et al., supra note 3, at 15 (describing involuntary commitment of Mrs. E.P.W. Packard in Illinois by petition of her husband and her subsequent campaign to reform mental health law).
40. See Slobogin, Rai & Reissner, supra note 3, at 703.
The pendulum swung back in the direction of paternalist interventionism during the Progressive Era of the late nineteenth and early twentieth centuries. Reformers, influenced by Dorothea Dix and others who had worked tirelessly to expand the resources available for the mentally disabled, showed a renewed interest in providing care and treatment for persons with mental illness, which in turn led to an increased "medicalization" of procedures for their civil commitment. These efforts gained momentum following World War II, as new treatment technologies—including modern psychotherapy, psychosurgery, electro-convulsive therapy, and early pharmacotherapies—became available. In the period between the 1940s and the early 1970s, consistent with this increasingly medicalized approach, hearings before juries or judicial officers were replaced in many jurisdictions by “lunacy commissions” or physicians’ boards, and dangerousness criteria disappeared in favor of “in need of treatment” standards for involuntary admission. More voluntary admission statutes also began to appear during this period.

Beginning in the late 1960s and early 1970s, the pendulum swung once more, this time in a libertarian direction. In addition, for the first time, a significant percentage of patients with mental disabilities were voluntarily admitted for inpatient treatment. Influenced by the civil rights movement and by a broad rethinking by both policymakers and the lay public of the mid-century commitment to a paternalistic “rehabilitative ideal,” the libertarian reforms of the 1970s produced widespread statutory revisions that refocused civil commitment criteria on dangerousness, as opposed to amenability to or need for treatment, and reintroduced legalized procedural protections such as a right to counsel, heightened burdens of proof, and the like.
The libertarian individual rights reforms of the 1970s combined with a dramatic increase in the use of psychotropic medications—which were first introduced in the 1950s—fueled a process of “deinstitutionalization” that reduced the census of public and private psychiatric hospitals throughout the final decades of the twentieth century. Thus, while more than 550,000 patients were in public mental hospitals on any given day in the mid-1950s, that number had dropped to roughly 130,000 patients by the mid-1980s. Contained within this declining population of inpatients receiving care and treatment for mental illness, however, was a steadily increasing percentage of voluntary patients. As the parens patriae rationale for hospital admission largely was excised from the laws and policies governing involuntary commitment, and as involuntary confinement essentially became limited to the dangerous mentally ill, it was only natural that clinicians, concerned family members, and others increasingly would come to rely on voluntary admissions as a means of providing access to inpatient services for patients with significant mental disabilities who did not present an imminent threat to themselves or to others. In addition, the underlying values of patient autonomy and self-determination that animated much of the effort toward civil commitment reform were consistent with a new emphasis on voluntary hospitalization that was said to enlist the patient in decision making about his or her own care, and that, at least superficially, permitted the patient to retain some control over the nature and duration of his or her hospitalization.

B. Voluntary Inpatient Admission: Coercion and Choice

In truth, while the restrictions imposed on involuntary commitment standards and the increased legalization of commitment procedures, on
the one hand, and the growth of voluntary admissions, on the other, were interconnected phenomena; the relationship between involuntary and voluntary admission was, and is, exceedingly complex. The simple notion that most voluntary patients act relatively free from coercion when they elect to enter the hospital, and the paired idea that involuntarily committed patients generally express a “knowing, overt resistance” to their admission, are both overstated.\(^{52}\) As Brakel points out, “[i]n the vast majority of involuntary and voluntary cases, it is the family or relatives who move toward, pressure for, or insist on commitment.”\(^{53}\) As a consequence, he reports, “[i]n many instances where it orders commitment, the state’s judicial machinery merely formalizes and sanctions a decision arrived at by the family and the family doctor.”\(^{54}\) At the same time, he explains, “[t]he phenomenon of a freely derived, fully conscious, voluntary decision to enter a mental facility (particularly a public facility) is as rare as knowing, overt resistance to involuntary commitment.”\(^{55}\) This is so, because in many cases, voluntary patients are “already in some form of official custody” when they “elect” hospitalization, and their choice in that respect is often made with “the threat of involuntary commitment as the principal means of persuasion.”\(^{56}\)

It is, perhaps, because of the coercive features present within many voluntary mental hospital admissions that state law generally imposes restrictions on the practice that are not ordinarily present in the case of voluntary hospital admissions for other health-care services. State laws typically contain detailed provisions setting out the process by which an individual may elect to enter an inpatient mental health treatment facility, the criteria under which that application for admission is to be assessed by hospital personnel, and the rights and restrictions (particularly on requesting discharge from the hospital) that apply once the voluntary pa-

\(^{52}\) See Brakel et al., supra note 3, at 32.

\(^{53}\) Id.

\(^{54}\) Id. Of course, this account assumes that the patient has an engaged family or support group and is being cared for by a family physician. Increasingly in recent years, however, many severely mentally ill patients come into (or re-enter) the public mental health system as homeless or nearly homeless individuals, often alienated from family and friends and without ongoing links to health care or other social service providers. In such cases, the patient is just as likely to come into the treatment system by way of the criminal justice system as through the efforts of family and friends. See Bernstein & Seltzer, supra note 48, at 143.

\(^{55}\) Brakel et al., supra note 3, at 32.

\(^{56}\) Janet A. Gilboy & John R. Schmidt, “Voluntary” Hospitalization of the Mentally Ill, 66 NW. U. L. REV. 429, 430 (1971); see also Stone, supra note 4, at 36 (“Individuals are taken from their home community and escorted through the door of the psychiatric facility accompanied by police, family members, or other interested individuals seeking inpatient psychiatric care and treatment for the patients. At that time, patients may be asked to avoid involuntary commitment and accept treatment on a voluntary basis. Hospital staff and other interested individuals may promise a quicker release date, a less adversarial posture, and general sentiments that this is best for all concerned.”).
tient is admitted. Even the designation of an admission as “voluntary” oversimplifies the variety of ways in which a patient can enter an inpatient psychiatric facility without going through the involuntary commitment process. Thus, in some states, a voluntary admission can result either from the applications of patients themselves, or from the application of a designated third-party decision maker, such as a court-appointed guardian or an individual authorized by a properly executed advanced directive. Moreover, in most states, the governing statutes distinguish between “pure” voluntary admissions, a status that permits the patient to leave the hospital whenever he or she wishes, and “conditional” voluntary admission (sometimes also referred to as “formal” admission), under which the patient is prohibited from leaving prior to the passage of a period of time (usually several days) after providing notice to hospital officials of an intention to leave.

C. The Zinermon Decision and the Problem of Informed Consent

However conceived, the notion of voluntarily seeking admission for inpatient treatment in a psychiatric facility fairly implies that the patient has exercised some sort of choice. In the case of other health-care services, including general medical inpatient treatment, patients must provide informed consent. The laws in some states governing voluntary admission into mental hospitals and other psychiatric inpatient settings reflect this general principle and require that informed consent be obtained before a patient can be admitted. In many other states, however, no explicit consent requirements are set out in the governing statutes. This patchwork of differing state law standards for consent also characterized the state of the law in 1990 and formed the background for the

57. See Brakel et al., supra note 3, at 177.

58. Brakel takes the position that admissions accomplished at the request of a third-party decision maker should not be labeled as “voluntary,” but that terminology does appear in other descriptions of the practice. See id.

59. See Gutheil & Appelbaum, supra note 1, at 49.

60. The requirement of informed consent for medical treatment is often traced to Justice (then Judge) Benjamin Cardozo’s opinion in Schloendorff v. Society of New York Hospital, 105 N.E. 92, 93 (N.Y. 1914), where the court stated: “Every human being of adult years and sound mind has a right to determine what shall be done with his own body; and a surgeon who performs an operation without his patient’s consent commits an assault, for which he is liable in damages.” Id. See generally A.D. Nieuw, Informed Consent, 12 MED. & L. 125 (1993).

61. See, e.g., 405 ILL. COMP. STAT. 5/3-400 (2014) (setting out consent requirement for voluntary admission in Illinois Mental Health and Developmental Disabilities Code); N.Y. MENTAL HYG. LAW § 9.17 (McKinney 2014) (governing requirement of informed consent); see also In re Gardner, 459 N.E.2d 17, 20 (Ill. App. Ct. 1984) (holding that individual who lacked capacity to provide informed consent for inpatient treatment could not be admitted by his guardian).

62. For a general discussion of the state statutes governing voluntary hospitalization, see infra notes 142–235 and accompanying text.
United States Supreme Court’s discussion of voluntary psychiatric admission in its decision in *Zinermon v. Burch.*

In *Zinermon,* the Court’s primary focus was on a technical question concerning whether a federal civil rights claim for damages under § 1983 could be brought against hospital officials who had treated the plaintiff as a voluntary patient, notwithstanding his disoriented psychotic condition upon admission. The underlying events had taken place in Florida, one of a minority of states whose law did (and still does) require that voluntary patients provide competent informed consent for admission. The narrow question addressed by the Justices was whether, given this statutory requirement, the patient Burch was entitled as a matter of federal due process to be screened for competency before being admitted as a voluntary patient. Given that he was “hallucinating” and believed he was “in heaven,” and thus, by all accounts, was not able to provide competent informed consent, Burch’s argument was that his admission should have been sought through Florida’s involuntary civil commitment process. That process would have provided him with an adversarial hearing, a right to counsel, and all of the other procedural safeguards associated with the legalized approach to commitment that had become the norm throughout the United States by 1990. Under then-existing § 1983 law, this narrow question turned on whether Florida’s requirement of informed consent had been foreseeably and routinely violated in practice, and whether a post-deprivation remedy could be made available or whether a pre-deprivation determination of competency was necessary to protect the underlying state-law-based liberty interest.

Writing for a five-person majority, Justice Blackmun held that “[i]t is hardly unforeseeable that a person requesting treatment for mental illness might be incapable of informed consent, and that state officials with the power to admit patients might take their apparent willingness to be admitted at face value and not initiate involuntary placement procedures.”


64. At the time *Zinermon* was decided, the Florida statute required that a patient voluntarily give “‘express and informed consent to evaluation or treatment,’” or, in the alternative, that “a proceeding for court-ordered evaluation or involuntary placement [be] initiated.” See id. at 122 (quoting Fla. Stat. § 394.463(1)(d) (1981)). Currently, the statute requires: “The patient . . . shall be asked to give express and informed consent to placement as a voluntary patient, and, if such consent is given, the patient shall be admitted as a voluntary patient; or [ ] A petition for involuntary placement shall be filed . . . .” Fla. Stat. § 394.463(1)(3)–(4) (2014).

65. See *Zinermon,* 494 U.S. at 118–19.

66. See *id.* at 124.


68. *Zinermon,* 494 U.S. at 136.
psychiatric inpatients. The majority did not detail what a constitutionally adequate preadmission procedure might look like, although the tenor of Blackmun’s discussion suggests that a judicial hearing likely would not be necessary and that a standardized assessment of competence by clinicians at the hospital might suffice.

The broader question lurking in the case—which the Justices chose not to resolve explicitly, given the procedural posture in which the matter came to the Court—was whether all voluntary patients in state facilities must provide some sort of informed consent and, if so, whether they therefore have a constitutional right to some preadmission process by which their competence to provide that consent is evaluated. Dicta in Justice Blackmun’s opinion has led one commentator to conclude that the Court’s view was that “all voluntary patients should be screened for competence before hospitalization.” This broader suggestion is grounded in Blackmun’s discussion of the liberty interest held by psychiatric inpatients. Perhaps, however, the majority’s suggestion—that a preadmission assessment of a patient’s capacity is required by the due process clause of the Fourteenth Amendment, even in states without an explicit statutory requirement of informed consent—was a function of the extent of the deprivation in Burch’s case. The majority opinion goes to some lengths to catalogue the specific circumstances of Burch’s hospitalization, including the fact that he remained in the state hospital for five months. As Justice Blackmun put the point: “Burch’s confinement at [Florida State Hospital] for five months without a hearing or any other procedure to determine either that he validly had consented to admission, or that he met the statutory standard for involuntary placement, clearly infringes on [his] liberty interest.”

In any event, in the period immediately following the Supreme Court’s decision in Zinermon, its implied message, however inchoate, that voluntary patients must be competent to consent to inpatient admission, attracted the attention of key stakeholders in the field. Most particularly, it caught the attention of the APA, which promptly put in place a task force to address the question. At the time, nearly three-quarters of all inpatient psychiatric admissions in the United States were voluntary.

69. See id.
71. See id.
72. One writer has referred to this as the “hidden agenda” of the majority opinion in Zinermon. Nidich, supra note 67, at 704. This reading of the case, he suggests, “as requiring informed consent before a state hospital can accept a request for voluntary admission is certainly on the minds of those involved in this process.” Id.
73. Zinermon, 494 U.S. at 131.
74. See Stone, supra note 4, at 42.
75. See Appelbaum, supra note 70, at 1060; see also Slobogin, Rai & Reisner, supra note 3, at 857.
Available research on the capacity of these patients to meet a rigorous standard of informed consent indicated that the great majority of these patients were incapable of understanding and rationally processing even fairly basic information about inpatient psychiatric admission.\textsuperscript{76} If the alternative route of involuntary civil commitment were to become the primary means by which inpatient psychiatric admissions were achieved, massive new adjudicative resources would be required. In their absence, the courts, lawyers, and administrative personnel available to operate a vastly expanded civil commitment system would be overwhelmed.\textsuperscript{77} In addition, many of the former voluntary patients who now would have to be involuntarily committed might be prevented from receiving inpatient treatment because, while less than fully competent, they would not likely be found imminently dangerous to themselves or to others.\textsuperscript{78}

The Task Force, motivated by all of these considerations, concluded that “strong policy interests support the establishment of a low threshold for competence in this situation.”\textsuperscript{79} The “undemanding threshold” they adopted was made up of two competencies, the ability to “communicat[e] choices” and the ability to understand “that he/she is being admitted to a psychiatric hospital or ward for treatment, and [ ] understand[ ] release from the hospital may not be automatic . . . .”\textsuperscript{80} No additional requirements often associated with competency determinations—such as the ability to rationally weigh costs and benefits or to evaluate a decision free from false beliefs—were included in the Task Force’s recommendations.\textsuperscript{81}

Even this diminished set of requirements endorsed by the Task Force might have been too demanding, had they been rigorously applied. However, they have not been, at least not consistently.\textsuperscript{82} Given the circumstances, it is easy to see why this has been so. In research conducted a few years after the Task Force recommendations were issued, Norman

---

\textsuperscript{76} One widely cited study, for example, found that “60 percent of the sample [of schizophrenic patients] consented to treatment even though only 15 percent fully understood what they were consenting to . . . .” Grossman & Summers, supra note 2, at 206. The authors concluded that their “results suggest that only a small portion of schizophrenics may be able to give fully informed consent as required by law,” and thus “there is a serious discrepancy between legal requirements and the capacity of schizophrenic patients.” Id.

\textsuperscript{77} See Appelbaum, supra note 70, at 1060; see also Gutheil & Appelbaum, supra note 1, at 50–51.

\textsuperscript{78} See Stone, supra note 4, at 37.


\textsuperscript{80} Cournos et al., supra note 25, at 299–300.

\textsuperscript{81} See Gutheil & Appelbaum, supra note 1, at 216.

\textsuperscript{82} Although written before Zinermon was decided, the following remains an apt description of ongoing practice: “Many clinicians are not aware of the right of competent patients to refuse treatment and will inform them that they ‘must’ sign themselves into the hospital. Few therapists on the other hand will reject the agreement of a patient to treatment, even if he is clearly incompetent.” Gutheil & Appelbaum, supra note 1, at 37 (citation omitted).
Poythress and his colleagues designed a study in which they distinguished between “weak” and “strong” models of informed consent. The question they addressed was similar to that confronted by the Task Force: “how stringent should the test of capacity be in screenings anticipated by the Zinermon dicta?”83 Instead of focusing on the strong model of consent—which was defined to include “fairly extensive disclosure” about the legal rights waived by voluntary admission, the procedures for discharge, and the “adverse social costs (e.g., stigma) potentially attendant to hospitalization”—Poythress and colleagues measured participants’ performance on a “relatively non-demanding” test that had been designed by the Treatment Competence Subgroup of the MacArthur Foundation’s Research Network on Mental Health and the Law.84 The study found that, even on this more limited test of competence, over 60% of the voluntary patients they tested “demonstrated impaired capacity to consent . . . .”85

D. The Alternative of Guardians and Other Third-Party Decision Makers

Perhaps in recognition of the tension between the law’s systemic preference for voluntary admissions on the one hand and the limited capacity of many psychiatric inpatients to provide fully informed consent on the other, both the Zinermon majority and the APA took the position that guardians or other properly authorized third-party decision makers might be relied upon to provide the consent needed for admission. Thus, in his majority opinion in Zinermon, Justice Blackmun noted that protections for due process might include “appointment of a guardian advocate to make treatment decisions . . . .”86 The APA had endorsed the use of third-party surrogate decision makers even before Zinermon, and in its assessment of the issue after Zinermon, the APA’s Task Force elaborated, by stating that such surrogates might include guardians, family members, or others authorized by state law to make decisions on behalf of the impaired patient.87

84. See id. at 441–42, 447.
85. See id. at 447. The research participants were 120 persons initially brought to “crisis stabilization units” (CSU) in Florida. See id. at 442–43. Half of the subjects had been involuntarily admitted following assessment by a psychiatrist and half had been “permitted to sign into the CSU as voluntary treatment patients [ ].” Id. at 443. Surprisingly, even more of the voluntary patients failed the minimal competency test used by the researchers than did the involuntary patients. See id. at 447. The authors speculate that this might have been the result, in part, of the psychiatrists’ practice of denying patients who refused to consent to take psychotropic medications the opportunity to be admitted as voluntary patients. See id. at 448–49. The suggestion here is that some number of these refusers may actually have had more capacity to process information than did those patients who agreed to treatment and were admitted voluntarily. See id. at 449.
87. See Halverson, supra note 4, at 172–73.
It is not difficult to understand why substitute decision makers are an attractive alternative in this setting. As noted, an expansive reading of the Zinermon majority opinion, that takes seriously the full implications of Justice Blackmun’s dicta, creates a potential gap in the treatment system of considerable moment. As Donald Stone has explained:

Mentally ill persons who are incapable of giving informed consent to admission may not necessarily meet the statutory standard for involuntary placement. . . . Therefore, some patients who are incapable of providing informed consent to psychiatric hospitalization will not meet the criteria for involuntary confinement and may be discharged.

By guarding against undue pressure and influence to accept patients lacking in capacity to consent, some mentally ill persons who want to receive inpatient care may be denied treatment as long as they can live safely outside an institution.88

One potentially effective solution to this problem is to rely on eligible third-party decision makers. These substitute decision makers could be appointed by a court (usually in the role of guardian or conservator), identified in advance by the patient through a properly executed advance directive,89 or authorized by state statutes that empower family members or others close to the individual.90 Advance directives are not widely used for mental health treatment; consequently, third-party decision makers arranged by the patient are infrequently available.91 A number of states permit other designated third-party decision makers to provide consent for psychiatric treatment without special court approval.92 An additional group of states expressly prohibits surrogates from making these treatment decisions, absent judicial authorization.93 In most jurisdictions, however, the law with respect to the authority of third-parties other than guardians or conservators is unsettled, at best.94 By a process of elimination, therefore, court-appointed guardians become an important potential resource in many jurisdictions for managing the decision-making process by which inpatient psychiatric treatment might be arranged for those persons with mental disabilities who are unable to provide adequate informed consent on their own. Unfortunately, with some notable exceptions, the law in many states governing the power of guardians to arrange for the psychiatric hospitalization of their adult wards without utilizing the full process for involuntary civil commitment is underdeveloped, confused, or

88. Stone, supra note 4, at 37.
89. See Winick, supra note 24, at 57.
90. See Halverson, supra note 4, at 173.
91. See id.
92. See id. at 174.
93. See id.
94. See id. at 173–75.
inadequate. Before turning to a consideration of that body of law, however, a brief discussion of guardianships more generally, and of the difficult issue of incompetency that ordinarily triggers the appointment of a guardian, is in order.

III. INCOMPETENCY AND GUARDIANSHIP

A. The Presumption of Competency

For the most part, adults over the age of eighteen are presumed competent to manage their own affairs, including making health-care treatment decisions. The most commonly employed mechanism for overcoming the standing presumption of competency is a judicial hearing, at which a finding of general incompetency or incompetency with respect to a specific decision or function may result in the court’s appointment of a general or limited guardian. This reliance on judicial hearings to determine incompetency is a relatively recent development. Before the 1960s, the legal status of incompetency was most often established by way of a different legal presumption contrary to the ordinary presumption of adult competence, which was triggered by the involuntary commitment (and sometimes even voluntary hospitalization) of an individual with a mental disability. In most jurisdictions, this presumption of incompetency was irrebuttable; thus, a determination that an individual was subject to civil commitment was the “equivalent of a finding of general incompetency . . . .” As time went along, some states adjusted this rule to make psychiatric hospitalization the basis for a rebuttable rather than irrebuttable presumption of incompetency. Nevertheless, the great majority of patients who were hospitalized for mental illness or other mental disabilities were deemed incompetent and thus lost the power to enter into contracts, to initiate lawsuits, to marry or divorce, to decide where to live, and to consent to medical care.

More recently, a great many clinicians and legal advocates have concluded that mental illnesses and other mental disabilities that are severe enough to warrant inpatient treatment do not necessarily render patients entirely incapable of making all significant decisions or otherwise participating in a variety of daily activities associated with managing their own affairs. As a result, the irrebuttable presumption that a person subject to involuntary commitment (or voluntary hospitalization) is generally in-

97. See Gutheil & Appelbaum, supra note 1, at 222.
98. See id.
99. See id.
100. See Slobogin, Rai & Reisner, supra note 3, at 940.
101. See Gutheil & Appelbaum, supra note 1, at 223; Weiner & Wettstein, supra note 96, at 116.
competent has been abrogated, replaced either by a rebuttable presumption of incompetence or, more commonly, by the ordinary background norm that all adults are presumed competent until found otherwise by a court of appropriate jurisdiction.\footnote{See Gutheil & Appelbaum, supra note 1, at 222–23.} As a rule, the determination of general incompetency is now legally distinct from the decision to civilly commit an individual for inpatient treatment.\footnote{See id. at 223.} As civil commitment increasingly has come to be centered on the state’s police power interest in restraining and treating patients whose mental disability poses a danger to themselves or to others,\footnote{See Brakel et al., supra note 3, at 26–27.} the state’s \textit{parens patriae} interest in caring for those who are incapable of caring for themselves increasingly has become concentrated in the judicial process by which incompetency (and guardianship) are determined.\footnote{See Gutheil & Appelbaum, supra note 1, at 222–23; see also Brakel et al., supra note 3, at 370.}

The now common presumption that patients subject to psychiatric hospitalization should be deemed competent unless a court makes contrary findings is, of course, relevant to the \textit{Zinermon} dicta suggesting that voluntary patients must be capable of providing consent for admission.\footnote{A very different approach pertained in the 1950s, when the National Institute of Mental Health took the position that voluntary psychiatric patients need not be competent to consent to hospitalization. See Nat’l Inst. of Mental Health, A Draft Act Governing the Hospitalization of the Mentally Ill, Public Health Serv. Pub. No. 51 (1951), available at http://catalog.hathitrust.org/Record/006203370. Paul Appelbaum reports that this decision led many states to adopt the same approach. See Appelbaum, supra note 70, at 1060. The governing assumption for this policy was that “the benefits of voluntary hospitalization, including a presumed acknowledgment of a need for treatment, a stronger alliance with treatment personnel, and avoidance of the stigma of court hearings, should not be denied non-objecting patients, even if they would not meet ordinary competence criteria.” Id.} Depending on the definition of competency one employs, however, research suggests that a significant portion of the population of voluntary patients in mental hospitals and specialized psychiatric units likely cannot engage fully in the cognitive process by which information relevant to the hospitalization decision must be evaluated in order to meet a standard of fully informed consent.\footnote{See Grossman & Summers, supra note 2, at 206.} Some states now routinely require voluntary patients, involuntary patients, or both to be screened for competency, and some routinely petition the appropriate court for the appointment of a guardian when the patient is determined to be incapable of providing the necessary consent either for admission, in the case of voluntary patients, or for other treatment decisions such as the administration of psychotropic medications, in the case of involuntary patients.\footnote{See Gutheil & Appelbaum, supra note 1, at 223.} There are significant limitations on the resources available to support such a widespread...
use of guardians for all patients who might be regarded as incompetent, at least according to some reasonably demanding conception of competency, not the least of which is the limited availability of persons willing and able to serve in that capacity.\textsuperscript{109} As noted above, there are also important legal restrictions on the authority of guardians in many jurisdictions to perform as substitute decision makers in order to provide the consent needed, at least in the case of approving voluntary psychiatric hospitalization.\textsuperscript{110}

\section*{B. Determining Incompetency}

Incompetency, which ordinarily is the standard by which a guardianship is authorized,\textsuperscript{111} is subject to significant contest, both as to its definition and its legal significance.\textsuperscript{112} The range of important legal consequences that flow from a finding of incompetency by a court turns on a constellation of criteria that may not be entirely appropriate when clinicians seek to evaluate the competency of patients for other purposes, including questions of medical management or treatment.\textsuperscript{113} A classic discussion of the factors relevant to evaluating competency—originally developed by Paul Appelbaum and Loren Roth for the purpose of determining the capacity of individuals with mental disabilities to consent to research—provides useful tools for considering both legal and clinical competency and has become a common starting point for commentators in this field.\textsuperscript{114} Originally conceived as a hierarchy of capacities, Professor Appelbaum subsequently suggested that each of the four categories of functionality he and his colleague identified might apply differently and might hold independent significance, depending on the individual circumstances presented by any particular individual whose competency is to be assessed.\textsuperscript{115} Logically, however, for most purposes, the four areas of capacity they describe form a set of criteria of apparent ascending importance.\textsuperscript{116}

The first is the capacity to evidence a choice.\textsuperscript{117} Ordinarily, even severely mentally disabled individuals are able to communicate their agree-

\begin{itemize}
  \item 109. See id. at 224–25.
  \item 110. For a further discussion of the various legal restrictions on the authority of guardians, see supra notes 9–14 and accompanying text.
  \item 111. Although, “[t]he trend in the law has been to abandon the term incompetent and refer to the person as incapacitated or disabled, since this is less pejorative.” Weiner & Wettstein, supra note 96, at 282.
  \item 112. See Gutheil & Appelbaum, supra note 1, at 215–16; Winick, supra note 95, at 6.
  \item 113. See Gutheil & Appelbaum, supra note 1, at 215; Weiner & Wettstein, supra note 96, at 116; cf. Roca, supra note 2, at 1177.
  \item 114. See generally Paul S. Appelbaum & Loren H. Roth, Competency to Consent to Research: A Psychiatric Overview, 39 Archives Gen. Psychiatry 951 (1982).
  \item 115. See Paul S. Appelbaum & Thomas Grisso, Mental Illness and Competence to Consent to Treatment, 19 L. & Hum. Behav. 105, 110 (1995).
  \item 116. See Slorogin, Rai & Reisner, supra note 3, at 932.
  \item 117. See Appelbaum & Roth, supra note 114, at 952–53.
\end{itemize}
ment or disagreement with a proposed decision, though the basis for that assent or refusal may not always be entirely rational or grounded in objective fact. Occasionally, however, a patient is either entirely incapable of communication, or provides a mix of verbal and/or nonverbal signals that are so inconsistent that a clear indication of the patient’s choice is impossible to discern. For the most part, though, relatively few incompetency determinations are based on this first sort of incapacity.118

The second set of functional capacities identified by Appelbaum and Roth are more consequential and govern the analysis in a greater number of cases. They relate to an individual’s ability to understand the facts central to a proposed decision or task.119 These capacities are largely cognitive in nature. Where a particular choice is contemplated, the required matters that must be understood include the options open to decision and the respective costs and benefits of electing one or another of these options.120 In the case of decision making with respect to medical care, this sort of factual understanding includes a clear comprehension of the proposed intervention, the likelihood of success, the risks and potential side effects associated with the treatment, and the available alternative options along with their risks and potential benefits.121

The level of detail and complexity of the information deemed essential for a competent decision varies considerably, depending on the subject under consideration and the judgment of the evaluator.122 Beyond the question of detail and complexity, Appelbaum and Roth helpfully divide this second category of functional capacity into two sub-categories that further refine the assessment process. The first sub-category goes to the individual’s basic “ability to understand” relevant facts, while the second concerns the individual’s more refined capacity to demonstrate “actual understanding” of those facts.123 An evaluator might determine that a patient has the basic ability to understand the facts bearing on a decision by exploring whether he or she has a grasp of facts of a similar order or level of complexity, or perhaps by asking him or her to identify relevant information on a written list of alternatives.124 The more demanding question is whether the patient can exhibit “actual understanding” by re-

118. See id.
119. See id. at 953–54.
120. See id.
121. See Gutheil & Appelbaum, supra note 1, at 219 (“The patient ought to have the ability to understand the proposed interventions, including their risks, benefits, and the possible alternatives. For example, an acutely psychotic patient should understand that psychotropic medication carries the risk of dystonic reactions . . . ; that the benefit is the probable resolution of the psychotic episode; and that alternatives include psychotherapy and milieu therapy, and possibly ECT, but that at least the two former alternatives carry a lower short-term success rate than does medication.”).
122. See Roca, supra note 2, at 1195.
123. See Appelbaum & Roth, supra note 114, at 953.
124. See Poythress et al., supra note 83, at 951.
peating back information he or she has been given, by paraphrasing that information in his or her own words, or by showing some comprehension of the consequences of the choices that are available.\textsuperscript{125}

Closely related to the demonstration of an actual understanding of essential facts is the third functional capacity described by Appelbaum and Roth, the ability to apply that information in a “rational” process of deliberation.\textsuperscript{126} Other writers, most prominently Michael Moore, have elaborated on the central role that “practical reasoning” plays in assessing an individual’s moral agency. In its most essential form, this practical reasoning—or rational deliberation—requires the decision maker to go through a process of weighing the relative costs and benefits of competing choices in order to arrive at a decision.\textsuperscript{127} This practical reasoning is said to have moral significance, in part because the very process of assigning weight to these various costs and benefits necessarily draws upon the individual’s foundational normative commitments and value structure.\textsuperscript{128}

In any event, a system in which competency is evaluated according to an individual’s capacity to engage in the rational manipulation of information must contend with several challenges. The first follows from the moral significance assigned by Moore and others to the process of practical reasoning. By definition, the assessment of another’s rationality is highly subjective, precisely because the judgment that another has arrived at a rational decision likely turns, at some level, on agreement between the assessor and the decision maker with respect to the relative weights assigned to the risks and benefits of the choices under consideration.\textsuperscript{129} Beyond the problem of subjectivity, the rational manipulation criterion also runs into potential difficulty when the patient being evaluated holds false beliefs with respect to one or more facts that enter into the practical reasoning process.\textsuperscript{130} Appelbaum and Roth point out that the law has long

\textsuperscript{125.} See Appelbaum & Roth, supra note 114, at 953–54.

\textsuperscript{126.} See id. at 954.


\textsuperscript{129.} See Appelbaum & Roth, supra note 114, at 954. This point has been pressed with particular vigor by Duncan Kennedy, who has argued that the decision to overrule another’s choice cannot be explained by a neutral conception of capacity, because “the question of capacity is hopelessly intertwined with the question of what the other wants to do in this particular case.” Duncan Kennedy, \textit{Distributive and Paternalist Motives in Contract and Tort Law, with Special Reference to Compulsory Terms and Unequal Bargaining Power}, 41 MD. L. REV. 563, 644 (1982). For Kennedy, the focus of some commentators and courts on the ability of an individual to undertake a rational decision-making process is effectively a cover for an underlying paternalism (perhaps necessary under the circumstances) of acting contrary to the wishes of the individual, in order to protect what another determines to be the incapacitated person’s best interests. See id. at 642–46.

\textsuperscript{130.} See Gutheil & Appelbaum, supra note 1, at 218 (stating that determining patient’s ability rationally to manipulate information requires examination of “the
treated the presence of “insane delusions” as adequate grounds for a finding that an individual lacks contractual or testimonial capacity, thus providing a basis to invalidate a contract or will executed on the basis of such false beliefs. But individuals with mental illness or other mental disabilities may suffer impairments that impact the rationality of particular choices or behaviors without affecting their cognitive processes more globally. Because “the impact of delusions, for example, may be limited to a discrete area of mental functioning,” the fact that a patient holds false beliefs relating to one set of issues may not be relevant to his or her capacity to assess the relative risks and benefits of a proposed course of action unrelated to his or her thought disorder. The question in such cases, then, is whether the individual should be regarded as competent to make the decisions and manage the portions of his or her affairs unaffected by the cognitive dysfunction caused by the mental disability.

The fourth competency identified by Appelbaum and Roth is the ability to “appreciat[e] the nature of the situation.” Appreciation, they explain, is distinct from factual understanding in that it requires the individual to possess the ability to apply the abstract information he or she has been provided to a concrete circumstance, most commonly the individual’s immediate situation. It is one thing, for example, to understand that the rules governing voluntary admission formally limit patients’ ability to leave the hospital without providing forty-eight hours’ notice; it is quite another for patients to appreciate that staff will stop them if they attempt to walk out the front door of the hospital. Beyond the application of abstract knowledge to concrete circumstances, the capacity of appreciating the basic components of the patient’s mental status: orientation, memory, intellectual functioning, judgment, impairment in rationality (hallucinations and delusions), and alterations of mood).
tion also entails the integration of cognition and affect. In a different context, the drafters of the American Law Institute’s (ALI) Model Penal Code provision governing the insanity defense recognized this insight when they altered the common law M’Naghten test for criminal defendants asserting a defense of insanity. Under the ALI test, the inquiry is not what a criminal defendant “knew” about the facts relevant to the offense, which was the prevailing common law standard, but whether the defendant was able to fully “appreciate” the significance of that information. In explaining this shift in emphasis, the Model Penal Code drafters observed that appreciation of the wrongfulness of one’s conduct entails a mature understanding of and emotional connection with its consequences. Clearly, requiring an individual to have this sort of appreciation, particularly regarding the decision whether to agree to enter an inpatient facility for mental health treatment, is a demanding standard for many patients with acute mental disabilities. If this capacity were part of the matrix necessary for legal competency, and if due process were understood to require informed consent measured in such a fashion, many patients who would otherwise be admitted as voluntary patients would lose that opportunity and would be limited to involuntary hospitalization, outpatient treatment, or no treatment at all.

The Appelbaum and Roth competency criteria, taken as a whole, are focused on the process by which a decision is made and not on the outcome of that decision-making process. One could also assess competency, however, by evaluating outcomes, with little or no attention to the process through which information supporting a choice has been obtained and weighed. This fundamental distinction between a process-based approach and an outcomes-centered approach to competency determinations is crucial in shaping the law governing adult guardianship.

136. See id.
137. See Model Penal Code § 4.01 (Official Draft 1962) (Mental Disease or Defect Excluding Responsibility).
139. See Slobozin, Rai & Reisner, supra note 3, at 947. In a slightly different context, Robert Roca has offered the following example, which helps to illuminate this key distinction:

Consider an elderly man who refuses to stop smoking despite severe progressive obstructive pulmonary disease. His physicians inform him that he is hastening his demise by continuing to smoke. He retorts that he does not believe smoking is harming him and that in fact he relaxes and breathes more comfortably when he smokes. His choice is unwise, unreasonable, and at odds with general knowledge about the relationship between smoking and pulmonary disease. A psychiatric evaluation is requested and reveals no evidence of dementia, major depression, or any other capacity-compromising psychiatric syndrome. His refusal to accept commonly-held beliefs about the relationship between smoking and lung disease might lead ardent anti-smokers to question his “competency,” and this challenge might be sustained if there were no requirement for the demonstration of a disabling psychiatric disorder.
The traditional state law approach governing the appointment of guardians for incompetent adults has focused on whether the individual is capable of making decisions that produce reasonable outcomes. More recently, a number of states have adopted an alternative approach, developed as part of the Uniform Probate Code (UPC), which shifts the emphasis from reasonable outcomes to the soundness of the individual's decision-making process itself. In response to a concern that the UPC's process-based approach is too subjective, a third alternative has appeared in some states that centers on the capacity of the individual to undertake specific tasks having to do with essential functions such as housing, health care, and nutrition.

An adult under guardianship could have considerable actual capacity to participate in the decision to seek inpatient treatment, could be only partially impaired, or could be largely incapable of usefully contributing to that decision-making process, depending upon which of these legal regimes for the appointment of a guardian the court has employed. Moreover, a ward might express agreement with the inpatient placement, express no opinion, or make clear his or her opposition to the plan. The authority of a guardian to arrange a voluntary admission ought to reflect an assessment of both the ward's capacity for engaging in the decision-making process and producing reasonable outcomes. In some instances, where the ward is capable of reasonable judgment and has communicated his or her agreement with inpatient admission, the guardian should be permitted to exercise considerable discretion. In other instances, where the ward is more severely impaired and/or has not conveyed agreement with the inpatient placement, judicial oversight will be essential. The legal formula for managing this complex relationship between adult wards and their guardians facing a hospitalization decision ordinarily will derive from the statutes governing voluntary admissions together with the state law setting out the powers of guardians. This Article now turns to a consideration of that body of law.

Roca, supra note 2, at 1188.

From a process-based point of view, it would be difficult, though not impossible, to conclude that the smoker in the above example is incompetent. His belief that smoking is not harming his health is a false belief judged against objective medical evidence, and under a rigorous version of the process approach, one might argue that his capacity to rationally assess the relevant evidence is therefore impaired. Virtually all of the other process-focused criteria identified by Appelbaum and Roth, however, point in the other direction. A more paternalistic outcomes-based perspective, on the other hand, might well support the conclusion that the elderly smoker is incompetent, given that his choice is clearly not in his objective best interest.
IV. SEEKING A BALANCE OF INTERESTS: A REVIEW OF STATE LAWS

The authority of a court-appointed guardian to consent to voluntary inpatient psychiatric treatment for his or her ward varies from state to state. In some jurisdictions, the guardian’s authority turns on whether the ward has objected to the proposed hospitalization, has indicated consent, or has failed to protest the admission. Often, the guardian’s ability to authorize inpatient treatment depends on how state law views the ward’s capacity following the guardian’s appointment. Thus, the ward may categorically be regarded as legally incompetent under state law by virtue of the court’s appointment of a general guardian, or the ward may be understood as having partial decision-making capacity to the extent that he or she can in fact make a reasonable choice with respect to hospitalization, notwithstanding the court’s prior guardianship decision. In addition, state laws vary on whether a patient subject to voluntary psychiatric hospitalization must be capable of giving informed consent. This variation in state law reflects, in part, ongoing uncertainty about whether the Supreme Court’s majority decision in the Zinermon case made the capacity to provide such consent (or some lesser degree of decision-making competence, as suggested by the APA Task Force recommendations) a requirement of federal due process.

In more than a half-dozen states, a guardian may provide consent for the voluntary psychiatric hospitalization of his or her adult ward. In Georgia, for example, guardians have broad decision-making authority under the guardian powers statutes and the voluntary commitment law that together appear to provide a sufficient foundation for providing consent to inpatient care. In Maine, a guardian may provide consent unless the adult ward objected at an earlier time when he or she had capacity, in which case court approval for inpatient psychiatric treatment is required. North Carolina’s guardian statute is especially broad, providing


141. See supra notes 71–74 and accompanying text.

142. See GA. CODE ANN. § 37-3-20 (2014) (authorizing guardian to make application for ward’s voluntary commitment); see also id. § 29-4-21(a)(3) (“[T]he appointment of a guardian shall remove from the ward the power to . . . [c]onsent to medical treatment . . . .”); id. § 29-423(a)(2) (allowing guardian to “give any consents or approvals that may be necessary for medical or other professional care, counsel, treatment, or service for the ward”).

143. Under Maine’s guardian powers statute, a guardian is authorized to provide consent for medical and other care or treatment: “A guardian may give or withhold consents or approvals related to medical or other professional care, counsel, treatment or service for the ward.” ME. REV. STAT. ANN. tit. 18-A, § 5-312(a)(3) (2013). However, if a guardian’s decision is against the wishes of his or her ward, as expressed by the ward when the ward had capacity, then court approval is necessary. See id. § 5-806; see also Guardianship of Boyle, 674 A.2d 912, 915 (Me. 1996) (“The guardian of an incapacitated person has the same powers that ‘a parent has respecting [the parent’s] unemancipated minor child,’ unless
the guardian with the power to “give any consent or approval that may be necessary to enable the ward to receive medical, legal, psychological, or other professional care, counsel, treatment, or service . . . .”144 In South Dakota, if the ward “presents” to a psychiatric facility, the ward’s guardian may provide consent for voluntary inpatient treatment.145

The statutory provisions in Wisconsin governing the provision of psychiatric treatment for an incompetent adult suggest a similar approach. With respect to healthcare generally, the Wisconsin guardian powers statute makes clear that a guardian has:

[T]he power to give an informed consent to the voluntary receipt by the guardian’s ward of a medical examination, medication, modified by court order. This authority includes the giving or withholding of consent or approval related to medical care.” (alteration in original) (citation omitted) (quoting Me. Rev. Stat. Ann. tit. 18-A, § 5-312(a)) (citing Me. Rev. Stat. Ann. tit. 18-A, §5-312(a)(3)).


(a) To the extent that it is not inconsistent with the terms of any order of the clerk or any other court of competent jurisdiction, a guardian of the person has the following powers and duties:

(3) The guardian of the person may give any consent or approval that may be necessary to enable the ward to receive medical, legal, psychological, or other professional care, counsel, treatment, or service; provided that, if the patient has a health care agent appointed pursuant to a valid health care power of attorney, the health care agent shall have the right to exercise the authority granted in the health care power of attorney unless the Clerk has suspended the authority of that health care agent . . . . The guardian of the person may give any other consent or approval on the ward’s behalf that may be required or in the ward’s best interest. The guardian may petition the clerk for the clerk’s concurrence in the consent or approval.

Id. § 35A-1241(a), (a)(3). Significantly, in directing the guardian’s decision making with respect to the ward’s place of abode, the statute sets out a preference for residences that are not treatment facilities and for treatment facilities that are community-based. Nevertheless, the authority to consent to inpatient psychiatric treatment is not expressly foreclosed by the statute:

(2) The guardian of the person may establish the ward’s place of abode within or without this State. . . . The guardian also shall give preference to places that are not treatment facilities. If the only available and appropriate places of domicile are treatment facilities, the guardian shall give preference to community-based treatment facilities, such as group homes or nursing homes, over treatment facilities that are not community-based.

Id. § 35A-1241(a)(2).

North Carolina’s voluntary admissions statute authorizes a health care agent/surrogate to provide proxy consent, but it does not address the authority of a guardian in that regard. See id. § 122C-211. Nevertheless, it appears that the guardian powers statute controls in the absence of an agent/surrogate.

145. See S.D. Codified Laws § 27A-8-18.1(1)–(4) (2014) (“If a person eighteen years of age or older presents for admission to an inpatient psychiatric facility and meets the requirements . . . . but the facility director or administrator determines that the person is incapable of exercising an informed consent to the admission, then the person may be admitted upon exercise of a substituted informed consent: [ ] [b]y a guardian . . . [or other possible surrogate decision makers].”).
2015]  “VOLUNTARY” INPATIENT TREATMENT  27

including any appropriate psychotropic medication, and medical
treatment that is in the ward’s best interest, if the guardian has
first made a good-faith attempt to discuss with the ward the vol-
tuntary receipt of the examination, medication, or treatment and
if the ward does not protest.\footnote{146}{WIS. STAT. ANN. § 54.25(2)(d) (West 2014).}

With respect to consent for inpatient treatment particularly, Wisconsin law
authorizes the guardian of an incompetent adult to provide consent for
the voluntary admission of the ward to an inpatient treatment facility in
cases where the ward does not indicate a desire to leave the facility, if
general statutory procedures for voluntary admission are followed.\footnote{147}{See id. § 51.10(8) (“An adult for whom, because of incompetency, a
guardian of the person has been appointed in this state may be voluntarily admitted to an inpatient treatment facility if the guardian consents after the requirements of sub. (4m)(a)1. are satisfied or if the guardian and the ward consent to the admission under this section.”).}

Statutes in Michigan permit the guardian of an adult ward to authorize
inpatient psychiatric treatment if the patient “assents,” although this
power is limited to “formal” voluntary admissions.\footnote{148}{See MICH. COMP. LAWS § 330.1415 (2014). Section 330.1411, on the other hand, deals with informal voluntary admissions and indicates that only patients themselves can apply to be admitted in this fashion. See id. § 330.1411. But see id. § 330.1415 (governing formal voluntary admissions). While almost identical, section 330.1415 provides: [A]n individual 18 years of age or over may be hospitalized as a formal voluntary patient if the individual executes an application for hospitalization as a formal voluntary patient or the individual assents and the full guardian of the individual, the limited guardian with authority to admit, or a patient advocate authorized by the individual to make mental health treatment decisions under the estates and protected individuals code executes an application for hospitalization and if the hospital director considers the individual to be clinically suitable for that form of hospitalization. Id. (citations omitted) (citing MICH. COMP. LAWS §§ 700.1101–700.8102; 1998 Mich. Pub. Acts 386).}

In Missouri, a guardian can consent to voluntary hospitalization for thirty days, after which a court order is required.\footnote{149}{See MO. ANN. STAT. § 632.120 (West 2014) (permitting guardian to make application for voluntary hospitalization of ward). But see id. § 475.120(5) (placing limit on guardian’s ability to have ward admitted: “No guardian of the person shall have authority to seek admission of the guardian’s ward to a mental health or intellectual disability facility for more than thirty days for any purpose without court order except as otherwise provided by law.”).}

In North Dakota, a guardian of an adult can consent to voluntary psychiatric hospitalization for up to forty-five days without an authorizing court order.\footnote{150}{See N.D. CENT. CODE § 30.1-28-12(2) (2013) (“[N]o guardian may voluntarily admit a ward to a mental health facility or state institution for a period of more than forty-five days without a mental health commitment proceeding or other court order.”). The negative implication is that a guardian may voluntarily admit a ward to a treatment facility for forty-five days or less. The Supreme Court of North Dakota has considered whether this guardian powers statute applies in...}
In more than a dozen states, by contrast, the law is settled that a guardian is not authorized to consent to voluntary psychiatric hospitalization if the ward either has not provided or is incapable of providing informed consent. In these states, the guardian seeking inpatient treatment for his or her ward ordinarily will be required to initiate an involuntary commitment process. States in this category include Alaska, Illinois, the case of a ward admitted to a care facility (a locked dementia unit), and has concluded, based on legislative history, that it does not. See In re Guardianship/Conservatorship of Van Sickle, 694 N.W.2d 212, 218–21 (N.D. 2005). The guardian in Van Sickle was permitted to place his ward in this facility without first obtaining a court order. See id. The clear implication of this holding is that, in circumstances in which the ward is to be admitted to a psychiatric facility as opposed to a care facility, for a period longer than forty-five days, the guardian will be required to obtain an authorizing court order.

151. See ALASKA STAT. § 13.26.150(e), (e)(1) (2014) (prohibiting guardians from placing wards in psychiatric facilities: “A guardian may not [ ] place the ward in a facility or institution for the mentally ill other than through a formal commitment proceeding under [Alaska Stat. §] 47.30 in which the ward has a separate guardian ad litem . . . .”).

152. See 755 ILL. COMP. STAT. ANN. 5/11a-17(a) (West 2014) (“A guardian of the person may not admit a ward to a mental health facility except at the ward’s request as provided in Article IV of the Mental Health and Developmental Disabilities Code and unless the ward has the capacity to consent to such admission as provided in Article IV of the Mental Health and Developmental Disabilities Code.”). This provision is consistent with the Illinois Appellate Court’s decision in In re Gardner, in which the court held that a guardian cannot authorize the admission of a non-consenting ward to a mental health treatment facility as a voluntary patient. See In re Gardner, 459 N.E.2d 17, 20 (Ill. App. Ct. 1984). The ward in Gardner lacked the capacity to provide informed consent for inpatient treatment, but was not dangerous. See id. at 18. The court, although recognizing that the ward would not be subject to hospitalization under the involuntary commitment statute, nevertheless refused to allow the guardian the authority to consent to voluntary inpatient treatment. See id. at 20; see also In re Guardianship of Muehlner v. Blessing Hosp., 782 N.E.2d 799, 802 (Ill. App. Ct. 2002) (“[A] trial court may not grant a guardian the power to admit a nonconsenting ward to a mental health facility for treatment as a voluntary patient.” (citing In re Gardner, 459 N.E.2d at 20)).
Maryland,\textsuperscript{153} Montana,\textsuperscript{154} New York,\textsuperscript{155} Pennsylvania,\textsuperscript{156} Texas,\textsuperscript{157} and the District of Columbia.\textsuperscript{158}

\begin{itemize}
\item \textsuperscript{153} See Md. Code Ann., Est. & Trusts § 13-708(b)(2) (LexisNexis 2014) (setting out functions of guardians and providing the “right to custody of the disabled person and to establish his place of abode within and without the State, provided there is court authorization for any change in the classification of abode, except that no one may be committed to a mental facility without an involuntary commitment proceeding as provided by law”).
\item \textsuperscript{154} See Mont. Code Ann. § 53-21-111 (2013). While Montana’s voluntary admission statute is silent as to whether a guardian can make an application to have a ward admitted, it does contain an informed consent requirement that suggests that the patient must provide informed consent. \textit{See id.} Additionally, the guardian powers statute appears to limit a guardian’s ability to place a ward in an inpatient psychiatric facility. \textit{See id.} § 72-5-321. Although it permits a guardian to give consent for medical and other care, and generally provides that the guardian of an adult ward has the same powers that parents have with respect to the care of their children, the statute then goes on to limit this power in the mental health context:
\begin{quote}
A full guardian or limited guardian may not involuntarily commit for mental health treatment or for treatment of a developmental disability or for observation or evaluation a ward who is unwilling or unable to give informed consent to commitment . . . unless the procedures for involuntary commitment set forth in Title 53, chapters 20 and 21, are followed.
\end{quote}
\textit{Id.} § 72-5-321(5).
\item \textsuperscript{155} See N.Y. Mental Hly. Law § 81.22(b)(1) (McKinney 2014) (“No guardian may: [ ] consent to the voluntary formal or informal admission of the incapacitated person to a mental hygiene facility . . . .” (informal admission refers to voluntary admission)); \textit{see also id.} § 9.13 (governing voluntary admissions but not authorizing a guardian to provide consent or make application); \textit{id.} § 9.17 (providing informed consent requirements but not authorizing a guardian to provide proxy consent).
\item \textsuperscript{156} See 20 Pa. Cons. Stat. § 5521(f) (2013) (setting out the “[p]owers and duties not granted to [a] guardian”). Section 5521(f) provides: “The court may not grant to a guardian powers controlled by other statute, including, but not limited to, the power: [ ] [t]o admit the incapacitated person to an inpatient psychiatric facility or State center for the mentally retarded.” \textit{Id.} In addition, the Pennsylvania statutes governing voluntary hospitalization do not contain any language that can be construed to authorize guardian consent. \textit{See 50 Pa. Cons. Stat. §§ 4402, 4403, 7201, 7203.}
\item \textsuperscript{157} See Tex. Estates Code Ann. § 1151.051(c)(4) (West 2014) (providing that guardians have “power to consent to medical, psychiatric, and surgical treatment other than the inpatient psychiatric commitment of the ward”). However, the statute qualifies this provision:
\begin{quote}
Notwithstanding Subsection (c)(4), a guardian of the person of a ward has the power to personally transport the ward or to direct the ward’s transport by emergency medical services or other means to an inpatient mental health facility for a preliminary examination in accordance with Subchapters A and C, Chapter 573, Health and Safety Code.
\end{quote}
\textit{Id.} § 1151.051(d).
\item Thus, under Texas law, a guardian can arrange for his or her ward to be transported to a psychiatric facility for evaluation, but is not empowered to consent to inpatient treatment. The voluntary admissions statutes are also consistent with the position that a guardian cannot provide the requisite consent. \textit{See Tex. Health & Safety Code Ann.} § 572.001, 572.002 (West 2013).
\item \textsuperscript{158} See D.C. Code § 21-2047(b)(4) (2012) (providing that, among other powers, guardians may “[c]onsent to medical examination and medical or other
In Washington state, guardians may not arrange inpatient psychiatric treatment for wards who are unwilling or unable to give informed consent, unless the procedures for involuntary commitment are followed. This statute, in effect, codifies the holding of the Court of Appeals of Washington in *In re Guardianship of Anderson*. In *Anderson*, the court construed an earlier guardian powers provision to require the consent of the ward for a voluntary admission, notwithstanding clear statutory language that “permit[ted] public or private health facilities to accept ‘any person . . . suitable . . . for care and treatment as mentally ill, or for observation as to the existence of mental illness, . . .’ who applies for admission individually or through their court-appointed guardian . . . .” The *Anderson* court explained that the hospitalization of an incapacitated individual under this section without his or her informed consent would constitute “involuntary incarceration,” which is only permissible under the state’s police powers through the involuntary civil commitment process. The guardian in *Anderson* had twice sought to have his ward committed under the involuntary civil commitment process, but he had failed because the lower court did not find that the ward was a danger “to self or others.” The Court of Appeals made clear that it would permit inpatient treatment under these circumstances only if a factual basis were established supporting either a finding of dangerousness or that the non-consenting ward was “gravely disabled.”

While the guardians of incompetent adults in Massachusetts at one time did have the authority to consent to inpatient mental health care, they no longer are permitted to do so. The evolution of Massachusetts professional care, treatment, or advice for the ward”). But see id. § 21-2047.01 (limiting guardians’ power in the mental health treatment context). Section 21-2047.01 states that:

A guardian shall not have the power: . . . [t]o consent to the involuntary or voluntary civil commitment of an incapacitated individual who is alleged to be mentally ill and dangerous under any provision or proceeding occurring under Chapter 5 of Title 21, except that a guardian may function as a petitioner for the commitment consistent with the requirements of [the involuntary civil commitment statute] . . . .”

Id. § 21-2047.01(4) (emphasis added).

159. See Wash. Rev. Code § 11.92.043(5) (2014) (providing additional duties of guardian: “No guardian, limited guardian, or standby guardian may involuntarily commit for mental health treatment, observation, or evaluation an alleged incapacitated person who is unable or unwilling to give informed consent to such commitment unless the procedures for involuntary commitment set forth in chapter 71.05 or 72.23 RCW are followed.” (emphasis added)).


161. Id. at 1192 (second, third, and fourth alterations in original) (quoting Wash. Rev. Code Ann. § 72.23.070 (repealed 1985)).

162. See id.

163. Id. at 1191.

164. Id. at 1192.

165. See Mass. Gen. Laws Ann. ch. 190B, § 5-509(f) (West 2014) (“No guardian shall be given the authority under this chapter to admit or commit an incapacitated person to a mental health facility or a mental retardation facility . . . .”).
law on this question is instructive, in part because of the shadow cast by a changing understanding over the past forty-five years of the rights associated with involuntary commitment in the state. 166 Before 1977, the guardian of an adult ward could commit the ward to a mental health facility without prior court approval, and such a commitment was treated as a voluntary admission. 167 In 1977, the state legislature enacted a statute “regulating the powers of guardians to admit or commit wards to mental health or retardation facilities without the consent of the wards.” 168 The new statute required judicial approval before an inpatient placement was permitted on the consent of the guardian and required the court to find that such a placement would be in the ward’s “best interests.” 169 The 1977 revisions also required that the ward be present at the hearing at which best interests were adjudicated and that counsel be provided if the ward was indigent. 170

In Doe v. Doe, 171 the Supreme Judicial Court of Massachusetts took up the question of what was required under this “best interests” standard, particularly in instances in which the ward had not joined in the guardian’s application for a voluntary admission. 172 Under the circumstances of that case, the court held that a showing of a “likelihood of serious harm,” which was also a required element for involuntary commitment in the state, would be necessary before judicial approval would be appropriate for a guardian seeking the voluntary admission of a ward. 173 The court arrived at this conclusion expressly because “commitment pursuant to [the guardian powers statute] produces the same loss of freedom and the same label of mental illness as commitment under [the involuntary commitment statute].” 174 Thus, in order to insure that a voluntary inpatient admission approved by a guardian without the ward’s consent would not be permitted as an end-run around the state’s involuntary commitment statute, the court concluded that essentially the same finding of dangerous

166. In 1970, the state’s involuntary commitment statute was completely revised. See generally MASS. GEN. LAWS ANN. ch. 123 (West 2014). A person is subject to commitment under the revised provision if he or she is found to be mentally ill and his or her discharge would create a “likelihood of serious harm.” Id. ch. 123, § 12.


169. See id.

170. See id.


172. See id. at 999.

173. See id. at 1000. In effect, the court held that the statutory requirement of “best interests” could only be established, in cases in which the ward had withheld consent, by a showing that his or her discharge from inpatient care would create a “likelihood of serious harm.”

174. Id. at 1001.
ness to self or others required for involuntary commitment was impliedly required by the guardian powers statute’s “best interest” standard. Eventually, this skeptical intuition, reflected in the court’s construction of the state’s guardian powers statute, was enacted into the statute itself, which now flatly prohibits guardians from consenting to the voluntary hospitalization of their wards without regard to any assessment of “best interests” or dangerousness.

Significantly, while it appears that a guardian may not place an adult ward in a mental health inpatient facility on a voluntary basis in Massachusetts, the state’s Supreme Judicial Court has indicated that an individual’s health care proxy agent may accomplish precisely the same result, at least under certain identified circumstances. In Cohen v. Bolduc, the court construed Massachusetts’s health care proxy statute to permit such a voluntary admission, so long as the principal does not object to the placement or revoke the proxy instrument. The Massachusetts statute permitting individuals to give advance directives for health-care decisions permits a proxy decision maker to arrange treatment for both physical and mental disorders. The Cohen court acknowledged that the statute is not explicit about whether “treatment” for mental disorders includes inpatient hospitalization, but it concluded that both the State’s interest in promoting patient autonomy and self-determination and the broad statutory terms support the conclusion that a health-care proxy agent can consent to such a voluntary placement if the principal does not object. Thus, in

---

175. See id.
176. See Mass. Gen. Laws Ann. ch. 190B, § 5-309(f) (West 2014). There remains some tension in the statutory provisions governing the voluntary hospitalization of adult wards in Massachusetts. The voluntary admissions statute contains language suggesting that guardians can apply for the voluntary commitment of their wards under certain circumstances. See Mass. Gen. Laws Ann. ch. 123, § 10 (West 2014) (“The application may be made . . . by the guardian of a person on behalf of a person under his guardianship.”); see also id. § 11 (suggesting that guardians have some authority to apply to have their wards voluntarily admitted).
177. See Mass. Gen. Laws Ann. ch. 190B, § 5-309(e) (providing that guardians may not revoke the ward’s health care proxy without court order and “[i]f a health care proxy is in effect, absent an order of the court to the contrary, a health-care decision of the agent takes precedence over that of a guardian”).
179. See id. at 718–23.
181. See Cohen, 760 N.E.2d at 720. The court pointed out that, under the Massachusetts statute, the principal is permitted to revoke the proxy decision maker’s authority even after a medical determination of incapacity has been made, and, in any event, the “principal’s wishes will always prevail over those of her agent, unless a judicial determination of her incapacity is obtained.” Id. at 722. In instances in which the principal does object to inpatient psychiatric treatment, the court was clear that the same interests in autonomy and self-determination exercised earlier in the proxy-granting document also operate in the withdrawal of the proxy decision maker’s authority to consent to the voluntary placement. See id. at 723–24. The court, however, left open the question of whether, under the Massachusetts statute, a judicial determination that the principal “lacks capacity to make
Massachusetts, a health care proxy agent may be authorized to consent to voluntary psychiatric hospitalization if the patient is non-objecting, but if the patient objects, the agent would have to obtain a court order or resort to the involuntary commitment process.182

This distinction between an objecting versus a non-objecting patient, which is key in Massachusetts for determining whether a health care proxy agent will be accorded the authority to consent to inpatient psychiatric treatment, is central as well to determining the authority of guardians of adult wards in a number of other states. In these states, the power of a guardian to authorize voluntary inpatient treatment turns on whether the ward has objected to the placement rather than on whether the ward has affirmatively provided informed consent. In Colorado, for example, a guardian who wishes to arrange for the voluntary admission of his or her ward must first demonstrate that the ward either agrees to or does not object to the placement. Even when the ward consents or fails to object, the guardian still must notify the court within ten days of placing the ward in an inpatient facility. If the ward does object, the guardian must utilize the state’s involuntary commitment process.183 In Ohio, a guardian may consent to voluntary psychiatric hospitalization unless the ward or another

health care decisions” should permit the proxy to provide substitute consent. Id. at 723. The court did indicate that, in the ordinary case, the proxy decision maker’s exercise of judgment should, in the first instance, be “in accordance with an ‘assessment’ of her ‘wishes,’ or, if her wishes are unknown, an ‘assessment’ of her ‘best interests.’” Id. at 721 (quoting MASS. GEN. LAWS ANN. ch. 201D, § 5).

As part of its discussion of this issue, the court reported on the approaches taken in a number of other states with advance directive/health care proxy statutes. In ten states, the law permits advance directives for mental health treatment but prohibits both voluntary and involuntary commitments under these directives. See id. at 718. Three states allow such directives but prohibit only involuntary commitment, and “[e]ight states allow commitment only if expressly authorized by the principal” in the proxy-granting document. See id. at 718–19. Four other states require that a separate document be executed in order to permit the health care proxy agent to consent to commitment. See id. at 719; see also Winick, supra note 24, at 57; cf. Dresser, supra note 24, at 777.


183. See COLO. REV. STAT. ANN. § 15-14-315(1)(d) (West 2014) (providing that a guardian may “[c]onsent to medical or other care, treatment, or service for the ward”). But see id. § 15-14-316(4) (providing limitation that “[a] guardian may not initiate the commitment of a ward to a mental health care institution or facility except in accordance with the state’s procedure for involuntary civil commitment”). Furthermore, section 13-14-316 makes clear that “[n]o guardian shall have the authority to consent to any such care or treatment against the will of the ward.” Id. If the ward consents, however, then his or her guardian appears to have authority to voluntarily admit the ward. See id. § 27-65-103(1) (“[A] ward . . . may be admitted to hospital or institutional care and treatment for mental illness by consent of the guardian for so long as the ward agrees to such care and treatment. Within ten days of any such admission of the ward for such hospital or institutional care and treatment, the guardian shall notify in writing the court that appointed the guardian of the admission.”).
“interested party” objects or the court has specifically withdrawn this authority.184

In a number of states, the statutes setting out the power of guardians to make decisions on behalf of their wards provide for the guardians’ authority to consent to health-care services, make housing and other residential decisions, and obtain other human services. Frequently, these statutes do not specifically address whether these general powers include the authority to consent to voluntary psychiatric hospitalization. States in this group include Alabama185 and West Virginia.186 Kentucky also falls within this general category, although guardians in Kentucky whose wards are developmentally disabled are expressly permitted to consent to voluntary hospitalization.187 In New Mexico, a guardian is authorized to consent to

184. See Ohio Rev. Code Ann. § 5122.02(B) (LexisNexis 2013) (authorizing guardian to provide consent: “[T]he application also may be made . . . on behalf of an adult incompetent person by the guardian or the person with custody of the incompetent person.”); id. § 2111.13(C) (“A guardian of the person may authorize or approve the provision to the ward of medical, health, or other professional care, counsel, treatment, or services unless the ward or an interested party files objections with the probate court, or the court, by rule or order, provides otherwise.”).

185. See Ala. Code § 26-2A-108 (2013) (conferring broad decision-making authority upon guardians: “[A] guardian of an incapacitated person is responsible for health, support, education, or maintenance of the ward . . . .”). In addition, this statute makes clear that the powers permitted to a guardian of an adult ward are the same as those of a guardian of a minor, which include the ability to “[c]onsent to medical or other professional care, treatment, or advice for the ward” and to “establish the ward’s place of abode within or without this state” (if consistent with court order). Id. § 26-2A-78(c)(2), (4) (detailing powers held by minor’s guardian); see also id. § 22-52-51 (confering upon guardians authority to make application to have wards admitted for psychiatric observation and diagnosis); id. § 22-52-55 (authorizing a guardian to request that his or her ward—who was voluntarily admitted at some earlier time—be discharged). Taken together, these statutes can reasonably be interpreted as conferring authority upon guardians to consent to the voluntary inpatient treatment of their wards.

186. See W. Va. Code Ann. § 44A-3-1(a) (LexisNexis 2014) (providing that guardians are “responsible for obtaining provision for and making decisions with respect to the protected person’s support, care, health, habilitation, education, therapeutic treatment, social interactions with friends and family”). Section 44A-3-1 does not specifically address the placement of wards in inpatient care or the requirement of their consent for medical/mental-health decisions. The state’s voluntary commitment statute, however, does suggest that an individual patient’s consent may be required and that a guardian cannot provide it. See 2013 W. Va. Acts ch. 128 (codified at W. Va. Code Ann. § 27-4-1); see also W. Va. Code Ann. § 27-4-4.

187. See Ky. Rev. Stat. Ann. § 387.065(3)(b) (West 2014) (allowing guardian to “[c]onsent to medical or other professional care, treatment, or advice for the ward”); see also id. § 387.660(3) (granting guardians authority to “give any necessary consent or approval to enable the ward to receive medical or other professional care, counsel, treatment or service”). Particularly relevant in this statute is a provision directing that a guardian must notify the court within thirty days if the guardian “places a ward in a licensed residential facility for developmentally disabled persons . . . .” Id. § 387.660(1). Another provision also provides that individuals with intellectual disabilities may be placed in an Intermediate Care Facility for Persons with Intellectual Disabilities by their guardians on a voluntary basis. See id. § 202B.021.
 voluntary admission for purposes of evaluation only.\textsuperscript{188} And in California, special statutory provisions governing wards with severe substance use disorders and others who are “gravely disabled” permit a conservator to provide consent for voluntary hospitalization,\textsuperscript{189} but the guardians of persons with mental illness or other mental disabilities that are chronic but not persistent are not so authorized.\textsuperscript{190}

Guardians in several other states are permitted to consent to voluntary hospitalization if they obtain specific authorization from the probate court that approved their appointment in the first instance or another court of appropriate jurisdiction. States in this category include Kansas\textsuperscript{191}

\begin{footnotesize}
\begin{itemize}
\item[188.] See N.M. STAT. ANN. § 45-5-312(B)(3) (LexisNexis 2014) (“[A] guardian may consent or withhold consent that may be necessary to enable the incapacitated person to receive or refuse medical or other professional care, counsel, treatment or service.”). However, a guardian may only present his or her ward for evaluation for inpatient mental health treatment. See id. § 43-1-14(B) (“A guardian appointed under the Uniform Probate Code, an agent or surrogate under the Uniform Health-Care Decisions Act or an agent under the Mental Health Care Treatment Decisions Act shall not consent to the admission of an individual to a mental health care facility. If a guardian has full power or limited power that includes medical or mental health treatment or, if the individual’s written advance health-care directive or advance directive for mental health treatment expressly permits treatment in a mental health care facility, the guardian, agent or surrogate may present the person to a facility only for evaluation for admission . . . .”).

\item[189.] See CAL. WELF. & INST. CODE § 6000(a)(1) (West 2014) (authorizing conservators to make application for voluntary admission: “[T]he application shall be made voluntarily by the person . . . or, if he is a conservatee with a conservator . . . by his conservator.”). “Any such person received in a state hospital shall be deemed a voluntary patient.” Id.; see also id. § 558(a) (“When ordered by the court after the hearing required by this section, a conservator appointed pursuant to this chapter shall place his . . . conservatee who is gravely disabled . . . in the least restrictive alternative placement, as designated by the court . . . . The placement may include a medical, psychiatric, nursing, or other state-licensed facility, or a state hospital, county hospital, hospital operated by the Regents of the University of California, a United States government hospital, or other nonmedical facility approved by the State Department of Health Care Services or an agency accredited by the State Department of Health Care Services, or in addition to any of the foregoing, in cases of chronic alcoholism, to a county alcoholic treatment center.”).

\item[190.] See CAL. PROB. CODE § 4652 (West 2014).

\item[191.] See KAN. STAT. ANN. § 59-3075 (2014) (conferring upon guardians general authority to consent to treatment when necessary). But, the statute makes clear that a guardian is not permitted “to place the ward in a treatment facility” unless the guardian obtains court authorization under the appropriate procedures. See id. § 59-3075(e)(9), see also id. § 59-3077 (outlining requisite procedures, referred to in section 59-3075(e)(9), for guardians to place wards in treatment facilities). One such procedural requirement is the provision of a lawyer for the incapacitated ward. See id. § 59-3077(c)(5). Kansas’s voluntary commitment statute confirms that a guardian must obtain court approval in order to consent to his or her ward’s placement in a mental health facility: “[I]f such person has a legal guardian, the legal guardian may make such application [for voluntary admission] provided that [ ] the legal guardian is [able] to obtain authority to do so pursuant to [Kan. Stat. Ann. §] 59-3077 . . . . then only in accordance with the provisions thereof.” See id. § 59-29b49; see also In re Guardianship & Conservatorship of Royston, 276 P.3d 838 (Kan. Ct. App. 2012) (confirming need for court order).
\end{itemize}
\end{footnotesize}
and Tennessee. Relatedly, a guardian in Mississippi may provide consent to inpatient psychiatric care but must first obtain the approval of the court and the facility director.

The legal standard (and criteria) governing the decision of a court to authorize a guardian to give consent to inpatient treatment often is not specified by statute. In practice, the standard is likely to be somewhat different from the standard (and associated criteria) governing involuntary civil commitment decisions in the relevant jurisdiction. In Arizona, however, the criteria for determining judicial approval of a guardian seeking inpatient mental health treatment for a ward are spelled out in some detail. Arizona Statute section 14-5312.01 elaborates the ability of a guardian to provide informed consent in the mental health context. It authorizes a guardian to consent to psychiatric treatment, provided it takes place outside a “level one” facility (i.e., a psychiatric hospital). For placement in a level one facility, the guardian must obtain a court order and comply with other notice requirements and procedural provisions. Before issuing such an order, the court must find on clear and convincing evidence that the ward is “incapacitated as a result of a mental disorder . . . and is likely to be in need of inpatient mental health care and treatment within the period of the authority granted . . . .” Moreover, the court

192. See Tenn. Code Ann. § 33-6-201 (2014) (listing decision makers who are eligible to consent to voluntary inpatient hospitalization where individual lacks capacity). This includes a conservator, provided the “court has expressly granted authority to apply for the person’s admission to a hospital or treatment resource for mental illness or serious emotional disturbance . . . .” Id. § 33-6-201(a)(3).

193. See Miss. Code Ann. § 41-21-103 (2013) (providing that guardians can make application for voluntary commitment of their wards, provided guardians obtain court authorization). In addition, to be admitted, the patient must meet the facility director’s requirements. See id. The relevant statutory language is as follows:

A person with an intellectual disability or with mental illness who is married or eighteen (18) years of age or older and who has a legal guardian or conservator may be admitted to a treatment facility upon application of his or her legal guardian or conservator if authorization to make the application has been received from the court having jurisdiction of the guardianship or conservatorship and the following has occurred:

(a) An investigation by the director that carefully probes the person’s social, psychological and developmental background; and
(b) A determination by the director that the person will benefit from care and treatment of his or her disorder at the facility and that services and facilities are available. The reasons for the determination shall be recorded in writing.

Id. § 41-21-103(3).

194. But see Doe v. Doe, 385 N.E.2d 996, 1001 (Mass. 1979). In Doe, the Supreme Judicial Court of Massachusetts held that the “best interests” standard, then applicable under the statute governing judicial authorization of a guardian to consent to inpatient treatment, required a showing of “likelihood of serious harm,” the same showing needed for involuntary civil commitment in the state. See id.


196. See id. § 14-5312.01(A).

197. Id. § 14-5312.01(B).
must consider “the cause of the ward’s disability and the ward’s foreseeable clinical needs” and must “limit the guardian’s authority to what is reasonably necessary to obtain the care required for the ward in the least restrictive treatment alternative.”198 Finally, the court must find that the evidence is supported by the opinion of a licensed mental health expert and must “limit the duration of the guardian’s authority to consent to inpatient mental health care and treatment and include other orders the court determines necessary to protect the ward’s best interests.”199

Once a guardian in Arizona has received judicial authorization to arrange inpatient treatment for the ward, a range of additional notice and procedural requirements comes into effect. Within forty-eight hours after placement, the guardian must give notice to the ward’s attorney.200 If requested by the attorney, the court must then hold a hearing within three days on the appropriateness of the placement.201 In addition, the treatment facility is directed to “assess the appropriateness of the ward’s placement every thirty days and [to] provide a copy of the assessment report to the ward’s attorney.”202 Within twenty-four hours after the treatment facility receives a written request from the ward seeking release, a change in placement, “or a change in the type or duration of treatment, the facility [must] forward this information to the ward’s attorney.”203 Finally:

The ward’s guardian shall place the ward in a least restrictive treatment alternative within five days after the guardian is notified by the medical director of the inpatient facility that the ward no longer needs inpatient care. The ward, a representative of the inpatient treatment facility, the ward’s attorney, the ward’s physician or any other interested person may petition the court to order the facility to discharge the ward to a least restrictive treatment alternative if the guardian does not act promptly to do so.204

This highly regulated system by which a guardian is permitted to obtain judicial approval for inpatient mental health treatment of a ward exists in Arizona alongside an equally restrictive judicial interpretation of the state’s guardian powers provisions as they relate to voluntary hospital admissions. In Pima County Public Fiduciary v. Superior Court for Pima County,205 the Court of Appeals of Arizona was faced with a statute then in place governing voluntary admissions that contained language that “[i]f

198. Id. § 14-5312.01(C).
199. Id. § 14-5312.01(B), (C).
200. See id. § 14-5312.01(D).
201. See id.; see also id. § 36-536 (entitling all persons facing court-ordered treatment to appointment of counsel).
202. Id. § 14-5312.01(E).
203. Id. § 14-5312.01(G).
204. Id. § 14-5312.01(I).
the person making voluntary application is under guardianship, the application shall be signed by the guardian.” 206 On the one hand, the court, relying on prior case law suggesting that an individual under guardianship should not be prevented from “performing the acts of which he is in fact capable,” 207 indicated that if such a person “is capable of making a decision to be admitted to a mental health treatment agency,” he or she should be permitted to do so, provided his or her guardian also signs the appointment. 208 This portion of the opinion is consistent with the statutory language expressly directing the guardian to sign the application for voluntary admission, although the court suggested that “voluntary” as used in the statute, “must refer to the voluntary act of the ward and not the voluntary act of the guardian.” 209

On the other hand, in the same opinion, the court went on to explain that a ward under guardianship is an “incapacitated person,” as defined under Arizona law, and therefore by definition is “not competent to make an application for admission to the state hospital and could not make a voluntary application” under the relevant statute. 210 A contrary conclusion, explained the court, would not be consistent with the requirements of due process and thus would not “pass constitutional muster.” 211 The court distinguished a California statute permitting a conservator to place his or her conservatee in a hospital, on the grounds that the California statute required prior judicial approval of the conservator’s authority and a determination that the conservatee was “gravely disabled.” 212 Absent such a prior judicial finding, explained the Arizona court, the decision of a guardian to approve inpatient treatment would not constitute an appropriate voluntary admission. 213 Thus, although the Pima County court’s analysis is somewhat ambivalent, the best understanding of the court’s position—and the conclusion most consistent with current statutory requirements—is that a guardian seeking inpatient treatment for his or her ward must obtain judicial approval, provide notice to the ward’s attorney, and follow the other procedural requirements designed to insure that the ward will be treated in the least restrictive placement available to meet his or her needs.

A. Legislative Balancing

The Arizona approach, which permits the hospitalization of an adult ward on the initiative of a guardian so long as substantive statutory criteria

207. Id. at 355 (quoting In re Sherrill’s Estate, 373 P.2d 353, 356 (Ariz. 1962) (in banc)).
208. Id.
209. Id.
210. Id. at 357.
211. Id. at 356.
212. See id. at 357.
213. See id.
are met and procedural requirements are satisfied, provides a sensible balance of the competing interests in this area. On the one hand, by requiring judicial approval of the placement based on clear criteria for decision, as well as notice to the ward’s counsel, the likelihood that a ward will be hospitalized unnecessarily over his or her objection is minimized. On the other hand, permitting incompetent adults to receive such care if the court finds that the statutory requirements have been met, even if they are unable to provide fully informed consent, insures that some individuals who may not meet the standard for involuntary commitment can still receive the benefits of inpatient treatment.

A similar balancing of interests is reflected in the approach adopted in New Hampshire. Section 464-A:25 of the state code sets out two alternative paths by which a guardian can arrange for inpatient care for a ward. Following the first path, a guardian “may admit a ward to a state institution with prior approval of the probate court if, following notice and hearing, the court finds beyond a reasonable doubt that the placement is in the ward’s best interest and is the least restrictive placement available.”

The second path permits the guardian to arrange an inpatient placement without prior court approval, but seeks to safeguard the ward’s autonomy, self-determination, and liberty interests by imposing a rigorous set of procedural protections to insure that the arrangement is monitored going forward. A guardian following the second path must obtain the written certification of a physician or psychiatrist and must submit written notice, with reasons for the placement, to the probate court within thirty-six hours of the inpatient admission. The probate court is obligated under the statute to evaluate the sufficiency of the notice by determining whether the admission is in the ward’s best interest and is the least restrictive placement available. The court then must appoint counsel for the ward and must notify the ward of that appointment. Counsel is given a limited period in which to file a report and request a hearing. If a hearing is scheduled, it is the guardian’s burden to show beyond a reasonable doubt that the inpatient admission is in the ward’s best interest and is the least restrictive placement available. Finally, the statute sets presumptive time limits on the inpatient hospitalization (no more than sixty days for any single admission and no more than ninety days in any twelve-month period), beyond which court approval is required, and authorizes the ward

217. See id. § 464-A:25(I)(a)(4) (giving counsel five days from date of appointment to file report and request hearing).
218. See id. § 464-A:25(I)(a)(5).
or the ward’s counsel to request a hearing “at any time” during the placement.\footnote{219}

In Oregon, guardians also are permitted to arrange inpatient mental health treatment without prior court approval, although they are required to provide notice to the court and to interested parties who may then request a court hearing. While Oregon’s guardian powers statute generally authorizes a guardian to “consent, refuse consent or withhold or withdraw consent to health care . . . for the protected person,” the statute makes clear that this power is subject to the State’s Health Care Decisions Act governing other health care agents.\footnote{220} In addition, the law sets out further limitations on a guardian’s authority with respect to inpatient admissions. It provides that “[b]efore a guardian may place an adult protected person in a mental health treatment facility, a nursing home or other residential facility, the guardian must file a statement with the court informing the court that the guardian intends to make the placement.”\footnote{221} The statute also imposes other procedural requirements, including the provision of notice to an enumerated list of interested persons with instructions on how they may object.\footnote{222} Once these notice requirements are met, “[t]he guardian may thereafter place the adult protected person in a...
mental health treatment facility, a nursing home or other residential facility without further court order.”

In Minnesota, the Uniform Probate Code, as adopted, charges guardians with the duty to consent to “necessary medical or other professional care, counsel, treatment, or service . . . [other than] psychosurgery, electroshock, sterilization, or experimental treatment of any kind” that a ward may require. The statute further provides, however, that a ward “may not be admitted to a regional treatment center” (i.e., a state mental health treatment facility) by the guardian except following a hearing under the state’s involuntary civil commitment law or for outpatient services.

In the mid-1990s, a Minnesota Supreme Court Task Force was charged with considering how to manage persons who, though incompetent to provide informed consent, were “not resisting the proposed [inpatient] treatment.” The Task Force recommended that “a new option, other than Civil Commitment, should be available for persons who are in need of mental health treatment, not resisting treatment, but [who] are incompetent to give informed consent to treatment or admission.” Specifically, the Task Force proposed, and the state legislature adopted, a system by which substitute consent can be provided, in the first instance by an individual designated as a health care power of attorney through an advance directive, or in the alternative by the local county mental health authority or its designee. By further statutory amendment in 2001, the state legislature authorized courts to appoint a different “substitute decision maker” if no health care power of attorney has been named and if the designated local agency is unable or unwilling to provide consent.

Presumably, under this provision, a guardian may receive court appointment as a “sub-

manner provided by [Or. Rev. Stat. §] 125.075, the court shall schedule a hearing on the objection as soon as practicable.

(f) The requirement that notice be served on an attorney for a protected person under paragraph (c)(A) of this subsection does not impose any responsibility on the attorney receiving the notice to represent the protected person in the protective proceeding.

Id.

223. Id. § 125.320(3)(e).


225. See id. § 524.5-313(c)(1).


227. Id. (quoting Minn. Civil Commitment Task Force, supra note 226, at 35) (internal quotation marks omitted).

228. See id.; see also Minn. Stat. Ann. § 253B.04, subd. 1a.

stitute decision maker” and thereby obtain the authority to consent to inpatient treatment for his or her ward.230

Neither the voluntary admissions statute nor the special provision governing substitute decision making sets out the duties of the entity—county agency, family member, or guardian—designated as the substitute decision maker.231 The statute does, however, provide criteria for determining whether such authority should be conferred by the court, as well as procedures for overseeing the ward’s inpatient admission. Thus, in cases where no health care power of attorney has been named in an advanced directive and in which the local mental health agency or its designee has not provided consent, “the person who is seeking treatment or admission, or an interested person acting on behalf of the person, may petition the court for appointment of a substitute decision maker who may give informed consent for voluntary treatment and services.”232 In evaluating such a petition, the court is directed to consider two criteria: first, whether “the person demonstrates an awareness of the person’s illness, and the reasons for treatment, its risks, benefits and alternatives, and the possible consequences of refusing treatment;” and second, “whether the person communicates verbally or nonverbally a clear choice concerning treatment that is a reasoned one, not based on delusion, even though it may not be in the person’s best interests.”233 Once the authority of a substitute decision maker to approve an inpatient admission is granted, “the person [the ward in the case of a guardian] or any interested person acting on the person’s behalf may seek court review within five days for a determination of whether the person’s agreement to accept treatment or admission is voluntary.”234 Moreover, the statute requires that the patient be informed in writing that he or she may leave the facility within twelve hours of making such a request.235

230. The Minnesota Voluntary admissions statute clearly contemplates that a guardian may receive court authorization to provide substitute consent for voluntary admission. See id. § 253B.04, subd. 1(c)(2) (“Legally valid substitute consent may be provided by a . . . guardian or conservator with authority to consent to mental health treatment . . . .”).
231. See Halverson, supra note 4, at 183–84.
233. Id. § 253.04, subd. 1a(b)(1)–(2). These are the same criteria for decision that the statute sets out to guide the decision of the designated local mental health agency. See id.; see also id. § 253.04, subd. 1b (noting same criteria should be used to evaluate petition for a substitute decision maker as laid out for evaluating designated mental health agencies in subdivision 1a(b)).
234. Id. § 253B.04, subd. 1a(c). This statutory right to seek judicial review of the voluntariness of a patient’s agreement to an inpatient placement applies both to admissions approved by a substitute decision maker and to admissions in which the patient himself or herself provided consent. See id.
235. See id. § 253B.04, subd. 2. Under this same section, an individual receiving inpatient treatment for chemical dependency is entitled to release within seventy-two hours of making such a request. See id.
B. Managing the Tension Between Incompetency and Partial Capacity

The ambivalence exhibited by the Arizona appellate court in *Pima County*, with respect to whether an adult under guardianship might possess sufficient agency to make a voluntary decision to enter inpatient treatment, is instructive. If it were to turn out that some wards are capable of participating in this sort of decision making—either by providing actual consent or by communicating agreement with a guardian’s decision short of fully informed consent—then perhaps the reluctance of many jurisdictions to permit the voluntary hospitalization of adults under guardianship might be ameliorated. Some writers take the position that all wards admitted for inpatient treatment by a guardian should be regarded as entering involuntarily, presumably because the very fact of their being under guardianship requires a judicial finding that the ward is incompetent.  

Ultimately, the Arizona court in *Pima County* seemed to settle on this position when it concluded that the decision of the guardian to approve inpatient treatment could not constitutionally be permitted under the state’s voluntary admissions statute, because the ward’s prior adjudication as an “incapacitated person” necessarily constituted a judicial finding that he lacked “sufficient understanding or capacity to make or communicate responsible decisions concerning his person.”  

Earlier in the opinion, however, the *Pima County* court appeared to recognize that a finding of legal incompetency, even if based on expert evidence of the ward’s impaired decision-making ability, need not be treated the same as a determination that the individual entirely lacked “clinical (de facto) competence or decision-making ability.”Acc to this understanding, even a ward with a judicially appointed guardian should not be prevented from “performing the acts of which he is in fact capable.” Moreover, said the court, if such an individual were to agree to inpatient treatment, that agreement would be “the voluntary act of the ward and not the voluntary act of the guardian.”

In order to make some sense of this tension between the categorical legal determination of incompetency often associated with the appointment of a general guardian and the partial functional understanding of capacity or incapacity more often employed by clinicians in practice, some
attention should be directed to the various legal tests that courts use in making guardianship determinations.241 One long-standing approach focuses on whether the individual is capable of making decisions that produce reasonable outcomes. State statutes that reflect this approach typically require a showing that the person is not capable of caring for himself or herself, or providing for his or her family, in order for a guardian to be appointed.242 A second approach, now followed in many states, was developed as part of the UPC. This approach shifts the court’s focus from reasonable outcomes to the regularity of the individual’s decision-making process itself.243 Under the UPC approach, a person with mental disabilities may be adjudged incompetent if shown to be impaired in the ability to engage in the cognitive process of rationally weighing the risks and benefits of competing choices.244 It was this standard, ultimately, that the Pima County court relied on in determining that an adult under guardianship in Arizona could not voluntarily elect inpatient treatment.245

The UPC process-based approach has been criticized for the subjective judgment it requires courts to make in determining whether an individual’s thought processes are “rational” and for its consequent susceptibility to abuse.246 In response to this critique, in recent years, some states have developed yet a third approach, which shifts the focus away from process and back toward a consideration of outcomes. Unlike the reasonable outcomes approach, however, this new “functional approach” directs the court’s analysis to specific decision-making tasks that an individual might be called upon to undertake with respect to housing, health care, and the like, and requires the individual’s dysfunction to be demonstrated through specific evidence indicating imminent risk.247

Clearly, the decision of a court to find an individual legally incompetent and in need of a guardian can carry a range of different meanings, depending upon which of these measures of incompetency the court has employed. Thus, an adult adjudicated incompetent under the functional approach because he or she suffers from a delusion specific to one important choice—say the decision whether to undergo treatment with psychotropic medications—but whose mental processes are otherwise not globally impaired, stands in a very different position than does another individual found by a court to be incompetent under the UPC approach because he or she is broadly impaired in his or her capacity to engage in a practical reasoning process. More to the point, the assumption that an individual under guardianship is categorically incapable of usefully con-

242. See id. at 743; see also Brakel et al., supra note 3, at 371.
244. See Slobogin, Rai & Reisner, supra note 3, at 946.
245. See Pima Cnty., 546 P.2d at 357.
246. See Slobogin, Rai & Reisner, supra note 3, at 946–47.
tributing to his or her guardian’s decision making with respect to inpatient treatment fails to take seriously the various distinct competencies that may (or may not) be impacted by even severe mental illness or other significant mental disabilities, as well as the different degrees of relevance that each of these competencies holds to such a decision.248

The substantive standards adopted in Arizona, New Hampshire, and Minnesota for judicial authorization or review of a guardian’s decision to arrange inpatient treatment, together with the notice and other procedural requirements reflected in these states’ laws, provide useful tools for evaluating and supervising this work.249 In reviewing these statutes, certain basic policy choices stand out. First, should judicial authorization be required in advance of any decision to place an adult under guardianship in an inpatient mental health setting, or is some combination of the patient himself or herself, the guardian, and involved mental health professionals a sufficient team for making the initial determination? Second, regardless of whether the determination is made by clinicians or by judicial actors, what level (or kind) of functional competency should be required of a patient who either seeks to provide consent or at least to assent to a guardian’s decision to elect inpatient treatment?250 And third, how should the law deal with adults under guardianship who withhold consent or perhaps even communicate a desire not to be admitted to an inpatient setting?

With respect to the first question, many of the same concerns identified by those who challenge the general preference for voluntary inpatient admission over involuntary commitment are relevant to an approach that permits guardians to elect inpatient treatment for their wards without a prior judicial hearing. Thus, worries over the potential for abuse, coercion, and the lack of an adversary process are all present in a system that permits guardians to provide consent without first being required to demonstrate through a judicial proceeding that such a placement is in the ward’s interest and is necessary for his or her well-being.251 On the other side, some argue that a requirement that the guardian obtain formal judicial approval in advance would be costly and unnecessary, particularly in jurisdictions in which the guardian’s judgment as to the necessity and utility of inpatient treatment must be supported by one or more mental health professionals, often through the submission of a written assessment to that effect.252 In addition, it is argued, a guardian has “already been

248. See Roca, supra note 2, at 1191.
249. See supra notes 195–219, 224–35 and accompanying text.
250. On the distinction between “assent” and “consent” to admission or treatment, see Stone, supra note 4, at 32 (explaining that assent “requires acquiescence, a tacit acceptance, or non-response such as silence,” while consent “requires a competent patient’s active and voluntary acceptance of a prescribed course of treatment”).
251. See Halverson, supra note 4, at 166; Stone, supra note 4, at 33.
appointed by a court [as] an appropriate person to provide informed consent," and thus should not be obliged to obtain “additional judicial review" for the hospitalization decision.\footnote{253} Finally, advocates for this position point out that the costs associated with delaying the provision of treatment are likely to outweigh the benefits of preventing what is likely to be a low potential for abuse.\footnote{254}

In light of these competing considerations, Donald Stone has concluded that a “middle ground” may well suffice to guard against the dangers of coercion or abuse he perceives as “common in psychiatric hospitals[\ldots]” inpatient admissions practices.\footnote{255} This middle ground does not mandate a prior judicial hearing, but it does require that the individual be provided counsel as a condition for being voluntarily admitted. Under Stone’s proposal, which would apply to all voluntary admissions and not just those approved by a guardian, the attorney would have an obligation to interview the individual and to conduct a thorough investigation. If, after this interview and investigation, the attorney were to conclude that his or her client “is capable of knowingly and voluntarily admitting himself [or herself] into the hospital, the patient should be permitted to exercise this option without judicial review.”\footnote{256}

Critics of Stone’s middle ground approach point out that attorneys “typically lack prior knowledge of the patient or the patient’s social or medical history,” and moreover that “a competency determination is hardly something that is within an attorney’s professional knowledge.”\footnote{257} These concerns may or may not be accurate, depending in part on whether the jurisdiction maintains a cadre of specialized attorneys with subject matter expertise and, at least for chronic patients, some prior relationship with their clients. In any event, the essential advantages of Stone’s proposal might be realized through a state law process something like that in place in New Hampshire and Oregon. Recall that in New Hampshire, a guardian is permitted to arrange an inpatient placement without prior court approval, but the guardian must obtain the written certification of a physician or psychiatrist and must submit written notice, with reasons for the placement, to the probate court within thirty-six hours of the inpatient admission.\footnote{258} This permits a guardian to arrange inpatient care without delay, but sets into motion a legal process that reduces the potential for abuse. The New Hampshire approach imposes some supervisory duties on the probate court, which must review the guardian’s filing in order to determine whether the admission is in the ward’s best interest and is the least restrictive placement available. As with Stone’s
approach, counsel must be appointed for the ward. Within a relatively short window of time, the lawyer must then interview the client and conduct an investigation, leading to the filing of a report and a possible hearing. If such a hearing is scheduled, the court must find beyond a reasonable doubt that the inpatient admission is in the ward’s best interest and is the least restrictive available.

On balance, this sort of hybrid approach, which permits the guardian to approve inpatient hospitalization but also triggers a judicial review process along with attorney consultation and investigation, appears to strike a sensible balance between the concerns present in this area. Even if such a system were put in place, however, a further question arises as to the kind of functional competency that should be required of a patient whose guardian has arranged inpatient treatment. Some jurisdictions, of course, require informed consent before a voluntary placement is permitted, even for individuals under guardianship. Stone defines such consent as requiring “a competent patient’s active and voluntary acceptance of a prescribed course of treatment following . . . full disclosure of associated risks and benefits . . . .” At the other extreme, some argue that mere “assent to retention by the facility” should suffice, in which case “acquiescence, a tacit acceptance, or non-response such as silence” could be a sufficient basis for permitting the inpatient admission. The best approach is one in which some reasonable assessment of the functional capacity of the ward to contribute to the decision-making process is used to determine the level of involvement required to permit the guardian to arrange an inpatient admission. Such a sliding scale of competency could be implemented as a way of operationalizing a best interests standard for approval of the guardian’s decision.

As noted above, in Arizona, the substantive criteria governing the decision of a court to authorize a guardian to approve inpatient treatment include whether the ward is “incapacitated as a result of a mental disorder . . . and is likely to be in need of inpatient mental health care and treatment within the period of the authority granted,” as well as “the cause

259. See N.H. REV. STAT. ANN. § 464-A:25(I) (a) (3).
260. See id. § 464-A:25(I) (a) (4), (5).
261. See, e.g., 405 ILL. COMP. STAT. ANN. 5/3-400 (West 2014) (setting out consent requirement for voluntary admission in Illinois Mental Health and Developmental Disabilities Code); N.Y. MENTAL HYG. LAW § 9.17 (McKinney 2014) (governing requirement of informed consent); see also In re Gardner, 459 N.E.2d 17, 20 (Ill. App. Ct. 1984) (holding that individuals who lack capacity to provide informed consent for inpatient treatment cannot be admitted by their guardians).
262. Stone, supra note 4, at 32.
263. See id.
of the ward’s disability and the ward’s foreseeable clinical needs.”

While these criteria are appropriate considerations for purposes of monitoring the decision of the guardian to elect inpatient hospitalization, they do not provide a basis for assessing whether the ward has contributed sufficiently to this decision in light of his or her actual ability to do so. Importantly, however, the Arizona statute also requires the court to “limit the duration of the guardian’s authority to consent to inpatient mental health care and treatment and include other orders the court determines necessary to protect the ward’s best interests.” This additional provision, centered on a best interests analysis, does begin to offer some basis for evaluating the capacity of the adult under guardianship to play a role in the decision-making process. Similarly, in New Hampshire, if a judicial hearing is held, it is the guardian’s burden to show beyond a reasonable doubt that the inpatient admission is in the ward’s best interest and is the least restrictive available.

A best interests requirement can be understood as: (1) an injunction to the decision maker to select a course of action based upon his or her own judgment of the best outcome for the subject, all things considered; (2) a requirement that the decision maker base the choice on an assessment of what the subject would want if he or she were able to make a competent decision; or (3) a requirement that the decision maker develop a choice that is consistent with the long-term values and commitments of the subject as revealed or expressed through conduct or prior statements.

In the Doe case, the Massachusetts Supreme Judicial Court construed the state’s then-existing statutory best interests requirement to contain an additional related feature, which is of particular assistance in thinking about the question of the ward’s capacity to participate in the decision-making process itself. The court explained that, in cases in which the ward is capable of formulating or expressing a preference, the ward’s “stated preference” with respect to the proposed inpatient admission “must be treated as a critical factor in the determination of his ‘best interests.’” Moreover, explained the court, this reading of the role that the ward’s expressed preferences should play in evaluating his best interests should apply notwithstanding the fact that the ward “failed to understand his mental condition and his need for treatment . . . .”

Returning to the four domains of competency identified by Appelbaum and Roth, we can now begin to construct a matrix for guiding the

---

266. See id. § 14-5312.01(C).
268. See Gutheil & Appelbaum, supra note 1, at 226; see also Frolik & Whitton, supra note 264, at 751–52.
270. See id.
271. For a discussion of the Appelbaum and Roth criteria, see supra notes 114–39 and accompanying text.
evaluation of the best interests of an adult under guardianship in light of his or her functional capacity to contribute to a decision to seek inpatient treatment. The first, most basic competency is the capacity to evidence a choice. While the number of adults under guardianship who are entirely unable to communicate their agreement or disagreement is likely to be small, such individuals do present from time to time. Clearly, these wards are not capable of contributing in any fashion to the decision to seek inpatient treatment. For those who do not meet the statutory standard for involuntary commitment because there is insufficient evidence of danger to others or to themselves, these individuals will be foreclosed from obtaining the advantages of inpatient care if their guardians are not permitted to provide consent on their behalf. While some argue that these individuals should receive care in the community as outpatients if they are not civilly committable, the reality in most localities is that adequate outpatient services often are unavailable or nonexistent. The position of the Massachusetts court in *Doe* with respect to such individuals is that a doctrine of substituted judgment may be appropriate. Of course, the basis for a substituted judgment can be relatively “objective” and based on the guardian’s own views about the costs and benefits of inpatient admission, or more closely tailored to the choice that the ward would have made were he or she able to do so. A properly devised statutory scheme in which the guardian’s judgment is subject to subsequent judicial scrutiny according to sensible substantive criteria—like those articulated in the Arizona statute—together with a rigorous requirement that the ward be placed in the least restrictive setting appropriate to his or her needs, is likely to be the most effective system for insuring that a process of substituted judgment is implemented appropriately.

A far greater number of adults under guardianship are likely to have the ability to evidence a choice and some further functional capacity within the second category of competency described by Appelbaum and Roth. These individuals may possess the cognitive capacity to understand the basic facts of their condition and of the proposed institutional placement, but may not be able entirely to comprehend the full import of those facts or the consequences of the choice that is being made by the guardian on their behalf to approve inpatient treatment. Whether these

275. See *Doe*, 385 N.E.2d at 1000.
278. As noted in the text, some writers take the not unreasonable position that patients who are incompetent in this sense should be admitted only as involuntary patients. Given the competing interests at stake, however, I have concluded that a system of judicially supervised voluntary admissions is the preferable approach.
279. See Appelbaum & Roth, *supra* note 114, at 952–56.
individuals’ cognitive capabilities are limited to a basic understanding of essential information or extend to an “actual understanding” of those facts and their likely consequences, their expressed preferences should be accorded some weight in the evaluation of their best interests. The more cognitive facility they demonstrate, the greater deference a reviewing court should accord their views. Thus, a guardian who is acting consistent with the articulated preferences of a ward with some cognitive purchase on the circumstances of the decision should be more likely to receive court approval than a guardian whose ward has been unable to engage even in that minimal level of participation in the decision-making process. This should be the case even in many instances in which the ward’s thinking includes some distorted or false beliefs, although the more centrally those false ideas implicate the very decision to be made, the more problematic they will be as counting in favor of the decision being approved by the court.

The next category of functional capacities identified by Appelbaum and Roth relates closely to the UPC’s approach to defining the line between competency and incompetency. This third category goes to the individual’s ability to undertake a “rational” process of deliberation. In the context of a decision to approve inpatient treatment, this sort of rational deliberation would include consideration of the relative advantages and disadvantages of inpatient care and of reasonably available alternatives. As noted earlier, an assessment of rationality may be difficult if the ward assigns values to the competing costs and benefits that are wildly divergent from those identified by others evaluating his or her competency. Here again, if the ward holds false beliefs that are directly relevant to his or her calculation of the competing options, the weight a reviewing court should accord his or her conclusions ought to be reduced accordingly. Nevertheless, an individual who demonstrates a basic ability to assess relevant information in order to generate a preference that reflects his or her essential values ought to be treated as a significant partner in the decision-making process, and, if that preference is consistent with the judgment of the guardian, it ought to shore up considerably the claim of the guardian to be acting in the ward’s best interests.

The final area of competency identified by Appelbaum and Roth is the ability to “appreciate[e] the nature of the situation.” This capacity—to apply information to new circumstances in order to formulate a plan of action or a solution to a problem and to engage that process of application in both the affective and cognitive domains—is the most de-

280. See id.
281. See UNIF. PROBATE CODE art. 5 (2010).
282. See Appelbaum & Roth, supra note 114, at 954.
283. For a discussion of the differences between the process-based and outcomes-centered approaches to competency, see supra note 139 and accompanying text.
284. See Appelbaum & Roth, supra note 114, at 954–56.
manding of the measures of competency in Appelbaum and Roth’s typography of functional abilities. It is not inconceivable that an individual with mental illness or other mental disability may have this capacity, even though he or she may also harbor distorted or false beliefs. While the presence of false beliefs may diminish somewhat the weight given to his or her judgment, the fact that an individual is able to appreciate in this rich sense the circumstances surrounding his or her hospitalization should count heavily in the court’s assessment. If an adult under guardianship possessed of this sort of appreciation of his or her situation expresses a preference that is in accord with the judgment of the guardian, the claim that the consent is voluntary is substantially strengthened.

Of course, a ward may also express either clear or implied opposition to a guardian’s plan to arrange inpatient care. In such cases, the tension between the judgment of the guardian and the wishes of the ward creates real problems in terms of the coherence of treating such an admission as voluntary. One could take the position that anything short of affirmative consent, or at least passive assent, renders the guardian’s choice in favor of inpatient treatment involuntary and therefore necessarily implicates the civil commitment standards and procedures that govern all involuntary inpatient commitments. The problem with this move, as noted earlier, is that it may make it impossible to provide inpatient treatment to a non-cooperating ward whose mental illness or other disability does not create a sufficient danger to self or others to satisfy the involuntary commitment standard. In the best of all worlds, this gap would be closed through the provision of clinically appropriate outpatient services. In most jurisdictions, however, the network of community-based mental health resources is less than fully adequate to meet this need. Arguably, a jurisdiction’s failure to provide a complete continuum of outpatient care should not drive the development of a distorted doctrinal approach to voluntary admissions. At least for adults under guardianship who are capable of expressing a preference and who indicate that they oppose an inpatient placement, therefore, important interests in patient autonomy and self-determination ought to counsel in favor of defaulting to the rigorous requirements of the involuntary civil commitment regime.

V. Conclusion

The conventional view is that there are two methods by which a guardian, family member, or other proxy decision maker might make a judgment on behalf of an impaired individual. First, the proxy decision maker could reach a decision based on the best interests of the ward. Ordinarily, this approach is characterized as “neutral or objective” because it is not based on the ward’s particular viewpoint regarding the matter for

---

286. See Stone, supra note 4, at 37.
287. See GUTHEIL & APPELBAUM, supra note 1, at 226–27.
decision. In truth, however, a best interests assessment is still a judgment resting on a particular perspective, precisely because it involves the proxy decision maker acting on the basis of his or her own best judgment as to the ward’s well-being. The alternative approach is more frankly subjective, in that it requires the proxy decision maker to determine what the ward would have wanted in the situation had he or she been able to decide for himself or herself.

For adults under guardianship facing the prospect of an inpatient admission, a better approach than either of these conventional alternatives standing alone is to integrate these formally distinct ways of thinking about proxy decision making, so that the best interests of the ward are construed to include a consideration of the ward’s likely preferences as well as a more objective determination of the advantages and disadvantages of the proposed inpatient placement. Taken in this fashion, a best interests determination would subsume an assessment of clinical factors with respect to the ward’s amenability to treatment, suitability for alternative outpatient care, and the like, as well as a consideration of the ward’s preferences expressed at the moment or in the past.

Statutes governing the “voluntary” admission of adults under guardianship should be revised to require that proxy decision makers and reviewing courts employ this modified conception of best interests. They should also contain procedural requirements that, while acknowledging the guardian’s authority to make an initial determination, insure effective subsequent judicial supervision of the inpatient placement, in order to guarantee that the ward’s needs are met in the least restrictive setting appropriate and in a fashion that is most consistent with the ward’s capacity for self-determination.

288. See Weiner & Wettstein, supra note 96, at 290.
290. See Weiner & Wettstein, supra note 96, at 291.
INSURING FLOODS: THE MOST COMMON AND DEVASTATING NATURAL CATASTROPHES IN AMERICA

CHRISTOPHER C. FRENCH*

I. INTRODUCTION

Along with houses and lives, Hurricanes Katrina and Sandy washed away the illusion for yet another generation of Americans that seawalls and levees can keep us safe from flood waters. Although the storm surge damage caused by Hurricanes Katrina and Sandy is indelibly etched upon our minds due to the 24-hour media coverage and haunting images of houses in ruins and people sitting on their roofs waiting to be rescued, smaller-scale flooding disasters that do not get national attention occur throughout the country every day.

Flooding is the most common catastrophe in America and the world, accounting for approximately 90% of all catastrophic losses annually. For the past thirty years in America, floods have on average caused more than $8 billion in damages annually, and most of the damage is uninsured. Hurricanes Katrina and Sandy alone caused over $160 billion in damage. Yet, only about 10% of the victims of Hurricane Katrina had flood insurance, and a little over 50% of Hurricane Sandy victims had flood insurance even though the communities devastated by these hurricanes are ones that obviously should be concerned about flooding because they are located on or near rivers and the coast. Even more sadly, much of the losses suffered by even the relatively fortunate few who had flood insur-

* Christopher C. French is a Visiting Assistant Professor at Penn State Law School; J.D., Harvard Law School; B.A., Columbia University. The Author gratefully acknowledges the legal research contributions of Pat Roncevich and Kyle Yevensky to the Essay, as well as the thoughtful comments provided by Ken Abraham, Jay Feinman, David Kaye, Kit Kimport, Leo Martinez, Stephen Ross, and the participants in a workshop held at Penn State Law School on earlier drafts of the Essay.


ance through the National Flood Insurance Program (NFIP) were not actually covered by such insurance.

Despite dramatic nationwide media coverage of major flooding events in large population centers such as New Orleans and New York, flooding is actually a problem throughout the entire country. Floods impact people in every state. And, floods will likely only become a worse problem, as climate change could increase water levels and create more devastating hurricanes and other storms. Yet, it is estimated that only about 7% of all homeowners in America have insurance for flood losses.4

The NFIP is the only significant source of flood insurance available to homeowners in America, and the program has numerous deficiencies. Every time there is a major flooding event, Congress revisits and attempts to address the deficiencies that are exposed by the most recent flood. This pattern has been repeated multiple times over the past forty-six years since the NFIP was created in 1968. And each time, the result is the same: most homeowners remain uninsured for flood losses and the insurance that is available to cover flood losses is inadequate.

If the way flood losses are insured (or uninsured) in America is not working, is there a better way? This Essay maintains that the answer is “yes.” The simple, but underappreciated, solution offered in this Essay is the elimination of the flood exclusion that is found in standard form “all risk” homeowners insurance policies.

How is it even possible that the biggest and most common disaster that can impact one’s home—flooding—is not already covered under “all risk” homeowners insurance? There are three principal theoretical bases that insurers historically have used to justify excluding coverage for flood damage under homeowners policies: (1) adverse selection, (2) moral hazard, and (3) correlated risks. Adverse selection is the idea that only the people most likely to have losses will purchase insurance.5 Moral hazard theory posits that a person who has insurance is less likely to take care to avoid losses due to the fact that insurance will cover the losses.6

4. See infra note 69 and accompanying text.

5. Adverse selection is “the disproportionate tendency of those who are more likely to suffer losses to seek insurance against those losses.” Kenneth S. Abraham & Lance Liebman, Private Insurance, Social Insurance, and Tort Reform: Toward a New Vision of Compensation for Illness and Injury, 93 COLUM. L. REV. 75, 102 n.82 (1993); see also Tom Baker, Containing the Promise of Insurance: Adverse Selection and Risk Classification, 9 CONN. INS. L.J. 371, 373, 375 (2003). Some critics of the concept of adverse selection have argued that insurers’ alleged concerns regarding the impact that adverse selection actually has on policyholders’ behavior are overblown. See, e.g., Peter Siegelman, Adverse Selection in Insurance Markets: An Exaggerated Threat, 113 YALE L.J. 1225 (2004).

lated risks are risks that may result in numerous losses concentrated in space and time. Due to the unpredictable nature of when correlated risks will result in losses, insurers historically have claimed that they cannot calculate and charge actuarially sound premiums for such risks.

When a risk simultaneously involves significant concerns related to adverse selection, moral hazard, and correlated risks, insurers historically have concluded the risk is either uninsurable or one they do not want to insure. Flooding is such a risk according to conventional insurance wisdom. First, only the people most likely to suffer flood losses will purchase insurance that covers only floods. Thus, the pool of insureds that would pay premiums to cover flood losses would be relatively small, and the premiums that insurers would need to charge in order to cover the potential losses would need to be extremely high, assuming actuarially-sound premiums could be calculated at all. Second, under moral hazard theory, a homeowner who has insurance against flooding would not bother to avoid or minimize flood losses or, at best, would take fewer measures to avoid or minimize flood losses. Third, flooding is a correlated risk because the losses associated with flooding events tend to occur in concentrated geographic locations at or about the same time.

the concept by stating that “[o]nce a person has insurance, he will take more risks than before because he bears less of the cost of his conduct.” W. Cas. & Sur. Co. v. W. World Ins. Co., 769 F.2d 381, 385 (7th Cir. 1985). The term “moral hazard” also generally encompasses situations where “[a] person . . . deliberately causes a loss . . . [or] exaggerates the size of a claim to defraud an insurer.” Mark S. Dorfman, Introduction to Risk Management and Insurance 480 (8th ed. 2005). Numerous scholars have written on moral hazard and have offered similar descriptions of the concept. See, e.g., Scott E. Harrington, Prices and Profits in the Liability Insurance Market, in Liability: Perspectives and Policy 47 (Robert E. Litan & Clifford Winston eds., 1988) (“Moral hazard is the tendency for the presence and characteristics of insurance coverage to produce inefficient changes in buyers’ loss prevention activities, including carelessness and fraud . . . .”); Robert H. Jerry, II & Douglas R. Richmond, Understanding Insurance Law 12 (5th ed. 2012) (“[T]he existence of insurance could have the perverse effect of increasing the probability of loss. . . . This phenomenon is called moral hazard.”); George L. Priest, The Current Insurance Crisis and Modern Tort Law, 96 Yale L.J. 1521, 1547 (1987) (“Moral hazard refers to the effect of the existence of insurance itself on the level of insurance claims made by the insured. . . . Ex ante moral hazard is the reduction in precautions taken by the insured to prevent the loss, because of the existence of insurance.”); Gary T. Schwartz, The Ethics and Economics of Tort Liability Insurance, 75 Cornell L. Rev. 313, 338 n.117 (1990) (“‘Moral hazard’ is sometimes distinguished from ‘morale hazard,’ the former referring to deliberate acts like arson, the latter to the mere relaxation of the defendant’s discipline of carefulness.” (citing C. Arthur Williams, Jr. & Richard M. Heins, Risk Management and Insurance 217 (4th ed. 1981))).

Although this Essay does not purport, or even attempt, to address the diverse political, environmental, and engineering issues that underlie the problem of how we can or should prevent or minimize flooding in America, it does address the problem of insuring against the risk of flood losses. It does so by questioning the soundness of the three historical justifications—adverse selection, moral hazard, and correlated risks—that have been used to explain why insurers should be allowed to refuse to cover flood losses under homeowners policies. This Essay concludes that, even if they were once sound, the foundations for these justifications are no longer solid. If the theoretical justifications for insurers’ refusal to insure flood losses are unsound, then a potential solution to America’s flood insurance problem emerges: the elimination of the flood exclusion in homeowners policies.

Homeowners insurance is a classic example of a situation where the pool of insureds across which the risk of loss is spread is so large that adverse selection is of little concern. Ninety-six percent of homeowners buy homeowners insurance because, in addition to risk aversion, the purchase of such insurance is effectively mandatory in America today. One cannot get a mortgage that is federally guaranteed, which most are, without homeowners insurance. Consequently, if flood losses were covered by homeowners insurance, then the pool of insureds across which the risk of loss would be spread would be so enormous that adverse selection regarding the purchase of insurance to cover flood losses would not be an issue.

Similarly, moral hazard concerns regarding flood losses are relatively low because of the enormous disruption and inconvenience that flooding causes a homeowner. Flooded homeowners not only lose irreplaceable items of sentimental value, but they are often rendered homeless while their houses are remediated and repaired. Thus, people already have incentives to avoid flood losses, regardless of whether they have insurance. Consequently, it stands to reason that, instead of putting the impetus on individual homeowners to prevent flood damage to their homes by excluding coverage for floods in homeowners policies, a more efficient means of avoiding or preventing flood losses would be through building codes and land development restrictions.

The potentially disastrous financial consequences associated with the correlated risk of flood losses for insurers also would be reduced, if not entirely eliminated, if homeowners policies covered floods. If coverage for flood losses were bundled together with the other risks of loss covered under homeowners insurance, then the risk of a flooding event causing an enormous drain on insurers’ capital reserves, that could result in insolvencies, would be minimal because the capital reserves generated by the premiums of 69 million homeowners with diverse risk profiles would be enormous instead of using just the premiums generated by the 5.5 million homeowners who currently are insured under the NFIP and primarily live
in known flood areas. In addition, through global reinsurance and catastrophe bonds, an insurer’s risk of insolvency due to flood events would be further reduced because each individual insurer’s risk can be spread to other insurers and investors.

This Essay addresses the problem of insuring flood losses in three parts. Part II provides a brief discussion of the history of flooding in America and the creation of the NFIP in 1968, after private insurers began refusing to cover flood losses. Part II also addresses the theoretical rationales that historically have been used to justify insurers’ refusal to cover flood losses. Part III discusses the problems with the NFIP and Congress’s ineffective attempts to solve those problems. Part IV discusses the arguments in favor of and against eliminating the flood exclusion that currently is contained in homeowners policies and considers how several European countries have addressed the problem of insuring flood losses. This Essay concludes that the elimination of the flood exclusion from homeowners policies would address the problems with the NFIP while increasing the number of American homeowners that have flood insurance from 5.5 million to 69 million, without causing undue financial hardship to insurers.

II. THE PROBLEM OF FLOODING AND THE ORIGINS OF THE NATIONAL FLOOD INSURANCE PROGRAM

A. The Flood Problem

Flooding is not a new problem in America that just emerged in the last decade with Hurricanes Katrina and Sandy. To the contrary, it is a problem we have been dealing with since the first building was constructed in this country. Every year, there is flooding somewhere. In the past thirty years alone, flooding has caused approximately $240 billion in damage, for an average of approximately $8 billion annually.8

Flooding only causes “property damage,” however, if there is property located in the area of the flooding. Largely due to waterways historically serving as a means of transportation, this country developed along waterways, with the largest cities and population centers located on the coasts or along rivers. With water comes periodic flooding despite mankind’s best efforts to prevent it. There are, of course, benefits that flow from Mother Nature’s will that rivers and other waterways flood periodically. Flooding enriches the soil by bringing nutrients to the flooded area.9 Indeed, the plains states are fertile farming lands because flooding from rivers such as the Mississippi River periodically enriches the soil in the surrounding areas.10 Flooding also replenishes the sediments in downstream areas and wetlands.11

8. See Flood Loss Data, supra note 2.
9. See Klein & Zellmer, supra note 3, at 1477; Scales, supra note 7, at 6.
10. See Klein & Zellmer, supra note 3, at 1477.
11. See id., at 1500–01.
Notwithstanding the benefits of flooding, in response to the many flooding events that have occurred since 1927, Congress has systematically passed numerous laws that were designed to prevent or control flooding.\textsuperscript{12} These laws in sum were intended to thwart Mother Nature’s will that rivers and other waterways flood periodically.

Consider, for example, the Mississippi River’s lengthy history of flooding, despite mankind’s unsuccessful attempts to control the natural tendency of the river to flood periodically. The Mississippi River is the largest river in North America at 2,300 miles, and America’s efforts to tame it have resulted in over 80\% of the river’s natural floodplain being “protected” by levees.\textsuperscript{13} Indeed, nearly 1,800 miles of the river are lined with levees.\textsuperscript{14} Nonetheless, the river still has flooded every decade since 1849.\textsuperscript{15}

One of the great floods of the Mississippi River that occurred despite miles and miles of levees was the Flood of 1927. From Illinois to Mississippi, 145 breaches in the levees occurred, and over 17 million acres of land were flooded in seven states with an inland sea 100 miles long created just north of Vicksburg, Mississippi.\textsuperscript{16} Billions of dollars in damage occurred.\textsuperscript{17}

Another major Mississippi River flooding event was the Flood of 1993. In this event, 40 of 226 federal levees and 1043 of 1345 nonfederal levees were overtopped or breached.\textsuperscript{18} Approximately 100,000 buildings were destroyed or severely damaged.\textsuperscript{19} Over 500 counties were declared federal disaster areas.\textsuperscript{20} The total damages were estimated to be as high as $20 billion.\textsuperscript{21}

In 2005, Hurricane Katrina was another major flooding event. With a twenty-one-foot storm surge, at least thirty levees surrounding New Orle-

\begin{itemize}
\item \textsuperscript{13} See Klein & Zellmer, supra note 3, at 1477, 1498.
\item \textsuperscript{14} See id. at 1480.
\item \textsuperscript{15} See id. at 1477–78.
\item \textsuperscript{16} See id. at 1481.
\item \textsuperscript{17} See id. at 1481–82.
\item \textsuperscript{19} See Klein & Zellmer, supra note 3, at 1494; see also Gerald E. Galloway, Jr., Corps of Engineers Responses to the Changing National Approach to Flood Plain Management Since the 1993 Midwest Flood, 130 J. CONTEM. WATER RES. & EDUC. 5, 5 (2005).
\item \textsuperscript{20} See Klein & Zellmer, supra note 3, at 1494; see also Susan Saulny, Development Rises on St. Louis Area Floodplains, N.Y. TIMES (May 15, 2007), http://www.nytimes.com/2007/05/15/us/15flood.html.
\item \textsuperscript{21} See Klein & Zellmer, supra note 3, at 1494; see also Galloway, supra note 19, at 5.
\end{itemize}
ans gave way, and at least 80% of New Orleans was submerged by up to twenty feet of water. 22 Approximately 300,000 homes were destroyed, and the property damage totaled approximately $100 billion. 23

Hurricane Sandy, which came ashore in New Jersey just south of New York City on October 29, 2012, was the second largest flooding disaster in America behind only Hurricane Katrina. 24 It had a twelve-foot storm surge in the New York City area, and the damage estimate exceeds $60 billion. 25

Given the repeated flooding events that have occurred decade after decade—despite levees, seawalls, and other flood prevention efforts—it is not a question of if levees and seawalls will fail. The question is only when they will fail. 26 Ironically, the construction of levees has resulted in greater, not lesser, damage because they have created a false sense of security that has encouraged people to develop land in floodplains. 27 Yet, the governmental agencies responsible for building and maintaining the levees that will inevitably fail are immune from liability for the damage that results when their levees fail. 28

Levees also have increased flood damage because they have served to funnel water downstream and, when the levees are breached, the escaping water gets trapped in the areas near the levee breaches. 29 Further, levees also have caused downstream land erosion because the natural replenishment of downstream areas with sediments dispersed during flood events has been frustrated or prevented by levees. 30 Louisiana, for example, has

23. See Klein & Zellmer, supra note 3, at 1499.
25. See id.; see also French, supra note 3, at 463.
27. See Klein & Zellmer, supra note 3, at 1487, 1489, 1495, 1503, 1507–08, 1520.
30. See Klein & Zellmer, supra note 3, at 1500–01, 1509; John McQuaid & Mark Schleifstein, In Harm’s Way: Surging Water Is the Biggest Threat to New Orleans,
lost 1.2 million acres of wetlands since the 1930s. As a result of the land erosion due to levees, it has been estimated that at the time of Hurricane Katrina, the Gulf of Mexico was, “in effect, probably 20 miles closer to [New Orleans] than it was in 1965 when Hurricane Betsy hit.” Consequently, research suggests that most of the flooding caused by Hurricane Katrina could have been avoided if the coastal wetlands downstream of New Orleans had still existed.

B. The Theoretical Concerns that Led to Private Insurers’ Refusal to Cover the Risk of Flooding

Unlike “named peril” policies, most homeowners insurance policies purport to cover “all risks” of loss, except for specific risks that are expressly excluded. Historically, named peril policies only covered one specified peril. The first non-marine named peril policy was a fire insurance policy, which originated in London following the Great Fire of 1666. In the 1940s and 1950s, insurers began to bundle coverages for multiple named perils together under one policy. These policies became known as “multi-peril” policies. Under both “named peril” and “multi-peril” policies, any perils that were not expressly listed as covered were excluded.

“All risk” policies were then developed from “multi-peril” policies. Unlike “multi-peril” and “named peril” policies, however, “all risk” policies cover all perils except for the perils that are specifically excluded. Historically, coverage for the risk of loss due to flooding could be purchased...
as an add-on coverage under named peril or multi-peril policies, and all risk policies covered flood losses unless flooding was expressly excluded.\textsuperscript{39}

Since the Flood of 1927, the losses caused by flooding have been staggering, with the total in the hundreds of billions of dollars.\textsuperscript{40} Indeed, flooding is the most common type of disaster in the world, with nine out of ten catastrophes each year related to flooding.\textsuperscript{41}

Consequently, by the 1960s, insurers had seen enough of flood losses, and they decided that insuring losses due to flooding generally was not a risk they wanted to accept. Almost uniformly, they have refused to insure flood losses for non-commercial entities despite selling “all risk” homeowners property policies.\textsuperscript{42} Thus, they have inserted a flood exclusion in all homeowners policies that is worded the same as or similarly to the following:

\begin{quote}
We do not insure for loss caused directly or indirectly by . . . Water Damage[, which] means: Flood, surface water, waves, tidal water, overflow of a body of water, or spray from any of these, whether or not driven by wind . . . .\textsuperscript{43}
\end{quote}

The three primary reasons given for private insurers’ refusal to cover the risk of losses due to flooding are: (1) the propensity of entities most likely to suffer from flooding losses to purchase coverage for flooding, while those unlikely to suffer such losses decline to purchase such coverage; (2) the near “certainty of losses in some areas,” and (3) “the ruinous, widespread nature of flooding events.”\textsuperscript{44} These concerns associated with insuring the risk of flooding generally fall under the theoretical headings of adverse selection, moral hazard, and correlated risks.

\subsection*{1. Adverse Selection}

Adverse selection is the tendency of people who are most likely to suffer losses to purchase insurance to protect against such losses, while the

\textsuperscript{39} See Scales, \textit{supra} note 7, at 7.
\textsuperscript{41} See Houck, \textit{supra} note 40, at 62.
\textsuperscript{42} See Warren Kriesel & Craig Landry, \textit{Participation in the National Flood Insurance Program: An Empirical Analysis for Coastal Properties}, 71 J. OF RISK & INS. 405, 405 (2004). Contrary to the conventional wisdom that no private insurers will provide coverage for flood losses, for the right price and for the right policyholders, there are some insurers that are still willing to provide limited coverage under commercial property policies for floods. \textit{See, e.g.}, Penford Corp. v. Nat’l Union Fire Ins. Co. of Pittsburgh, Pa., 662 F.3d 497 (8th Cir. 2011) (litigating whether policyholder, whose manufacturing facility was located on banks of Cedar River, had $50 million or $20 million in coverage for flood losses under “all risk” commercial property policy issued by AIG and ACE).
\textsuperscript{44} Kriesel & Landry, \textit{supra} note 42, at 405; \textit{see also} Scales, \textit{supra} note 7, at 8–9.
people who are the least likely to suffer losses decline to purchase insurance.45 The foundational premise of the theory of adverse selection is that policyholders have an informational advantage over insurers about their risks and that people who know they are high risk buy more insurance than people who are low risk.46

If insurance is sold on a peril-by-peril basis, the idea of adverse selection makes some sense. It stands to reason that a person who thinks his house may be flooded because he lives along a riverbank is more likely to want to purchase flood insurance than someone who lives on top of a mountain.

The concept of adverse selection begins to break down, however, if all of the most common types of risk of loss are bundled together and are covered under the same insurance policy being sold. So, for example, if the homeowners policies available for purchase cover all risks of loss—including flood, wind, fire, vandalism, etc.—then a person who lives along a riverbank that is concerned about flooding will not have a greater incentive to purchase the policy than a person who lives on a mountain top and is more concerned about wind damage than flooding. If the policyholder wants or is required to have homeowners insurance to cover any of the types of risk of loss he may suffer that are covered under all risk policies, then he will buy the policy regardless of whether he is concerned about all of the types of risks of loss covered by the policy.

2. Moral Hazard

The concept of moral hazard captures the intuitive idea that people who have insurance are less likely to take steps to avoid or minimize losses because the losses will be paid by someone else—the insurer.47 As previously noted, Judge Easterbrook has described the theory underlying the concept by stating that, “[o]nce a person has insurance, he will take more risks than before because he bears less of the cost of his conduct.”48 Another commentator has described moral hazard as follows: “What moral hazard means is that, if you cushion the consequences of bad behavior, then you encourage that bad behavior. The lesson of moral hazard is that less is more.”49

Although the concept of moral hazard has a lot of intuitive appeal, the pervasiveness of the moral hazard problem in connection with natural disasters such as flooding is tenuous. The idea that someone will not bother to take steps to prevent his house from being flooded simply because he has insurance ignores the lack of control people have over

45. See supra note 5.
46. See Siegelman, supra note 5, at 1247.
47. See supra note 6.
floods, the real and significant inconvenience of being flooded out of one’s home, and the significant deductibles contained in many homeowners policies. For most people, the fact that their house may be repaired in the future with insurance monies does not mean they would not mind being homeless for weeks or months while they wait for the insurance check to arrive and for their house to be repaired. Most people also do not relish the prospect of having to call contractors and arrange for the removal of mud-caked drywall, carpeting, furniture, and personal belongings from their homes. Similarly, most people also would prefer that their personal belongings and irreplaceable items of sentimental value, such as family heirlooms and photos, not be ruined by muddy flood waters. So, the idea underlying the concept of moral hazard—that people simply would not bother to take preventative measures to avoid being flooded if they have insurance—may have some theoretical appeal, but it overlooks the reality of actually living through a flood.50

3. Correlated Risks

Correlated risks are risks of loss that happen to numerous people in concentrated areas at approximately the same time.51 Insurers do not want to insure correlated risks of loss because they do not believe they can accurately predict the frequency or severity of such losses or collect enough premiums to adequately spread the risk of loss across a large enough pool of policyholders to cover the losses when they occur.52

Although flood events do present correlated risks of loss, it does not necessarily mean flood losses are uninsurable today. Correlated risk concerns are most legitimate when an insurer only sells insurance in a limited geographic area, because the pool of policyholders is limited and all of the policyholders are likely to face the same types of natural hazards that occur at the same time (e.g., people who live in the same area share the

50. A better way to address moral hazard concerns is for the insurance policy to cover, and require the policyholder to take, preventative measures intended to avoid or minimize a loss. This coverage is already provided under homeowners policies, the “sue and labor” provisions of commercial property policies, and even NFIP policies to a limited extent. See, e.g., John S. Clark Co. v. United Nat’l. Ins. Co., 304 F. Supp. 2d 758, 767 (M.D.N.C. 2004) (“[T]o be covered as reimbursable sue and labor expenses [under a commercial property policy], those expenditures must be made for the benefit of the insurer in mitigating or preventing a covered loss.” (first alteration in original) (quoting Swire Pac. Holdings, Inc. v. Zurich Ins. Co., 139 F. Supp. 2d 1374, 1385 (S.D. Fla. 2001))); Sample Homeowners Policy §§ E.2.a (Additional Coverages), I.B.4 (Conditions) (requiring homeowner to protect damaged property against further damage with insurer agreeing to pay costs incurred to do so), reprinted in Abraham, supra note 43, at 199, 207; FEMA, Standard Flood Insurance Policy § III.C.2 [hereinafter FEMA, Standard Policy], available at http://www.fema.gov/media-library-data/1398950546439-c78022c2faa410c9902076b4a0a5a/F-122_NFIP_DwellingForm_June2014.pdf (covering up to $1,000 of policyholder’s costs incurred to avoid or minimize losses).

51. See supra note 7.

52. See Bruggeman, Faure & Heldt, supra note 7, at 187; Cummins, supra note 7, at 342–45.
common risks of flooding, earthquakes, and tornadoes). Consequently, when insurance companies only sold policies locally or regionally, correlated risk concerns regarding flooding were understandable.

The negative financial consequences for insurers associated with the correlated risk of flooding have substantially diminished since the NFIP originally was created, however, due to the emergence of multinational insurance companies that sell insurance to policyholders across America and throughout the world.53 Similarly, the financial impact of correlated risks also has been diminished by reinsurance, which is now a worldwide business in which global reinsurers insure all of or portions of another insurer’s portfolio of business (known as “treaty” reinsurance).54

Insurers today also can further spread the risk of loss associated with catastrophic events through the sale of catastrophe bonds, pursuant to which insurers sell bonds for specific types of catastrophes, such as earthquakes and hurricanes, to investors who receive interest payments on the bonds and the return of their principal at the end of the bond term unless the specified catastrophe occurs, in which case, the insurer keeps the principal and ceases to make interest payments on the bonds.55 Since 1996, insurers have spread their risks through the issuance of $51 billion in catastrophe bonds while incurring only $682 million in losses for the catastrophes covered by the bonds (only 1.3% of the total amount issued).56

Thus, the development of multinational insurance companies, global reinsurance, and catastrophe bonds means that the pool of insureds and investors across which the risk of a flood loss in Long Island, for example, can now be spread are located not only on Long Island, but also across the entire state of New York and in more distant places throughout America, Australia, and Europe.

In addition, as is the case with adverse selection, the correlated risk problem is much greater when perils are insured separately, as opposed to when coverage for numerous perils is bundled together under the same policy. If insurance for flooding is covered only under named peril policies and sold only regionally, then people who live in the same geographic flood areas would have a correlated risk of suffering flood losses at the same time. If the policies being sold, however, cover floods along with the other most common perils such as wind, fire, and vandalism, then the


54. See Cummins, supra note 7, at 343.


56. See Yoon & Scism, supra note 55.
pool of policyholders who would choose to purchase such insurance would be much more diverse, both geographically and from a risk profile perspective. Consider again the example of two policyholders, one who lives along a river with flooding concerns and the other who lives on a mountain top with wind damage concerns. Because the perils of wind and flooding would be bundled together and sold under the same policy form, the risks of a flood loss for the two policyholders would not be correlated. Nor would the risks of wind damage for the two policyholders be correlated. Thus, bundling coverage for multiple perils together in a single policy decreases: (1) correlated risk concerns because the pool of policyholders likely will have more diverse risk profiles and (2) adverse selection concerns because more policyholders with diverse risk profiles are likely to purchase the insurance.

C. The Creation of the National Flood Insurance Program

As a result of insurers’ refusal to cover flood losses due to adverse selection, moral hazard, and correlated risk concerns, the National Flood Insurance Program was created through the National Flood Insurance Act of 1968 to fill the void.\textsuperscript{57} The NFIP also was created as an attempt to recoup some of the monies the federal government was spending on post-disaster relief efforts by getting the people who were most likely to be victims of flooding to pay premiums to offset some of the post-disaster relief costs the government was incurring.\textsuperscript{58}

Although purchasing flood insurance under the NFIP is considered voluntary, a property owner must have flood insurance if the property is located in a high risk flood area (i.e., the 100-year flood plain) and the loan used to purchase the property is secured by a federally insured lender, which currently encompasses the vast majority of loans.\textsuperscript{59}

As deficiencies in the NFIP have been revealed by various flood disasters since 1968, the NFIP has been amended reactively in attempts to address the deficiencies. For example, in 1994, following the Midwest Flood of 1993, which resulted in the NFIP becoming insolvent due to the combination of low participation rates and high loss rates, the program was amended to increase the number of properties that are required to purchase flood insurance.\textsuperscript{60} In 2004, when it was realized that many


\textsuperscript{59.} See 42 U.S.C. § 4012a.

\textsuperscript{60.} See id. § 4012a(b)(4)(B).
properties repeatedly are flooded and repaired using NFIP funds, the program was amended to attempt to “disincentivize property owners from living in areas repeatedly flooded” by providing these property owners assistance in either elevating the properties or moving.\(^6\) In 2012, the program, again insolvent in the amount of approximately $17 billion due to Hurricane Katrina,\(^6\) was amended to address the fact that: (1) about 28% of the policies are sold at substantially subsidized premium rates, and (2) the flood maps are outdated and inaccurate.\(^6\) It remains to be seen whether actuarially-sound premium rates will ever be charged, because several coastal states successfully lobbied against the new premium rates that would be charged to their residents such that Congress recently passed the Homeowner Flood Insurance Affordability Act of 2014, which delays the full implementation of actuarially sound premium rates.\(^6\)

The NFIP currently is administered by the Federal Emergency Management Agency (FEMA), which uses private insurance companies to: (1) sell the NFIP policies on behalf of FEMA\(^6\) in exchange for a 30% sales commission and (2) handle flood claims when they arise for additional claims handling fees.\(^6\) In 2012, the private insurers that sell NFIP policies and handle NFIP claims for FEMA, without actually insuring any of the policyholders’ risks of flood losses, were paid a total of 43.6% of the premiums collected.\(^6\)

For a property owner to qualify for flood insurance under the NFIP, the property owner’s community must agree to adopt and enforce ordinances that meet or exceed FEMA requirements to reduce or minimize the risk of flooding.\(^6\) Currently, only 5.5 million of the approximately 72

---


65. See Scales, supra note 7, at 14.

66. See 42 U.S.C. § 4012a; see also Scales, supra note 7, at 14.

67. See GAO OVERVIEW OF PAST WORK, supra note 58, at 12.

million homeowners in America purchase flood insurance from the NFIP.69

1. The Coverage Provided Under NFIP Policies

The maximum amount of NFIP coverage that can be purchased for a residential property is $250,000 for the building itself and $100,000 for personal property.70 Coverage for basements, which includes any spaces below ground level, is limited to only a few items including fuel tanks, furnaces, and water heaters, which means many things commonly found or stored in basements—such as drywall, ceilings, carpeting, clothing, electronic equipment, furniture, and other types of personal property—are not covered.71 NFIP policies also do not cover property and belongings that are located outside of the insured building, such as trees, plants, wells, septic systems, decks, patios, fences, hot tubs, and swimming pools.72

In valuing losses, unlike “guaranteed replacement cost” policies that pay the cost needed to rebuild the home regardless of the limit of liability or “valued” policies that pay the full limit of the policy in the event of a total loss, the policyholder receives the lesser of the replacement cost or actual cash value (i.e., the depreciated value) of damaged items under NFIP policies.73 In addition, NFIP policies do not cover the cost of temporary housing.74

To recover any amount of a loss under an NFIP policy, the policyholder must prepare and submit a “proof of loss” form within sixty days that includes, among other things, bills, receipts, and related documents even if the claims adjuster assigned by the NFIP program to investigate the loss does not furnish a proof of loss form or help the policyholder com-


70. See 42 U.S.C. § 4013(b)(2), (3); GAO Overview of Past Work, supra note 58, at 3.


72. See id. at 2.

73. See id. at 4.

74. See id. at 2.
complete it. Generally, the requirement to submit a proof of loss form in accordance with the terms of NFIP policy is strictly enforced by courts.

In addition, unlike insurance sold by private insurers, the claims adjuster assigned to handle the policyholder’s loss under an NFIP policy does not have authority to approve or disapprove the claim or to even tell the policyholder whether his claim has been approved or disapproved.

2. Litigating Claims Under NFIP Policies

One of the unpleasant surprises policyholders who were unfamiliar with the terms of NFIP policies discovered in the wake of Hurricane Katrina was that the NFIP flood policies did not actually cover much of their damages such as landscaping and personal property located in their yards and basements. Another unpleasant surprise for policyholders was that suing the NFIP or the NFIP’s claims handlers for poor or improper handling of flood claims was generally fruitless. Unlike insurance policies issued by private insurers, FEMA and the private insurers who sell NFIP policies and then adjust the claims on FEMA’s behalf are not subject to liability for underpaying claims, handling claims poorly, or even acting in bad faith, because the NFIP does not allow for recovery for such misconduct and preempts state laws that do provide such relief.

If it becomes necessary to sue under an NFIP policy in order to collect, the policyholder can only file a lawsuit in federal court and must do so within one year. Failure to file the lawsuit within a year results in the forfeiture of the claim.

III. The Problems with the National Flood Insurance Program

Poor coverage, low participation rates, and insolvency are three of the biggest problems with the NFIP. Hurricane Katrina caused over $100 bil-
lion in losses. Hurricane Sandy caused over $60 billion in losses. These figures account for only $160 billion of the $274 billion in flood damage caused since 1978. So, how much of the $274 billion has been paid by the NFIP? As a percentage of the total amount, the NFIP has not paid much. As discussed above, very few homeowners actually have flood insurance. And, even for the few people who buy it, NFIP policies often only cover a portion of a person’s flood losses. Consequently, through March 31, 2014, of the approximately $274 billion of flood losses caused since 1978, the NFIP has paid a total of $50.6 billion, or about 18%. Yet, despite only paying a small fraction of the total damages related to floods, the NFIP was insolvent with a deficit of $24 billion as of December 2013.

A. Poor Coverage and Unfavorable Laws Regarding Claims Handling

The insurance coverage provided by NFIP policies is inadequate. The maximum coverage available for a house is $250,000 and $100,000 is the maximum amount available for a home’s contents, which account for only a fraction of the actual value of many properties that are located in large population centers such as Boston, New York, and California. There also is no coverage under NFIP policies for the vast majority of things commonly found in basements and yards. Thus, it is not unusual for significant portions of a policyholder’s loss to simply not be covered under NFIP policies.

Even those portions of a policyholder’s loss that are covered, however, are paid only on an actual cash value basis (i.e., depreciated value), which means the policyholder is paid pennies on the dollar for his ruined personal property. And, in order to recover those pennies on the dollar, one must submit a proof of loss form with supporting documentation within sixty days and then sue within one year if the policyholder disagrees with the amount paid.

83. See Klein & Zellmer, supra note 3, at 1499.
84. See French, supra note 3, at 463.
85. See Flood Loss Data, supra note 2.
86. See supra note 69 and accompanying text.
87. See supra notes 70–77 and accompanying text.
88. See supra notes 70–77 and accompanying text.
91. See 42 U.S.C. § 4013(b) (2), (3) (2012); see also supra note 70 and accompanying text.
92. See supra note 73 and accompanying text.
93. See supra notes 75, 81 and accompanying text.
Unlike many states' laws regarding compliance with provisions in policies sold by private insurers, a policyholder’s failure to strictly comply with the NFIP policy’s proof of loss and suit limitation provisions results in a forfeiture of coverage.94 Also unlike a policyholder’s common law and statutory rights against private insurers regarding poor claims handling practices or bad faith conduct, the federal government and its agents that administer the NFIP are immune from liability for such misconduct under the NFIP.95

In sum, the insurance coverage provided by NFIP policies is limited, a lot of flood-damaged property is not covered by it, there are numerous procedural hurdles that the policyholder must clear in order to avoid the forfeiture of a claim such as filing a proof of loss with supporting documentation within sixty days, and the NFIP administrators and claims handlers are immune from liability for poor or improper claims handling conduct.

B. Low Participation Rates and Insolvency

In addition to the problems discussed above, the actual administration of the NFIP could be used as the poster-child to support the argument that the federal government should not be in the insurance business. The NFIP historically has used outdated floodplain maps due to a lack of funds to create accurate ones, so in many instances the wrong homes were insured or uninsured.96 For example, the flood map for the New York City area that was being used at the time of Hurricane Sandy was based on data and modeling that were over thirty years old.97 The new draft of the map for the New York City area that was released in June 2013 essentially doubled the number of houses that are located in the high-risk flood zones (i.e., the 100-year flood plain).98

The NFIP also is actuarially unsound by design because it intentionally has been charging subsidized premium rates for old homes grandfathered into the program for decades, which has led to frequent periods of insolvency.99 The NFIP currently owes the U.S. Treasury more than $24 billion.100

94. See supra notes 76, 82 and accompanying text.
95. See supra notes 79–80 and accompanying text.
96. See Cummins, supra note 7, at 358; Beth A. Dickhaus & Darrin N. Sacks, Recent Developments in Insurance Regulation, 42 TORT TRIAL & INS. PRAC. L.J. 571, 582 (2007); GAO Overview of Past Work, supra note 58, at 37; Kriesel & Landry, supra note 42, at 406–07.
97. See Dixon et al., supra note 78, at 1.
98. See id. at 2.
100. See Extreme Weather Events, supra note 89. The NFIP program also insures a lot of recidivist policyholders. See id. Many homes covered under the program get flooded over and over again because policyholders were not required to take preventative measures to avoid flooding under the NFIP unless the repair costs exceeded 50% of the value of the home. See id.; see also Klein & Zellmer, supra note
In addition, participation in the NFIP has been poor because the NFIP is a voluntary program and people often do not understand the risk of flooding for their homes due, at least in part, to: (1) poor communication of such risks by the NFIP and (2) the outdated and inaccurate flood maps. Many people do not seem to appreciate that a 1% annual chance of being flooded (i.e., the property is located in a 100-year flood plain) actually means you have a 26% chance of being flooded during the course of a 30-year mortgage. In the New Orleans area, for example, only 10% of the homes flooded by Hurricane Katrina had flood insurance even though the city sits below sea level and is surrounded by water. In the New York City area, a little more than 50% of the homes flooded had flood insurance even though Manhattan and Staten Island are islands.

In short, when given a choice, very few people actually purchase NFIP policies. Indeed, most homeowners who have flood insurance purchase it only because they are required to do so in order to get a mortgage if their home shows up on the 100-year flood plain map. At the time of Hurricane Sandy, only about 20% of the homeowners located in areas considered at high risk for flooding who were not required to have flood insurance actually had flood coverage, and only approximately 1% of homeowners located outside of areas designated at high risk for floods nationwide purchase NFIP policies. Consequently, as of the end of 2012, only 5.5 million of the 72 million homeowners in America had a NFIP policy, which means the vast majority of homes in America do not have flood coverage.


101. See supra notes 96–98 and accompanying text.

102. See GAO OVERVIEW OF PAST WORK, supra note 58, at 6, 25; Scales, supra note 7, at 18; Pham, supra note 69, at 652.

103. See Klein & Zellmer, supra note 3, at 1502; Scales, supra note 7, at 15; Pham, supra note 69, at 639.

104. See Dixon et al., supra note 78, at xvii.

105. See id.

106. See id.; GAO OVERVIEW OF PAST WORK, supra note 58, at 22 (stating that 2006 Rand study estimated that only 1% of homes not located in areas designated high flood risk areas purchased flood insurance).

107. See supra note 69 and accompanying text.
IV. THE ELIMINATION OF THE FLOOD EXCLUSION IN HOMEOWNERS POLICIES AS ONE PART OF THE SOLUTION TO THE FLOOD PROBLEM

If the NFIP is inadequate despite Congress’s continuing efforts over the past fifty-six years to make it work, is there a better way to insure flood losses? Yes, homeowners insurance. This part shows that, despite impediments to doing so, the flood exclusion should be eliminated from homeowners insurance policies.

A. Homeowners Insurance as a Social Necessity

Insurance plays a critical role in society today. Insurance is a “social instrument” because it protects the limited assets of individuals by spreading and transferring, through an insurer intermediary, the risk of losses due to natural and unnatural disasters from the individual to a larger population. Insurance is even more important now than it was when America was first founded. Homeownership is not even possible for most people today without insurance because anyone who wants to purchase a house using a bank to finance the purchase is required to have homeowners insurance in an amount adequate to cover the mortgage.

Yet, despite the importance of insurance, over the past fifty years, insurers have undermined the purpose that insurance serves by hollowing

108. See Jeffrey W. Stempel, The Insurance Policy as Social Instrument and Social Institution, 51 WM. & MARY L. REV. 1489, 1489 (2010); see also Deborah A. Stone, Beyond Moral Hazard: Insurance as Moral Opportunity, 6 CONN. INS. L.J. 11, 26–29 (1999) (“Because virtually every adult citizen participates in various forms of mandatory insurance, from automobile liability insurance to unemployment insurance, old-age pensions and disability insurance, everyone is exposed to two of the moral assumptions of these programs: collective responsibility for the well-being of individuals and individual responsibility for the well-being of others.”).

109. See JERRY & RICHMOND, supra note 6, at 18; see also JAY M. FEINMAN, DELAY, DENY, DEFEND: WHY INSURANCE COMPANIES DON’T PAY CLAIMS AND WHAT YOU CAN DO ABOUT IT 21 (2010); Christopher C. French, The Role of the Profit Imperative in Risk Management, 17 U. PA. J. BUS. L. (forthcoming) (manuscript at 1–2) (on file with author) (arguing that social purpose of insurance is being marginalized by private insurers’ pursuit of profits).

110. See French, supra note 109, at 1–2.

111. See MARTIN F. GRACE ET AL., CATASTROPHE INSURANCE: CONSUMER DEMAND, MARKETS AND REGULATION 83 (2003) (discussing demand for homeowners insurance, stating that “homeowners insurance . . . is essentially mandatory”); French, supra note 109, at 11; Stempel, supra note 108, at 1497.
out the coverages that are provided under numerous lines of insurance, including homeowners insurance, through the addition of exclusions for some of the most common catastrophic risks such as floods. The trend toward less and less coverage under “all risk” policies accelerated when publicly traded stock companies began replacing mutual companies as the dominant providers of insurance during the past two decades.

The trend of insurers hollowing out the coverage provided under “all risk” policies should be reversed, starting with the exclusion for flood losses. Homeowners insurance, as a social necessity today, should be required to cover losses caused by floods. Indeed, catastrophic risks such as floods are exactly the types of risk of loss that should be covered under “all risk” homeowners insurance. “Removing coverage for the very types of losses that are most common and have the most devastating impact on people, businesses and communities is antithetical to the risk transferring” and social safety net purposes of insurance.

Although not being done in a coordinated or comprehensive way, in recent years, we have seen some legislatures moving in the direction of requiring the insurance industry to actually fulfill the purpose of insurance by requiring that it provide products that serve insurance’s purpose as a social safety net. For example, with respect to health insurance, the Affordable Care Act has: (1) removed insurers’ ability to refuse to insure certain people, (2) removed insurers’ ability to cancel insurance for people who become sick, and (3) reduced insurers’ ability to use reverse adverse selection to charge certain risk classifications prohibitively expensive premiums.

Other examples of legislatures enacting statutes to override policy exclusions or to require insurers to cover risks that they would otherwise refuse to cover can be found in the auto insurance context, where drivers who insurers refuse to insure because the insurers do not consider the drivers adequately profitable generally can still get a minimum amount of liability coverage through state residual risk insurance pools. Another example is states’ passage of statutes that nullify the “innocent co-insured” exclusion. Insurers add that exclusion to property policies in an attempt to prevent innocent co-insureds from recovering under policies for losses that were expected or intended by another insured such as when an estranged spouse who is no longer living in the couple’s house destroys the

112. See French, supra note 109, at 4.
113. Id. at 1–3; see also Ottogontsetseg Erhemjamts & J. Tyler Leverty, The Demise of the Mutual Organizational Form: An Investigation of the Life Insurance Industry, 42 J. MONEY, CREDIT & BANKING 1011 (2010).
114. French, supra note 109, at 39.
116. See Abraham, supra note 43, at 771.
house to prevent the spouse from having it.\textsuperscript{117} Yet another example is the state of California’s refusal to enforce exclusions that purport to eliminate coverage for losses that are concurrently caused by both an excluded peril and a covered peril if the efficient proximate cause of the loss is a covered peril.\textsuperscript{118}

In short, when an issue has been deemed sufficiently important, there are numerous examples of states exercising their regulatory power to require insurers to cover perils that insurers otherwise would not agree to cover.

B. \textit{Insuring Flood Losses Under Homeowners Policies Would Address the Problems with the NFIP}

All of the problems with the NFIP—poor coverage, lack of accountability for poor claims payment practices, low participation rates that leave millions of homeowners uninsured, and the periodic insolvency of the program—would be addressed if the flood exclusion were eliminated from homeowners policies.\textsuperscript{119} Indeed, the NFIP as it currently exists would no longer even be necessary.

First, standard homeowners insurance policies provide much more favorable coverage terms than NFIP policies.\textsuperscript{120} For example, coverage is not limited to $250,000 for the house itself and $100,000 for the contents of the house. In addition, homeowners insurance policies typically pay losses at replacement cost instead of actual cash value.\textsuperscript{121}

Second, unlike under the NFIP,\textsuperscript{122} insurers’ claims handling practices under homeowners policies are governed by state law, not federal law, and most states provide a remedy for an insurer’s improper handling of a policyholder’s claim.\textsuperscript{123} Similarly, unlike the federal law governing the NFIP, inadvertent claims submission errors by the policyholder such as failing to

\textsuperscript{117.} See French, supra note 109, at 40 n.174; see also JERRY & RICHMOND, supra note 6, at 425–26.


\textsuperscript{119.} See Scales, supra note 7, at 43–44. Due to many of the problems with the NFIP discussed in this Essay, Professor Scales also proposes in his article regarding flood insurance that flood losses be covered under homeowners insurance. See id.

\textsuperscript{120.} See, e.g., Sample Homeowners Policy § I.C.2 (Conditions), reprinted in ABRAHAM, supra note 43, at 207 (covering cost to repair or replace damaged property without deduction for depreciation, unlike NFIP policies that only pay lesser of replacement cost or actual cash value which includes depreciation); see also Abbott, supra note 62, at 134–55.

\textsuperscript{121.} See supra note 120.

\textsuperscript{122.} 44 C.F.R. pt. 61, app. A(1), art. IX (“This [flood] policy and all disputes arising from the handling of any claim under the policy are governed exclusively by the flood insurance regulations issued by FEMA, the National Flood Insurance Act of 1968, as amended, and Federal common law.” (citation omitted)); see also Abbott, supra note 62, at 144–47; Scales, supra note 7, at 32.

\textsuperscript{123.} See supra notes 94–95 and accompanying text.
file a proof of loss timely often do not result in the forfeiture of coverage under state law.\textsuperscript{124}

Third, including coverage for floods under homeowners policies would eliminate the participation problem that plagues the NFIP. Currently, in contrast to the only 5.5 million homeowners who have NFIP policies, 96\% of the 72 million homeowners in the United States have homeowners insurance (69 million).\textsuperscript{125} A related benefit of covering floods under homeowners policies is that the accuracy of flood maps would be much less important because a person’s decision whether to purchase flood insurance would not depend upon whether the person’s house is included within the 100-year floodplain on a map.

Fourth, because the risk of flooding would be bundled together under homeowners insurance with all of the other types of coverages such as fire, wind, and vandalism, the cost of insuring against floods would be spread across 69 million policyholders instead of 5.5 million.\textsuperscript{126} Consequently, the total amount of premiums generated to cover flood losses would be much greater than it is under the NFIP. Thus, the capital reserves generated by premiums available to pay flood losses would be much greater, and the risk of insurer insolvency would be much lower. Nonetheless, if insurer insolvency were still a concern despite: (1) the much larger pool of insureds paying premiums, (2) the availability of private global reinsurance, and (3) the availability of catastrophe bonds, then the federal government could act as a reinsurer for losses that exceed a certain stated amount as it currently does for the risk of losses due to terrorism.\textsuperscript{127}

Finally, as discussed above in part III.B, in situations where the various types of perils that are being insured are bundled together under the same policies—as they are under homeowners policies—the arguments regarding adverse selection, moral hazard, and correlated risks that have been used to justify private insurers’ historical refusal to cover flood losses do not carry much weight today. Because nearly all homeowners carry homeowners insurance, adverse selection concerns would be negligible. Moral hazard concerns regarding flooding are also overstated. Regardless of whether they have flood coverage, homeowners cannot control the weather that causes flooding and homeowners generally are highly motivated to protect their homes due, among other reasons, to the risk of loss of irreplaceable items of sentimental value and the inconvenience of be-

\begin{footnotesize}
\vspace{-1pt}
\begin{itemize}
  \item[124.] See Abbott, \textit{supra} note 62, at 144–47; Scales, \textit{supra} note 7, at 34–35.
  \item[125.] See \textit{supra} note 69 and accompanying text.
  \item[126.] See \textit{supra} note 69 and accompanying text.
\end{itemize}
\end{footnotesize}
ing homeless. And, unlike in the years prior to 1968, when the NFIP was created, because the pool of insureds across which the risk of loss is spread today is worldwide, the risk of flood losses by the entire pool of insureds is not correlated. Thus, due to the creation of multinational insurance companies, the development of the global reinsurance market, and the emergence of catastrophe bonds, the risk of insuring against flood losses in, for example, Cedar Rapids, would now be spread across the world to policyholders and investors in places such as Maine, Switzerland, and New Zealand.128

C. Insuring Flood Losses Under Homeowners Policies Would Eliminate Wind Versus Water Damage Litigation

If homeowners insurance covered flood damage, then another added benefit would be much less litigation following water-related catastrophes such as hurricanes. As things currently stand, “policyholders and their insurers spend millions of dollars engaged in endless litigation whenever catastrophes hit. For example, whenever a hurricane comes ashore and brings flooding with it, litigation ensues regarding whether the damage to homes was caused by flooding (an excluded peril) versus wind (a covered peril).”129 More than 6,600 lawsuits were filed in federal court in New

128. The impact of correlated risks on large, multinational insurers has diminished, as insurers have become more global and implemented additional risk transferring mechanisms, such as reinsurance and catastrophe bonds. There still are, however, a substantial number of smaller insurers that sell homeowners insurance on a statewide or regional level. These smaller, regional insurers would still have legitimate correlated risk concerns if they were to sell homeowners insurance that is required to cover correlated catastrophic risks such as floods. To address the risk of insurer insolvency for these smaller insurers in the event of a catastrophic flooding event, it may need to be mandatory that such insurers purchase reinsurance or sell catastrophe bonds. Indeed, it is possible that smaller insurers could no longer sell homeowners insurance because they may not be able to do so at competitive rates due to the additional costs they would need to incur to ensure solvency as a result of the fact they have smaller pools of capital available to cover the losses associated with catastrophic flooding events.


Notably, because private insurers handle claims under the NFIP on behalf of the federal government, many insurers that issued homeowners policies were responsible for estimating the repair costs for both the federal flood policies under the NFIP program and the insurers’ homeowners policies. See Feinman, supra note 109, at 163. When estimating the cost to remove and replace drywall, for example, because the insurer was paying the claim with NFIP money instead of its own
Orleans in connection with Hurricane Katrina. In Mississippi, due to insurers’ position that most policyholders’ claims should be denied due to the flood exclusion, one insurer filed a motion to transfer the lawsuits out of Mississippi on the basis of a survey that showed that 49% of the people in southern Mississippi “believe that insurance executives are on the same level as child molesters.” The Louisiana Department of Insurance received 20,000 complaints per month during the six-month period following the storm. Thus, it is an understatement to say there is a high degree of dissatisfaction with the current state of affairs regarding the payment of claims following hurricanes in America, largely due to the presence of the flood exclusion in homeowners policies, and a significant amount of litigation could be avoided if homeowners policies covered both wind and water damage.

D. European Countries’ Handling of Insurance for Flood Losses

In considering whether it would be viable for floods to be covered under homeowners policies in America, one can look to Europe for an answer. The United Kingdom, France, and Belgium are three examples to consider.

In the United Kingdom, insurance for flood losses is mandatory, and it is covered by private insurers. As opposed to selling insurance that covers just the risk of flooding, the risk of loss due to flooding is bundled together with multiple types of risk of loss that are covered under the policies. The insurance is available to all property owners and the premium cannot exceed one-half percent of the amount insured regardless of where the house is located. Despite the fact that insurers’ ability to refuse to sell insurance to properties located in flood areas is quite limited, private insurers in the United Kingdom continue to be financially viable even though the United Kingdom is located on an island that is known for its rainy weather.

France is another country where the risk of flood losses is covered by private insurers and, again, the risk is bundled together with other hazards instead of being sold as standalone insurance. Coverage for floods is

money, one insurer calculated the cost at $.76 per square foot when the costs would be charged to its homeowners policy, and at $3.31 per square foot when it would be charged to the federal flood program. Presumably, such practices would cease if insurers were responsible for paying the entire loss.

130. Feinman, supra note 109, at 147.

131. Id. at 145 (quoting Eugene Benick, The Flood After the Storm: The Hurricane Katrina Homeowners’ Insurance Litigation, 4 Bus. L. Brief 49, 51 (2007)).

132. Id. at 147.

133. See Majmudar, supra note 61, at 199.

134. See id.

135. See id.

136. See id. at 200–01.

137. See id. at 201–02; see also Olivier Moréteau, Policing the Compensation of Victims of Catastrophes: Combining Solidarity and Self-Responsibility, 54 Loy. L. Rev. 65,
mandatory, and there is no differentiation in the amount of the premium charged based upon the policyholder’s risk of flood losses.\(^{138}\) To incentivize loss prevention measures, there is a premium discount if the municipality in which the property is located has adopted a “prevention of risk plan.”\(^{139}\) A state-sponsored program, known as the Caisse Centrale de Reassurance (CCR), provides reinsurance to the insurers, receives 50% of the premiums collected by the insurers, and likewise pays 50% of any losses.\(^{140}\) France provides an unlimited guarantee of the CCR.\(^{141}\)

In Belgium, it also is mandatory that homeowners insurance cover natural catastrophes, including floods.\(^{142}\) Unlike the French system, the insurer can adjust the premium based upon the risk of loss presented by individual policyholders.\(^{143}\) Approximately 50% of the insurers nonetheless charge the same premium for all of their policyholders.\(^{144}\) Each individual insurer’s risk of losses has a limit, and a state-sponsored “Disaster Fund” covers any losses above the insurer’s limit.\(^{145}\)

In short, the United Kingdom, France, and Belgium are just three examples of countries that are successfully insuring the risk of flooding in a way similar to the proposal in this Essay.

E. Impediments to Change

To avoid insuring flood losses under homeowners policies when faced with the elimination of the flood exclusion, insurers likely will point to:

1. the concept of freedom of contract and the notion that they should be free to cover whatever risks they choose,\(^{146}\)

86 (2008). Germany is another country where flood insurance is covered by policies sold by private insurers, and the risk of loss due to flooding is bundled together with coverage for other natural disasters such as earthquakes. See Majmudar, supra note 61, at 203.

\(^{138}\) See Majmudar, supra note 61, at 202.

\(^{139}\) See Michael Faure & Véronique Bruggeman, Catastrophic Risks and First-Party Insurance, 15 CONN. INS. L.J. 1, 44 (2008).

\(^{140}\) See id. at 45; Moréateau, supra note 137, at 89.

\(^{141}\) Faure & Bruggeman, supra note 139, at 45; Moréateau, supra note 137, at 90.

\(^{142}\) See Faure & Bruggeman, supra note 139, at 45–46.

\(^{143}\) See id. at 46.

\(^{144}\) See id.

\(^{145}\) See id.

\(^{146}\) This argument ignores, of course, the fact that insurance policies are contracts of adhesion sold on a take it or leave it basis in which the policyholder has no input regarding the language contained in the policy. See Jeffrey W. Stempel, Law of Insurance Contract Disputes § 4.06[b], at 4–37 (2d ed. 1999) (“In a sense, the typical insurance contract is one of ‘super-adhesion’ in that the contract is completely standardized and not even reviewed prior to contract formation.”); Michelle Boardman, Insuring Understanding: The Tested Language Defense, 95 IOWA L. REV. 1075, 1091 (2010) (describing “hyperstandardization” of insurance policies); James M. Fischer, Why Are Insurance Contracts Subject to Special Rules of Interpretation?: Text Versus Context, 24 ARIZ. ST. L.J. 995, 996 (1992) (“The only part of the standard policy that is generally customized to the consumer-insured is the Declaration
itability of insuring flood losses due to adverse selection, moral hazard, and correlated risk concerns. Such arguments have carried the day for the past sixty years, which is why we still have the NFIP program. As discussed in part III.B, although such arguments have much less force today than they did in the past, the insurance industry’s lobby is strong—as evidenced by the fact that insurance regulators have repeatedly approved policy forms that contain the flood, and other, exclusions for decades. Indeed, private insurers have great financial incentive to avoid covering floods because they currently are being paid more than 40% of the premiums collected by the NFIP in some years without assuming any of the risk to pay for flood losses.147

Insurers also likely will argue that they would not be permitted to charge actuarially sound premiums if policies are required to cover flood losses. Insurers made such arguments in recent years when they were forced to continue selling homeowners policies that covered wind damage caused by hurricanes in coastal states.148 The correct response to such arguments is, of course, to allow insurers to charge a higher premium to reflect the fact that the policies are now insuring an additional risk. However, in light of the fact that the additional risk assumed by insurers would be spread across approximately 69 million policyholders with a wide range of risk profiles, instead of just the 5.5 million homeowners who currently are covered by the NFIP and live almost exclusively in areas known to be at a high risk for flooding,149 the average additional premium that should be charged to each individual policyholder when the risk pool is viewed as a whole should be quite limited.150

Sheet . . . .  [T]here is little, if any, freedom to negotiate the standardized language of the insurance contract that determines the scope of coverage.”); Susan Randall, Freedom of Contract in Insurance, 14 CONN. INS. L.J. 107, 125 (2007) (“[I]n some lines of insurance, all insurance companies provide identical coverage on the same take-it-or-leave-it basis.”); Kent D. Syverud, The Duty to Settle, 76 VA. L. REV. 1113, 1153 (1990) (“[P]roperty owner’s liability insurance contracts are standardized across insurers in a form few insureds have the power or experience to bargain around.”).

147. See GAO OVERVIEW OF PAST WORK, supra note 58, at 12.
149. See supra note 69 and accompanying text.
150. Ideally, to accomplish the broadest spreading of the risk of flood losses, nationwide premium rates would be charged with only modestly higher, but not actuarially based, premium rates being charged to high risk properties, similar to the way the Affordable Care Act allows only limited premium variation based upon
Requiring private insurers that sell homeowners insurance to cover flood losses also likely would meet strong political opposition in this politically polarized country. To some people, requiring homeowners policies to cover flood losses, in combination with the existing laws that require homeowners to purchase homeowners insurance, would be viewed as a form of socialism. The reality, however, is that insurance already is a highly regulated industry where public policy and social concerns such as the compensation of innocent victims play a large role. States already regulate insurers by, among other things, approving policy language, approving premium rates, and establishing capital surplus requirements. States also already manage guaranty associations to cover claims submitted to insolvent insurers, and states already have insurance programs for auto drivers that insurers refuse to accept as policyholders. Consequently, eliminating the flood exclusion from homeowners policies would simply be another example, in a long list of examples, of states exercising their regulatory power over the insurance industry.

Other people will object to including flood coverage in homeowners insurance policies because it would result in premium subsidization. Unless there were numerous risk classifications based upon where a person’s home is located and the flood prevention measures in place—which would in some respects defeat the purpose of bundling flood coverage together with other perils in a single policy form—then people who live in areas with a low risk of flooding would to some extent be subsidizing the premiums of the people in higher risk areas. That already occurs, however, under homeowners policies because coverage for all risks is provided under such policies regardless of whether the policyholder actually needs or wants coverage for all of the risks. So, policyholders are already paying for some coverages they do not need but other people do. Indeed, that

---

*151.* See Abraham, supra note 43, at 119, 122–24, 153–37; Jerry & Richmond, supra note 6, at 90–94.


*153.* The reality, of course, is that subsidization already occurs under all lines of insurance and, ironically, taxpayers currently are subsidizing the NFIP in the approximate amount of $24 billion, even though many people are uninsured for flood losses and many people who are insured for flood losses are not paid in full for such losses when they occur. See Kristen Uhlenbrock, *Despite Hazard of Sea Level Rise, Senate Halts Flood Insurance Reforms*, THINKPROGRESS.ORG (Jan. 31, 2014, 10:37
is one of the very reasons why flood coverage should be included in homeowners policies. It would make flood coverage more affordable and dramatically increase the participation rates. Yet, in a country that highly values individual rights, many people likely will object to being required to subsidize their fellow Americans’ insurance rates even if: (1) the cost of doing so is only a nominal amount, and (2) they already are subsidizing the premium rates of other policyholders without realizing it.154

To crystallize the difficulties associated with changing the flood insurance landscape in America in significant ways, one need only consider what has happened with the Affordable Care Act. It was passed only when Democrats held the majority in both houses of Congress, and it has been subject to more than fifty attempts to repeal or change it since Republicans obtained a majority in the House of Representatives.155 If something as important as ensuring that everyone in this country is able to obtain health care has met so much political resistance, then it is doubtful that the political will currently exists to pass legislation that would result in everyone being covered for something undoubtedly of less importance—flood losses.

V. Conclusion

Despite accounting for 90% of all natural disasters and annually causing billions of dollars in damages, only about 7% of homeowners in America have insurance coverage for flood losses. And the insurance coverage that the 7% does have from the NFIP, effectively the sole provider of AM), http://thinkprogress.org/climate/2014/01/31/3230141/senate-flood-insurance/. Thus, the current system requires mandatory subsidization by all taxpayers for a flood insurance program that provides only poor coverage to the small percentage of homeowners who actually participate in the program.

154. One example of a state in which inland homeowners currently subsidize the homeowners insurance premium rates of coastal homeowners, due to the greater risks the coastal homeowners face, is Florida. See Cassandra R. Cole et al., *The Use of Postloss Financing of Catastrophic Risk*, 14 Risk MGMT. & INS. REV. 265, 267–71 (2011). The dominant homeowners insurance provider in Florida is Citizens Property Insurance Corporation, a state-sponsored program. See id. at 269. It offers property insurance to homeowners in Florida at premium rates, significantly lower than private insurers, because: (1) it does not need “to provide adequate returns to investors;” (2) it is tax exempt; (3) it does not need to raise excessive amounts of capital to pre-fund losses because it has the ability to do post-loss assessments; and (4) it is reinsured by a state-sponsored reinsurer, the Florida Hurricane Catastrophe Fund. See id. at 267–71. The premium rates charged to coastal residents are also subsidized by inland residents in varying amounts depending upon the location of the insured property. See id. at 280–85. Despite problems with the program in the past, and the fact that the state of Florida is periodically pounded by hurricanes, the program has successfully provided affordable property insurance to the state’s residents where private insurers generally have refused or failed to do so. See id. at 269.

flood insurance for homeowners in America, is not very good. It does not cover a significant amount of the losses that a typical flood victim incurs and the administrators of the NFIP, as well as their private claims handlers, are immune from liability no matter how poorly they handle claims or treat policyholders. The NFIP is also insolvent without a current path toward solvency.

Homeowners insurance, instead of the NFIP, would be a better vehicle through which to insure against the risk of flood losses. Far more Americans would have flood coverage at reasonable rates if coverage for flood losses were included in homeowners insurance and the coverage itself would be much better.

The justifications historically used to allow insurers to refuse to cover the risk of flood losses—adverse selection, moral hazard, and correlated risks—have much less force today than they did in 1968 when the NFIP was created. Insurance is now a global business by which the risk of loss of any individual policyholder or group of policyholders in any given geographic area is spread across a worldwide pool of premium-paying insureds and investors. Consequently, the time is ripe for the federal government to get out of the flood insurance business and for insurers to get back in the business.
TAKING BACK TAKINGS CLAIMS: WHY CONGRESS GIVING JUST COMPENSATION JURISDICTION TO THE COURT OF FEDERAL CLAIMS IS UNCONSTITUTIONAL

MICHAEL P. GOODMAN, PH.D.*

I. INTRODUCTION

THE federal government’s response to the global financial crisis of 2008 has led to a series of some of the largest dollar-value lawsuits ever filed against the federal government. One of those cases involves Fannie Mae and Freddie Mac.1 Both government sponsored enterprises faced a loss of investor confidence during the crisis that led to their placement into conservatorship and ultimately to the U.S. Treasury investing more than $100 billion in a new class of stock that guaranteed the government preferred status if they again became profitable.2 Another case involves General Motors and the Chrysler Corporation, which the government assisted by acquiring a 60.8% ownership interest in each. The government also, allegedly, required them to terminate agreements with franchisees as a condition of the car manufacturers receiving financial assistance.3 Perhaps the most notable of the bailout cases involves the American International Group, once a member of the Dow Jones Industrial Average, better known by its ticker symbol, AIG. In the midst of the crisis, AIG experienced a 95% plummet of its share price and was experiencing a liquidity crunch that threatened, as it just had for Lehman Brothers, to collapse the company. Then chairman of the Federal Reserve, Ben Bernanke, declared that AIG’s bankruptcy could have “triggered a 1930’s-style global financial and economic meltdown . . . .”4 The United States Government

---

* Frank H. Marks Visiting Associate Professor of Law, George Washington University School of Law. University of Mary Washington, B.A; Duke Law School, J.D.; Emory University, M.A./Ph.D. I would like to thank Robert Brauneis, Gregory Dolin, Kristina Caggiano Kelly, and the members of the University of Maryland’s junior faculty works in progress workshop for their comments to various drafts of this paper, to thank John Whealan for his support, and to thank my wife Melody for everything.

bailed out AIG; in exchange for financial assistance, the United States became a controlling lender of the company and acquired 80% of its stock. Not one of these cases has yet been resolved.

As many Americans expressed frustration with the bailouts, Congress got involved. Congress held numerous hearings during which Treasury Secretary Timothy Geithner presented the executive branch’s account of the crisis. While the Treasury and Congress defended the bailouts on the basis that the taxpayers got something in return for assisting those institutions, it was the creditors of Fannie and Freddie, the dealerships who lost their franchises, and the shareholders of AIG who each felt they had lost more than they gained. Accordingly, each group filed a complaint alleging that the government owes them compensation for having taken their property. Each of the bailout cases just discussed is based upon the Takings Clause and seeks “just compensation” pursuant to that constitutional provision. The Takings Clause states: “nor shall private property be taken for public use without just compensation.”

Despite presenting constitutional questions, those complaints could not be filed in the ninety-four federal district courts within the independent federal judiciary. Instead, the citizens were forced to file those claims in the United States Court of Federal Claims, a unique court created by and subject to the very governmental entities responsible for the bailouts.

The thesis of this Article is that while Congress may be able to relegate certain types of claims to a non-Article III court, such as the Court of Federal Claims, relegate takings claims to that entity is unconstitutional. This Article demonstrates why claims based upon that provision must be brought before Article III judges.

The next section of this Article introduces the Court of Federal Claims and explores how that court’s consideration of takings claims violates the values underlying Article III of the Constitution. The Article explores the various rationales the Court has used to justify the use of non-Article III courts and demonstrates why none of the rationales justify Congress’ current use of such a court for takings cases. There is no more

5. See Starr, 106 Fed. Cl. at 57.
7. See, e.g., Starr Complaint, supra note 4, at ¶¶ 10–12.
8. See City of Monterey v. Del Monte Dunes at Monterey, Ltd., 526 U.S. 687, 717 (1999) (“When the government repudiates this duty, either by denying just compensation in fact or by refusing to provide procedures through which compensation may be sought, it violates the Constitution.”).
10. This Article generally describes these entities as “Article I” courts, using that term interchangeably with “legislative” or “non-Article III,” although these courts are not always created pursuant to Congress’ Article I power, and are, at times, created pursuant to a specific provision found elsewhere, such as in Canter, which permitted the creation of territorial courts pursuant to Article IV, § 3, cl. 2. See Am. Ins. Co. v. Canter, 26 U.S. 511, 546 (1828).
essential time to evaluate the scheme Congress has established for considering large takings claims, as the Court of Federal Claims is now considering many claims, like the bailout cases, involving billions of dollars and great social import.

The third section of this Article explores how the current unconstitutional situation came to pass. The Supreme Court has never considered whether the Court of Federal Claims’ adjudication of takings claims is consistent with Article III of the Constitution. While the Court previously approved the Court of Federal Claims’ predecessor, the Court of Claims, considering takings claims,\(^\text{11}\) that entity was an Article III court, not a legislative court.\(^\text{12}\) In addition, whereas the Supreme Court once held that sovereign immunity principles justified Congress dealing with claims against the United States using a legislative court,\(^\text{13}\) the Court’s takings jurisprudence has established that sovereign immunity is inapplicable to claims brought under the “self-executing” Takings Clause.\(^\text{14}\) Indeed, Congress appears to have relegated takings claims to a non-Article III court as an inadvertent byproduct of other decisions it made when creating the Federal Circuit, and that effect is inconsistent with Congress’ expressed intent at the time.

The fourth section of the Article briefly explores some possible solutions to resolve this unconstitutional situation, including by granting takings case jurisdiction only to federal district courts. Those courts, staffed by Article III judges, are currently entrusted to protect each of the guarantees provided for in the Bill of Rights. Takings claims should receive no less protection.

II. Why Takings Claims Belong in Article III Courts

A. The Court of Federal Claims, Claims for Just Compensation, and Article III Values

The Court of Federal Claims operates much like a federal district court, but it deals exclusively with claims against the United States.\(^\text{15}\) Al-

\(^{11}\) See United States v. Causby, 328 U.S. 256, 267 (1946) (“If there is a taking, the claim is ‘founded upon the Constitution’ and within the jurisdiction of the Court of Claims to hear and determine.”).


\(^{13}\) See Williams v. United States, 289 U.S. 553, 580–81 (1933) (holding sovereign immunity principles justify Congress dealing with claims against United States using legislative courts).


though the court has the authority to adopt its own rules of procedure, in practice it has adopted many of the rules applicable in the federal district courts, and Congress has required some of those rules to be identical. Like district court decisions, the decisions of the Court of Federal Claims are final judgments.

The Court of Federal Claims is not entirely like the federal district courts, however. It is a specialized court with the unique responsibility, described in the Tucker Act:

[T]o render judgment upon any claim against the United States founded either upon the Constitution, or any Act of Congress or any regulation of an executive department, or upon any express or implied contract with the United States, or for liquidated or unliquidated damages in cases not sounding in tort.

Beyond that specialized jurisdictional grant, there are important differences between the Court of Federal Claims and the federal district courts. One major difference is that there is no possibility of a jury hearing citizens’ complaints in the Court of Federal Claims. Rather, the judges on the court only conduct bench trials.

Moreover, Congress did not create the Court of Federal Claims as an independent “constitutional” court pursuant to Article III of the Constitution. Instead, Congress explicitly provided, when creating it, that the new Court of Federal Claims is a “legislative court,” created pursuant to Article I. The distinction is one with a profound difference.

Article III of the Constitution, which establishes an independent judiciary, is one of the three pillars of the triumvirate federal government, based upon the concept of separation of powers. As the Supreme Court

---

17. See, e.g., id. § 460 (making applicable to Court of Federal Claims provisions for federal courts and judges described in 28 U.S.C. §§ 452–59, 462); Anderson v. United States, 344 F.3d 1343, 1350 n.1 (Fed. Cir. 2003) (recognizing that Court of Federal Claims applies Article III’s standing requirements).
19. Id. In the fiscal year ending in September 2013, the Court of Federal Claims issued decisions in 586 cases. See Admin. Office of the U.S. Courts, 2013 Annual Report of the Director: Judicial Business of the United States Courts tbl. G-2A (2013), available at http://www.uscourts.gov/uscourts/Statistics/JudicialBusiness/2013/appendices/G02ASep13.pdf. Nearly 30% of those cases involved contract disputes with the government or protests of government contracts; approximately 25% involved claims related to federal employee pay; more than 15% involved tax, copyright, patent claims, or cases filed by Native American tribes; and 20% fell into other categories. See id. The remaining 10% or so, 64 decisions, involved takings cases, such as the bailout cases, which fall within the Court of Federal Claims’ Tucker Act jurisdiction because they are “founded upon the Constitution.” See id. The court also decided 1,030 vaccine cases, which are not included in these statistics because of the different way in which they are handled. See id.
21. Id. § 171(a).
recently noted, Article III is “an inseparable element of the constitutional system of checks and balances” that “both defines the power and protects the independence of the Judicial Branch.” 22 The Court stressed that:

[T]he basic concept of separation of powers . . . that flow[s] from the scheme of a tripartite government adopted in the Constitution, the judicial Power of the United States . . . can no more be shared with another branch than the Chief Executive, for example, can share with the Judiciary the veto power, or the Congress share with the Judiciary the power to override a Presidential veto. 23

The entire purpose of Article III was to truly separate the judiciary from the other branches when we fear those other branches’ influence:

In establishing the system of divided power in the Constitution, the Framers considered it essential that “the judiciary remain[ )] truly distinct from both the legislature and the executive.” As Hamilton put it, quoting Montesquieu, “there is no liberty if the power of judging be not separated from the legislative and executive powers.” 24

To ensure that separation, and the independence of the courts, Article III creates two particular requirements, both of which came out of the Declaration of Independence’s complaints against King George, who had “made Judges dependent on his Will alone, for the tenure of their offices, and the amount and payment of their salaries.” 25 As incorporated into the Constitution, the requirements state: “[t]he Judges, both of the supreme and inferior Courts, shall hold their Offices during good Behaviour, and shall, at stated Times, receive for their Services, a Compensation, which shall not be diminished during their Continuance . . . .

---

23. Id. (alterations in original) (quoting United States v. Nixon, 418 U.S. 683, 704 (1974)) (internal quotation marks omitted); see also Gordon v. United States, 117 U.S. 697, 701 (1864) (“[T]o insure its impartiality it was absolutely necessary to make it independent of the legislative power, and the influence direct or indirect of Congress and the Executive. Hence the care with which its jurisdiction, powers, and duties are defined in the Constitution, and its independence of the legislative branch of the government secured.”).
24. Stern, 131 S. Ct. at 2608 (citation omitted) (quoting The Federalist No. 78, at 466 (Alexander Hamilton) (Clinton Rossiter ed., 1961)); see also Gordon, 117 U.S. at 706 (“In this distinct and separate existence (says Blackstone) of the judicial power in a peculiar body of men, nominated indeed but not removable at pleasure by the crown, consists one main preservative of public liberty, which cannot subsist long in any State unless the administration of common justice be in some degree separated from the legislative and executive power.”) (quoting 1 William Blackstone, Commentaries *268, *269) (internal quotation marks omitted)).
25. The Declaration of Independence, para. 10 (U.S. 1776).
in Office." The Good Behaviour Clause provides that judges who hear lawsuits in the Federal Judiciary serve lifetime appointments with no term limits. The Court has held that the Clause guarantees that judges can only be removed through impeachment. Of like importance, the Compensation Clause provides that judges can never have their salaries cut by those who control the other branches of government. As the Federalist states:

Next to permanency in office, nothing can contribute more to the independence of the judges than a fixed provision for their support. . . . In the general course of human nature, a power over a man’s subsistence amounts to a power over his will.

The Supreme Court has described the purpose of the prohibition upon reduction in salary by stating:

[T]he primary purpose of the prohibition against diminution was not to benefit the judges, but, like the clause in respect of tenure, to attract good and competent men to the bench and to promote that independence of action and judgment which is essential to the maintenance of the guaranties, limitations, and pervading principles of the Constitution and to the administration of justice without respect to persons and with equal concern for the poor and the rich.

In short, the Compensation Clause guarantees that a judge’s salary can never be reduced.

There are sixteen active judges who serve on the Court of Federal Claims. Those judges must live within fifty miles of the District of Co-

26. U.S. Const. art. III, § 1, cl. 2; see also Freytag v. Comm’r, 501 U.S. 868, 907 (1991) (“Like the President, the Judicial Branch was separated from Congress not merely by a paper assignment of functions, but by endowment with the means to resist encroachment—foremost among which, of course, are life tenure (during ‘good behavior’) and permanent salary. These structural accoutrements not only assure the fearless adjudication of cases and controversies, they also render the Judiciary a potential repository of appointment power free of congressional (as well as Presidential) influence.” (citations omitted)).

27. See The Federalist No. 78 (Alexander Hamilton); see also Stern, 131 S. Ct. at 2609 (describing requirement that federal judges be permitted to serve “without term limits”).


lumbia, but they may conduct proceedings anywhere within the United States. Because of Congress’ decision to establish that body as a legislative court, the Court of Federal Claims judges, unlike their federal district court brethren who serve lifetime appointments, serve only for fifteen years. Article III federal judges can only be removed by impeachment, which necessarily involves the legislative branch removing that judge. Judges of the Court of Federal Claims, however, can be removed by a majority vote of the judges of the appellate court that reviews the Court of Federal Claims’ decisions: the United States Court of Appeals for the Federal Circuit. The chief judge of the Court of Federal Claims, selected from amongst its members, serves, quite literally, at the pleasure of the President, who can replace that judge for any reason whatsoever. Each of the judges can be removed from the judgeship "for incompetency, misconduct, neglect of duty, engaging in the practice of law, or physical or mental disability."

The salaries of the active judges who serve on the Court of Federal Claims are currently tied to the salaries of district court judges, although there is no guarantee that they will continue to be, and Congress could reduce their salaries if it chooses. Because of the foregoing features of their employment, the judges who serve on the Court of Federal Claims are precisely what Congress labeled them: Article I judges unprotected by the guarantees of Article III of the Constitution.

While Article III creates a limit upon Congress’ authority to create non-Article III adjudicative bodies, it does not entirely prohibit Congress

---

33. Id. § 175.
34. Id. § 173.
35. See id. § 171(a) (“The court is declared to be a court established under article I of the Constitution of the United States.”); id. § 172(a) (“Each judge of the United States Court of Federal Claims shall be appointed for a term of fifteen years.”).
37. See 28 U.S.C. § 176. Retired members of the Court of Federal Claims are authorized to continue to hear cases as “senior” judges; there are, as of this writing, seven senior judges serving on the court. See Judges—Biographies, U.S. CT. OF FED. CLAIMS, http://www.uscfc.uscourts.gov/judges-biographies (last visited Nov. 4, 2014).
39. Id. § 176(a). Removal is effectuated by vote of a majority of the judges of the Court of Appeals for the Federal Circuit. Id. For a discussion of this unique power of an Article III court over the judges whose decisions it reviews, see Elizabeth I. Winston, Differentiating the Federal Circuit, 76 MO. L. R EV. 813, 830 (2011).
41. But see GREGORY C. SISK, L ITIGATION WITH THE FEDERAL GOVERNMENT 231–32 (4th ed. 2006) (arguing “that the Court of Federal Claims should be integrated more fully into the Judicial Branch by formally [being given] Article III status,” and contending that “[g]iven that a judge of the Court of Federal Claims upon expiration of his or her fifteen-year term may become a senior judge and thereby continue to act in a judicial capacity and receive a full salary, the court already has been given de facto Article III status by Congress”).
from creating courts that stray from Article III’s requirements. Although a literal reading of the text of Article III might suggest that Congress can never create non-Article III courts, at this late date in the jurisprudence of this area, “virtually no one considers a literal interpretation possible.” Indeed, Chief Justice Marshall first approved Congress’ authority to create courts and establish judgeships outside the boundaries of Article III in 1828 in *American Insurance Co. v. Canter*, when assessing the use of non-Article III courts in the territories that were not yet states. That decision relies upon the notion that in some circumstances, the policies underlying Article III are not implicated by Congress’ formation of an Article I court, or at least are not greatly curtailed. In an effort to explain the *Canter* holding, the Court later opined that the outcome of the case flowed from “the character of the early territories and some of the practical problems arising from their administration . . . .” The Court explained:

[T]he realities of territorial government typically made it less urgent that judges there enjoy the independence from Congress and the President envisioned by that article. For the territories were not ruled immediately from Washington; in a day of poor roads and slow mails, it was unthinkable that they should be. Because the other branches were unlikely to interfere with the running of the territorial courts, the Court reasoned, the territorial courts did not need the protections of Article III. The Court thus acknowledged the role of Article III in preserving the independence of the judiciary from the other branches of government, but suggested that the practical realities of administering the territories made the protections of Article III less necessary in that particular situation.

Every subsequent decision in which the Court has addressed whether a non-Article III entity impermissibly encroaches upon Article III has included at least some discussion of that provision’s purposes and values. Nonetheless, this factor has not always been given controlling weight, and after the Court’s approval of a pair of administrative structures in the mid-1980s, commentators openly questioned whether “the original structure

---

43. 26 U.S. 511 (1828).
44. See generally id.
46. Id. at 546.
47. See id.
and the values embodied in [Article III] are still regarded as important."50 Recently, in *Stern v. Marshall*,51 the Court said the “short but emphatic answer is yes.”52 After exploring the dual purposes of the constitutional command, separation of powers and protection of the individual, the Court explained that “Article III could neither serve its purpose in the system of checks and balances nor preserve the integrity of judicial decisionmaking if the other branches of the Federal Government could confer the Government’s ‘judicial Power’ on entities outside Article III.”53 Describing Article III as “the guardian of individual liberty and separation of powers,” the Court emphasized that “[a] statute may no more lawfully chip away at the authority of the Judicial Branch than it may eliminate it entirely.”54 Those values, the Court has explained, are to put into effect the concept of separation of powers and to guarantee the impartiality of judges, to the benefit of litigants.55 For takings claims, both of those values are not only fully implicated, they are at their apex.56

The separation of powers principle is, in a nutshell, an attempt “to protect each branch of government from incursion by the others.”57 As Chief Justice Marshall recognized in *Canter*, the risk of undue influence is not equal in all situations. In that case, he apparently felt that concern that the other branches would try to influence the courts was minimized because the other branches were unlikely, and indeed in all likelihood were unable, to interfere with the running of the territorial courts.58

---

52. *Id.* at 2620.
53. *Id.* at 2609.
54. *Id.* at 2615, 2620.
55. See *id.* at 2609 (“Article III protects liberty not only through its role in implementing the separation of powers, but also by specifying the defining characteristics of Article III judges.”); see also N. Pipeline Constr. Co. v. Marathon Pipe Line Co., 458 U.S. 50, 58 (1982) (plurality opinion) (characterizing Article III as “an inseparable element of the constitutional system of checks and balances” that “both defines the power and protects the independence of the Judicial Branch”).
56. This is in disagreement with Professor Sward, who has said: “[T]here is virtually no encroachment on Article III values because sovereign immunity would have shunted such claims to the legislature prior to the waiver of sovereign immunity and because an Article III court reviews the legislative court’s judgment. Thus, Congress’s determination to give citizens with claims against the government a relatively expeditious judicial determination of those claims in a non-Article III court is a reasonable one.
57. *Stern*, 131 S. Ct. at 2609 (quoting Bond v. United States, 131 S. Ct. 2355, 2365 (2011)).
practice White opined that in his view, the bankruptcy courts were more permissible for just this reason, as they “deal with issues likely to be of little interest to the political branches . . . .” In contrast to that slight risk of influence upon territorial courts, there are reasons to believe there is a significant risk that the other branches will care deeply about the outcomes of takings cases in the Court of Federal Claims. There are three factors that would tend to increase the government’s interest in a case and therefore the risk that the government will want to influence the case’s outcome: when the case involves the government, the Constitution, and the government’s money. All three are involved in takings claims.

First, the separation of powers principle is implicated more in cases involving the government as a party than in cases in which the government has no direct interest. As Justice Brennan recognized in *Northern Pipeline Construction Co. v. Marathon Pipe Line Co.*, “doubtless it could be argued that the need for independent judicial determination is greatest in cases arising between the Government and an individual.” Indeed, it is a truism that the government’s interests are most directly at issue in those cases in which the government is a party. But it is not only the fact that takings cases arise between the government and an individual that makes the possibility of pressure from the legislative or executive branches such a real concern. Many decisions adverse to the government can simply be overturned by legislative fiat. Because takings cases have a Constitutional basis, however, the elected bodies cannot overturn a takings decision, even if they want to. The concern that the elected branches would exert pressure over the judiciary, even if subtle, is even greater for constitutional cases between the government and an individual. Finally, the cases decided by the Court of Federal Claims, including takings cases, all involve money judgments. More particularly, they involve money that, if not used for judgments of the court, could be used for other congressional purposes. Even the most casual observer of Washington would agree that battles over money dominate beltway politics. The history of the Court of Federal Claims is a history of Congress attempting to maintain its influence over money judgments. Before there was a Court of Claims, Congress decided for itself whether to pay claims against the government. After the court was established, Congress did not cease its attempts to exert influ-

59. *See Northern Pipeline*, 458 U.S. at 115 (White, J., dissenting).
60. 458 U.S. 50 (1982) (plurality opinion).
61. Id. at 68 n.20.
62. *See Martin H. Redish, Legislative Courts, Administrative Agencies, and the Northern Pipeline Decision*, 1983 Duke L.J. 197, 224 (1983) (“The threat of domination by the political branches of government, as well as of imposition of majoritarian tyranny, is greatest in such cases, for it is only such decisions which the political branches are unable to overrule through simple legislative action. It is therefore those decisions which the political branches are most likely to attempt to influence.”).
63. For a discussion of Congress’ attempts to maintain its influence over the Court of Federal Claims, see *infra*, section III.
ence over those decisions. The well-known Pocono Pines Assembly Hotels Co. v. United States cases are instructive on this point.

Pocono Pines involved a claim filed in the Court of Claims by a property owner after its hospital was damaged in a fire during a government lease of the building. The government defended the case by arguing that the company had not met its burden to prove that the fire was the government’s fault. The Court of Claims ruled against the government and issued a final judgment in the amount of $227,239.53. Congress did not simply pay that judgment, however.

Instead, after it received the Comptroller General’s recommendation that Congress direct the Court of Claims to grant the government a new trial, Congress referred the case back to the Court of Claims with “instructions to find the facts and report them to the Senate, so that the Senate might conclude whether or not it would make an appropriation in this case.” The Court of Claims responded by docketing the case as a congressional reference matter. The property owner then filed a writ of mandamus in the Supreme Court to stop further proceedings in the Court of Claims—a request that was denied.

After that denial, the Court of Claims retried the case, again finding against the government, and reported the result to Congress. While the judgment was not altered by Congress’ actions, that case demonstrates just how important money judgments against the government can be to Congress. The entity that began its life as an institution reporting directly to Congress about whether to pay monetary claims against the government has never really shaken that role.

Pocono Pines and other examples like it demonstrate that Article III’s purpose of ensuring “that the acts of each [branch of government] shall never be controlled by, or subjected, directly or indirectly, to the coercive influence of either of the other departments,” is as much implicated by the Court of Federal Claims as it is any

64. Pocono Pines Assembly Hotels Co. v. United States (Pocono Pines II), 73 Ct. Cl. 447, 493 (1932); Pocono Pines Assembly Hotels Co. v. United States (Pocono Pines I), 69 Ct. Cl. 91, 99–100 (1930).
67. See id. at 106.
68. See id. at 110.
69. See Shimomura, supra note 65, at 675–76 (quoting 74 CONG. REC. 6076 (1931)) (internal quotation marks omitted).
70. See Pocono Pines II, 73 Ct. Cl. 447, 449 (1932).
72. See Pocono Pines Assembly Hotels Co. v. United States (Pocono Pines III), 76 Ct. Cl. 334, 352 (1932); see also 76 CONG. REC. 40, 60 (1932).
73. For a discussion of the Court of Claims’ institutional role in reporting to Congress regarding whether to pay monetary claims against the government, see infra notes 270–95 and accompanying text.
legislative entity. As the Supreme Court recently explained, those issues about which the other branches care deeply are precisely when an independent judiciary is most needed.\textsuperscript{75}

The Supreme Court has also said that in addition to maintaining the “checks and balances of the constitutional structure,” Article III also works “to guarantee that the process of adjudication itself remained impartial.”\textsuperscript{76} In other words, it provides “judges who are free from potential domination by other branches of government.”\textsuperscript{77} The \textit{Pocono Pines} judgment was for $227,239.53, which in today’s dollars would be a little more than $3.2 million dollars.\textsuperscript{78} One can only imagine how much interest Congress might show if the Court of Federal Claims were to award the AIG shareholders the $25 billion dollars they are seeking.

Still, the more insidious influence by the other branches is not the unlikely possibility that they would take any overt action, such as that taken in \textit{Pocono Pines}. It is that the other branches’ influence will be more subtle, perhaps even invisible to the judges themselves. Consider the power structure of the Court of Federal Claims. As noted earlier, the President can designate or remove the chief judge of the Court of Claims at will. The chief judge, in turn, has authority to decide which judge will hear any particular case and can replace the judge assigned to any case at will. Though the possibility of a replacement might be remote in any particular case, there is at least some concern that judges, aware that they might be replaced, may want to please their bosses and “get it right,” which may mean deciding in favor of the government. Even the appearance that such concerns might come into play may already affect the court’s credibility.

Focusing on the AIG shareholder example, the plaintiffs are seeking an enormous amount of money from the government and accusing many government officials, including former Treasury Secretary Timothy Geithner, of impropriety. The other branches’ interest in this case is great. Various members of Congress and the President have publicly discussed the bailouts, including the AIG bailout, multiple times. The Court of Federal Claims sits adjacent to the offices of the very individuals who are being accused of impropriety in the lawsuit. One can fairly say that the judge who conducts the proceeding sits both literally and figuratively in the shadow of the White House.

Even if there is no undue influence exerted by any members of Congress or executive officers, the structure lends itself to at least the appearance of impropriety. As the Court has recognized:

\begin{itemize}
\item \textsuperscript{75} See \textit{Stern v. Marshall}, 131 S. Ct. 2594, 2609 (2011).
\item \textsuperscript{76} \textit{N. Pipeline Constr. Co. v. Marathon Pipe Line Co.}, 458 U.S. 50, 58 (1982) (plurality opinion).
\item \textsuperscript{77} \textit{Id.} (quoting \textit{United States v. Will}}, 449 U.S. 200, 217–18 (1980)) (internal quotation marks omitted).
\end{itemize}
The sole function of the [Court of Claims] being to decide between the government and private suitors, a condition, on the part of the judges, of entire dependence upon the legislative pleasure for the tenure of their offices and for a continuance of adequate compensation during their service in office, to say the least, is not desirable.79

Because the values underlying Article III are strongly implicated by takings claims for large amounts of money, those claims should not be heard in non-Article III courts. The bailout cases currently being considered by the Court of Federal Claims are completely unlike the types of claims Chief Justice Marshall thought did not need to be heard by the Article III judiciary. Instead, they involve takings claims that strongly implicate the purpose of Article III. Accordingly, they are precisely the types of cases that must be heard in Article III courts.

B. None of the Justifications for the Use of Legislative Courts Validates the Court of Federal Claims’ Consideration of Takings Claims

The Supreme Court’s Article III jurisprudence is not a model of consistency, and the Court does not always speak with one voice. The Court’s consideration of when Congress may permit adjudication of a particular type of claim by a non-Article III body has generally involved cases with multiple dissents, has rarely achieved a strong majority in a particular case, and has caused many scholars and judges to suggest that the cases are incoherent.80 This Article does not attempt to criticize or evaluate the various rationales that the justices have relied upon and does not take sides in the debate about which of those factors, if any, should be determinative. Rather, this section reviews each of the factors the Court has used to permit Congress to stray from Article III’s requirements and applies those factors to the Court of Federal Claims’ consideration of takings claims. This review demonstrates that while the justices have not always agreed about the bounds of when Congress can permit non-Article III courts to adjudicate particular claims, none of those various frameworks or rationales that have ever been adopted by individual justices would permit the current framework that forces the bailout cases to the Court of Federal Claims. In short, it is not only poor policy for the Court of Federal Claims to consider takings claims, it is unconstitutional.

80. See Northern Pipeline, 458 U.S. at 90 (Rehnquist, J., concurring) (describing this jurisprudence as involving “frequently arcane distinctions and confusing precedents”); id. at 91 (“The cases dealing with the authority of Congress to create courts other than by use of its power under Art. III do not admit of easy synthesis.”); id. (describing Court’s Article III precedents as “landmarks on a judicial ‘darkling plain’ where ignorant armies have clashed by night”); Saphire & Solimine, supra note 50, at 85 (describing Court’s decisions about legislative courts as “amorphous and arcane”). Saphire and Solimine also describe the factors in the Court’s decisions as inarticulate and incoherent. See id. at 86.
1. Determining Just Compensation Is Not a Specific Congressional Power

The first case in which the Supreme Court endorsed Congress’ creation of non-Article III courts was the previously discussed case of *Canter*, wherein Chief Justice Marshall addressed Congress’ creation of “territorial courts.”  

*Canter* involved a cargo of cotton purchased through a judicial sale that had been ordered by the territorial court then established in Key West, Florida. Because that court was established by the territorial legislature and not established pursuant to Article III, the insurers claimed that the order was void. In ruling against the insurance company, the Chief Justice explained that the courts were not “constitutional Courts,” but were instead “legislative Courts” that need not comply with the requirements of Article III. The Court offered little by way of explanation for that holding, but what was stated provides the foundation for the first two justifications for the use of non-Article III courts. The first justification, already discussed, was the Court’s view that the failure of a claim to implicate Article III values weighs against the necessity of employing Article III courts. The second justification, the Court stated, was simply that the territorial courts were “created in virtue of the general right of sovereignty which exists in the government, or in virtue of that clause which enables Congress to make all needful rules and regulations, respecting the territory belonging to the United States.”

The *Canter* decision was the first in a series of cases holding that congressional authority to create non-Article III courts is derived from those congressional powers specifically enumerated in the Constitution. In *Canter*, the enumerated power was the power to “make all needful Rules and Regulations respecting the Territory or other Property belonging to the United States.” Shortly thereafter, the Court relied upon that same rationale when it sustained Congress’ creation of military courts pursuant to Congress’ specifically delineated Article I powers “to provide and maintain a Navy,” and “to make rules for the government of the land and naval forces.” The Court similarly approved Congress’ creation of the United States Court in the Indian Territory upon the basis that “[C]ongress possesses plenary power” over the tribes, and it affirmed that Congress may

---

82. Id. at 546.
83. Id.; see also U.S. Const. art. IV, § 3, cl. 2 (giving Congress power to “make all needful Rules and Regulations respecting the Territory or other Property belonging to the United States”).
84. See Saphire & Solimine, supra note 50, at 89 (collecting cases in which Court has justified Congress’ creation of non-Article III courts based upon its Article I powers).
85. See U.S. Const. art. IV, § 3, cl. 2. This rationale was also used by the Court to uphold the creation of the Court of Private Land Claims in 1894. See United States v. Coe, 155 U.S. 76, 85–86 (1894).
87. See Stephens v. Cherokee Nation, 174 U.S. 445, 478 (1899) (“The United States court in the Indian Territory is a legislative court, and was authorized to
create non-Article III consular courts based on its enumerated power to enter into treaties and deal with foreign countries. 88

In modern times, the Court has continued to consider whether Congress is effectuating a particular constitutional grant of power when deciding whether a legislative court is permissible. The Court’s most recent explicit reliance upon that rationale was in 1973, in *Palmore v. United States*, 89 when the Court reaffirmed that Congress may create non-Article III courts to adjudicate disputes within the District of Columbia based upon its Article I power to: “exercise exclusive Legislation in all Cases whatsoever, over such District (not exceeding ten Miles square) as may, by Cession of particular States, and the Acceptance of Congress, become the Seat of the Government of the United States . . . .”90

The Court later explained that this rationale applies when the subject matter considered by the courts at issue “involves a constitutional grant of power that has been historically understood as giving the political Branches of Government extraordinary control over the precise subject matter at issue.”91 The first rationale that emerges from the Court’s Article III jurisprudence is thus: if the subject with which an adjudicative body deals is one wholly within Congress’ purview, such as the rules governing military conduct, Congress need not concern itself with Article III.

While the Court found that rationale applicable in cases involving congressional power over the territories, the military, the tribes, and the District of Columbia, the Court explicitly rejected the notion that takings claims are the province of the legislature back in 1893. In *Monongahela Navigation Co. v. United States*, 92 the Court explained:

[W]hen the taking has been ordered, then the question of compensation is judicial. It does not rest with the public, taking the property, through congress or the legislature, its representative, to say what compensation shall be paid, or even what shall be the rule of compensation. The constitution has declared that just

exercise jurisdiction in these citizenship cases as a part of the machinery devised by Congress in the discharge of its duties in respect of these Indian tribes, and, assuming that congress possesses plenary power of legislation in regard to them, subject only to the constitution of the United States, it follows that the validity of remedial legislation of this sort cannot be questioned unless in violation of some prohibition of that instrument.”); see also Wallace v. Adams, 204 U.S. 415 (1907).


90. U.S. CONST. art. I, § 8, cl. 17; see also *Palmore*, 411 U.S. at 397–98 (relying upon this Article I provision in holding that Congress may create non-Article III courts to adjudicate disputes within District of Columbia).


92. 148 U.S. 312 (1893).
compensation shall be paid, and the ascertainment of that is a judicial inquiry.\textsuperscript{93}

The Takings Clause thus cannot be said to be “historically understood as giving the political Branches of Government” any control at all over the determination of just compensation. Rather, the historical understanding is that it grants that authority to the judiciary. The first rationale the Supreme Court used to permit Congressional use of legislative courts therefore does not appear to apply to takings claims.

2. \textit{Takings Claims Are Not “Public Rights”}

An early attempt to define the line between those types of controversies that implicate Article III and those that do not was the Supreme Court’s 1856 decision in \textit{Murray’s Lessee v. Hoboken Land and Improvement Co.}\textsuperscript{94} In that case, the Court held that a treasury official’s determination that certain property would be sold in order for the United States to collect a debt did not, at that time, involve a “judicial controversy” at all. An Article III judge was not, therefore, necessary.\textsuperscript{95} To contrast those types of cases that require an Article III judge with the types of adjudications that do not, the Court stated that Congress could not:

[B]ring under the judicial power a matter which, from its nature, is not a subject for judicial determination. At the same time there are matters, involving public rights, which may be presented in such form that the judicial power is capable of acting on them, and which are susceptible of judicial determination, but which congress may or may not bring within the cognizance of the courts of the United States, as it may deem proper.\textsuperscript{96}

\textsuperscript{93} Id. at 327.
\textsuperscript{94} 59 U.S. 272 (1855). This case did not precisely involve a legislative court. Rather, it involved the actions of an administrative official. Nonetheless, in assessing the scope of the executive official’s authority and what limitation, if any, Article III places upon the Executive, the Court referred to Article III’s limitations as if they apply equally to both legislative courts explicitly established by Congress as well as executive actions that indirectly implicate statutory commands originating from Congress. Some commentators have suggested that there is no reason why the mandate of Article III should apply any differently to an administrative agency or to a legislative court. \textit{See}, e.g., Redish, \textit{supra} note 62, at 201 (“[T]heir work cannot be functionally or theoretically distinguished.”). In \textit{Stern}, the Court suggested, however, that there may be reason to treat administrative agencies and legislative courts differently. The Court discussed the public rights doctrine as limited to “cases in which the claim at issue derives from a federal regulatory scheme, or in which resolution of the claim by an expert government agency is deemed essential to a limited regulatory objective within the agency’s authority.” \textit{Stern v. Marshall}, 131 S. Ct. 2594, 2613 (2011). As so described, this exception permits adjudication by a non-Article III body of some claims decided by an administrative agency but does not permit any adjudication by a legislative court.
\textsuperscript{95} See \textit{Murray’s Lessee}, 59 U.S. at 281.
\textsuperscript{96} Id. at 284.
Through that statement, the Court provided what is perhaps the single most important justification that the Supreme Court has offered for permitting Congress to establish Article I courts: sovereign immunity. 97 The rationale is that because the federal government generally enjoys sovereign immunity from suits, Congress need not permit its citizens to file lawsuits against the sovereign in the first place. Congress may choose to prohibit such lawsuits in any forum, and may therefore, if it chooses to permit the lawsuits at all, control the forum in which such suits may be brought. It may even relegate such lawsuits to a non-judicial forum. 98

One of the most important cases in which the Court expressly relied upon the sovereign-rights based “public rights” doctrine is the *Ex parte Bakelite Corp.* 99 decision of 1929. In Bakelite, the Court upheld Congress’ authority to establish the Court of Customs Appeals as an Article I court. The Court explained that Article I courts:

> [M]ay be created as special tribunals to examine and determine various matters, arising between the government and others, which from their nature do not require judicial determination and yet are susceptible of it. The mode of determining matters of this class is completely within congressional control. Congress may reserve to itself the power to decide, may delegate that power to executive officers, or may commit it to judicial tribunals. 100

As is clear from that passage, the Court viewed the fact that a matter is one “arising between the government and others” as a necessary, but not sufficient, condition to conclude that a right is a “public right.” The Court’s rationale for the public rights distinction was sovereign immunity: “The mode of determining matters of this class is completely within congressional control. Congress may reserve to itself the power to decide . . . .” 101

97. Professor Pfander asserts that *Murray’s Lessee* has been misread and does not actually stand for this proposition at all. See Pfander, supra note 42, at 731–38. Whatever may have been the Court’s intention in *Murray’s Lessee*, the case has come to stand for this proposition as applied by later decisions of the Court.

98. *See Murray’s Lessee*, 59 U.S. at 284. Like most—if not all—of the justifications the Court has deemed sufficient to legitimize legislative courts, this rationale has received its fair share of criticism. See, e.g., Redish, supra note 62, at 212–13. Professor Redish asserts that even if Congress is not obliged to permit a suit, the unconstitutional conditions doctrine suggests that once Congress has permitted a lawsuit, it may not condition the right to sue upon the waiver of Article III protections. See id. He also argues that it is not clear that Congress could, either legally or practically, decide all of the public rights issues. See id. at 213. Others have been less critical of the sovereign immunity rationale. See, e.g., Sward, supra note 56, at 1123 (“[T]his makes some sense given that in the absence of a court for such claims, citizens asserting a claim against the government would have to go to Congress itself, seeking private legislation to pay the claim.”). 99. 279 U.S. 438 (1929).

100. *Id.* at 451.

101. *Id.*
The public rights doctrine has been extensively discussed and applied in the Court’s recent Article III decisions. In *Northern Pipeline*, Justice Brennan’s plurality decision explicitly recognized that the primary justification for excluding public rights from Article III is “the traditional principle of sovereign immunity, which recognizes that the Government may attach conditions to its consent to be sued.” Justice Brennan also explained that, in his view, the public rights rationale is consistent with and “draws upon the principle of separation of powers, and a historical understanding that certain prerogatives were reserved to the political Branches of Government.” Thus, he explained, the public rights doctrine applies “only to matters that historically could have been determined exclusively by those departments.” This assessment is really no different from the rationale captured in Justice Brennan’s discussion of Congress’ Article I power, which, he explained, permits establishing non-Article III entities for areas that have “been historically understood as giving the political Branches of Government extraordinary control over the precise subject matter at issue.” Whether the justification can be categorized as falling within either the public rights exception to Article III or the first exception ultimately does not matter. The rationale is that when a matter is one that “the Framers expected that Congress would be free to commit . . . completely to nonjudicial executive determination . . . there can be no constitutional objection to Congress’ employing the less drastic expedient of committing their determination to a legislative court or an administrative agency.”

In sum, “public rights” are those matters involving disputes between an entity and the federal government to which sovereign immunity applies, such that Congress need not have permitted the lawsuit in the first place. They are “matters that could be conclusively determined by the Executive and Legislative branches,” as contrasted with “matters that are ‘inherently . . . judicial.’” The public rights rationale for permitting the

103. *Id.*
104. *Id.* at 68 (emphasis added).
105. *Id.* at 66.
106. *Id.* at 68.
107. The concept of sovereign immunity applied to the federal government has been roundly and repeatedly criticized. See, e.g., Eric Berger, *The Collision of the Takings and State Sovereign Immunity Doctrines*, 63 WASH. & LEE L. REV. 493, 550 (2006) (“Current sovereign immunity doctrine is very hard to square with a federalism premised on ‘the protection of individuals.’” (quoting *New York v. United States*, 505 U.S. 144, 181 (1992))). This Article takes no position about whether sovereign immunity should apply generally but only that it does not apply to takings claims against the federal government.
108. *Northern Pipeline*, 458 U.S. at 68 (alteration in original) (quoting *Ex parte Bakelite Corp.*, 279 U.S. 438, 458 (1929)); see also *Granfinanciera, S.A. v. Nordberg*, 492 U.S. 33, 68 (1989) (Scalia, J., concurring in part) (“[C]entral to our reasoning was the device of waiver of sovereign immunity . . . .”); *Ex parte Bakelite*, *Id.* at 68 (emphasis added).
creation of non-Article III courts has been heavily criticized on a number of fronts, with some commentators suggesting that the doctrine should be abandoned altogether.109

Nonetheless, the Court has not abandoned the sovereign immunity based rationale, declaring: “The rule that the United States may not be sued without its consent is all-embracing.”110 Still, despite such a broad proclamation, the public rights doctrine cannot justify the Court of Federal Claims considering takings claims.

For the majority of the cases that fall within the Court of Federal Claims’ jurisdiction, the sovereign immunity rationale arguably has some force.111 The majority of the cases heard in the Court of Federal Claims involve waivers of sovereign immunity and therefore fit into the “public rights” category that the Supreme Court has described.112 With respect to the majority of the cases for monetary compensation filed in the Court of Federal Claims, that rationale for permitting Congress to elect to have an Article I legislative court decide them has some common-sense appeal. After all, although the practice has been criticized, if the alternative is that Congress could choose not to permit the case in the first place, it is not irrational that Congress gets to choose its own forum when it magnani-

279 U.S. at 451 (explaining that public rights are matters “which from their nature do not require judicial determination and yet are susceptible of it. The mode of determining matters of this class is completely within congressional control. Congress may reserve to itself the power to decide, may delegate that power to executive officers, or may commit it to judicial tribunals.”).

109. For example, Professors Saphire and Solimine have decried the sovereign immunity factor as not “speak[ing] to the theoretical limits on Congress’s ability to assign public rights cases to non-article III tribunals.” Saphire & Solimine, supra note 50, at 116. Describing the factor as one that “merely suggests that the government has some leeway in determining whether Congress can assign the determination of rights it creates, and which it could adjudicate itself, to non-article III tribunals,” they contend that “[t]he separation of powers rationale begs the questions of whether and when congressional reliance on non-article III tribunals encroaches on the article III judicial power.” Id. Because they find the doctrine to be unhelpful, and lacking in historical justification, they propose that the Court “abandon the doctrine altogether.” Id. at 120. Professor Redish asserts that even if Congress is not obliged to permit a suit, the unconstitutional conditions doctrine suggests that once Congress has permitted a lawsuit, it may not condition the right to sue upon the waiver of Article III protections. See Redish, supra note 62, at 212–13.


111. Some commentators have suggested that even contract cases, the bread and butter of the Court of Federal Claims, do not involve public rights but should instead be treated as private rights. See, e.g., Ellen E. Sward & Rodney F. Page, The Federal Courts Improvement Act: A Practitioner’s Perspective, 33 Am. U. L. Rev. 385, 412 (1984) (“In its proprietary capacity, the government deals with citizens in the same way that individual citizens deal with each other. For example, the government enters into contracts and compensates persons for damages it causes. The mere fact that the government is a party to a contract does not reasonably suggest that public rights are implicated. In fact, the rights at issue are more in the nature of private rights.”).

112. See supra note 19.
mously allows its citizens to sue. But that rationale does not work for cases
that Congress could not have prevented a citizen from filing. If a citizen
could file a claim regardless of Congress’ permission to do so, then the
sovereign immunity rationale has no force, and Congress may not rely
upon this rationale to relegate a case to an Article I court. Such is the case
for takings claims, which do not involve waivers of sovereign immunity and
are therefore not public rights.

The first evidence that takings claims are not public rights is the
Court’s 1893 decision in Monongahela. As discussed earlier, the Court
therein held that determining just compensation is not a task for Congress
but is instead a “judicial inquiry.” That statement directly undermines
the notion that takings claims are public rights, which, the Court has said,
“do not require judicial determination . . . .” The Court has since made
even clearer that a waiver of sovereign immunity is not necessary for citi-
zens to file a takings claim.

In 1919, the United States requisitioned land to store supplies for the
Army and agreed to pay the owner, Seaboard Air Line Railway Company,
$235.80 plus 6% interest. Unhappy with that amount, Seaboard sued and
was awarded $6,000, as determined by a jury, plus 7% interest. The
government appealed the award of interest, arguing that the United States
had not consented to pay interest on takings, or any other, claims. The
court of appeals accepted the government’s position, holding that the
United States could not be forced to pay interest because, as a sovereign, it
is immune from payments other than those it has agreed to pay: such
“conditions necessarily arise in dealing with the sovereign, and for which
there is no redress.” The Supreme Court reversed, rejecting the gov-
ernment’s premise that the United States needed to waive its sovereign
immunity as to that claim. While recognizing that there was no statute
authorizing interest awards, the Court held that “[j]ust compensation is
provided for by the Constitution and the right to it cannot be taken away
by statute. Its ascertainment is a judicial function.” Finding that “[a]n
promise to pay is not necessary,” the Court ordered the government to pay
interest, which the district court had determined would be just com-

113. See Monongahela Navigation Co. v. United States, 148 U.S. 312, 327
(1893).
(1923).
116. See id. at 303.
117. See id.
119. See Seaboard, 261 U.S. at 304–06.
120. Id. at 304.
121. See id. at 304–05.
an express waiver of sovereign immunity," the Takings Clause creates an exception to that rule.\textsuperscript{122} The decision regarding Seaboard Air Line Railway Company stands for the proposition that sovereign immunity does not apply to takings claims.\textsuperscript{123}

Were there any doubt of that holding, the Court reiterated it a decade later, in 1933, in \textit{Jacobs v. United States}.\textsuperscript{124} After the Fifth Circuit Court of Appeals had denied claims for interest against the government on the basis that the claims were filed pursuant to the Tucker Act rather than as part of condemnation proceedings, the Court held:

This ruling cannot be sustained. The suits were based on the right to recover just compensation for property taken by the United States for public use in the exercise of its power of eminent domain. That right was guaranteed by the Constitution. The fact that condemnation proceedings were not instituted and that the right was asserted in suits by the owners did not change the essential nature of the claim. The form of the remedy did not qualify the right. It rested upon the Fifth Amendment. Statutory recognition was not necessary. A promise to pay was not necessary. Such a promise was implied because of the duty to pay imposed by the amendment. The suits were thus founded upon the Constitution of the United States.\textsuperscript{125}

Since 1933, it has thus been clear that takings claims filed pursuant to the Tucker Act do not require waivers of sovereign immunity. The Court has not strayed from that conclusion. Indeed, that takings claims are not limited by sovereign immunity makes perfect sense. For an action to constitute a taking in the first place, the government must exercise its power pursuant to its sovereign power of eminent domain.\textsuperscript{126} If the government could do so while simultaneously asserting that, as sovereign, it may not be sued to collect compensation for that taking, then the constitutional command, “nor shall private property be taken for public use, without just compensation,” would have no meaning whatsoever.\textsuperscript{127}

The Court has described this principle, that the Takings Clause has independent force without the government’s permission or waiver of sov-

\textsuperscript{123} The various states began to understand their respective state constitutions’ just compensation provisions as abrogating sovereign immunity around the same time \textit{Seaboard Air Line Railway Company} was decided in the 1920s. \textit{See} Robert Brauneis, \textit{The First Constitutional Tort: The Remedial Revolution in Nineteenth-Century State Just Compensation Law}, 52 \textit{VAND. L. REV.} 57, 138–39 (1999).
\textsuperscript{124} 290 U.S. 13 (1933).
\textsuperscript{125} \textit{Id.} at 16.
\textsuperscript{126} \textit{See, e.g.,} Berman v. Parker, 348 U.S. 26, 33 (1954).
\textsuperscript{127} \textit{See} U.S. \textit{CONST.} amend. V.
ereign immunity, as the “self-executing” nature of the Clause. The Court has explained that while the government can file a claim to condemn or formally take a property, the government may also take property by “physically entering into possession and ousting the owner,” in which case owners can also file “inverse condemnation” cases to seek the compensation to which they are entitled by the Fifth Amendment. The Court has explained that “[t]he owner’s right to bring such a suit derives from ‘the self-executing character of the constitutional provision with respect to condemnation . . . .’”

In a landmark decision, in which the Court explained that the Tucker Act is a jurisdictional statute that does not itself waive sovereign immunity as to the types of cases for which it grants jurisdiction, the Supreme Court was careful to distinguish takings claims from other types of claims against the United States. The Court explained the general rule that “[i]n a suit against the United States, there cannot be a right to money damages without a waiver of sovereign immunity,” and rejected a contrary rule based upon cases applying the Takings Clause. The Court distinguished the takings cases from the general rule by explaining that “[t]hese Fifth Amendment cases are tied to the language, purpose, and self-executing aspects of that constitutional provision, and are not authority to the effect that the Tucker Act eliminates from consideration the sovereign immunity of the United States.”

To the extent there was any lingering doubt that the Takings Clause trumps the government’s assertion of sovereign immunity, the Court removed that doubt in 1987, in First English Evangelical Lutheran Church of Glendale v. County of Los Angeles, California. In that case, the government argued, as amicus curiae, that the “Constitution did not work a surrender of the immunity of the States, and the Constitution likewise did not with-

129. Kirby Forest Indus., Inc. v. United States, 467 U.S. 1, 5 (1984); see, e.g., United States v. Lee, 106 U.S. 196 (1882). Lee is a bedrock case that involved the former estate of General Robert E. Lee, where Arlington Cemetery now sits. In that case, the United States had acquired the land for nonpayment of taxes, even though the taxes had in fact been paid. The owners filed an ejectment action against the government officials who held the land, and the Court held that the action was not one against the sovereign and therefore was not barred by sovereign immunity. See Malone v. Bowdoin, 369 U.S. 645, 645–46 (1962) (describing Lee). As the Court later described the Lee case, it demonstrated “the constitutional exception to the doctrine of sovereign immunity.” Larson v. Domestic & Foreign Commerce Corp., 337 U.S. 682, 696 (1949).
130. Kirby, 467 U.S. at 6 n.6 (alternation in original) (quoting Clarke, 445 U.S. at 257).
132. Id.
133. Id. at 401 (citation omitted) (citing Reg’t Rail Reorganization Act Cases, 419 U.S. 102 (1974); Jacobs v. United States, 290 U.S. 13, 16 (1933)).
hold this essential 'attribute[ ] of sovereignty' from the Government of the United States.”135 Rejecting that argument, the Court held that its cases “make clear that it is the Constitution that dictates the remedy for interference with property rights amounting to a taking.”136

It is worth noting that Congress established the Court of Federal Claims in 1982, five years before the First English decision. To the extent that Congress, like the government, misunderstood this aspect of the Takings Clause before that decision, there can be no misunderstanding now.137 After First English, it is now explicit that property owners enjoy the right to bring takings claims, not because Congress has consented to their doing so, but because the Constitution guarantees that right. It is the recognition of that self-executing provision that forecloses the Court of Federal Claims’ consideration of takings cases.

3. Review by the Federal Circuit Does Not Justify the Court of Federal Claims Deciding Takings Claims

After Bakelite created a fairly clear line between cases that can be considered by a non-Article III body (the public rights) and those that cannot (the private rights), the Court quickly complicated matters by holding that even private rights need not necessarily be considered by an Article III court. In Crowell v. Benson,138 the Court adopted a new rationale for the use of non-Article III judges, establishing that such tribunals may sometimes be permissible if the decisions of those tribunals are reviewed by Article III courts such that they are merely adjuncts of the Article III court.

Crowell involved a claim brought by an employee against his employer before the United States Employees’ Compensation Commission, a non-Article III entity created pursuant to a federal statute.139 Starting with the “public rights” distinction, the Court first noted that the claim was not...
against the government and was therefore a “private right.” Nonetheless, the Crowell Court upheld the authority of the United States Employees’ Compensation Commission to adjudicate such disputes between private parties, because an Article III judge of a U.S. District Court conducted a de novo review of the legislative agency’s decision. The Court explained that as long as an Article III court has the authority to ultimately decide the issue, then it is permissible for that “adjunct” legislative court to first decide the question.

The Court applied this doctrine in 1980, in United States v. Raddatz, when assessing the constitutionality of a provision of the Federal Magistrates Act. The Act permits a federal district court to submit a case to a magistrate, to conduct an evidentiary hearing and prepare findings of fact and recommendations that the district court judge then uses to rule upon a motion to suppress evidence, rather than conducting another evidentiary hearing. The Court upheld that procedure after recognizing that “the magistrate acts subsidiary to and only in aid of the district court. Thereafter, the entire process takes place under the district court’s total control and jurisdiction.” As the Court noted:

[T]he magistrate’s proposed findings and recommendations shall be subjected to a de novo determination “by the judge who . . . then exercise[s] the ultimate authority to issue an appropriate order.” Moreover, “[t]he authority—and the responsibility—to make an informed, final determination . . . remains with the judge.”

The Court explained that a magistrate’s role is not different from that of “masters in chancery or commissioners in admiralty where the proceeding is ‘constantly subject to the court’s control.’”

The task of describing the “essential attributes” of judicial power that must be retained in the Article III courts continued in Northern Pipeline, as Justice Brennan, writing for the plurality, applied this factor to the bankruptcy courts. He noted that, as the Court held in Crowell, for this rationale to apply, “the functions of the adjunct must be limited in such a way that ‘the essential attributes’ of judicial power are retained in the Art[icle]
III court."\textsuperscript{150} He also explained that the Court had permitted such adjunct fact finding in \textit{Raddatz} because the work of a magistrate judge is \textit{subject to de novo} review by the district court, which was free to rehear the evidence or to call for additional evidence.\textsuperscript{151}

In rejecting the notion that the bankruptcy courts were mere adjuncts, Justice Brennan listed some of the powers enjoyed by those courts, emphasizing their power to enforce their own judgments, rather than—like the agency the Court assessed in \textit{Crowell}—issuing orders that "could be enforced only by order of the district court."\textsuperscript{152} He also suggested that an Article III court reviewing the findings of the non-Article III entity under a "substantial evidence" standard weighs in favor of permitting the system, whereas the "clearly erroneous" review conducted by Article III courts of bankruptcy judges’ factual findings brought that adjudicatory body’s actions too close to the inherently judicial line.\textsuperscript{153}

In a concurring opinion, Justice Rehnquist noted that whatever the boundaries of the "adjunct" doctrine that originated in \textit{Crowell}, "traditional appellate review by Art. III courts" is not sufficient to make a court a permissible "adjunct."\textsuperscript{154} In dissent, Justice White explained that in his view, \textit{Crowell} rested upon the proposition that as long as questions of law were preserved for determination by Article III courts, the delegation of factual determinations to non-Article III judges is permissible.\textsuperscript{155}

In \textit{Stern}, the majority of the Court agreed with the notion "that \textit{Crowell} by its terms addresses the determination of \textit{facts} outside Article III."\textsuperscript{156} The \textit{Stern} majority also noted that the \textit{Northern Pipeline} decision had confined the permissible work of adjuncts to making "only specialized, narrowly confined factual determinations regarding a particularized area of law . . . ."\textsuperscript{157} In addition, the Court reaffirmed that the power to enter a final judgment, at least, is an essential attribute of the Article III judicial power.\textsuperscript{158} The Court found objectionable the bankruptcy courts’ "power to enter 'appropriate orders and judgments'—including final judgments—subject to review only if a party chooses to appeal."\textsuperscript{159} The Court

\textsuperscript{150.} \textit{Id.} at 81.
\textsuperscript{151.} \textit{Id.} at 79.
\textsuperscript{152.} \textit{Id.} at 85.
\textsuperscript{153.} \textit{See id.}
\textsuperscript{154.} \textit{See id.} at 91 (Rehnquist, J., concurring). Indeed, as Professor Resnik has noted, permitting appellate review to justify the use of non-Article III courts might support a transformation of Article III judges into administrators who supervise inferior judges. \textit{See} Judith Resnik, "Uncle Sam Modernizes His Justice": Inventing the Federal District Courts of the Twentieth Century for the District of Columbia and the Nation, 90 Geo. L.J. 607, 668–69 (2002).
\textsuperscript{155.} \textit{See Northern Pipeline}, 458 U.S. at 110 (White, J., dissenting).
\textsuperscript{156.} \textit{Stern v. Marshall}, 131 S. Ct. 2594, 2612 n.6 (2011).
\textsuperscript{157.} \textit{Id.} at 2618 (quoting \textit{Northern Pipeline}, 458 U.S. at 85) (internal quotation marks omitted).
\textsuperscript{158.} \textit{See id.} at 2619.
\textsuperscript{159.} \textit{Id.} (quoting 28 U.S.C. § 257(b)(1)).
concluded, “[g]iven that authority, a bankruptcy court can no more be deemed a mere ‘adjunct’ of the district court than a district court can be deemed such an ‘adjunct’ of the court of appeals.”

The fact that Court of Federal Claims decisions are reviewed by the Federal Circuit—which is an Article III appellate court—does not render the current structure permissible. The Court of Federal Claims judges exercise more judicial power than the bankruptcy judges that the Court held are not adjuncts. While Justice White noted, “a scheme of Art. I courts that provides for appellate review by Art. III courts should be substantially less controversial than a legislative attempt entirely to avoid judicial review in a constitutional court,” less controversial does not equal proper. Indeed, while Justice Brennan suggested that review by an Article III court may be a necessary condition for adjudication by a non-Article III body, no Justice has suggested that it is sufficient. And while commentators who have proposed appellate review theories of Article III have gone so far as to suggest that “sufficiently searching review of a legislative court’s or administrative agency’s decisions by a constitutional court will always satisfy the requirements of article III,” even the most ardent supporters of that theory would require a level of review from the Article III courts that is unavailable under the current scheme, wherein the Federal Circuit reviews the Court of Federal Claims.

Under any of the competing standards described in the Court’s recent decisions, the Court of Federal Claims cannot be considered an adjunct of the Federal Circuit. Under either the Northern Pipeline plurality’s or Stern majority’s rationale, the fact that Court of Federal Claims judges issue final decisions is fatal to their constitutionality. In contrast to the Stern majority’s description of permissible entities as those adjuncts making “only specialized, narrowly confined factual determinations regarding a particularized area of law,” the Court of Federal Claims issues judg-

160. Id. The Court has since reaffirmed that holding, stating that a bankruptcy court may enter proposed findings of fact and conclusions of law only if they are reviewed de novo by district courts. See Exec. Benefits Ins. Agency v. Arkison, 134 S. Ct. 2165, 2173 (2014).

161. This point has also been succinctly made by Sward and Page, who recognized that the Claims Court was much more like the bankruptcy courts that the Supreme Court invalidated in Northern Pipeline than a permissible adjunct. See Sward & Page, supra note 111, at 414–15.

162. Northern Pipeline, 458 U.S. at 115 (White, J., dissenting).

163. Richard H. Fallon, Jr., Of Legislative Courts, Administrative Agencies, and Article III, 101 HARV. L. REV. 915, 933 (1988). Professor Fallon would require de novo review of constitutional issues, including “the application of constitutional norms to particular facts . . . .” Id. at 976; see also Pfandner, supra note 42, at 743 (proposing unifying theory of “inferiority” of non-Article III tribunals and agreeing that “inferiority should require de novo review of any claim that Article I adjudication violates constitutional rights”).

164. See Stern, 131 S. Ct. at 2619; Northern Pipeline, 458 U.S. at 85.

165. Stern, 131 S. Ct. at 2618 (quoting Northern Pipeline, 458 U.S. at 85) (internal quotation marks omitted) (describing permissible entities).
ments dependent upon both fact and law in any area of law that arises during its consideration of takings claims. The Northern Pipeline plurality also found it impermissible that bankruptcy judges’ factual findings were reviewed only under a “clearly erroneous” standard. That is the same standard the Federal Circuit applies to review the Court of Federal Claims’ factual findings in takings decisions.

Even those commentators who have suggested that the primary consideration of whether an adjudication scheme comports with Article III should be the level of review available in an Article III court have suggested that review of fact finding should be done under a less deferential standard than is currently applied to the Court of Federal Claims’ fact finding. The Federal Circuit review that is currently available would also fail to meet the standard of the Northern Pipeline concurring opinion, which stated that “traditional appellate review by Art. III courts” is not sufficient to make a court a permissible “adjunct.”

Finally, while the Northern Pipeline dissent suggested that the bankruptcy judges qualified as adjuncts because questions of law were reserved to Article III judges, Court of Federal Claims judges decide both issues of fact and law, and therefore would not meet even that standard. Indeed, the Court of Federal Claims is not an adjunct of the Federal Circuit any more than all district courts are adjuncts of courts of appeal. Thus, like the other rationales, this one is unavailable to justify the current scheme.

4. Takings Claims Are Neither Created by Congress nor Closely Intertwined with a Federal Regulatory Program Congress Has Enacted

Another factor, which is an extension of the previously described “public rights” doctrine, is the notion that Congress has augmented authority to use non-Article III entities to resolve disputes when they involve rights Congress has itself created. This is the consideration that justifies much of the administrative state.

When discussing the public rights doctrine in his Northern Pipeline plurality decision, Justice Brennan stressed that one limitation upon the public rights doctrine is that it involves matters that “at a minimum arise

166. See Northern Pipeline, 458 U.S. at 85–86.
167. See, e.g., City of El Centro v. United States, 922 F.2d 816, 819 (Fed. Cir. 1990).
168. See, e.g., Saphire & Solimine, supra note 50, at 144 (“While the ‘clearly erroneous’ standard may be appropriate for appellate review of the factual findings of an article III trial judge, it seems entirely too deferential for other contexts, and would reduce article III review—the only time a litigant will have her case before an article III tribunal—to little more than a formality.”). In professors Saphire and Solimine’s view, the Bankruptcy Act’s reliance on that standard is itself reason enough to invalidate the Act. See id. at 147.
169. Northern Pipeline, 458 U.S. at 91 (Rehnquist, J., concurring).
170. See id. at 110 (White, J., dissenting).
He also opined, however, that “it is clear that when Congress creates a substantive federal right, it possesses substantial discretion to prescribe the manner in which that right may be adjudicated—including the assignment to an adjunct of some functions historically performed by judges.”

Elaborating upon that principle, he stated:

[W]hen Congress creates a statutory right, it clearly has the discretion, in defining that right, to create presumptions, or assign burdens of proof, or prescribe remedies; it may also provide that persons seeking to vindicate that right must do so before particularized tribunals created to perform the specialized adjudicative tasks related to that right. Such provisions do, in a sense, affect the exercise of judicial power, but they are also incidental to Congress’ power to define the right that it has created.

Justice Brennan contrasted that situation with one in which the rights at issue are not created by Congress. In that case, he explained, Congress has less discretion to assign fact finding related to that issue to a non-Article III entity.

This factor is not unlike the sovereign immunity rationale underlying the public rights exception. In the public rights context, the rationale is that if Congress permits suits against the government when it had no obligation to do so, Congress may specify how those lawsuits proceed. Similarly, the rationale here is that if Congress creates rights between private parties, again when it had no obligation to do so, Congress may specify how lawsuits involving those rights proceed. This doctrine is thus an extension of the public rights rationale described earlier, and in modern cases, the courts describe the public rights doctrine as including both categories of cases, that is, both those disputes that involve the government as a party as well as those involving private entities’ dispute over a federally created right.

171. Id. at 69 (plurality opinion) (quoting Ex parte Bakelite Corp., 279 U.S. 438, 451 (1929)).
172. Id. at 80 (emphasis added).
173. Id. at 83 (footnote omitted).
174. See id. at 81–82; see also id. at 84 (“Congress’ authority to control the manner in which [a non-congressionally created] right is adjudicated, through assignment of historically judicial functions to a non-Art. III ‘adjunct,’ plainly must be deemed at a minimum.”); id. at 71 (describing public rights as involving “congressionally created benefits”). Although Justice Brennan’s discussion of the distinction between congressionally created, statutory rights and rights not of congressional creation took place within the context of the “adjunct” exception, the distinction has had independent applicability to cases not involving entities purporting to be adjuncts.
175. See, e.g., Thomas v. Union Carbide Agric. Prods. Co., 473 U.S. 568, 596 (1985) (Brennan, J., concurring) (recognizing that public rights doctrine, “as that concept had come to be understood,” includes disputes “arising from the Federal Government’s administration of its laws or programs”). For a discussion of the
The Court overtly relied upon this rationale in *Thomas v. Union Carbide Agricultural Products Co.*,176 when it upheld Congress’ decision to require binding arbitration within the federal regulatory scheme involving pesticides.177 There, the Court, describing the doctrine as the public rights doctrine, stated:

In essence, the public rights doctrine reflects simply a pragmatic understanding that when Congress selects a quasi-judicial method of resolving matters that “could be conclusively determined by the Executive and Legislative Branches,” the danger of encroaching on the judicial powers is reduced.178

The Court then noted that the dispute in that case involved a right that was created by Congress,179 and explained that, accordingly, “Congress, without implicating Article III, could have authorized EPA to charge follow-on registrants fees to cover the cost of data and could have directly subsidized [ ] data submitters for their contributions of needed data.”180 Because the dispute therefore involved a function that the Court described as “essentially legislative,” the Court held that it could be resolved by non-Article III entities.181

Shortly after that decision, in *Commodity Futures Trading Commission v. Schor*,182 the Court extended the doctrine beyond those rights actually created by Congress to also permit a non-Article III entity to consider additional counterclaims that were not created by Congress, as long as adjudication of those additional claims is necessary in order to effectively adjudicate the congressionally created right.183 The Court stated:

Congress, acting for a valid legislative purpose pursuant to its constitutional powers under Article I, may create a seemingly pri-

evolution of the public rights doctrine, see Gordon G. Young, *Public Rights and the Federal Judicial Power: From Murray’s Lessee Through Crowell to Schor*, 35 Buff. L. Rev. 765 (1986); see also Granfinanciera S.A. v. Nordberg, 492 U.S. 33, 68 (1989) (Scalia, J., concurring in part) (“It is clear that what we meant by public rights were not rights important to the public, or rights created by the public, but rights of the public—that is, rights pertaining to claims brought by or against the United States. For central to our reasoning was the device of waiver of sovereign immunity, as a means of converting a subject which, though its resolution involved a ‘judicial act,’ could not be brought before the courts, into the stuff of an Article III ‘judicial controversy.’ Waiver of sovereign immunity can only be implicated, of course, in suits where the Government is a party. We understood this from the time the doctrine of public rights was born, in 1856, until two Terms ago . . . .”).

177. See id. at 589–94.
178. Id. at 589 (quoting *Northern Pipeline*, 458 U.S. at 68).
179. See id.
180. Id. at 590.
181. See id. (citing St. Joseph Stockyards Co. v. United States, 298 U.S. 38, 49–53 (1936)).
183. See id. at 857.
vate right that is so closely integrated into a public regulatory scheme as to be a matter appropriate for agency resolution with limited involvement by the Article III judiciary.\footnote{184}{Id. at 840 (quoting \textit{Thomas}, 473 U.S. at 593) (internal quotation marks omitted).}

In \textit{Granfinanciera, S.A. v. Nordberg},\footnote{185}{492 U.S. 33 (1989).} the Court again described this extension of the public rights doctrine.\footnote{186}{See id. at 54–56.} While the dispute was between two private parties and did not technically involve the federal government, Justice Brennan, writing for the majority, described the public rights doctrine as broad enough to encompass litigation between private parties if that litigation is pursuant to a complex regulatory scheme.\footnote{187}{See id.} The Court described the doctrine as involving those rights that are “closely intertwined with a federal regulatory program Congress has power to enact . . . .”\footnote{188}{Id. at 54.}

Takings claims do not fit into this category. Unlike those subject matters that Congress may have a justifiable basis for keeping under its thumb, Congress did not create takings claims. Unlike the “essentially legislative” task at issue in \textit{Thomas}, deciding takings claims is a “judicial inquiry.”\footnote{189}{Compare \textit{Thomas}, 473 U.S. at 590, with Monongahela Navigation Co. v. United States, 148 U.S. 312, 327 (1893).}

Commentators have previously expressed reasonable concern when Article III has been interpreted rigidly, such that it could be said to “delegitimize[ ] many of the institutions of the modern administrative state.”\footnote{190}{Saphire & Solimine, supra note 50, at 112.} It is worth mentioning, therefore, that although this Article calls into question the ability of the Court of Federal Claims to adjudicate takings claims, it has no effect on the myriad administrative agencies that the Court has approved or not yet addressed. Most, if not all, of those agencies are primarily tasked with adjudicating statutorily created rights and fall within this exception to Article III. In addition, their consideration of issues beyond their narrow area of expertise, is, as the Court explained in \textit{Schor}, “incidental to, and completely dependent upon, adjudication of [ ] claims created by federal law, and in actual fact is limited to claims arising out of the same transaction or occurrence as the [ ] claim.”\footnote{191}{Commodity Futures Trading Comm’n v. Schor, 478 U.S. 833, 856 (1986).} As some commentators have described them, they are “adjudicatory schemes designed to solve unique and unforeseen [sic] problems.”\footnote{192}{Saphire & Solimine, supra note 50, at 112.}
In contrast, the Court of Federal Claims is exceptional in that it is currently the only non-Article III entity being asked to adjudicate a constitutional, as opposed to statutory, right. The constitutional problem identified in this Article therefore does not encompass Congress’ use of agency adjudicators in the various areas in which they are currently employed. Unlike the various administrative entities that fit within this exception, the justification does not apply to the Court of Federal Claims considering takings cases.

5. The Court of Federal Claims’ Deciding Takings Claims Fails the Modern Balancing Tests

While this Article has already touched upon many of the Court’s recent decisions when describing the different factors that the Court has considered in this area, some of those decisions warrant slightly more expansive consideration, as they demonstrate how the Court has combined and weighed the factors in recent years.

In Northern Pipeline, the Court considered whether it was constitutional for Congress to give bankruptcy court judges the authority to consider state law claims related to bankruptcy, in addition to their more traditional role in considering the bankruptcy itself. The bankruptcy judges in question were appointed by the President to fourteen-year terms with the advice and consent of the Senate, were subject to removal by a “judicial council of the circuit,” and were issued salaries subject, at least in theory, to periodic adjustment. The judges thus clearly were not Article III judges.

The Justices took very different approaches to deciding whether the bankruptcy court was acting contrary to constitutional command. In the plurality opinion, Justice Brennan offered what has been described as

---


194. See id. at 53.

195. See id. at 60.

196. Continuing the widespread criticism of the Court’s Article III jurisprudence, Professor Redish described all of the opinions as adopting guidelines that are “without legitimate basis in logic or in constitutional language, history, or policy.” Redish, supra note 92, at 200; see also Eric G. Behrens, Stern v. Marshall: The Supreme Court’s Continuing Erosion of Bankruptcy Court Jurisdiction and Article I Courts, 85 AM. BANKR. L.J. 387, 396–97 (2011) (describing Northern Pipeline decision as “controversial” and “analytically excoriated”); Erwin Chemerinsky, Ending the Marathon: Is It Time to Overrule Northern Pipeline, 65 AM. BANKR. L.J. 311, 311–12 (1991).
a “categorical approach” to the problem. 197 Justice Brennan first explained that the grant of power Congress had entrusted to bankruptcy judges did not derive from an “exceptional grant of power,” such as Congress’ authority over the District of Columbia, the Territories, or the nation’s military forces. 198 To fall within that category, Justice Brennan explained, the subject matter relegated to a non-Article III entity must be one that “involves a constitutional grant of power that has been historically understood as giving the political Branches of Government extraordinary control over the precise subject matter at issue.” 199 The plurality also rejected the possibility that the grant of power fits within the public rights exception to Article III, noting that public rights are those that arise “between the Government and persons subject to its authority in connection with the performance of the constitutional functions of the executive or legislative departments.” 200 Finally, the plurality rejected the notion that the bankruptcy courts were permissible “adjuncts” to Article III courts. 201 Considering the breadth of the powers of the bankruptcy court and the lack of de novo review of that court’s conclusions by Article III courts, the plurality concluded that they were not acting solely as adjuncts. 202

In dissent, Justice White rejected the plurality’s attempts to approach the problem based upon the type of cases that Congress sent to a non-Article III court, suggesting that “[t]here is no difference in principle between the work that Congress may assign to an Art. I court and that which the Constitution assigns to Art. III courts.” 203 Nonetheless, Justice White did not believe that Congress may relegate cases to legislative courts whenever it wants. Instead, he proposed a balancing test in which the Court would weigh the values underlying Article III against the legislative interest in having such cases heard by non-Article III courts:

I do not suggest that the Court should simply look to the strength of the legislative interest and ask itself if that interest is more compelling than the values furthered by Art. III. The inquiry should, rather, focus equally on those Art. III values and ask whether and to what extent the legislative scheme accommodates them or, conversely, substantially undermines them. The burden


198. See Northern Pipeline, 458 U.S. at 70–71.

199. Id. at 66.


201. See id. at 87.

202. See id. at 84–86.

203. Id. at 113 (White, J., dissenting).
on Art. III values should then be measured against the values Congress hopes to serve through the use of Art. I courts.\textsuperscript{204} Justice White then proposed two considerations relevant to conducting that balancing test. First, legislative courts reviewed by Article III courts should be “less controversial” than legislative courts that “entirely avoid judicial review in a constitutional court.”\textsuperscript{205} Second, Article I courts would generally be permissible when they are “designed to deal with issues likely to be of little interest to the political branches, [as] there is less reason to fear that such courts represent a dangerous accumulation of power in one of the political branches of government.”\textsuperscript{206} Applying those considerations to the bankruptcy courts, Justice White noted that there was substantial opportunity to appeal bankruptcy judges’ decisions both to district courts and courts of appeals, that “[b]ankruptcy matters are, for the most part, private adjudications of little political significance,” and that the “tremendous increase in bankruptcy cases” combined with the “extreme specialization” of bankruptcy proceedings justified Congress’ creation of legislative courts to deal with those issues.\textsuperscript{207}

Shortly after \textit{Northern Pipeline}, the Court issued \textit{Thomas}, wherein it considered whether Congress violated Article III when it decided to require binding arbitration as the mechanism for resolution of disputes between manufacturers who registered their pesticides with the government.\textsuperscript{208} In the majority decision, the Court held that resolution of the question of the appropriateness of using a non-Article III body to resolve disputes turns upon “practical attention to substance rather than doctrinaire reliance on formal categories . . . .”\textsuperscript{209} The Court therefore applied a balancing test, as Justice White had proposed in his dissent in \textit{Northern Pipeline}.\textsuperscript{210} On the one hand, the Court weighed “the nature of the right at issue,” namely that the dispute involved a right Congress had created, as well as its opinion that the con-

\textsuperscript{204} Id. at 115. The first criticism of the balancing test was the \textit{Northern Pipeline} plurality decision, which describes the balancing approach as one “in which Congress can essentially determine for itself whether Art. III courts are required.” Id. at 70 n.25 (plurality opinion). The problem with assessing the values underlying Article III on a case-by-case basis is that, when those values are weighed against the immediate legislative interests involved when Congress creates any particular non-Article III body, the concrete values that drove Congress to that decision will necessarily be more recognizable than the Article III interests, which are designed to prophylactically protect against subtle pressure on judges. “Thus, any case-by-case balancing process will always tend to find the benefit of maintaining these protections illusory.” Redish, supra note 62, at 222.

\textsuperscript{205} Id., 458 U.S. at 115 (White, J., dissenting).

\textsuperscript{206} Id.

\textsuperscript{207} Id. at 116–18.


\textsuperscript{209} Id. at 587.

\textsuperscript{210} See id. at 589–90.
cerns motivating Congress to choose a non-Article III method of resolution were reasonable, in particular Congress’ concern that delaying pesticide registrations would endanger public health. Against those considerations, the Court weighed the purpose of Article III, concluding that the system Congress had established did not threaten the independent role of the judiciary or “diminish the likelihood of impartial decision-making, free from political influence.”

The next year, the Court extended its application of the “pragmatic” approach used in Thomas when it decided Schor. The case raised the question of whether the Commodity Futures Trading Commission’s (CFTC) consideration of state law counterclaims violates Article III. Holding that it did not, a majority of the Court adopted a balancing approach similar to that Justice White had described in Northern Pipeline, explaining that “the constitutionality of a given congressional delegation of adjudicative functions to a non-Article III body must be assessed by reference to the purposes underlying the requirements of Article III.” That inquiry, the Court explained, is in turn guided by “practical attention to substance rather than doctrinaire reliance on formal categories . . . .”

The Court began its inquiry by noting that one of the purposes of Article III is to protect a litigant’s personal interest in having his or her claim “decided before judges who are free from potential domination by other branches of government.” The Court explained that because that interest is a “personal” interest, it can be waived, and that, in the instant case, the litigant had “indisputably waived any right he may have possessed to the full trial of [the commodity future broker’s] counterclaim before an Article III court” because he had “expressly demanded that [the commodity future broker] proceed on its counterclaim” before the CFTC. The Court also held that Article III serves the separate structural interest of “safeguard[ing] the role of the Judicial Branch in our tripartite system by barring congressional attempts ‘to transfer jurisdiction [to non-Article III tribunals] for the purpose of emasculating’ constitutional courts,” an interest that cannot be waived. Despite that personal

211. See id. at 590.
212. Id.
214. See id.
215. Id. at 847.
216. Id. at 848 (quoting Thomas, 473 U.S. at 587) (internal quotation marks omitted).
218. Id. at 848–50.
219. Id. at 849.
220. Id. at 850 (second alteration in original) (quoting Nat’l Ins. Co. v. Tide-water Co., 337 U.S. 582, 644 (1949) (Vinson, C.J., dissenting)); see also id. at 851 (”[T]he parties cannot by consent cure the constitutional difficulty for the same
waiver, the question of whether Congress has exceeded its authority in
designating an issue for non-Article III adjudications cannot be waived.
Therefore, the Court then turned to that question.

As the Court described the problem, it is one of determining “the
extent to which a given congressional decision to authorize the adjudica-
tion of Article III business in a non-Article III tribunal impermissibly
threatens the institutional integrity of the Judicial Branch . . . .”221 To
reach that question, the Court purported to weigh many of the considera-
tions the Court has recognized over the years, bundling them into three
general factors.

The first factor is “the extent to which the ‘essential attributes of judicial
court’ are reserved to Article III courts, and, conversely, the extent to
which the non-Article III forum exercises the range of jurisdiction and
powers normally vested only in Article III courts . . . .”222 Applying that
first factor to the CFTC, the Court found multiple features to weigh in
favor of permitting that body to decide state law counterclaims:

- It deals only with a “particularized area of law.”223 This is in
  contrast to the bankruptcy courts’ consideration of all claims
  “related to” cases under the bankruptcy code, which the Court
  invalidated in Northern Pipeline.224

- Its orders “are enforceable only by order of the district court,”
in contrast to the bankruptcy court’s authority to enforce its
  own orders.225

- Its orders are reviewed under the less deferential “weight of the
evidence” standard, as opposed to the more deferential clearly
  erroneous standard.226

- It does not exercise “all ordinary powers of district courts.”227

reason that the parties by consent cannot confer on federal courts subject-matter
jurisdiction beyond the limitations imposed by Article III, § 2. When these Article
III limitations are at issue, notions of consent and waiver cannot be dispositive
because the limitations serve institutional interests that the parties cannot be ex-
eted to protect.” (citation omitted) (citing United States v. Griffin, 303 U.S. 226,
229 (1938)).

221. Id.
222. Id.
223. Id. at 852 (quoting N. Pipeline Constr. Co. v. Marathon Pipe Line Co.,
458 U.S. 50, 85 (1982) (plurality opinion)) (internal quotation marks omitted).
224. See id. at 852–53 (quoting 28 U.S.C. § 1471(b); Northern Pipeline, 458 U.S.
at 85).
225. Compare id. at 853, with Northern Pipeline, 458 U.S. at 85–86.
226. Compare Schor, 478 U.S. at 853, with Northern Pipeline, 458 U.S. at 85.
227. Schor, 478 U.S. at 853 (quoting Northern Pipeline, 458 U.S. at 85) (internal
quotation marks omitted).
The second factor the Court considered is “the origins and importance of the right to be adjudicated . . . .”\(^{228}\) Applying that factor, the Court recognized that the state law counterclaim involved a private right of the type that was historically subject to resolution by Article III courts and therefore warranted a searching inquiry.\(^{229}\) Conducting that “searching” inquiry, the Court found it important that Congress had not entirely removed adjudication of the right at issue from the Article III courts but that, instead, “Congress gave the CFTC the authority to adjudicate such matters, but the decision to invoke this forum is left entirely to the parties and the power of the federal judiciary to take jurisdiction of these matters is unaffected.”\(^{230}\) The Court concluded: “Congress may make available a quasi-judicial mechanism through which willing parties may, at their option, elect to resolve their differences.”\(^{231}\)

The third, and last, factor the Court considered was “the concerns that drove Congress to depart from the requirements of Article III.”\(^{232}\) Here, the Court noted that in authorizing the CFTC’s jurisdiction, Congress’ goal was “to create an inexpensive and expeditious alternative forum,” and the “primary focus was on making effective a specific and limited federal regulatory scheme, not on allocating jurisdiction among federal tribunals.”\(^{233}\) Also weighing in favor of permitting the adjudicatory scheme was “the perception that the CFTC was relatively immune from political pressures,” and “the obvious expertise that the Commission possesses” in the area in which it adjudicates.\(^{234}\) Finally, the Court stressed that the objectionable portion of the CFTC’s jurisdiction, its authority to decide common law counterclaims, “is incidental to, and completely dependent upon, adjudication of reparations claims created by federal law, and in actual fact is limited to claims arising out of the same transaction or occurrence as the reparations claim.”\(^{235}\)

The Court considered another challenge to the constitutionality of a non-Article III entity in its 2011 decision in *Stern v. Marshall*.\(^{236}\) In *Stern*, the question was whether a bankruptcy court’s consideration of a counter-

\(^{228}\) Id. at 851.
\(^{229}\) See id. at 853–54.
\(^{230}\) Id. at 855.
\(^{231}\) Id. The Court’s decision appears to rely upon the fact that the district courts retained jurisdiction to consider the same claims. See id. at 862–65 (Brennan, J., dissenting) (describing majority’s approach as relying upon fact that “Congress does not altogether eliminate federal-court jurisdiction over ancillary state-law counterclaims,” and that “the CFTC shares in, rather than displaces, federal district court jurisdiction over these claims . . . .”).
\(^{232}\) Id. at 851 (majority opinion).
\(^{233}\) Id. at 855.
\(^{234}\) Id. at 855–56.
\(^{235}\) Id. at 856.
\(^{236}\) See Stern v. Marshall, 131 S. Ct. 2594 (2011). In the interim, the Court held that it was unconstitutional for an Article IV judge, the Chief Judge of the District Court for the Northern Mariana Islands, to sit by designation on the Ninth Circuit Court of Appeals. See Nguyen v. United States, 539 U.S. 69, 72 (2003).
claim of tortious interference violates Article III. The majority began and ended its analysis of that question by reflecting upon the purpose of that constitutional provision. The Court also assessed whether the case fit within any of the categories outlined above.

First, the Court discussed the suggestion that the proceeding involved public rights. The Court began by rejecting the applicability of the original understanding of the public rights doctrine to the counterclaim, noting that it is “not a matter that can be pursued only by grace of the other branches,” or “one that ‘historically could have been determined exclusively by’ those branches,” but was instead one that “does not ‘depend[] on the will of congress;’ Congress has nothing to do with it.”

The Court then concluded that it also did not fall into one of the extensions of the public rights doctrines, as it “does not flow from a federal statutory scheme,” and “is not ‘completely dependent upon’ adjudication of a claim created by federal law . . . .” The Court concluded that the counterclaim “does not fall within any of the varied formulations of the public rights exception in this Court’s cases.” In so concluding, the Court also noted that bankruptcy courts do not fit within the extension of the public rights doctrine that the Court had described in Schor:

We deal here not with an agency but with a court, with substantive jurisdiction reaching any area of the corpus juris. This is not a situation in which Congress devised an “expert and inexpensive method for dealing with a class of questions of fact which are particularly suited to examination and determination by an administrative agency specially assigned to that task.” The “experts” in the federal system at resolving common law counterclaims such as [the one at issue] are the Article III courts, and it is with those courts that [the] claim must stay.

The Court next considered whether the bankruptcy courts’ role could be described as that of an adjunct to the district courts. Because the bankruptcy courts did more than fact finding, instead deciding “[a]ll matters of fact and law in whatever domains of the law to which’ the parties’ coun-

---

237. See Stern, 131 S. Ct. at 2600.
238. See id. at 2608.
239. Id. at 2614 (alteration in original) (quoting N. Pipeline Constr. Co. v. Marathon Pipe Line Co., 458 U.S. 50, 66 (1982) (plurality opinion); Murray’s Lessee v. Hoboken Land & Improvement Co., 59 U.S. 272, 284 (1855)).
241. Id.
242. Id. at 2615 (citations omitted) (quoting Crowell v. Benson, 285 U.S. 22, 46 (1932)).
terclaims might lead,” and the bankruptcy courts have power to enter final judgments, the Court concluded that they were not adjuncts.243

In dissent, Justice Breyer applied the balancing approach, which he described as addressing the question whether Congress’ use of a non-Article III body poses “a genuine and serious threat that one branch of Government sought to aggrandize its own constitutionally delegated authority by encroaching upon a field of authority that the Constitution assigns exclusively to another branch.”244 The dissent concluded that the bankruptcy courts’ consideration of the counterclaim at issue in that case was permissible based upon a variety of the Schor factors.245 Weighing in favor of constitutionality, the dissent explained, was that bankruptcy judges “enjoy considerable protection from improper political influence.”246 To explain this statement, Justice Breyer contrasted the earlier, impermissible scheme that had “provided for the appointment of bankruptcy judges by the President with the advice and consent of the Senate,” with the fact that “current law provides that the federal courts of appeals appoint federal bankruptcy judges.”247 The dissent also found solace in the fact that “Article III judges control and supervise the bankruptcy court’s determinations” and that district court judges may even “withdraw, in whole or in part, any case or proceeding referred [to the Bankruptcy Court] . . . on its own motion or on timely motion of any party, for cause shown.”248 The dissent next disagreed with the majority’s conclusion that the litigant challenging the forum had nowhere else to go, asserting “he could have litigated it in a state or federal court after distribution.”249 Finally, the dissent addressed the nature and importance of the legislative purpose. Because Congress is specifically given authority to establish uniform bankruptcy laws by Article I, Section 8 of the Constitution, the dissent weighed that factor in favor of the constitutionality of Congress’ decisions with respect to the extent of the bankruptcy courts’ jurisdiction.250

As described in the previous sections, none of the primary rationales the Court has applied to justify the use of Article I courts weigh in favor of the Court of Federal Claims considering takings claims. The current scheme therefore fails under the categorical approach of the majority in Northern Pipeline. The scheme also fails to pass muster under the dissent’s approach in Northern Pipeline, which described bankruptcy courts as involving “issues likely to be of little interest to the political branches,” with “li-

243. Id. at 2618–19 (alteration in original) (quoting Northern Pipeline, 458 U.S. at 91 (Rehnquist, J., concurring)).
244. Id. at 2624 (Breyer, J., dissenting).
245. See id. at 2626–28.
246. Id. at 2620–27.
247. Id. at 2627.
248. Id. (alterations in original) (quoting 28 U.S.C. § 157(d)) (internal quotation marks omitted).
249. Id. at 2628 (citing 11 U.S.C. § 523(a)(6)).
250. Id. at 2629.
As outlined above, no one could seriously argue that AIG’s shareholders’ claim—that they are owed billions of dollars from the Treasury as a result of the government’s bailout of that company—is not of interest to the political branches. As the Schor decision represents the lowest bar to date for permitting Congress to employ non-Article III tribunals to adjudicate claims, it is particularly telling to apply the Schor criteria to the Court of Federal Claims’ consideration of takings claims. The result is striking, as every factor the Court has looked to, even while it affirmed the CFTC’s status as a non-Article III body, militates against the Court of Federal Claims considering takings cases. That court’s continuing to do so does not even come close to clearing the low bar set in Schor and the dissent in Stern.

Both the Schor majority and Stern dissent noted that in the bankruptcy context, the litigants could have chosen to have the claims at issue decided by an Article III court but elected the bankruptcy forum instead. In contrast, plaintiffs seeking to file a takings claim for greater than $10,000 have only one forum available to them: the Court of Federal Claims. Thus, the AIG shareholders, and any other party bringing takings claims, indisputably have “nowhere else to go.” Whatever the force of that rationale in the context of the bankruptcy courts, it has none to the Court of Federal Claims.

With respect to the first Schor factor, “the extent to which the non-Article III forum exercises the range of jurisdiction and powers normally vested only in Article III courts,” the Court of Federal Claims is much more like an Article III district court than the CFTC, the entity whose jurisdiction the Schor majority found unobjectionable. Indeed, as federal district courts have concurrent jurisdiction to consider takings claims seeking less than $10,000, the Court of Federal Claims can be said to be exactly like a district court with respect to the takings portion of its jurisdiction. Rather than dealing with a particularized area of law, the Court of Federal Claims’ consideration of takings claims directly overlaps claims the district courts have developed expertise in considering. Moreover, unlike

252. See Sapphire & Solimine, supra note 50, at 120 (describing balancing analysis adopted in that decision as having "questionable capacity to impose any principled limitations on Congress’s power to use non-article III adjudicatory institutions to implement federal policies").
253. See Stern, 131 S. Ct. at 2628 (Breyer, J., dissenting) (“[H]e could have litigated it in a state or federal court . . . .”); Commodity Futures Trading Comm’n v. Schor, 478 U.S. 833, 849 (1986) (concluding that Schor indisputably waived any right he may have possessed to the full trial of [the commodity future broker’s] counterclaim before an Article III court because he had “expressly demanded that [the commodity future broker] proceed on its counterclaim” before CFTC”).
255. See Stern, 131 S. Ct. at 2614.
256. See Schor, 478 U.S. at 851.
the CFTC, which does not exercise “all ordinary powers of district courts,” and whose orders “are enforceable only by order of the district court,” the Court of Federal Claims exercises all of the same authority as a district court, up to and including the power to enforce its own orders and issue final judgments.\textsuperscript{257} Finally, as noted previously, while the CFTC’s decisions are reviewed under the less deferential “weight of the evidence” standard, the Federal Circuit reviews the Court of Federal Claims’ decisions under the more deferential “clearly erroneous” standard that the \textit{Schor} Court described as objectionable.\textsuperscript{258}

The second \textit{Schor} factor, “the origins and importance of the right to be adjudicated,” could hardly apply more strongly than in a case dealing with the Takings Clause of the U.S. Constitution.\textsuperscript{259} In \textit{Northern Pipeline, Schor}, and \textit{Stern}, the Court found that even state law claims trigger this factor.\textsuperscript{260} If the state law claims at issue in those cases were important enough for this factor to weigh against non-Article III adjudication, then surely a federal right included in the Bill of Rights is entitled to at least that much weight.\textsuperscript{261} The \textit{Schor} Court found this factor counteracted by the fact that Congress had not entirely removed adjudication of the right at issue from the Article III courts, and instead “the decision to invoke this forum is left entirely to the parties and the power of the federal judiciary to take jurisdiction of these matters is unaffected.”\textsuperscript{262} In stark contrast, litigants such as the AIG shareholders have no option but to file their takings claims in the Court of Federal Claims. The “origins and importance of the right” factor thus weighs heavily against permitting takings claims to be adjudicated outside an Article III court.

The final \textit{Schor} factor, “the concerns that drove Congress to depart from the requirements of Article III,” does not save Congress’ choice in this instance.\textsuperscript{263} As noted earlier, the creation of the Court of Federal Claims as an Article I court appears to have largely been done with little study and almost as an afterthought to the creation of the Federal Circuit. To the extent Congress considered the new body’s status at all, Congress was not focused upon that entity’s Takings Clause jurisdiction, and it appears not to have recognized, or even considered, that the new structure it

\begin{itemize}
  \item \textsuperscript{257} See id. at 853 (quoting N. Pipeline Constr. Co. v. Marathon Pipe Line Co., 458 U.S. 50, 85 (1982) (plurality opinion)).
  \item \textsuperscript{258} See id.
  \item \textsuperscript{259} See id. at 851.
  \item \textsuperscript{260} See \textit{Stern}, 131 S. Ct. at 2611; \textit{Schor}, 478 U.S. at 852; \textit{Northern Pipeline}, 458 U.S. at 84.
  \item \textsuperscript{261} See \textit{William Howard Taft, Popular Government: Its Essence, Its Permanence and Its Perils} 90 (Kraus Reprint Co., 1971) (1913) (“\textit{[N]ext to the right of liberty, the right of property is the most important individual right guaranteed by the Constitution and the one which, united with that of personal liberty, has contributed more to the growth of civilization than any other institution established by the human race.”).
  \item \textsuperscript{262} \textit{Schor}, 478 U.S. at 855.
  \item \textsuperscript{263} See id. at 851.
\end{itemize}
was creating would remove takings claims from an Article III court. A sen-
ate report captures Congress’ thoughts about why it departed from the
requirements of Article III:

The court will be established under Article I of the Constitution of the United States. Because 28 U.S.C. 2509 of existing law gives the trial judges of the Court of Claims jurisdiction to hear congressional reference cases, which are not ‘cases and controversies’ in the constitutional sense, and because the cases heard by the Claims Court are in many ways essentially similar to the limited jurisdiction cases considered by the tax court, judges of the Claims Court are made Article I judges rather than Article III judges.264

Thus, according to Congress, the Court of Federal Claims is an Article I court so that it can continue to hear congressional reference cases and because the cases heard by the newly created court are similar to those heard by the tax court. While those justifications may or may not be valid with respect to some aspects of the court’s jurisdiction, they do not justify Congress’ decision to give the court jurisdiction to consider takings cases. Congress’ desire to maintain some entity to consider congressional reference cases is understandable, as Congress reasonably wants an expert body to assist it in determining when to issue private bills.265 That Congress may create such a body says nothing about whether it must be the same body that considers takings claims. If Congress wants to be able to send congressional reference cases—which are not “cases or controversies” and therefore do not require adjudication by an Article III court—to some entity of its creation, it may, of course, do so. By sweeping up takings cases in the same basket Congress created to deal with private bills, one can fairly say that Congress “sought to aggrandize its own constitutionally delegated authority by encroaching upon a field of authority that the Constitution assigns exclusively to another branch,” which is precisely what the Stern dissent said it may not do.266

Congress’ second rationale for making the Court of Federal Claims an Article I court, its recognition that the Court of Federal Claims “in many ways” resembles the tax court, also does not encompass takings cases. Instead, it appears that Congress was referring to the other subject matters the court considers, such as contract disputes and tax cases. While the public rights rationale could be applied both to those aspects of the court’s jurisdiction as well as to the tax court, this Article has already demonstrated why it does not justify the Court of Federal Claims considering takings cases. Takings claims thus fall outside the “many ways” those

265. See generally Matthew G. Bisanz, Current Development, The Honor of a Na-
266. See Stern v. Marshall, 131 S. Ct. 2594, 2624 (Breyer, J., dissenting).
courts are similar. The “concerns that drove Congress to depart from the requirements of Article III” therefore simply do not apply to that aspect of the court’s jurisdiction.

Finally, the Schor Court weighed in favor of non-Article III adjudication that deciding the otherwise impermissible claim was “incidental to, and completely dependent upon, adjudication of [] claims created by federal law.” The Court noted that the situation wherein the CFTC decided claims not created by federal law “in actual fact is limited to claims arising out of the same transaction or occurrence as the [created by federal law] claim.” In contrast, the Court of Federal Claims considers takings claims not only as counterclaims or when they are incidental to other claims but also entirely independent of its other jurisdictional grants.

C. How the Court of Federal Claims Came to Be an Article I Court that Decides Takings Claims

The previous section of this Article demonstrates that none of the rationales the Supreme Court has relied upon to justify Congress placing adjudicative responsibilities in a non-Article III body applies to justify the Court of Federal Claims considering takings claims. This section demonstrates that the status quo is actually a very recent development. Taking a longer view, Congress has generally expanded citizens’ ability to sue the sovereign, and the history of the predecessors of the Court of Federal Claims reveals a series of expansions upon citizens’ ability to hold the sovereign to task. In addition, the recent reduction of citizens’ rights to sue the sovereign in a constitutional court appears to have been inadvertent and inconsistent with Congress’ intentions when it created the Court of Federal Claims.

In the early years of the nation, at least since 1789, citizens who wanted to file claims for money against the government did so through petitions filed directly with Congress. But not everyone thought that system worked well, including President John Quincy Adams who, while a congressman, stated:

There ought to be no private claims business before Congress. There is a great defect in our institutions by the want of a Court of Exchequer or a Chamber of Accounts. It is judicial business, and legislative assemblies ought to have nothing to do with it. One half of the time of Congress is consumed by it, and there is no common rule of justice for any two of the cases decided. A

267. Schor, 478 U.S. at 856.
268. Id.
269. For a history, see Shimomura, supra note 65; see also Charles Chauncey Binney, Origin and Development of Legal Recourse Against the Government in the United States, 57 U. PA. L. Rev. 372, 381 (1909).
deliberative Assembly is the worst of all tribunals for the administration of justice. 270

In order “to relieve the pressure on Congress caused by the volume of private bills,” Congress created the U.S. Court of Claims in 1855. 271 Initially, the individuals who served on that court, and considered monetary claims against the federal government, were administrators who reported directly to Congress, having no real authority to decide cases at all. 272 As the Supreme Court has described the original Court of Claims, it was “originally nothing more than an administrative or advisory body . . . .” 273 Appropriately, as a creature of the legislature, that body was housed in the Capitol Building itself. 274

The Court of Claims was reorganized in 1863, after President Lincoln opined:

The investigation and adjudication of claims in their nature belong to the judicial department . . . . While the court has proved to be an effective and valuable means of investigation it in great degree fails to effect the objects of its creation for want of power to make its decisions final. 275

In response to that and similar criticisms, in 1863, Congress passed “An Act to amend an Act to establish a court for the investigation of claims against the United States,” which provided that appeals from the Court of Claims could be taken to the Supreme Court. 276

That Act provided the Court with its first opportunity to address the status of the first Court of Claims in *Gordon v. United States*. 277 In that
decision, the Court noted that based upon the powers Congress granted
the Court of Claims, that body was not actually deciding any cases.278 Instead, the statute called upon the Court of Claims, and the Supreme Court
upon review, to provide advisory opinions to the Secretary of the Treasury,
who would then provide Congress with an estimate, which Congress would
then decide to fund or not to fund.279 The Court noted: “Congress may
undoubtedly establish tribunals with special powers to examine testimony
and decide, in the first instance, upon the validity and justice of any claim
for money against the United States, subject to the supervision and control
of Congress, or a head of any of the Executive Departments.”280 Such
authority, the Court noted, is “like to that of an Auditor or Comptroller,”
and is not the “judicial power in the sense in which those words are used
in the Constitution.”281 As the Court succinctly recognized, the Court of
Claims, as it was structured in 1864, was not a court at all, and the values
underlying Article III were not undermined by its actions.282 That case
offers little explicit guidance with respect to the question of the circum-
stances under which legislative courts may be established, because the
court being considered was not yet even an Article I court at that time.

Beginning in 1865, Congress removed the requirement that Treasury
prepare an estimate, began to make appropriations in advance of the
Court of Claims’ decisions, and enacted various statutes requiring the
Court of Claims to resolve the multitude of claims that came out of the
Civil War.283 Along with that increased authority, the court was also given
a new location, moving out of the Capitol into its own building near the
White House, on Pennsylvania Avenue.284 After those changes, the Court
of Claims began deciding actual cases, while also providing non-binding
opinions to Congress in what are known as congressional reference

“[i]respect of its intrinsic value, it has an interest for the court and the
bar . . . .” Id.

278. See Gordon, 117 U.S. at 698–99.
279. See id.
280. Id. at 699.
281. Id.
282. See id. at 700. As the Court recognized in Glidden, the Gordon decision is
not technically the decision of the Court in that proceeding. See Glidden Co. v.
Zdanok, 370 U.S. 530, 569 (1962). The Glidden Court stressed that the Court’s
refusal to hear the Gordon appeal was due solely to the fact that the Treasury
Department could revise the decision, not due to the fact that an appropriation
might not be forthcoming. See id. (citing United States v. Jones, 119 U.S. 477, 478
(1886)). Regardless, after Congress both removed Treasury’s ability to revise the
decisions and began to issue appropriations in advance of judgments, the Court
began to hear appeals from the Court of Claims. See id.
283. See An Act in Relation to the Court of Claims, Sess. I, ch. 19, 14 Stat. 9
(1866). After those changes, the Supreme Court agreed to hear appeals from the
Court of Claims, which it did first in De Groot v. United States and explicitly approved
in 1871. See United States v. Klein, 80 U.S. 128, 133 (1871); De Groot v. United
States, 72 U.S. 419, 427 (1866); see also Shimomura, supra note 65, at 663 n.307
(listing and detailing history of related cases).
284. See Flanders, supra note 270, at 113.
cases. The Supreme Court agreed to consider appeals from the court’s case-decisional authority, even while refusing to consider appeals from the congressional reference cases in what has been described as an “uneasy accommodation” based upon an understanding that the Court of Claims’ participation in reference cases was only an “ancillary” function. With those changes, the Court of Claims went straight from a non-adjudicatory body to an Article III court.

From 1866, when the Supreme Court first began to hear Court of Claims appeals, through 1925, the Court consistently treated that body as an Article III court. Because the Court of Claims was perceived to be an Article III court during this timeframe, the Supreme Court’s pronouncements tell us little about the circumstances necessary for Congress to permissibly establish the Court as an Article I court. Also, throughout this period, perhaps because of the history of claims against the United States proceeding only through presentation to Congress, the Supreme Court perceived the Court of Claims, and its jurisdictional basis, through

285. See 28 U.S.C. §§ 1492, 2509 (2012) (providing details of congressional reference cases). Congressional reference is a procedure that allows either house of Congress to refer a matter to the chief judge of the Court of Claims with directions to report findings back to the referring house. See id. § 2509. Under that procedure, the chief judge designates a hearing officer to make factual findings and a reviewing panel, consisting of three judges of the Court of Claims. See id. This procedure is still available to Congress, and reference cases are occasionally transmitted to the U.S. Court of Federal Claims, which will provide an advisory decision after reviewing the reference. See id. Under that procedure, the hearing officer must make findings in accordance with the rules of the Court of Federal Claims, and “[h]e shall append to his findings of fact conclusions sufficient to inform Congress whether the demand is a legal or equitable claim or a gratuity, and the amount, if any, legally or equitably due from the United States to the claimant.” Id. § 2509(c).

286. See In re Sanborn, 148 U.S. 222, 226 (1893) (describing Court of Claims’ advisory duties as “ancillary” function); Shimomura, supra note 65, at 662 (describing the “uneasy accommodation”); see, e.g., Muskrat v. United States, 219 U.S. 346, 362–63 (1911) (affirming that Court would not consider appeals of reference cases); United States v. Jones, 119 U.S. 477 (1886) (giving one example of Supreme Court hearing appeal from Court of Claims during this period).

287. The first case in which the Supreme Court agreed to hear an appeal from the Court of Claims appears to be De Groot. See De Groot, 72 U.S. at 427. The Court explicitly approved doing so in 1871. See Klein, 80 U.S. at 133 (deciding case on appeal from Court of Claims); see also Shimomura, supra note 65, at 662 n.307 (detailing this history and listing additional cases during this time period). The last case in which the Court did so was Miles v. Graham. See Miles v. Graham, 268 U.S. 501, 505 (1925). In Glidden Co. v. Zdanok, the Court reviewed the history of the formation of the Court of Claims, noting that “there are substantial indications in the debates that Congress thought it was establishing a court under Article III.” Glidden, 370 U.S. at 553. The Court explained that in its view, Congress had, throughout the history of that court, treated the Court of Claims as an Article III court. Id. at 555. The dissent in that case thought otherwise. See id. at 594 (Black, J., dissenting) (“Congress, however, has always understood that it was only establishing Article I courts when it created the Court of Claims . . .”). In Glidden, the Court went so far as to say that “[t]he creation of the Court of Claims can be viewed as a fulfillment of the design of Article III.” Id. at 558 (majority opinion).
the prism of sovereign immunity.288 Thus, when a plaintiff in the Court of Claims, Thomas McElrath, objected to that court’s consideration of a government counterclaim seeking $6,106.53 against him without providing him the opportunity to have that case heard by a jury as provided for by the Seventh Amendment,289 the Court ruled in favor of the government in a decision explicitly grounded upon the principle of sovereign immunity:

The government cannot be sued, except with its own consent. It can declare in what court it may be sued, and prescribe the forms of pleading and the rules of practice to be observed in such suits. . . . If the claimant avails himself of the privilege thus granted, he must do so subject to the conditions annexed by the government to the exercise of the privilege. Nothing more need be said on this subject.290

In 1887, Congress again changed the character of the court with the passage of the Tucker Act, which broadened the types of claims that could be submitted to the court, as opposed to directly to Congress.291 Because the Tucker Act provided that “claims founded upon the Constitution” would be heard by the Court of Claims, takings claims were also, for the first time, to be brought in the Court of Claims.292 In so doing, Congress created a court with much more authority than it previously enjoyed. Along with its upgraded status, the Court of Claims was placed in a new location in 1899, in what was then known as the Court of Claims Building.293 From the time the court was moved out of the halls of Congress and given authority to decide cases, the court apparently thought of itself, and the Supreme Court considered it to formally be, part of the federal judiciary as an Article III or “constitutional” court.294

288. See, e.g., Cohens v. Virginia, 19 U.S. 264, 411–12 (1821) (“The universally received opinion is, that no suit can be commenced or prosecuted against the United States; that the judiciary act does not authorize such suits.”). At least since 1821, the Supreme Court had described the United States’ sovereign immunity as absolute. See id.
289. See U.S. CONST. amend VII.
291. See CHARLES ALAN WRIGHT ET AL., STATUTORY EXCEPTIONS TO SOVEREIGN IMMUNITY—ACTIONS UNDER THE TUCKER ACT, 14 FED. PRACTICE & PROCEDURE § 3657 (3d ed. 2014).
292. See An Act to Provide for the Bringing of Suits Against the Government of the United States, ch. 359, 24 Stat. 505 (1887). This appears to be the first time takings claims could be filed in federal courts. See William Michael Treanor, The Original Understanding of the Takings Clause and the Political Process, 95 COLUM. L. REV. 782, 794–95 n.69 (1995) (discussing lack of judicial remedy for most federal takings until enactment of Tucker Act); see also Langford v. United States, 101 U.S. 341, 343 (1879) (“It is to be regretted that Congress has made no provision by any general law for ascertaining and paying this just compensation.”).
293. See FLANDERS, supra note 270, at 113.
During that entire time period, the Supreme Court continued to describe the Court of Claims as an entity whose existence was entirely based upon waivers of sovereign immunity. As the Court stated, “the United States as a government may not be sued without its consent, yet with its consent it may be sued, and the judicial power of the United States extends to such a controversy. Indeed, the whole jurisdiction of the Court of Claims rests upon this proposition.”

During the twentieth century, there were various changes to the makeup of the Court of Claims. The primary change occurred in 1925, when Congress reorganized the court by designating a group of seven commissioners who would hear evidence and decide factual questions and whose findings would then be reviewed by judges who served as an appellate body. Congress had created a trial division made up of commissioners, who conducted trials, and a separate appellate division to review those findings. There was no indication that Congress, or anyone else, considered the Court to be anything other than an Article III court. Under the Article III analysis described above, the old Court of Claims would have passed constitutional muster to consider takings claims, and there would be no need for this Article if things had stayed that way.

But in 1929, the Supreme Court altered its treatment of the Court of Claims, beginning to describe it not as subject to Article III, but instead as a legislative court established pursuant to Congress’ Article I power. The Court first did so in dicta in the 1929 *Bakelite* decision involving the Court of Customs Appeals, opining that the Court of Claims:

> [W]as created, and has been maintained, as a special tribunal to examine and determine claims for money against the United States. This is a function which belongs primarily to Congress as an incident of its power to pay the debts of the United States.

---


296. See An Act to Authorize the Appointment of Commissioners by the Court of Claims and to Prescribe Their Powers and Compensation, Pub. L. No. 68-451, ch. 301, § 1, 43 Stat. 964 (1925); Wilson Gowan, Philip Nichols, Jr. & Marion T. Bennett, The United States Court of Claims, A History, Part II 89 (1978) (“[T]his legislation authoriz[ed] the court to appoint a number of persons, not to exceed seven, to be known as commissioners who would have, and who would perform, in general, the same powers and duties as those pertaining to special masters in chancery.”).

But the function is one which Congress has a discretion either to exercise directly or to delegate to other agencies.\(^{298}\)

While the Court recognized that “[o]ther claims have since been included in the delegation,” the Court declared that “the court is still what Congress at the outset declared it should be—‘a court for the investigation of claims against the United States.’ The matters made cognizable therein include nothing which inherently or necessarily requires judicial determination.”\(^{299}\)

Interestingly, even as the Justices changed course about whether the Court of Claims was an Article I or Article III court, the Supreme Court did not waver in its understanding that the basis of the Court of Claims’ jurisdiction is a waiver of sovereign immunity. Indeed, as revealed in the \textit{Bakelite} decision, the entire justification for determining that the Court of Claims was an Article I court was that the Court decided that Congress could have, if it had elected to do so, done all of the work then being done by the Court of Claims. This is explicit in the Court’s statement that the Court of Claims’ jurisdiction includes “nothing which inherently or necessarily requires judicial determination.”\(^{300}\) The \textit{Bakelite} Court also noted the long history of the Court of Claims’ issuing advisory opinions to Congress and noted that “[a] duty to give decisions which are advisory only, and so without force as judicial judgments, may be laid on a legislative court, but not on a constitutional court established under article 3.”\(^{301}\) The \textit{Bakelite} Court recognized that the Court of Claims was “undoubtedly and completely under the control of Congress.”\(^{302}\)

The Court expressly adopted the sovereign immunity/public rights rationale as its justification for determining that the Court of Claims could conduct its proceedings as a legislative court in the 1933 decision \textit{Williams v. United States}.\(^{303}\) The question presented in that case was whether Congress could reduce the salary of a judge serving on the old Court of Claims.\(^{304}\) The Court held:

Since all matters made cognizable by the Court of Claims are equally susceptible of legislative or executive determination, they are, of course, matters in respect of which there is no constitutional right to a judicial remedy, and the authority to inquire into

\(^{298}\). \textit{Ex parte Bakelite Corp.}, 279 U.S. 438, 452 (1929).
\(^{299}\). \textit{Id.} at 452–53.
\(^{300}\). \textit{Id.} at 453. As one commentator aptly described the Supreme Court’s understanding of the role played by the Court of Claims, it is “that the Court of Claims is most properly viewed as completely under the control of the legislative department.” Comment, \textit{The Distinction Between Legislative and Constitutional Courts and Its Effect on Judicial Assignment}, 62 \textit{COLUM. L. REV.} 133, 145 (1962).
\(^{301}\). \textit{Bakelite}, 279 U.S. at 454.
\(^{302}\). \textit{Id.} at 455.
\(^{303}\). 289 U.S. 553 (1933).
\(^{304}\). \textit{Id.} at 561.
and decide them may constitutionally be conferred on a nonjudicial officer or body.\[305\]

The Court explained that the reason for its decision was the dicta presented in Bakelite that the Court of Claims' function, “to examine and determine claims for money against the United States. . . . is one which Congress has a discretion either to exercise directly or to delegate to other agencies,” and that “none of the matters made cognizable by the court inherently or necessarily requires judicial determination . . . .”\[306\] The Court explicitly held that the earlier utterances describing the Court as subject to Article III had been dicta. The Court never questioned, but instead explicitly relied upon, the proposition that the entire basis of the jurisdiction of the Court of Claims involved waivers of sovereign immunity.\[307\]

Congress had never actually asked for authority to downgrade the court’s status to an Article I court, however. Congress therefore attempted to undo the Williams holding in 1953, by expressly declaring that the Court of Claims was “established under Article III of the Constitution of the United States.”\[308\] In 1962, the Supreme Court again considered the Court of Claims’ status in Glidden Co. v. Zdanok,\[309\] this time affirming that the Court of Claims was indeed a true constitutional court.\[310\] The context in Glidden was the constitutionality of a Court of Claims judge sitting by designation on an Article III court, which was only permissible if the Court of Claims judge was an Article III judge.\[311\] Explaining that the

\[305\] Id. at 579–80 (citations omitted) (citing Bakelite, 279 U.S. at 452, 458; United States v. Babcock, 250 U.S. 328, 331 (1919)).

\[306\] Id. at 569 (quoting Bakelite, 279 U.S. at 438) (internal quotation marks omitted).

\[307\] Id. at 568 (collecting cases supporting conclusion).

\[308\] An Act to Amend Title 28, United States Code, Pub. L. No. 83-158, § 1, 67 Stat. 226 (1953) (“Such court is hereby declared to be a court established under article III of the Constitution of the United States.”); see also Note, The Constitutional Status of the Court of Claims, 68 Harv. L. Rev. 527, 527 (1955) (“Congress has recently attempted to reverse the result of the Williams case . . . .”).


\[310\] See id. at 584. Because of that change in status to an Article III court, it was apparent that judges of the Court of Claims could no longer decide congressional reference cases. See id. In response, in 1966, Congress created a procedure whereby it would direct congressional reference cases to the chief commissioner of the Court of Claims rather than to the judges of the court. See Act to Amend Title 28, Entitled “Judiciary and Judicial Procedure,” of the United States Code to Provide for the Reporting of Congressional Reference Cases by Commissioners of the United States Court of Federal Claims, Pub. L. No. 89-681, 80 Stat. 958, 958–59 (1966). Apparently Congress believed that in doing so it was preserving the independence of the judges as Article III judges and the commissioners as Article I judges. See Office of Pers. Mgmt. v. Richmond, 496 U.S. 414, 431 (1990). For a discussion of congressional reference cases, see Bisanz, supra note 265.

\[311\] Glidden, 370 U.S. at 532–33. The Court was asked to address whether it was constitutional for Judge J. Warren Madden, then a judge on the Court of Claims, to have sat by designation on the U.S. Court of Appeals for the Second Circuit. See id.
Court would not disregard Congress’ declaration, a plurality of the Court held that the Court of Claims was, after all, an Article III court.312

In some ways, this decision is no more helpful to assessing whether the current Court of Federal Claims may constitutionally hear takings cases as an Article I court than was the Court’s Gordon decision; in both cases, the Court determined that the entity it was assessing was not an Article I court and the Court, therefore, need not have explored the boundaries of what Congress could have done had Congress decided to establish a legislative court. The Glidden decision is instructive, however, not because of any question it answered, but because of one it raised. The Glidden decision was quite critical of the earlier Williams decision, which it described as of “questionable soundness.”313 The Glidden Court recognized that the earlier Williams decision equated Congress’ perceived ability to establish the Court of Claims as an Article I court with Congress having actually done so.314 As the Court noted, however, those propositions need not logically follow.315 Without addressing the extent of Congress’ power to do something it had not done, the Court explicitly disagreed with the conclusion that the Court of Claims was an Article I court, holding instead that Congress had actually established that entity as an Article III court.316

Most importantly for this Article, while so holding, the Court in Glidden not only did not agree with the Williams holding—that Congress could commit the work then being performed by the Court of Claims to a non-Article III tribunal—but also explicitly refused to assess that question.317 The Court instead took the opportunity to indicate that it had “certain reservations about” the accuracy of the Williams Court’s description of the jurisdiction of the Court of Claims.318 The Court explicitly recognized that, contrary to the Williams Court’s description of that jurisdiction, the grant of authority “to award just compensation for a governmental taking, empowered [the Court of Claims] to decide what had previously been described as a judicial and not a legislative question.”319 Because the Court determined that the Court of Claims was an Article III court, it never had an occasion to correct that error.320

312. See id. at 541.
313. Id. at 543 (noting criticism of Williams decision).
314. See id. at 549–50.
315. See id. at 549 (“But because Congress may employ such tribunals assuredly does not mean that it must.”).
316. See id. at 552.
317. See id. at 549 (“Nor need we now explore the extent to which Congress may commit the execution of even ‘inherently’ judicial business to tribunals other than Article III courts.”).
318. See id.
319. See id. at 549 n.21 (citing Williams v. United States, 289 U.S. 553, 581 (1933); Monongahela Navigation Co. v. United States, 148 U.S. 312, 327 (1893)).
320. As noted above, in 1987, the Supreme Court reaffirmed, in First English, that takings claims do not require waivers of sovereign immunity. It is therefore not accurate to describe takings claims as involving waivers of sovereign immunity, and the Glidden Court’s criticism of the Williams decision has been validated. The
In 1966, Congress reaffirmed its intention that the Court of Claims be designated an Article III court. The structure of the Court of Claims, from 1966 through to its dissolution in 1983, had many features that the Supreme Court has found to weigh in favor of permitting adjudication by adjuncts to Article III judges. While the Supreme Court did not assess that arrangement, the rationale from the Court’s Crowell decision would likely justify the structure of the Court of Claims. In many ways the Court of Claims was analogous to the structure of the magistrates that the Court assessed in Raddatz. Like magistrate judges, the old commissioners, even when they were called trial judges, did not issue final judgments. Rather, their decisions were reviewed de novo by Article III judges before the decisions were issued. Also like a magistrate judge, in the old Court

Williams decision was based upon the erroneous factual predicate that the cases considered by the Court of Claims only included those “which Congress has a discretion either to exercise directly or to delegate to other agencies.” See Williams, 289 U.S. at 569 (quoting Ex parte Bakelite Corp., 279 U.S. 438, 452 (1929)) (internal quotation marks omitted). It turns out that was an overstatement. In contrast to the Williams Court’s statement that “none of the matters made cognizable by the court inherently or necessarily requires judicial determination,” the Takings Clause requires judicial determination. See id. At the time the Glidden decision was issued, the Court would have had to overturn Williams to correct that decision’s error. The new Court of Federal Claims is, however, a different entity, with different rules governing the appointment of judges and their authority. Accordingly, the Court need not overturn Williams to correct, or allow a correction of, this error; all that is needed is that when the Supreme Court assesses the Court of Federal Claims’ jurisdiction in the future, it recognize that, in addition to the contract claims and other matters considered by that court, it also hears takings claims, which do not require a waiver of sovereign immunity. See Monongahela, 148 U.S. at 327. But see Phillip R. Trimble, Foreign Policy Frustrated—Dames & Moore, Claims Court Jurisdiction and a New Raid on the Treasury, 84 COLUM. L. REV. 317, 381 n.291 (1984) (“There is no precedent for the proposition that the Constitution requires judicial determination for a takings claim or any aspect of it.”). Because the rationale underlying Williams is that the federal government’s immunity from suit is absolute, commentators point to the logic of sovereign immunity when they explain why the current Court of Federal Claims is an Article I court. See, e.g., Pfander, supra note 42, at 658.

321. See An Act to Provide for the Appointment of Two Additional Judges for the United States Court of Claims, and for Other Purposes, Pub. L. No. 89-425, 80 Stat. 139 (1966) (modifying 28 U.S.C. § 171 to state: “[T]he President shall appoint, by and with the advice and consent of the Senate, a chief judge and six associate judges who shall constitute a court of record known as the United States Court of Claims. Such court is hereby declared to be a court established under Article III of the Constitution of the United States.”).


323. See Seamon, supra note 322, at 545.
of Claims, the non-Article III judges worked “subsidiary to and only in aid of” the Article III judges, and “the entire process [took] place under the [Article III judge’s] total control and jurisdiction.”\(^324\) They were very much adjuncts of the Article III judges. If that structure had not changed, the court that hears takings claims would still be an Article III court, and questions of that court’s constitutionality vis-à-vis Article III would not be implicated.

That structure changed when Congress passed the Federal Courts Improvement Act of 1982: “An Act to establish a United States Court of Appeals for the Federal Circuit, to establish a United States Claims Court, and for other purposes.”\(^325\) When Congress created that new structure, it was focused upon the creation of a new appellate court, the United States Court of Appeals for the Federal Circuit, which would, for the first time, consolidate patent appeals into a single court.\(^326\) To create the new appellate court, Congress, after ten years of study, decided to merge the Court of Customs and Patent Appeals with the appellate division of the Court of Claims.\(^327\) Congress simply moved the Court of Claims judges to the newly created appellate court. But Congress still had to create a trial court to deal with the cases that had been decided by the old Court of Claims. In what appears to have been almost an afterthought, in the same act that created the Federal Circuit, Congress “elevated” the commissioners of the old Court of Claims, making those judges the first to serve on a new trial court, the United States Claims Court, which would later come to be called the United States Court of Federal Claims.\(^328\) To explain the structure of the new trial court responsible for deciding claims against the federal gov-

---


\(^326\). See Sward & Page, supra note 111, at 386–87. The Act’s consolidation of patent appeals jurisdiction in a single federal appeals court is generally considered to be “[t]he most significant aspect of the Act . . . .” See id. at 386.


\(^328\). See Federal Courts Improvement Act of 1982, Pub. L. No. 97-164, §§ 127, 133, 96 Stat. 25. As one commentator describes the process, “because the FCIA [Federal Courts Improvement Act] was directed to appellate reform, neither the FCIA’s supporters nor Congress gave sustained attention to reforms needed at the trial level. The FCIA created the COFC, out of necessity, to take over the Court of Claims’ trial functions . . . .” Seamon, supra note 322, at 545–46 (describing Court of Federal Claims’ formation as “incidental to the creation of the Federal Circuit”).
The establishment of the Claims Court accomplishes a much needed reorganization of the current system by assigning the trial function of the court to trial judges whose status is upgraded and who are truly independent. Presently, the commissioners of the Court of Claims are appointed by the Article III judges of that court and do not have the power to enter dispositive orders; final judgment in a case must be made by the Article III judges after reviewing findings of fact and recommendations of law submitted by a commissioner. The creation of the United States Claims Court will reduce delay in individual cases and will produce greater efficiencies in the handling of the court’s docket by eliminating some of the overlapping work that has occurred as a result of this process.\(^{329}\)

The creation of that new entity did not have the effect that Congress intended. Congress believed that the role of the Court of Federal Claims judges was to serve as “upgraded” versions of the commissioners from the old Court of Claims. That is not an accurate description of the solution that was implemented, however. While the status of the individual commissioners was upgraded from their previous positions as adjuncts to Article III judges, those newly elevated judges do not fill the role of the old commissioners. Unlike the commissioners who they replaced, the Court of Federal Claims judges issue final judgments that are not reviewed by Article III judges before the final judgments are issued.\(^{330}\) In addition, while the traditional role of the old Court of Claims was limited to the award of monetary damages, the court now has expanded authority “[t]o provide an entire remedy and to complete the relief afforded by the judgment,” which involves the granting of some equitable relief.\(^{331}\) The judges of the Court of Federal Claims therefore exercise authority that is more akin to the authority of the Article III judges of the Court of Claims. From the perspective of litigants bringing claims in the Court of Federal Claims, the Court of Federal Claims judges stepped into the shoes of the old Article III judges of the Court of Claims, not the commissioners of that dissolved body. When the appropriate comparison is made, the new trial court judges have a reduced, not upgraded, status.

In addition, while Congress believed that in creating the Court of Federal Claims it was creating judges who are “truly independent,” in ef-


ffect Congress reduced the independence of the body that decides monetary claims against the United States. While the judges who had previously heard such claims were Article III judges, the judges who now hear those cases are appointed for limited terms and do not enjoy Article III’s protections. Because the central purpose of Article III is to ensure the independence of the judiciary, by substituting Article I judges into the role once filled by Article III judges, Congress has actually reduced the independence of that body. Thus, through the 1982 restructuring, Congress, in effect, changed its mind about its decision to permit citizens to bring lawsuits for money to the judiciary. Instead it has decided that when citizens sue the federal government and believe they are entitled to more than $10,000, they may only pursue those large claims in legislative courts controlled by Congress. But it is not at all clear that Congress intended that effect.

D. Options Moving Forward

The Supreme Court has not discussed whether the current structure of the Court of Federal Claims comports with the requirements of Article III. As shown above, none of the rationales that the Supreme Court has used to uphold non-Article III entities’ authority to adjudicate claims applies to the Court of Federal Claims deciding takings claims. That is not to say, of course, that no such justification could be conceived. As the review of those rationales demonstrates, they have been added over time. Further, as Justice Scalia noted, they seem to have been added “almost randomly.” Indeed, in the Court’s most recent discussion of those factors, Justice Scalia proposed what may be a new one, stating: “in my view an Article III judge is required in all federal adjudications, unless there is a firmly established historical practice to the contrary.”

332. The Tucker Act creates exclusive jurisdiction to consider monetary claims, other than tort claims, in which plaintiffs seek greater than $10,000 from the federal government. For claims seeking less than $10,000, the federal district courts have concurrent jurisdiction pursuant to the “Little Tucker Act.” See 28 U.S.C. § 1346(a)(2). Tort claims may be filed in a federal district court, pursuant to the Federal Tort Claims Act. See id. § 1346(b).

333. The Supreme Court has granted certiorari to takings cases originating in the Court of Federal Claims on multiple occasions. See, e.g., Ark. Game & Fish Comm’n v. United States, 133 S. Ct. 511 (2012). When the Court has granted certiorari, it appears to have assumed that jurisdiction is proper based upon its statutory source, the Tucker Act, without assessing the Act’s constitutionality. See, e.g., Ruckelshaus v. Monsanto Co., 467 U.S. 986, 1016–17 (1984) (citing United States v. Causby, 328 U.S. 256, 267 (1946) (“If there is a taking, the claim is ‘founded upon the Constitution’ and within the jurisdiction of the Court of Claims to hear and determine.”)) (“[A]n individual claiming that the United States has taken his property can seek just compensation under the Tucker Act.” (citation omitted)).


335. Id.
TAKING BACK TAKINGS CLAIMS

2015

no such firmly established historical practice. Barring conception of a new justification, it appears that under the Supreme Court’s jurisprudence the status quo is unconstitutional.

As an aside, it is worth noting that even if it were not unconstitutional for Congress to insist that takings claims be filed in an Article I court, it is still a good idea for Congress to ensure that takings claims are decided by an Article III court for policy reasons. Because the Court of Federal Claims, as a result of its structure and history, is a political court, the current system lends itself to at least the appearance of unfairness. Worse still, by forcing only high dollar claims into its legislative court, Congress creates the appearance not only that the system is fixed, but also that it is fixed only for those cases that Congress really does not want to pay. It is essential that the body that hears lawsuits as expensive and important as the AIG shareholders’ lawsuit is, and is perceived to be, independent from the Treasury and those who orchestrated the bailout. Congress recognized the value of the court’s independence when it created the Court of Federal Claims. Because the Court of Federal Claims’ independence would be augmented by making it an Article III court, which would be consistent with Congress’ expressed intent, Congress should consider taking that step even if the Constitution did not require it and even if takings claims were no longer considered by that court.

Although the Court of Federal Claims, as it is currently structured, cannot constitutionally consider takings claims, those claims must be heard in some forum.336 There are several potential solutions that would rectify the problem identified in this Article. First, in perhaps the most direct solution to the problem, Congress could formally elevate the court’s status to that of an Article III court.337 Similarly, Congress could create a specialized Article III court just for takings claims, perhaps calling it the Court of Takings Claims. Both solutions would leave the existing state of affairs intact in that takings claims would continue to be decided by a specialized court. Another potential solution would be to eliminate the concurrent jurisdiction of the Court of Federal Claims and the district courts for takings claims, making the district courts the sole forum for citizens to bring takings actions. Of course, Congress could also dissolve the Court of Federal Claims entirely.338 To some extent, then, choosing a solution to the problem involves assessing how important it is to maintain a special-

336. It would raise a different “serious constitutional question” if Congress were to deny any judicial forum for a colorable constitutional claim. See, e.g., Webster v. Doe, 486 U.S. 592, 603 (1988).

337. Indeed, at least one commentator has proposed just that solution, arguing “that the Court of Federal Claims should be integrated more fully into the Judicial Branch by formally [being given] Article III status.” See Sisk, supra note 41, at 231–32.

ized court: a question Congress has already answered with respect to takings cases.

While the history of the old Court of Claims involves an unbroken line of specialist courts deciding many types of claims against the government, since that time Congress has "woven a broad tapestry of authorized judicial actions against the federal government," which "fit together into a reasonably well-integrated pattern of causes of action covering most subjects of dispute between the government and its citizens."339 Times have changed since the Tucker Act first allowed citizens to file takings claims against the federal government in the old Court of Claims. Congress met the challenge of the nineteenth century by creating a special forum to allow citizens to sue the sovereign, but citizens now routinely do so, not only in a special court in Washington, but in all of the district courts across the nation. There appears, therefore, to be at least somewhat less motivation for Congress to keep that court specialized than there once was.

Proponents of specialized courts generally identify efficiency, subject matter expertise, and uniformity as the three benefits offered by judicial specialization.340 With respect to takings claims, Congress has already eschewed a strong preference for those qualities, and indicated that it does not believe a specialized court is necessary, by granting generalist district court judges concurrent jurisdiction to consider takings claims for less than $10,000. While specialist courts may or may not have utility in certain complex legal areas,341 Congress does not appear to think them necessary for deciding takings claims. If specialization is considered important, Congress could resolve the problem identified in this Article without creating an entirely new court; Congress could simply direct large takings claims to a single, existing, Article III entity, such as the D.C. District Court.

The question about how best to deal with the problem identified in this Article also implicates the question of whether policymakers should simultaneously address other existing concerns with the Court of Federal Claims. Various commentators have presented concerns with the current structure of the Court of Federal Claims. For example, commentators have questioned the lawfulness of the Court of Federal Claims considering counterclaims in fraud.342 Another commentator has suggested that because the public rights doctrine no longer appears to justify Congress’ reliance upon Article I tribunals:


Reliance appears less a permissible exercise of Congress’s power of the purse than an attempt to shift matters within the judicial power of the United States to Article I tribunals. It thus seems appropriate to suggest some provision for the Article III judicial determination of all claims for money and property against the federal government.343

Other commentators have recognized that Congress’ authority to pay a debt,344 which is the only one of Congress’ explicitly defined powers that could be used to justify that court’s creation as an Article I court generally, is unlike the other powers that the Supreme Court had permitted Congress to rely upon to create Article I courts.345 Those commentators reason that, whereas the other Article I powers are those unique to a sovereign, the payment of a debt “is an obligation of any existing entity—a sovereign, a corporation, or an individual,” and “[a]lthough article I courts may be appropriate courts to consider issues that are unique to Congress’ sovereign powers, they are less appropriate for cases in which Congress is merely exercising its inherent authority, without regard to its status as sovereign.”346 That logic undermines the Court of Federal Claims’ ability, as an Article I court, to consider most of the cases currently within its jurisdiction. Each of those criticisms of the Court of Federal Claims affects the majority of the types of cases heard by the Court of Federal Claims, including contract and federal employee pay disputes, which comprise the majority of the claims in that body.347

While it is, of course, preferable to eliminate any potential concerns that commentators have identified in the course of addressing the one identified in this Article, this Article is concerned primarily with takings claims and how best to ensure their consideration by Article III judges. Moreover, as the Court of Federal Claims is currently exercising authority over a great number of claims related to the financial crisis that involve large sums of money and serious questions of public policy, developing a solution to the immediate problem identified in this Article, rather than waiting to see how the Supreme Court might address a challenge to the Court of Federal Claims’ consideration of a takings claim, seems a prudent course. It is important, therefore, to keep simplicity in mind when assessing the available options. While Congress could eliminate the Court of Federal Claims, as some have proposed, that is a more dramatic solution than necessary to solve the problem identified herein. In addition, while Congress could elevate the court’s status to an Article III court, that would affect many types of claims that need not be, and perhaps cannot be, decided by Article III courts. For example, the Court of Federal Claims cur-

343. Pfander, supra note 42, at 762.
345. See Sward & Page, supra note 111, at 411.
346. Id. at 411–12.
347. See supra note 19.
rently continues to issue advisory opinions in congressional reference cases, which the court could not do if Congress were to make the court an Article III entity.

Moreover, while elevating the court to an Article III entity would be straightforward, it is not clear that simply elevating the existing court to Article III status would entirely solve even the problem identified in this Article. There is some question whether Congress would also have to provide for a jury trial, pursuant to the Seventh Amendment, which is currently unavailable in the Court of Federal Claims. In Granfinanciera, the Supreme Court held that Congress may avoid the Seventh Amendment’s jury trial requirement only “where the Government is involved in its sovereign capacity under an otherwise valid statute creating enforceable public rights.” The Court explained that the question of whether adjudication of a right must be done by an Article III court, and the question whether a right to trial by jury is associated with that adjudication, are answered together:

Indeed, our decisions point to the conclusion that, if a statutory cause of action is legal in nature, the question whether the Seventh Amendment permits Congress to assign its adjudication to a tribunal that does not employ juries as factfinders requires the same answer as the question whether Article III allows Congress to assign adjudication of that cause of action to a non-Article III tribunal. For if a statutory cause of action . . . is not a “public right” for Article III purposes, then Congress may not assign its adjudication to a specialized non-Article III court lacking “the essential attributes of the judicial power.” And if the action must be tried under the auspices of an Article III court, then the Seventh Amendment affords the parties a right to a jury trial whenever the cause of action is legal in nature. Conversely, if Congress may assign the adjudication of a statutory cause of action to a non-Article III tribunal, then the Seventh Amendment poses no independent bar to the adjudication of that action by a nonjury factfinder.

348. U.S. CONST. amend. VII (“In Suits at common law, where the value in controversy shall exceed twenty dollars, the right of trial by jury shall be preserved . . . .”).


Thus, whether Congress can submit a legal issue to a legislative court and whether that court can dispense with a civil jury on that legal issue are answered by the same analysis.352 The Court stressed, however, that the two correspond only when the issue being adjudicated is a legal, as opposed to an equitable, question.353 In a particularly interesting analysis, Eric Grant marshaled a great deal of support for the proposition that takings claims are legal actions, or, in the words of the Seventh Amendment, “suits at common law.”354 If he is correct, then a litigant who files a Fifth Amendment takings claim may be entitled to a jury trial pursuant to the Seventh Amendment.355

The conclusion that a takings claim brought against the federal government requires the availability of a jury trial was strengthened by the Supreme Court’s decision in City of Monterey v. Del Monte Dunes at Monterey, Ltd.356 There, the Supreme Court held that a takings claim brought pursuant to section 1983 requires the availability of a jury trial.357 In so finding, the Court noted, based upon its analysis of historical precedent, that not all takings claims require jury trials: “[b]ecause the jury’s role in estimating just compensation in condemnation proceedings was inconsistent and unclear at the time the Seventh Amendment was adopted, this Court has said ‘that there is no constitutional right to a jury in eminent domain proceedings.’”358 The Court explained, however, that “[e]arly authority finding no jury right in a condemnation proceeding did so on the ground that condemnation did not involve the determination of legal rights because liability was undisputed . . . .”359 In contrast, when an inverse condemnation proceeding is filed, liability is in question. Consider the AIG

352. See id.
353. See id. at 42 n.4. (“The Seventh Amendment protects a litigant’s right to a jury trial only if a cause of action is legal in nature and it involves a matter of ‘private right.’”).
354. See Grant, supra note 350, at 208. But see Sward, supra note 56, at 1125 (“The easy answer to the question whether suits against the government are suits at common law is that they are not.”).
355. While that conclusion conflicts with the Court’s categorical statement in 1880, that “[s]uits against the government in the Court of Claims . . . are not controlled by the Seventh Amendment,” the decision that announced that rule, McElrath v. United States, like the Williams decision, was based upon the notion that all suits in the Court of Claims were permissible only because the federal government waived its sovereign immunity. See McElrath v. United States, 102 U.S. 426, 440 (1880) (“The government cannot be sued, except with its own consent. It can declare in what court it may be sued, and prescribe the forms of pleading and the rules of practice to be observed in such suits. . . . If the claimant avails himself of the privilege thus granted, he must do so subject to the conditions annexed by the government to the exercise of the privilege.”). As this Article has already demonstrated, the sovereign immunity defense does not apply to takings claims.
357. See id. at 722.
358. See id. at 711 (quoting United States v. Reynolds, 397 U.S. 14, 18 (1970)) (citing Bauman v. Ross, 167 U.S. 548, 593 (1897)).
359. Id. at 712–13.
shareholder lawsuit. The government disputes not only the amount of compensation the shareholders seek, but also the issue of whether the government’s actions amount to a taking. Under the Court’s analysis in City of Monterey, there is at least some question whether a takings claim requires the availability of a jury.

With all of those considerations in mind, it appears that the simplest way to ensure that takings claims are heard and decided by Article III judges is for Congress to grant federal district courts jurisdiction to consider all takings cases, not just the takings cases seeking less than $10,000. In that way, Congress could address the particular issue identified in this Article without completely altering the status quo. An added benefit to that solution is that it would place takings claims in the same court that is currently entrusted to consider claims based upon the remainder of the Bill of Rights. There appears to be little downside to this solution. It would not significantly add to the workload of the district courts, as the number of takings claims filed in the Court of Federal Claims over the last decade amounts to the equivalent of less than one case per district court per year.360

III. Conclusion

When a citizen wants to sue the federal government to collect money pursuant to the Takings Clause of the Fifth Amendment to the Constitution, that takings claim can be filed in one of the Article III district courts if the citizen is seeking less than $10,000. If someone believes they are owed more than that amount, however, the citizen must file that takings claim in the Court of Federal Claims, an Article I court. A court not entirely independent from the political branches of government thus currently decides all large takings claims. Because Article III is designed to prevent interference by the political branches in those cases in which they might have some interest, the current scheme is precisely backward with respect to the policies underlying Article III. Moreover, the current scheme is not permissible upon any of the grounds that the Supreme Court has sanctioned for Congress’ use of non-Article III entities. The current situation is, therefore, unconstitutional. That situation is also a recent development, appears to have been created accidently, and is inconsistent with Congress’ intent to create a truly independent court to hear claims against the federal government. Now is the time to fix this error and bring takings claims back to Article III courts.

Right now, the Court of Federal Claims is deciding a number of takings claims involving the recent financial crisis in which the political branches have a profound interest. The tension between the policies of Article III and the current structure has thus extended from the theoretical realm to the tangible. Congress should act now to ensure that the judges who decide those and similar cases are truly independent by pro-

360. For statistics on cases in the Court of Federal Claims, see supra note 19.
TAKING BACK TAKINGS CLAIMS

2015]  

viding Article III district courts with jurisdiction to consider takings claims. While Congress could fix the problem in any number of ways, the simplest solution is for Congress to simply permit the federal district courts—which already consider takings claims—to also hear the more expensive, and important, takings claims. This approach would demonstrate that the United States is serious about ensuring that it complies with the constitutional command that private property not be taken for public use without paying just compensation.361

361. See U.S. CONST. amend. V.
144  VILLANOVA LAW REVIEW  [Vol. 60: p. 83
2015]

CHANGING COURSE IN THE ANTI-DOMESTIC VIOLENCE LEGAL MOVEMENT: FROM SAFETY TO SECURITY

MARGARET E. JOHNSON*

"[In the field of cybersecurity] resilience is what allows a system to endure security threats instead of critically failing. A key to resilience is accepting the inevitability of threats and even limited failures in your defenses. It is about remaining operational with the understanding that attacks and incidents happen on a continuous basis. . . . [T]he objective should be to prioritize resources and operations, protect key assets and systems from attacks, and ultimately restore normal operations."1

MARY

MARY met Todd in Florida. Todd moved back to his home state of Washington, D.C.; after a few months, he asked her to come and join him. Although she had no family or friends in D.C., Mary followed Todd willingly. She loved Todd and, flat broke and unsuccessful in finding a job in Florida, hoped she would have better prospects up north.

After Mary moved in with Todd, however, he would not let her look for a job. He would not let her out of the house without his permission. If he thought she had done something that was against the many rules he established for the household, he would hit her. He told her that if she ever left, he would find her. She had no access to people or a telephone. Mary’s necessary medicine for her long-term illness ran out, but Todd refused to let her refill the prescription. As the days went on, she got more and more sick. She knew she needed to get out and get to a hospital.

* Associate Professor, Director, Family Law Clinic, and Co-Director, Center on Applied Feminism, University of Baltimore School of Law. I appreciate the comments I have received on this piece during the UB Law Faculty Research and Development Series, the Law and Society Annual Meeting, the Converge! Conference at the University of Miami Law School, and the Clinical Law Review Writing Workshop at NYU Law School. Special thanks to Dionne Koller, Michele Gilman, Leigh Goodmark, Karen Czapanski, Donna Coker, Margaret Drew, Phyllis Goldfarb, James Ptacek, Fred Brown, and Mae Quinn. I especially thank Matt Fraidin for his amazing assistance in helping me think through the ideas in this Article. I am grateful to my research assistants Lana Castor, Amy Lazas, Nicole Whitaker, and Addie Crawford, and the University of Baltimore Law Library staff, including David Matchen, Bijal Shah, and Adeen Postar. I appreciate the support of the University of Baltimore School of Law Summer Research Fellowship. All errors are of course mine.


2. Mary is a fictional name, but her story is based on the experiences of a client in my law clinic.
One day when Todd left the home to run an errand, Mary quickly left the house and asked a stranger to take her to a hospital. The hospital discharged Mary after treating her for a week.

Homeless, penniless, and afraid to return to Todd, Mary sought and won a temporary protective order and with that successfully applied for a thirty-day stay in an emergency shelter. Two weeks later, the court issued a one-year order that enjoined Todd from further abusing her, coming near her, or contacting her. Both orders gave Mary the power to call the police to enforce them through arrest and criminal prosecution.3

While in the shelter, Mary received limited transportation vouchers. The vouchers covered her travel from the distant shelter to the court, but they did not cover her travel to the library where she wanted to apply for jobs on the computer, to a nonprofit organization offering professional attire she wanted for her job interviews, or to the job interviews she obtained. She accessed meager emergency grocery-store vouchers, with which she purchased food and personal items not provided by the shelter, such as toilet paper and tampons. When Mary’s thirty-day emergency shelter time came to an end, she could not access any long-term housing because it was in short supply. She no longer had her transportation or food vouchers. With no money of her own, and feeling like she had no other option, Mary reluctantly returned to Todd.

I. INTRODUCTION

Protection of women in the name of safety from domestic violence has gone too far. This Article argues that domestic violence law and practice overemphasize women’s short-term safety in ways that deprive women

3. But see, e.g., Town of Castle Rock, Colo. v. Gonzales, 545 U.S. 748, 768 (2005) (holding that, under U.S. Constitution’s Due Process Clause, wife had no constitutionally protected property interest in police enforcement of her restraining order against her husband). Subsequent to the Supreme Court’s decision, Ms. Lenahan (formerly Gonzales) brought a successful claim of human rights violations under the American Declaration on the Rights and Duties of Man against the Castle Rock Police Department for failing to protect her and her children, who were in fact killed by her estranged husband during the pendency of her protective order, and human rights violations by the U.S. courts for failing to remedy her situation. See Caroline Bettinger-Lopez, Introduction: Jessica Lenahan (Gonzales) v. United States: Implementation, Litigation, and Mobilization Strategies, 21 Am. U. J. GENDER SOC. POLY & L. 207, 219–20 (2012).
of dignity and agency, counterintuitively, make women less safe. Promoting a different paradigm, I argue women need security, both short-term security that permits women to determine the best response to the domestic violence, as well as options for long-term security—economic security, housing security, health security, and relationship security—that will support their own resilience and allow them to grow, thrive, and direct their lives after experiencing intimate partner abuse.

With federal and state domestic violence funding devoted largely to law enforcement, and civil remedies emphasizing only physical violence and primarily short-term remedies, women subjected to abuse get most of our attention only in moments of crisis. Emergency measures temporarily

4. Human dignity is the capacity of persons to be decision makers using “their own conscience and conviction” regarding “the most fundamental questions touching the meaning and value of their own lives.” Margaret E. Johnson, Balancing Liberty, Dignity, and Safety: The Impact of Domestic Violence Lethality Screening, 32 Cardozo L. Rev. 519, 546 (2010) [hereinafter Johnson, Balancing Liberty] (quoting Ronald Dworkin, Unenumerated Rights: Whether and How Roe Should be Overruled, 59 U. Chi. L. Rev. 381, 426 (1992)) (internal quotation marks omitted). Dignity recognizes that women subjected to abuse have the right to decide how best to address the intimate partner abuse without unwanted state intrusion or coercion. See id. at 547. For a discussion of the importance of dignity in domestic violence law and policy, see id. at 519, 543–46.

5. Agency is the “capacity for self-definition and self-direction” despite subordination based on gender.” Margaret E. Johnson, Redefining Harm, Reimagining Remedies, and Reclaiming Domestic Violence Law, 42 U.C. Davis L. Rev. 1107, 1114 n.24 (2009) [hereinafter Johnson, Redefining Harm] (quoting Kathryn Abrams, Subordination and Agency in Sexual Harassment Law, in Directions in Sexual Harassment Law 112–14 (Catharine A. MacKinnon & Reva B. Siegel eds., 2004)). It is important that the domestic violence legal system and policy respect the agency of women subjected to abuse, because research shows that women are the best predictors of their risk of future abuse and are able to make decisions that best address the abuse. See id. at 1124–29.

6. To discuss domestic violence, I use the terms domestic violence and intimate partner abuse interchangeably to mean all the forms of abuse inflicted on an intimate partner in an attempt to exert power and control over her. I focus on male-on-female abuse because it is the most prevalent form of intimate partner abuse. See Johnson, Redefining Harm, supra note 5, at 1110 n.4 (citing Joan B. Kelly & Michael P. Johnson, Differentiation Among Types of Intimate Partner Violence: Research Update and Implications for Interventions, 46 Fam. Ct. Rev. 476, 481–82 (2008)). Recognizing that men can be subjected to abuse as well, I often discuss persons subjected to abuse, because research shows that women are the best predictors of their risk of future abuse and are able to make decisions that best address the abuse. See id. at 1124–29. As I discuss in this Article, domestic violence law and policy need to do a better job of accounting for the whole person, and thus, I prefer the term woman or person subjected to abuse. For the same reason, I do not call the person perpetrating the acts of abuse a batterer as the term may reduce his personhood in ways that may not resonate with the woman in relationship with him. See id. I will use the term victims when discussing how the criminal justice system views women subjected to abuse. And in discussing court cases, I use the terms designated by the court system, such as petitioner, for the person subjected to abuse, and respondent or defendant, for the person who committed the abuse.
separate women from their partners who are abusing them, offering women a sometimes desperately-needed refuge.

But then what? Like Mary, many women who experience domestic violence find themselves homeless and jobless when the dust settles. Like Mary, they may have found temporary, emergency housing in a shelter for victims of domestic violence. They may have received crisis counseling. But then, life marches on.

Persons subjected to intimate partner abuse may have been forced from their homes because they lacked ownership rights and to avoid future abuse. They may have lost their jobs because injury or stress caused them to miss work, or because their partner locked them in the house, or because their partner’s on-the-job harassment frightened a boss or co-workers. In some instances, women may lose their children to foster care, accused of “failing to protect” the children from the man who has endangered them. Persons subjected to abuse may have lasting physical and mental health damage from the intimate partner abuse to which they were subjected.

Displaced from home, employment, and education; exiled from their community; and separated from their children, women who experience abuse may be safe when the most conspicuous crisis passes. Or the responses taken through physical separation may have made them less safe. Either way, are they living their fully-actualized life? Not yet.

And so the anti-domestic violence movement is at a choice moment. It is time for a change. It is time for the anti-domestic violence movement to consider taking a critical look at the state and institutional response to domestic violence and the current goal of safety, asking whether each established response or new initiative is addressing the needs of persons subjected to abuse in terms of their personal goals, resilience, agency, and dignity. The idea is for the response to be centered on the person’s experience of the abuse and not centered on the state or institution. This shift is one that needs to move from a primarily short-term safety focus to a short and long-term security focus. Security would include increasing access to resources to increase economic, housing, health, and relationship security of persons subjected to abuse. Security would require that all new legislative, policy, and funding initiatives address these areas.


This Article argues that the vast majority of current domestic violence law and practice is oriented almost obsessively on short-term safety. Domestic violence policy has accordingly prioritized the narratives of physical violence, crime, and danger over women’s experience of the broad-range of abuse to which they are subjected, their experienced harms, and their goals for greater satisfaction, including reduction or elimination of the abuse.\(^{10}\)

I argue that the criminal justice system has been the axis around which much current anti-domestic violence policy has been formulated. As the iconic manifestation of state power, domestic violence policy and practices have focused on crimes, physical violence, and the pursuit of the physical safety of community members.\(^{11}\) Examples of policies and practices focused on physical safety include laws mandating the arrest of suspected perpetrators of domestic violence and laws forbidding prosecutors from dismissing prosecutions—despite the fact that both approaches have very mixed results in decreasing violence and sometimes increase it.\(^{12}\) Other examples include expansive funding for these criminal justice system initiatives, thereby taking resources away from civil legal system re-

Every anti-domestic violence law, policy, funding priority, administrative rule, you name it, should be subjected to a material resources test, which asks: What is the impact, what is the effect of this law, policy, regulation, et cetera, on the material resources of the women who are likely to come in contact with this law, policy, regulation, et cetera? Will this have an impact that either primarily or secondarily gives women greater access to material resources? Because women’s circumstances differ in ways that dramatically affect their access to material resources, further, I argue, the standard for determining whether or not a given law, policy, et cetera, passes muster under a material resources test should be the situation of the women who are in the greatest need, those who are dramatically affected by the inequalities of gender, race, and class.

\(^{10}\) In addition, as Beth Richie recently commented, the
Emerging feminist analysis [of the domestic violence movement] did not include race and class . . . [as the] white women ha[d] the power to define which problems are real. And from those definitions came very narrow definitions of who is entitled to protection . . . more married you are, the more American, the more legal, the more temporarily poor, the fewer felony convictions the more you will be entitled to the attention and resources of this movement. And these very narrow definitions of the problem aligned with state interest.

\(^{11}\) See infra notes 52–119 and accompanying text.

\(^{12}\) See id.
sponses and grassroots organizations that provide necessary short and long-term responses to domestic violence;\textsuperscript{13} a civil legal system that prioritizes physical abuse and crimes, thereby failing to remedy the full range of abuse that society has an interest in eradicating;\textsuperscript{14} not enough funding for emergency shelters and very little funding for longer-term housing,\textsuperscript{15} despite the importance of housing in decreasing violence and stabilizing the life of a person subjected to abuse;\textsuperscript{16} and not enough funding or support responses to increase the economic, housing, health, and relationship security of persons subjected to domestic violence.\textsuperscript{17} In addition, the response has been focused on short-term assistance, such as arrests, prosecutions, incarcerations, short-term civil protection orders, and emergency shelters, aimed at the immediate safety risk with far less attention to the longer-term.\textsuperscript{18}

The negative consequences of the safety paradigm include the inability of persons subjected to abuse to be heard in the legal system because they do not experience severe enough violence,\textsuperscript{19} because what they want is not deemed “safe” by the institutions and the state from which they may be seeking assistance,\textsuperscript{20} and because they have not sought assistance but are coerced or ordered to receive assistance to be “safe,” as deemed by the institutions and the state.\textsuperscript{21} Being safe and free from violent harm is important. Yet there is an important difference between the woman subjected to abuse choosing to employ actions to be safe and free of violent harm and the state requiring the woman to physically separate from her abusive partner. In addition, there is an important difference between the

\textsuperscript{13} See id.; see also Coker, Economic Rights, supra note 9, at 188.

\textsuperscript{14} See Johnson, Redefining Harm, supra note 5, at 1131–38.


\textsuperscript{16} See id. at 11–17.

\textsuperscript{17} See infra notes 52–119 and accompanying text.

\textsuperscript{18} See Erika A. Sussman, The Civil Protection Order as a Tool for Economic Justice, Advoc. Q., 2006 Issue 3, at 1 (citing to legislative history for state civil protection order statutes and showing purpose is to promote safety).

\textsuperscript{19} See Johnson, Redefining Harm, supra note 5, at 1145–48.


many ways that she may act in order to create safety for herself—such as having her own finances, or continuing the relationship but eradicating the abuse—and the almost one-size-fits-all way in which the state says she should act to create safety—physically separating from her partner. The state and institutions too often use physical separation reflexively in ways that are not responsive to the situation of a woman who is experiencing or has experienced abuse. As a result, a woman’s own agency and dignity can be undermined in pursuit of what is deemed safe by others. The harm of this undermining may be greater than the harm that the safety measures would address. These institutional and state responses can cause significant negative consequences as shown by the research that says that women subjected to abuse are the best predictors of violence, that non-physical violence is more damaging to women subjected to abuse than the physical violence, and that the most effective response to help women subjected to abuse is personalized. A personalized response would ensure that their goals are listened to and that there are many options from which they can evaluate and choose as opposed to having decisions made for them.

Moreover, the safety paradigm reinforces the problematic victim paradigm. This construction of safety is problematic for several reasons. It reinforces gender stereotypes of helpless women who stay with abusers because they cannot think rationally and take actions to protect themselves. The safety paradigm also constructs women’s worlds as being

23. See Johnson, *Redefining Harm*, supra note 5, at 1113 n.20 (stating that women identified “threats and verbal abuse as more devastating than the physical” (quoting ELIZABETH M. SCHNEIDER, *Battered Women & Feminist Lawmaking* 66 (2000)) (internal quotation marks omitted)).
24. See Donna Coker, *Shifting Power for Battered Women: Law, Material Resources, and Poor Women of Color*, 33 U.C. DAVIS L. REV. 1009, 1017–20 (2000) [hereinafter Coker, *Shifting Power*] (“It is a cruel trap when the state’s legal interventions rest on the presumption that women who are ‘serious’ about ending domestic violence will leave their partner while, at the same time, reducing dramatically the availability of public assistance that makes leaving somewhat possible. . . .  The second problem with equating separation with safety is that legal actors frequently believe a corollary presumption: women’s use and full cooperation with legal remedies increases their safety. . . .  The third problem with equating separation with safety is that frequently the laws and services based on a separation premise devalue women’s connections with their partner and their investment in building family.” (footnotes omitted)); Johnson, *Balancing Liberty*, supra note 4, at 578–80.
26. One reason offered for this framing is that it is rooted in Dr. Lenore Walker’s work on learned helplessness and battered women’s syndrome, that women subjected to abuse are coerced into passivity and non-action. See Leigh Goodmark, *A Troubled Marriage: Domestic Violence and the Legal System* 57–61 (2013); Kohn, *supra* note 20, at 206–09. Another reason, I argue, is that this construction of women as victims is much broader than originating in domestic violence theory and is rooted in gender oppression writ large. For instance, such a construction of women as victims is seen in such early writings as Blackstone’s commentaries discussing coverture and women’s necessary loss of legal identity upon marriage. See William Blackstone, *Commentaries* *430* (“By marriage, the hus-
filled with dangers and havens—the dangers are identified and removed by the state or institutions, while the havens are identified and provided by the state or institutions. Because the state and institutions construct the “danger” as physical violence, laws and funding streams prioritize this danger as the “problem.” This leaves fewer resources for the broader range of intimate partner abuse. Relying on the safety construct, lawmakers funnel a large percentage of public funds into the criminal justice system in response to domestic violence.27 The preference for criminal justice measures contributes to long-standing systemic issues plaguing persons subjected to intimate abuse. The issue is illustrated in the housing context where, even though domestic violence is the leading cause of family homelessness, a relatively small percentage of public funds is spent on housing for persons subjected to abuse.28 Despite research that shows women are best able to predict their own risk, the state and institutions construct a “haven” based on their view of the problem, thereby prioritizing their ideas of how best to save women from the violence, while not prioritizing women’s own perception of what is “safe” for themselves today, tomorrow, and in the future. As a result, the state and institution construction of danger and havens does not always effectuate even short-term safety because it misdiagnoses the problem and generates inappropriate options. The state and institutional construction of safety deters the creation of laws and funding that support women’s security—their resiliency, agency, and dignity—over the long-term, not just in the short-term.

This Article argues that the safety paradigm misses the forest for the trees. Arguing for a reorientation toward “security,” this Article suggests that the current “safety paradigm” is harmful to women and fails even to achieve its stated objective of ensuring women’s safety. A broadened, realistic vision of women’s strengths and needs can provide women with the support, tools, and resources needed to fully repair their lives from the harm inflicted in abusive relationships. Measures designed to promote security recognize women’s interests in sustained growth, health, and agency. Reorienting societal responses toward ensuring women’s health, housing, and economic viability can provide the stability women need to make choices consistent with their own definitions of dignity and self-worth while, not coincidentally, keeping them safer in the long run. The reframing of the domestic violence movement from the predominant goal of safety to security will help reshape a legal system that can be more re-

27. See infra notes 52–119 and accompanying text.
responsive to the needs and harms of persons subjected to abuse, such as Mary, while also addressing domestic violence more effectively.

This Article proceeds in three parts. Part II examines the prevailing short-term safety paradigm goal of anti-domestic violence legal and funding systems. This Part also analyzes the different policies and laws used to effectuate the safety paradigm, such as the criminal justice system itself, mandatory arrest and prosecution policies, mandatory no contact orders, civil protection orders (CPO) mandating physical separation, domestic violence shelter mandated separation, lethality assessment programs, and the lack of robust long-term options for housing, employment, and financial stability. This Part concludes by discussing how these policies affect women subjected to abuse as well the epidemic of domestic violence generally.

In Part III, I propose that security, not safety, serve as the philosophical driving force behind domestic violence law and practice. I argue that law, funding, policy, and practice should promote long-term assistance intended to support resilience, agency, and dignity. Such a response would include long-term housing options, including the ability to be within a supportive community; long-term employment or other income-generation options, with long-term employment supports such as career counseling, education, quality child care, and available transportation; increased income and asset-building opportunities without the barriers of discrimination; access to public assistance; long-term physical and mental health care, preventative, and treatment options; long-term civil protective and other court orders that provide injunctive, family, housing, and economic remedies; and the development of enhanced social capital and community connection for persons subjected to abuse. An approach oriented toward long-term measures, such as these, would provide the security and stability women need to determine how best to address the violence in their relationships.

Finally, in Part IV, I offer examples of how the domestic violence legal and funding systems could do more to address security. For example, the legal and funding systems could address all forms of intimate partner abuse, end mandatory responses to domestic violence, address barriers to employment and access to secure and livable wages, address coerced debt, expand opportunities for asset building, increase housing options, support social capital development through time banking, and provide monetary damages for domestic violence.

II. DOMESTIC VIOLENCE: SAFETY PARADIGM

To understand the prevailing safety paradigm, it is important to understand what domestic violence is and how our current laws and policies address domestic violence.
A. Domestic Violence

Research has demonstrated that domestic violence, or “intimate partner abuse,” can take a wide range of forms, in addition to physical aggression by one partner against another. According to social work professor Judy Postmus, for example, intimate partner violence is:

[A] pattern of behavior in a relationship by which the batterer attempts to control his victim through a variety of tactics. These tactics may include fear and intimidation, physical and/or sexual abuse, psychological and emotional abuse, destruction of property and pets, isolation and imprisonment, economic abuse, and rigid expectations of sex roles.29

This pattern of behavior has been called “[c]oercive [c]ontrolling [v]iolence”30 or “coercive control.”31 Such abuse may involve one of these tactics alone, such as physical abuse, or employ several or all tactics.32 As psychologists Mary Ann Dutton and Lisa Goodman have documented in their studies, there are nine areas of coercive control used by abusers on their partners: “personal activities/appearance,” “support/social life/family,” “household,” “work/economic/resources,” “health,” “intimate relationship,” “legal,” “immigration,” and “children.”33 Of the various forms of domestic violence, women subjected to intimate partner abuse “frequently describe the threats and verbal abuse as more devastating than the physical.”34 Further, research shows that psychological abuse often can develop into physical abuse.35

Focusing on economic abuse, law professor Dana Harrington Conner catalogues three forms: (1) targeting women of limited economic means, in poverty, in vulnerable situations because of immigration status, limited English proficiency, connection to their community, or limited social capital; (2) creating financial dependence by precluding access to money and

30. Kelly & Johnson, supra note 6, at 481.
32. See Kelly & Johnson, supra note 6, at 481 (“Abusers do not necessarily use all of these tactics, but they do use a combination of the ones that they feel are most likely to work for them. Because these nonviolent control tactics may be effective without the use of violence (especially if there has been a history of violence in the past), Coercive Controlling Violence does not necessarily manifest itself in high levels of violence.”).
35. See Johnson, Redefining Harm, supra note 5, at 1113 (citing Kelly & Johnson, supra note 6, at 483).
property resources, sabotaging one's economic stability and property ownership, precluding or sabotaging employment and education opportunities, and/or isolating from people; and (3) limiting access to economic resources, destroying whatever economic resources she may have, and damaging social capital. The coercion around work and finances is a significant form of intimate partner abuse because it is so powerful. Such coercion can include “[m]aking or attempting to make a person financially dependent, e.g., maintaining total control over financial resources, withholding access to money, forbidding attendance at school or employment.” Further tactics include acts such as “sabotag[ing] a woman’s efforts to find a job or attend a job training . . . [by] turning off her alarm clock so she will be late for work, cutting off her hair to cause her great embarrassment, inflicting visible injuries or creating conflicts before crucial events, and hiding or destroying her books, homework, or clothing.”

Even if the woman becomes successfully employed, the abuse may continue with the intimate partner “disrupting her transportation or childcare arrangements or harassing her at work.”

As discussed above, domestic violence can be perpetrated in multiple ways, with varying tactics to exert coercive control, beyond simply physical violence.

B. The Safety Paradigm

The current safety paradigm of anti-domestic violence law and policy is the result of responses to the state’s initial absence of a response to domestic violence. Domestic violence advocacy in the 1960s was a grassroots movement, focused on all forms of abuse, and on women’s empowerment, agency, and dignity as a way to address the abuse. Women who experienced abuse were provided counseling focused primarily on consciousness-raising about domestic violence as a form of gender oppression. Most of the efforts were expended to create a network of grassroots shelters to which women could escape from abusive relationships.

In the 1970s and 1980s, feminists focused on the void of state responses to the continued epidemic levels of abuse and lobbied the state
for formal laws and policies to address intimate partner violence.\footnote{See id. at 196–97.} That lobbying focused on domestic violence as physical violence and especially criminal acts.\footnote{Cf. Buzawa & Buzawa, supra note 34, at 4–5 (“[C]riminal justice case processing requirements may have directly defined the parameters of permissible contact by criminalizing certain violent conduct, while tacitly condoning harassment or other strategies of coercive control, actions that in practice rarely result in arrest or prosecution.”).} Finding that the state’s response was nonetheless still often lacking, advocates successfully lobbied for new laws to remove state actors’ discretion by mandating the arrest and prosecution of suspected batterers.\footnote{See Kohn, supra note 20, at 196–97, 199; Miccio, supra note 21, at 278.} States enacted civil protection order laws, permitting women to seek protection primarily from physical abuse that would constitute a crime regardless of the outcome of criminal proceedings against a batterer.\footnote{See Kohn, supra note 20, at 196–97.} Through this engagement of the state, philosophical and funding priorities changed from women’s own empowerment responses addressing broad forms of intimate partner abuse to the state and other outsiders’ responses addressing primarily criminal physical abuse. In 1994, the Violence Against Women Act\footnote{Violence Against Women Act of 1994, Pub. L. No. 103-322, §§ 40001–40703, 108 Stat. 1902.} (VAWA) was enacted, and 1.6 billion dollars\footnote{See Robert F. Friedman, Note, Protecting Victims from Themselves, but Not Necessarily from Abusers: Issuing a No-Contact Order over the Objection of the Victim-Spouse, 19 WM. & MARY BILL RTS. J. 235, 239 (2010) (citing Julie Goldscheid, United States v. Morrison and the Civil Rights Remedy of the Violence Against Women Act: A Civil Rights Law Struck Down in the Name of Federalism, 86 CORNELL L. REV. 109, 113 (2000)).} were allocated to support legal, court, and community-based responses to domestic violence, which were focused on combatting physical violence and used short-term physical separation as the remedy.\footnote{See Kohn, supra note 20, at 200 (“Most individuals in the justice system hope to protect battered women from further violence by seeking to remove them from abusive relationships.”). VAWA also contained a civil rights provision for persons subjected to abuse that permitted them to get injunctive and declaratory relief as well as compensatory and punitive damages for crimes of violence motivated by gender. See Goldscheid, supra, at 113–15. This provision was later struck down by the Supreme Court as unconstitutional. See United States v. Morrison, 529 U.S. 598 (2000).}

As a result of these efforts, there are police who will now arrest persons for abusing their partner, there are prosecutors who will finally prosecute these crimes, there are civil protection order laws, and there is a developed coordinated response of criminal and civil courts dedicated to domestic violence.\footnote{See generally Deborah Epstein, Effective Intervention in Domestic Violence Cases: Rethinking the Roles of Prosecutors, Judges, and the Court System, 11 YALE J. L. & FEMINISM 3 (1999).} There are also many services developed to respond to domestic violence, including civil legal services for civil protective orders, emergency shelters, counseling, transitional housing, and more.\footnote{See Goodmark, supra note 26, at 25–28.}
These developments have occurred in large part due to the success of the anti-domestic violence movement and its alliance with the state, including the large amounts of government funding that support the government-designed response and, to a much lesser extent, the non-legal response. Many of these developments constitute the safety paradigm.

C. The Safety Paradigm in Contemporary Law, Policy, and Practice

The safety paradigm has four components. First, it is focused primarily on physical and sexual violence as domestic violence rather than also looking at psychological, emotional, and economic abuse and coercive control. Second, its goal is principally to address the immediate crisis and defuse the situation through short-term remedies. Third, its goal is to ensure that the government ends future physical and sexual violence through the physical separation of the parties. Fourth, it provides that the state, and not the person subjected to abuse, is the decision maker as to how best to address the domestic violence. There are many examples of the safety paradigm in the current domestic violence law and funding schemes.

The criminal justice system demonstrates a safety paradigm as it is generally only concerned with physical violence, injuries caused by physical violence, and remedies ensuring physical separation through arrest, criminal prosecution, incarceration, fines, criminal stay away orders, and batterers intervention programs. Moreover, the state is the decision maker as to the remedy pursued and not the person subjected to abuse.

52. See Buzawa & Buzawa, supra note 34, at 4–5 (“[C]riminal codes are typically rather blunt instruments, defining violence as individual acts, usually a physical assault or threat of physical harm intended to cause physical harm.”).

53. See id. at 5 (explaining criminal justice system has “implicit requirement for the identification of a crime with a defined victim and offender in the context of a recognized applicable criminal statute”); Jennifer G. Long et al., Aequitas, Model Policy for Prosecutors and Judges on Imposing, Modifying and Lifting Criminal No Contact Orders 2 (2010), available at http://www.aequitasresource.org/model_policy.pdf (stating that goals of criminal justice system regarding domestic violence are “to seek justice, protect the victim and the community, hold the offender accountable for his crimes, prevent and deter future crime, and rehabilitate the abuser”).

54. See Buzawa & Buzawa, supra note 34, at 4 (discussing ascendancy of criminal justice system domestic violence response: “the United States currently has a propensity to use coercive legal powers to ‘solve’ social problems”).

55. It should be noted that prosecutors’ offices do staff victim advocates who work with the persons subjected to abuse and address some short-term needs such as housing and food. Primarily, these victim advocates are in place to help obtain the cooperation of the person who was subjected to abuse in the prosecution of the abuse. In other words, these services are offered to fulfill the criminal justice system’s goals and not any incompatible goals of the person subjected to abuse. See generally Thomas L. Kirsch II, Problems in Domestic Violence: Should Victims Be Forced to Participate in the Prosecution of Their Abusers?, 7 Wm. & Mary J. Women & L. 383, 398–406 (2001); Kohn, supra note 20.

56. See Buzawa & Buzawa, supra note 34, at 4, 126–28.
One clear example of this point is the mandatory arrest laws. As of December 2011, twenty-two states plus Washington, D.C. had mandatory arrest laws that required the responding police officer to arrest at least the primary aggressor if an assault had occurred. This requirement exists even if the person subjected to abuse opposes or does not request the arrest. In addition, thirty-three states had mandatory arrest laws that required a police officer to arrest a respondent who violated a civil protection order. Six states preferred arrest laws “encouraging” police officers to make arrests if an assault occurred. Such laws permit police officers discretion in deciding whether to arrest, including some consideration of the crime victim’s opinion on the arrest.

“No-drop” prosecution policies provide another example of the state acting as the sole decision maker as to the remedy pursued. Many states have developed mandatory prosecution, or no-drop prosecution policies, as a response to the early outcry by domestic violence advocates that the state was not seriously addressing domestic violence. In addition to the police’s failure to arrest for domestic violence, advocates cited the prosecutors’ failure to proceed with prosecutions of those who were in fact arrested. No-drop prosecution policies may be contained in legislation or issued by state attorney general offices and are encouraged and financially supported by federal law. Under “hard” no-drop policies, neither prosecutors nor the person subjected to abuse have any discretion as to whether a case supported by evidence will go forward. Prosecutors must prosecute the case even when the crime victim, the person subjected to abuse, requests that the case be dismissed and/or refuses to cooperate with the prosecutor by being interviewed or testifying at trial. “Soft” no-drop prosecution policies permit prosecutors to dismiss criminal cases based on evidence.

58. See id.
59. See id.
62. See Buzawa & Buzawa, supra note 34, at 194–95 (“[Sixty-six percent] of a sample of local prosecutors in jurisdictions with populations of over 250,000 had no-drop policies.”); see also Miccio, supra note 21, at 265.
64. See Miccio, supra note 21, at 266.
the victim’s desires if the victim has physically separated from the defendant.65 One stated justification for a hard no-drop policy is victim safety. Such safety is intended to result from the prosecution and eventual incarceration, criminal stay away orders, or rehabilitation training of defendants that will in turn reduce future battering through separation, stigmatizing the relationship, and notifying the victim and defendant that the victim cannot be controlled by the defendant. Moreover, the no-drop policy informs victims who desire to stay with the defendant that their wishes are the result of coercion and not agency.66

Criminal no contact orders are another example of the safety paradigm. Criminal no contact orders are instituted almost routinely during the prosecution of a domestic violence crime as part of pretrial release and are often included as part of the sentence.67 These stay away orders are ordered even against the will of the person subjected to abuse.68

Civil protection order laws permit persons subjected to abuse to petition a civil court for various forms of injunctive relief, such as no further abuse, stay away, no contact, and vacate the home orders; family-based relief, such as a temporary child custody and support order; and other relief, such as domestic violence counseling and control over some forms of personal property.69 All United States jurisdictions permit a petition for a protective order to be based upon a criminal or violent act between intimate cohabitating partners.70 CPOs contain both safety and security elements, but the predominant elements are the safety elements. Only one-third of all jurisdictions recognize or remedy harms from emotional and economic abuse, and only one-third of the jurisdictions recognize and provide a remedy for coercive control.71 Further, only thirty-seven CPO laws provide a catch-all provision that could provide an economic remedy beyond child or spousal support to include compensatory and punitive damages.72 On the other hand, all CPOs contain directives specifically based on the physical separation safety paradigm and last for a short time, usually one year in length.73 In some jurisdictions, the petitioners’ choice

65. See Friedman, supra note 48, at 242 (citing BUZAWA & BUZAWA, supra note 34, at 194).
66. See id., at 242–43.
67. See BUZAWA & BUZAWA, supra note 34, at 235; Friedman, supra note 48, at 246–47.
68. See Friedman, supra note 48, at 247.
70. See Johnson, Redefining Harm, supra note 5, at 1131–32; Klein & Orloff, supra note 69, at 848–58.
71. See Johnson, Redefining Harm, supra note 5, at 1133–38.
72. New Jersey has provided compensatory and punitive damages under its catch-all provision. See Klein & Orloff, supra note 69, at 912 n.695.
73. Jurisdictions vary on the duration of CPOs: three months (two states); six months (five states); one year (twenty-two states); eighteen months to two years (nine states); three to five years (eight states); and a few with permanent protective
of how best to address the abuse is severely limited, in that she may not request an injunction against further abuse without also requesting a stay away and no contact order.\textsuperscript{74} Even when the governing statute does not mandate physical separation, courts deciding CPO matters have required it even against the will of the petitioner, the person subjected to abuse requesting the CPO. In certain jurisdictions, if the petitioner chooses to vacate her CPO after it is issued, the court will not permit her to do so.\textsuperscript{75} One court in New Jersey stated that it would not vacate the order because its duty was to protect the victim of domestic violence.\textsuperscript{76} The cycle of violence theory—that is related to battered women’s syndrome—states that after violent incidents, the abusive partner apologizes and creates a “honeymoon”-like atmosphere in the relationship, thus, the women subjected to abuse are lulled into a false sense of security that the violence is over.\textsuperscript{77} Courts refusing to vacate CPOs override women’s wishes, relying on the cycle of violence to justify the courts’ decisions.\textsuperscript{78} This reliance exists despite the fact that researchers have questioned the theory’s validity and universal applicability.\textsuperscript{79}

Similarly, crime victim compensation fund laws may discourage payment to a person subjected to abuse if she has not physically separated from or continues to reside with the person who committed the abuse. For instance, under West Virginia law, no payment of monies may be made to a person subjected to abuse who chooses to maintain her home with her partner who was abusive because the monies should not serve to support the offender or cause him unjust enrichment.\textsuperscript{80} Accordingly, the person subjected to abuse who chooses to maintain her relationship with the person who had abused her could be denied funds to support her housing and daily living expenses that she would otherwise be able to collect if she were physically separated from her partner.

Domestic violence shelters also illustrate aspects of the safety paradigm. First, many shelters make decisions for residents as to how they spend their time in the shelter and require certain behaviors to gain access to the shelter. For instance, many shelters require residents to undergo orders. See Jane K. Stoever, \textit{Enjoining Abuse: The Case for Indefinite Domestic Violence Protection Orders}, 67 Vand. L. Rev. 1015, 1046–50 (2014).


\textsuperscript{75} See Kohn, \textit{supra} note 20, at 225–26; Kuennen, \textit{supra} note 20, at 45–46.


\textsuperscript{77} See Goodmark, \textit{supra} note 26, at 31–33.

\textsuperscript{78} See id. at 33.

\textsuperscript{79} See id. at 32–33.

therapy and substance abuse counseling, abide by bureaucratic rules, apply for public benefits, and conform their parenting to certain guidelines as a condition of remaining in the shelter.81 These requirements are dictated in large part by the funding they receive. Second, many domestic violence shelters require separation by locating in confidential locations to prevent women’s partners from being able to find them.82 Many shelters also require that the woman subjected to abuse cease all contact with the partner who had been abusive.83 Third, most shelters exclude men or transgendered women, especially those who are not postoperative, on the assumption that men and transgendered women are potential perpetrators of abuse.84

Many of the programs described above are funded by the Violence Against Women Act and other federal laws. The largest category of federal appropriations to combat domestic violence was granted to the criminal justice system. A large category of these criminal justice system funds go towards arrest and prosecution policies, such as mandatory arrest and no drop prosecution policies.85 The federal government allocated a total of $944 million to the criminal justice system response to domestic violence, including the Services, Training, Officers, Prosecutors (STOP) program, the Arrest Policies Program, and VOCA grant programs.86 The 2012 funding for the STOP grants was $189 million.87 The funding is used to establish protocols and special units in prosecutors’ offices and law enforcement agencies. The Grants to Encourage Arrests (Arrest Policies Program) was granted $50 million in 2012.88 These grants are given to support “criminal justice policies, practices and procedures regarding arrest and protection order laws to enhance victim safety and ensure offender accountability.”89 A significant amount of the federal appropriations go to crime victims funds that provide financial assistance to persons subjected to physical domestic violence or other domestic violence crimes. The Victims of Crime Act Fund received $705 million in 2012.90 While the funds may reimburse a person subjected to abuse for

81. See Johnson, Home, supra note 15, at 44.
82. See id. at 38 (citing Goodman & Epstein, supra note 8, at 102).
83. See id. at 44.
87. See id. at 8.
88. See id. at 18.
89. Id.
90. See id. at 39.
certain types of relief necessary because of the domestic violence—such as physical and mental health treatment—a person who is subjected to solely psychological, emotional, or economic abuse would not qualify because a potential recipient needs to prove that she was subjected to a crime.91

Certain short-term responses to domestic violence are also funded at a high level. The federal government allocated $173.7 million to predominately short-term safety and security initiatives.92 Under the Family Violence Prevention and Services Act, the federal government appropriated $129.5 million in 2012 for domestic violence shelters, counseling, and hotline services.93 The federal government granted $3.2 million for the National Domestic Violence Hotline in 2012.94 It granted $41 million in civil legal assistance, which covered short and long-term responses to domestic violence.95 If these short-term services were provided to effectuate goals that were not those of the person subjected to abuse, they would qualify as safety paradigm services under this Article’s working definition of safety.

The federal government appropriated much less money, $31.4 million, for long-term solutions to domestic violence. The following are the monies spent for such security items as economic, housing, health, and relationship security. For instance, in 2012, the federal government granted $25 million in transitional housing monies.96 The Transitional Housing Assistance Grants provide monies to persons subjected to abuse for longer-term housing costs, including rent and utilities, as well as other support services, like transportation and childcare; and the grants provide monies to service providers managing the transitional housing programs.97 The government also appropriated $5.4 million in grants to the Domestic Violence Prevention Enhancement and Leadership Through Alliances (DELTA) program.98 DELTA “supports statewide projects to integrate primary prevention principles and practices into local coordinated community responses that address and reduce the incidence of intimate partner violence,” such as peer education for men about families and relationships.99 The federal government granted $1 million in 2012 for the National Resources Center on Workplace Responses to Assist Victims of Domestic and Sexual Violence.100

93. See id. at 29.
94. See id. at 32.
95. See id. at 19.
96. See id. at 20.
97. See id.
98. See id. at 36.
99. Id.
100. See id. at 26.
The safety paradigm is equally prominent in state legislation.101 In
the 2012 legislative sessions, the fifty states and the District of Columbia
passed 249 laws regarding domestic violence. Twenty-six states passed a
total of ninety-seven criminal laws regarding domestic violence that rein-
forced the safety paradigm.102 Nineteen jurisdictions passed thirty-nine
family law provisions regarding custody and child welfare issues that ad-
dressed both the safety paradigm as well as relationship security. Eighteen
jurisdictions passed thirty-eight laws instituting procedures for orders of
protection, which blended features of the safety and security paradigms,
and addressed issues such as confidentiality, procedure, firearm removal,
and actionable abuse. Finally, three jurisdictions passed four laws address-
ing treatment of offenders, exhibiting elements both of the safety and re-
lation security paradigms.103

Only thirteen jurisdictions addressed purely security issues, such as programs and policies relating to teen dat-
ing violence,104 unemployment insurance for persons subjected to
abuse,105 in-state tuition and education waivers for persons subjected to
abuse and their children,106 expanded housing options,107 expanding ac-
cess to crime victims compensation fund monies,108 confidentiality protec-
tion,109 and studies examining the treatment and prevention of domestic

101. For instance, the National Council of Juvenile and Family Court Judges
stated that “the 2012 legislative session [across all states] reflects steadfast focus on
crime victim safety, increased perpetrator accountability, and ambitious preven-
tion efforts.” See Nat’l Council of Juv. & Fam. Court Judges, Family Vio-

102. See id. at 10–17 (tallying numbers of states that have enacted domestic
violence legislation).

103. See id. at 10–13 (showing Florida, Kansas, and West Virginia have passed
“prevention and treatment” legislation in addition to criminal measures).

104. See id. at 14–17 (noting Connecticut, Delaware, Illinois, and Oregon
passed “adolescent partner violence” legislation).

105. See id. at 59–60. In Maryland, new legislation permits a person subjected
to abuse to receive unemployment insurance if she has left her job because of the
abuse. See id.

106. See id. at 26, 72 (discussing legislation in California and New Jersey). In
California, legislation exempted persons subjected to abuse from paying nonresi-
dential tuition at state universities and community colleges, permitted them to ap-
ply for financial aid, and waived community college fees. See id. at 26. In New
Jersey, new legislation requires the state to pay for any school tuition inured be-
cause a child is in a domestic violence shelter or transitional housing. See id. at 72.

107. See id. at 26, 40 (discussing amendments in California and Florida). In
California, new legislation permits persons subjected to abuse to use CPOs as evi-
dence of domestic violence to terminate a tenancy due to the abuse. See id. at 26.
In Florida, legislation was passed providing relocation assistance for victims of sex-
ual battery. See id. at 40.

108. See id. at 44. In Illinois, a new law permits a person subjected to stalking
and who obtained a no stalking order to receive crime victims’ compensation fund-
ing. See id. This new law is notable because it does not require the person to be
subjected to physical violence.

109. See id. at 15–17 (showing ten states passed confidentiality laws in 2012).
For example, in New York, health insurers must honor requests by persons sub-
violence. Of the twenty-two pieces, only one piece of legislation, Oregon’s legislation regarding teen dating violence programs, actually provides money to implement the programs. Three other laws, those dealing with in-state tuition, waiver of tuition, and access to crime victims compensation fund monies, provide access to resources or lower the cost of the resources. Accordingly, as seen above, the safety paradigm is the prominent paradigm for legislation as well as funding of domestic violence responses.

One final nationwide program that exemplifies the safety paradigm is the use of the lethality assessment screening program to evaluate women’s risk of being killed because of domestic violence. Law enforcement officers use this program when they respond to 911 calls involving intimate partner abuse. It is an eleven question instrument that is administered by the officer and is designed to elicit responses by the woman. After answering all of the questions, the police officer scores the woman’s risk of lethality, and if high, contacts a hotline or shelter, recommending that the woman speak to a counselor. Law enforcement officials do not give the woman being screened a choice regarding whether to be screened or to receive her score. If she is reluctant to speak to the counselor, law enforcement protocol instructs them to continue to persuade her to talk to the counselor. In addition to law enforcement, the program is being implemented by court clerks, judges, legal service providers, shelter providers, and other system actors to evaluate the severity of abuse and to determine the level of services that should be provided to the woman. Importantly, the screening tool has yet to be validated. The program started in Maryland and has been implemented in other jurisdictions. As of 2013, thirty jurisdictions are using the program.

110. See generally Johnson, Balancing Liberty, supra note 4.
111. See id. at 532–42.
112. See id. at 533, 535 n.75 (discussing questions and listing questions asked).
113. See id. at 536–37 (noting that, although police officers should only suggest counselor at first, they are instructed to eventually “coerce” victim into compliance).
114. See id. (discussing mandatory nature of protocol).
115. See id.
116. See id. at 539–42.
117. See id. at 558. Researchers are currently studying the effectiveness of the lethality assessment screen in Oklahoma. See, e.g., Jill T. Messing et al., Collaborating with Police Departments: Recruitment in the Oklahoma Lethality Assessment (OK-LA) Study, 17 VIOLENCE AGAINST WOMEN 163 (2011).
118. See Johnson, Balancing Liberty, supra note 4. at 539–40; see also LAP: Nationally, supra note 118 (noting Alabama, Arizona, California, Colorado, Connecticut, Delaware, Florida, Georgia, Indiana, Kentucky, Maryland,
D. Evaluation of the Safety Paradigm

There is good reason to be concerned about the safety of persons subjected to abuse. Deprivations of liberty and dignity by assaults against a person’s bodily integrity or emotional well-being are antithetical to the values of a well-ordered society. Accordingly, a view that recognizes the crisis and danger of domestic violence is important. The state is uniquely well-situated to respond to an emergency or a crime related to the health, safety, and welfare of its citizens. In addition, the network of heavily state-funded, professionalized service providers that offer crisis hotlines, emergency shelters, and legal services provide important options for persons subjected to abuse to address any crisis moment of intimate partner abuse.

There are disadvantages to the safety paradigm, however. Emergency-driven measures designed to provide short-term physical protection ignore the range of supports necessary for a woman’s long-term future. Indeed, some women who experience abuse may not be in physical danger, and are therefore served poorly or not at all by the panoply of remedies designed to provide physical protection.

A safety orientation also often impinges on women’s agency and dignity. For instance, there are many reasons why women subjected to abuse do not take the course that the state deems the most safe—engaging the state to intervene through the criminal justice system. One reason is that mandatory arrest policies can deter women from calling 911 to help with the immediate situation because they may not want to subject their partners to automatic arrest and prosecution. For women of color, as Beth Richie writes, such a decision is a recognition of the “Prison Nation” that is at the intersection of race, gender, and domestic violence, where young black men are incarcerated at alarmingly high rates as a result of a racist penal system.120

Women also choose not to engage the state because, once they do, their opinion as to what would be the safest option for them and what would be the best outcome for them is often ignored. Separating from one’s partner who is abusive can be extremely dangerous.121 For some women, they choose not to engage the state because they believe it to be the safest course of action. Research has shown that women are the best predictors of future assault.122

---

120. See BETH E. RICHIE, ARRESTED JUSTICE: BLACK WOMEN, VIOLENCE, AND AMERICA’S PRISON NATION 46 (2012) (discussing how black women subjected to abuse avoid involving formal legal channels so they are not seen as adding to arrests of black men).


122. See Johnson, Balancing Liberty, supra note 4, at 558–61 (discussing research showing women are best predictors of their future risk of assault).
Further, women do not engage the state because they do not want their entire family to suffer the consequences of their partner’s arrest or incarceration.\textsuperscript{123} Such actions could result in the loss of his job or the ability to get future employment, diminishment of the family income, loss of a partner in parenting the children, and loss of an intimate partner.\textsuperscript{124} Some of these women would prefer a way to end the abuse without ending the relationship through the criminal justice system.\textsuperscript{125}

Another reason women choose not to involve or cooperate with the criminal justice system is because mandatory arrest policies make them less safe. Mandatory arrest laws became popular in the mid-1980s, upon the release of the Minneapolis Domestic Violence Experiment (MDVE) study that found arrests decreased the number of re-offenses of domestic violence.\textsuperscript{126} However, subsequent studies did not find the same results.\textsuperscript{127} The authors of the MDVE study showed that arrest might slightly decrease the rate of re-offense for those offenders who are married or employed, while increasing the rates of re-offense if the offender was unmarried or unemployed.\textsuperscript{128} The studies also showed that if the offender had prior assaults against his partner, or was drinking or using drugs at the time of the abuse, he was more likely to reoffend regardless of whether he was arrested.\textsuperscript{129} Therefore, arrest does not necessarily make persons subjected to abuse safer.\textsuperscript{130} Nonetheless, the mandatory arrest laws remain on the books and in practice based, in part, on the publicity around MDVE and its conclusion that such laws decreased future abuse.\textsuperscript{131} More-
over, mandatory arrest policies interfere with the ability of persons subjected to abuse to make choices for themselves about how to address the violence. 132 This is a negative consequence because research shows that when women subjected to abuse are permitted to make their own decisions for how to address the violence, they identify greater life satisfaction and reduced violence in their lives. 133 Another negative outcome of mandatory arrest is that police often arrest both parties when one party might not have committed a crime. For instance, one party may have assaulted the other in self-defense. 134 Therefore, while having a criminal justice system response to domestic violence is critical and necessary to protect and support persons subjected to abuse, 135 such a response also needs to address the needs and desires of persons subjected to abuse. 136

There are similar downsides to no-drop prosecution policies as to mandatory arrest policies. First, the prosecution will go forward against the objections of the person subjected to abuse, despite the fact that it might increase her risk of being subjected to further abuse. 137 Second, such policies result in fewer economic opportunities for the person being prosecuted as well as the person subjected to abuse. 138 Third, because mandatory arrest policies often result in incarceration or no contact orders, such policies disrupt the ability of the person subjected to abuse to continue to maintain a relationship with the person who caused the abuse. 139 By granting the person subjected to abuse the power to drop


133. See Johnson, Redefining Harm, supra note 5, at 1125–27. But see Univ. of Ky. Ctr. for Research on Violence Against Women, supra note 57, at 3 (citing study showing vast majority of women identified mandatory arrest as benefit because it took responsibility away from women to pursue arrest).

134. See Univ. of Ky. Ctr. for Research on Violence Against Women, supra note 57, at 3.

135. As Buzawa and Buzawa state, persons subjected to abuse “consistently wanted the police to respond to incidents, even if arrest was not always desired.” Buzawa & Buzawa, supra note 34, at 139 (noting women wanted offender controlled more often than arrested).

136. Data shows that given mandatory arrest policies, “police refuse[d] to follow the victim’s request more frequently in domestic assault cases (25%) than nondomestic assault cases (4%). The victim disagreed with arrests made in 60% of domestic assaults compared with only 12% of nondomestic assaults . . . .” Id. at 140.

137. See id. at 183. (describing study finding 50% of “victims reported that their assailants had physically threatened them if they preceded further and attempted court [proceedings] . . . .”).

138. See id. at 183–84 (citing victims’ fears of financial loss potentially incurred through participating in prosecution of assailant or resulting from assailants’ incarceration or termination from employment).

139. See id. at 182 (citing research showing that victims chose not to cooperate in prosecution because they are “far less concerned with deterrence as an esoteric concept than with using the criminal justice system as a whole to accomplish personal goals of enhancing safety, maintaining economic viability, protecting chil-
charges, the government would equip her with "the threat of continued prosecution as a 'victim power resource.'"140

Further, a criminal trial results in punishment that may provide no meaningful remedy for the person subjected to abuse. The defendant may be incarcerated or fined, however, none of these outcomes provide an immediate or tangible benefit for the person subjected to abuse, and in fact, they can be harmful.141 The no-drop policy intentionally dehumanizes the person subjected to abuse and discounts her agency and dignity by requiring the prosecutor to proceed against her will.142 Studies on no-drop prosecution show that "enhancing the empowerment of the victim [including permitting victims to drop charges], rather than the actual prosecution, may be a key factor" in reducing re-abuse.143 Accordingly, while it is important to have a criminal justice system available to prosecute domestic violence crimes, such prosecution, similar to the arrest decision, should be informed by and responsive to the needs and desires of the person subjected to abuse.144

There are also similar positive and negative consequences to criminal no contact orders. If the person subjected to abuse wants the stay away order, it can be a useful response to the violence. However, once issued, only the state can control whether the order is lifted prior to its expiration. A response that provides more control to the person subjected to abuse would be a stay away or no contact order in the civil protective order case, which should remain more in her control.

Other negative consequences of criminal no contact orders arise when they are issued against the wishes of the person who was subjected to the abuse at issue in the criminal case. For instance, research shows that criminal no contact orders do not stop future abuse and can harm women, or having an opportunity to force participation in batterers' counseling programs").

140. Id. at 201.

141. See UNIV. OF KY. CTR. FOR RESEARCH ON VIOLENCE AGAINST WOMEN, supra note 57, at 4.

142. See BUZAWA & BUZAWA, supra note 34, at 198 ("As one prosecutor explained, if the victim recants, the proper prosecution for domestic abuse cases is similar to a homicide . . . ."). Many scholars trace victims' refusal to cooperate with the prosecution to the Battered Women's Syndrome and Learned Helplessness theories as well because such theories reinforce the state's idea that women subjected to abuse are "passive [and] victimized" and should "not be seen as active partners in domestic violence interventions." Kohn, supra note 20, at 208 (citing Elizabeth M. Schneider, Particularity and Generality: Challenges of Feminist Theory and Practice in Work on Woman-Abuse, 67 N.Y.U. L. REV. 520, 561 (1992)); see also Friedman & McCormack, supra note 85, at 1188–89 (noting no-drop policies assume that women subjected to abuse suffer from false consciousness regarding their risk of danger).

143. BUZAWA & BUZAWA, supra note 34, at 202.

144. See id. at 201 ("[M]any women may be safer if they can freely drop actions but ultimately may be less safe if they do . . . . [V]ictims need to be treated as full partners in the prosecution of the case.").
As discussed in a model policy for courts and prosecutors regarding no contact orders, “research has shown that, in some cases, a victim’s separation from an abuser actually increases the risk of lethality. The decision by the prosecutor or the court to restrict contact between a batterer and his victim does not always achieve the system’s goal of victim protection and safety.” In addition, research shows that, because a mandatory criminal no contact order issued against the will of the person subjected to abuse is detrimental to her agency, it also negatively impacts her safety. Finally, criminal no contact orders issued against the will of a married victim create a *sua sponte* divorce of the spouses by the court.

The civil protection order laws permit a person subjected to abuse to obtain an injunction against the offender in the form of a stay away order, no contact order, or some form of exclusion from the home. A positive consequence of the CPO laws is that they can support the agency of persons subjected to abuse: petitioners can choose to file the case, they choose what acts of abuse to include in the petition, they choose what remedy to select, and they choose for how long they will request the order be in place. A negative consequence of the CPO laws is that they constrain the extent of the petitioner’s control, as they are often limited by statute to a period of one year, and they do not cover all forms of abuse. In addition, as discussed above, courts may intervene and undermine the agency of persons subjected to abuse by ordering them to separate from the person who has abused them, even if that was not the relief requested. Further, courts may refuse to permit the person subjected to abuse to dismiss her petition for a protective order, even if she believes it is in her best interest to do so. Some studies have shown that the CPO can be an effective tool in reducing domestic violence or achieving...
ing the goals of the petitioner for the CPO.155 Other studies show that CPOs are only effective for a segment of petitioners, similar to mandatory arrest.156 Moreover, CPOs are most effective when they effectuate the goals of the petitioner herself.157

The availability of emergency domestic violence shelters can be critical for persons subjected to abuse. It is important to have housing available for a person subjected to abuse who decides that she needs to leave her home in order to address the abuse. Shelters offer a necessary option for persons experiencing abuse and trying to decide how best to address it. In this way, shelters are the quintessential short-term remedy for physical violence through physical separation and show how necessary certain options in this safety model can be.

However, as discussed above, there are limits to the benefits of the shelter system’s operating model of total physical separation. A 2001 study shows that separation of the parties does not reduce partner homicide.158 In fact, some of the findings showed that resources focused on physical separation can be incredibly dangerous to persons subjected to abuse.159 This dangerousness is due to the fact that focusing only on physical separation ignores the reality of the partners’ interconnected lives—sharing children, property, friends, and communities—and, therefore, separation may not be a true option.160 Further, the risk of violence may continue after leaving the shelter. In one study, almost half of all women experienced continued domestic violence within a short time after leaving a shelter.161 Accordingly, the shelter model of total physical separation can be lethal because separation may not be possible, safe, or desired at that particular time.

Other aspects of the shelter system are also detrimental for persons subjected to abuse. The confidential nature of many shelters disconnects persons subjected to abuse from their support networks. While in the shelter, women cannot disclose their location and are often unable to maintain contact with the person who abused them.162 What research

---

155. See Johnson, Redefining Harm, supra note 5, at 1128–29 (showing that CPOs are effective in decreasing future violence).

156. See Jane Aiken & Katherine Goldwasser, The Perils of Empowerment, 20 CORNELL J.L. & PUB. POL’Y 139, 158–59 (2010) (explaining CPOs are less effective on “men with prior criminal histories, men who are unemployed, and men who have substance abuse problems” (footnotes omitted)).

157. See Johnson, Redefining Harm, supra note 5, at 1124–29.

158. See Johnson, Balancing Liberty, supra note 4, at 573.

159. See id.

160. See GOODMAN & EPSTEIN, supra note 8, at 102.


162. See GOODMAN & EPSTEIN, supra note 8, at 112.
shows, however, is that because hidden shelters isolate women from their community and the security that comes from being a part of a neighborhood, women’s risk of experiencing violence increases.163 “As Professors Epstein and Goodman explain, ‘research shows that women in hidden locations are no safer during their stay than women in open shelters where community members can participate in keeping residents safe.’”164 The negative effect of separation from a shelter stay was documented in a study of Puerto Rican women, which found that the women “‘don’t want to leave their community and come to a new place. They may have language problems. They don’t drive. They may never have paid bills or done a budget. . . . They’re not used to living collectively or sharing apartments like white women do.’”165

Shelters often require that a resident be female, thereby excluding men and transgendered women, resulting in fewer temporary homes for persons subjected to abuse and fortifying a “worthy victim” stereotype.166 To rectify this gap in services, some advocates have suggested creating shelters for men who experience intimate partner violence.167 Similarly, anti-discrimination laws protecting transgendered persons have helped to open up some shelters to transgendered women.168

Finally, it is problematic that shelters operating under the safety paradigm are the party dictating the appropriate response to domestic violence rather than the person subjected to abuse. One study has shown that women subjected to abuse found services provided in a woman-centered manner that permitted the women to be the decision makers produced “‘good social support networks, fewer feelings of isolation, and better access to childcare after receiving assistance.’”169 When services were woman-centered, included active listening, and respected the woman’s ability to make decisions, the services were most effective.170 The least helpful were services that told the woman what to do in response to the abuse and criticized her decisions.171

163. See Johnson, Home, supra note 15, at 38.
164. Id. (quoting Goodman & Epstein, supra note 8, at 102).
165. Id. at 38 n.159 (quoting Schechter, supra note 41, at 59).
167. See id.; see also Crick, supra note 84, at 1059 (discussing lack of services for men in Washington State and proposals for addressing them).
168. See Goodmark, supra note 84, at 70–71.
170. See Johnson, Balancing Liberty, supra note 4, at 572–73.
171. See id. at 572.
As seen in the previous section, federal funding is heavily skewed towards safety paradigm programs. As a result, it is no surprise that the majority of state-legislated programs also focus on funding safety paradigm programs. It is important that violence against women is recognized by the state and federal governments and that the responses to the abuse are well-funded. However, as seen in the above discussion, the funding of certain programs, such as mandatory arrest and no-drop prosecution, does not necessarily effectively address the state’s interest in “safety;” as such, programs may increase or have no effect on future abuse. Moreover, increased safety programs that override the choices of persons subjected to abuse, such as the lethality assessment program, may cause significant harm to women by decreasing their safety and negatively impacting their dignity and agency.172

Mary’s story173 from the beginning of this Article demonstrates some of the positive aspects of the safety paradigm while highlighting the very problematic results of not having a security paradigm. First, from Mary’s story, we learn that only her physical abuse was recognized under the CPO law and under the crime victims compensation fund (CVC), which provided the monies for the shelter, transportation vouchers, and food vouchers. As a result, the other forms of abuse to which she was subjected—emotional, psychological, economic, and medical—were not recognized or addressed by the court or the CVC officers. Because of the physical violence, however, she was able to gain emergency short-term relief based on physical separation, which she desperately needed, such as emergency shelter, emergency medical treatment, temporary transportation and food vouchers, and a CPO for one year. This relief is one of the great successes of the current domestic violence movement. As a result, Mary initially felt “safe” through this court-ordered and community-based support for her physical separation from her ex-partner.

But this almost textbook safety was not at all secure in the short- or long-term. As Mary recounted, yes she had left Todd, and she had an order that said he could not harm her, but she had no job, no housing after thirty days174 no money for food after thirty days, no medicine, and no clothes other than what she left the home wearing. After her shelter stay, she was left with no choice but to return to Todd.

The question raised by Mary’s experience is whether there is any other way to construct the response to intimate partner abuse that can be better at effectuating choices of women subjected to abuse to lead secure lives. This Article argues that the security paradigm offers such a construction.

172. See generally id.
173. See supra notes 2–3 and accompanying text.
174. Similar to Mary’s experience, most shelters limit residents’ stays to thirty days in length. See Johnson, Home, supra note 15, at 43.
III. TOWARDS A SECURITY PARADIGM

There is another paradigm to address domestic violence that maintains the benefits of the safety paradigm while mitigating its harms. Premised on a broad, research-based understanding of “domestic violence” as “intimate partner abuse,” this is the security paradigm.

A. The Security Paradigm

As reflected above, a “safety-at-all-costs” orientation impinges on women’s agency, dignity, and well-being in a way that may inadvertently and counterintuitively make women less safe. A different approach to addressing domestic violence would be to re-orient the system to prioritize women’s overall well-being, in addition to their physical safety, and to do so in a women-centered manner, based on their views of what will increase their well-being.

The security paradigm has four main components. First, the security paradigm recognizes all forms of intimate partner abuse, including physical, sexual, psychological, emotional, and economic abuse. It recognizes the coercive use of power and control in an intimate partner relationship as key to the abuse. In other words, the security paradigm recognizes actions that coercively constrain or aim to constrain an intimate partner’s liberty, dignity, and agency. Second, the security paradigm recognizes all forms of harm that result from the range of abuse. Therefore, the security paradigm recognizes physical injury as well as psychological trauma, lack of economic resources, and isolation. Third, the security paradigm seeks to create multiple options to address the abuse and harm. These options are both short-term responses focused on true physical safety as well as short-term and long-term responses focused on economic, housing, health, and relationship security. Fourth, the person subjected to abuse remains the decision maker in responding to all of the circumstances in her life, including decisions in the context of the abuse she has experienced. The state and other institutions’ role will be to help generate and support options consistent with the security paradigm and help facilitate her decision making. To this end, the security paradigm recognizes the range of motivations and goals she may have in addressing the abuse, including the fact that ending the relationship may not be her goal and addressing the abuse may be weighed against other important goals, such as the economic security of her children.

Security can be accomplished by implementing measures designed to promote women’s economic security, housing security, health security, and relationship security, in addition to the physical safety of a woman subjected to abuse. For example, economic security would entail having a steady full-time job with a reliable and livable income, reliable public benefits, predictable and manageable living expenses, and the accumulation of some assets, such as a home, a car, and some money in savings. If she does not have a job currently, economic security includes career counsel-
ing, education, job training, and child care to assist the woman in obtaining job security. Economic security would also entail access to monetary damages and restitution for any and all harm resulting from the abuse, as well as adequate family support if the woman lives separately from her partner because of the abuse.

Economic security would also include eliminating and remedying employment discrimination against persons who have been subjected to abuse. When women subjected to abuse are unemployed because of the abuse, unemployment benefits would provide economic security.

Economic security permits a woman subjected to abuse to make true choices about how she wants to address the abuse in her relationship. With economic security, she need not choose to stay in an abusive relationship simply because her partner earns the family’s money and without him she will not be able to keep her home or support her children. With economic security as well as job security, she may have more mental, emotional, and physical strength to problem solve around the abuse in her relationship.

Housing security is the ability to control one’s housing, whether it be having a home titled in her name, being the mortgagor for a home, or being the tenant in an apartment. Such housing should be affordable and stable, such that it can be on a long-term basis if that is what she wants. Housing security would address and remedy any housing discrimination resulting from the abuse. Similarly, housing security includes, after abuse occurs, the ability to decide whether to remain living in her home, whether to vacate her partner from the home, and whether options for short- and long-term affordable alternative housing are available if she chooses to leave the home. Housing security offers stability that has the potential to reinforce her agency in determining how best to rearrange her relationship to address satisfactorily the violence, while either ending or continuing her relationship.

Similarly, health security means that a woman has access to adequate medical and mental health care for herself and her dependents, including regular preventive care, access to affordable medications, access to specialists when necessary, and transportation to and from appointments. A woman who has these things can maintain her physical and mental health, engage in typical activities of daily living, including self-care, and maintain the independence, agency, and vitality required to satisfy her responsibilities and contribute to the best of her abilities.

Relationship security is having relationships that are stable, healthy, respectful, and supportive of each individual’s agency and dignity. The
relationships include relationships with intimate partners, children, other family members, friends, and neighbors. Relationship security includes the ability to choose to stay in her community, connected to all supportive aspects of her community, such as her place of worship, her neighborhood and neighbors, and other social networks or resources. Relationship security acknowledges the importance of social capital, the interdependence of people, and the benefits of building and maintaining a social network upon which she can rely for assistance as she continues to build her future life.

B. Focusing on Security Will More Effectively Address All Types of Intimate Partner Abuse and Their Harms

A focus on security looks at all well-being indicia for the person subjected to abuse—her agency, dignity, liberty, and health. A focus on security can thereby take into account all types of intimate partner abuse, not just physical, but also economic, emotional, and psychological abuse. A focus on security can generate options to address the broad spectrum of coercive abuse while supporting a woman’s effort to obtain economic, housing, health, and relational security. For instance, as Dana Harrington Conner describes, a woman may be targeted by her partner for economic abuse because she is economically unstable, with limited income or no home of her own.¹⁷⁶ The security paradigm would foster and support state, institutional, and community-based options—funded in part by state money that is currently allocated to ineffective safety initiatives. These provisions would support empowerment-focused career counseling, while also encouraging social capital creation through time banking. This initiative would also identify potential jobs and address obstacles to obtaining employment, ranging from transportation to child care.

The security paradigm can look at policy through the lens of addressing and remediying the harms of intimate partner abuse and, specifically, coercive control, as identified by the person subjected to abuse. As such, security permits us to focus on the varied actions taken or contemplated by persons subjected to abuse to address the abuse rather than measures taken on her behalf by others. The actions taken by a person subjected to abuse include her proactive and responsive deployment of the state, her community, and local service providers as well as her decisions about her relationships. On the other hand, in the safety paradigm, a person subjected to abuse has no choice regarding the mandatory arrest or prosecution following a 911 call and no control over a criminal no contact order issued by the court. Under the security paradigm, the person subjected to abuse may choose to have the person committing the acts of abuse arrested to stop the abuse in the instant and then decide she either does not want him to be prosecuted or does not want a no-contact order to be issued. She may make these decisions for a myriad of reasons, perhaps be-

¹⁷⁶. See Conner, supra note 36, at 359–60.
cause she wants to continue living and managing her family together or because the criminal justice system response will not meet any of her goals. In fact, the criminal justice system response may exacerbate problems such as poverty or racial oppression through incarceration and its effect on his economic livelihood.\textsuperscript{177}

All types of intimate partner abuse can—and often do—affect the economic life of the person experiencing the abuse. For instance, domestic violence is the leading cause of family homelessness.\textsuperscript{178} The cause of this is as multifaceted as the abuse itself. For instance, “[m]any abusers so isolate and demoralize their victims that the victims are left with few support systems, and hence have great difficulty getting on their feet . . . .”\textsuperscript{179} Often, the woman experiencing abuse is economically dependent on her abuser’s income, lacks access to other sufficient resources, and suffers from decreased physical and mental health that negatively affects her ability to gain adequate employment.\textsuperscript{180} Also, if the abuser is separated from the woman through the civil legal system, “his income often leaves with him despite orders that he pay child support and maintenance. Some women are reluctant to insist upon the payments because they believe that they may be at greater risk of retaliation and violence if they make such demands.”\textsuperscript{181} If the abuser is convicted and incarcerated for his abuse, the criminal court has no jurisdiction to issue child or spousal support to the woman who experienced the abuse. During his incarceration, he is no longer employed and earning an income to support her, and his future employment is jeopardized with his criminal record.\textsuperscript{182} If he is an immigrant, his conviction can result in deportation, which further jeopardizes his ability to offer financial support to her and their children.\textsuperscript{183} As a result, the woman subjected to abuse may be in or on the edge of poverty,\textsuperscript{184} and as mentioned above, at risk of becoming homeless.\textsuperscript{185} Further, if she was isolated through the abuse, she may not have had access to or sought out medical or mental health professionals to help her with the

\textsuperscript{177.} See Riché, supra note 120, at 121–22 (explaining black women’s lack of reliance on criminal legal system).


\textsuperscript{179.} Aiken & Goldwasser, supra note 156, at 160.

\textsuperscript{180.} See Johnson, Redefining Harm, supra note 5, at 1123; see also Aiken & Goldwasser, supra note 156, at 160 (noting women’s employment is often at great risk because of abuse).

\textsuperscript{181.} Aiken & Goldwasser, supra note 156, at 160.

\textsuperscript{182.} See Bailey, supra note 21, at 1794–96.

\textsuperscript{183.} See id.

\textsuperscript{184.} See id. at 1794 (“A domestic violence victim faces a fifty percent chance her income will fall below the poverty level if she leaves her batterer.”).

\textsuperscript{185.} See Johnson, Redefining Harm, supra note 5, at 1123; see also Aiken & Goldwasser, supra note 156, at 160 (noting women’s employment is often at great risk because of abuse).
abuse, any resulting depression, and other related issues. Instead, if she turned to substance abuse to treat her mental health issues, any “substance abuse issues may hinder the woman from being able to maintain full-time work, provide adequate care for children, or seek appropriate services that will assist in her recovery from the abuse.” Accordingly, focusing on security is important to address all harms from all forms of abuse.

C. Focusing on Security Will Support the Agency, Dignity, and Resilience of Persons Subjected to Abuse

A security focus more effectively addresses domestic violence because it more comprehensively supports the agency and dignity of the person subjected to abuse. Agency is making “choices, acts of resistance, self-direction, and self-definition.” “Dignity is the inherent nature that renders human beings capable of autonomous action and thought.” Dignity recognizes people as separate from the state, with the rights and responsibilities to exercise their agency to address “fundamental questions touching the meaning and value of their own lives.” To support dignity, society must respect and support individuals’ capabilities such as rationality; “life; bodily health; bodily integrity; senses, imagination and thought; emotions; practical reason; affiliation; respect for other species; play; and control over one’s environment.” A security focus on security is an important endeavor because it is linked to the greater satisfaction and happiness of the person subjected to abuse, even when she continues in a relationship with her abuser. Security lets persons subjected to abuse be an agent—to identify their own goals regarding the abusive relationship and their life satisfaction and to make informed decisions to achieve their goals without outsiders defining dangers and havens for them. When we shift the frame to security from safety, we permit the goals of the person subjected to abuse to drive the options that exist externally and internally. When these goals come into focus, we can see that they are complex and multiple, driven by the woman’s whole life experience, not just a decontextualized act of abuse. Research shows that violence may decrease when women subjected to abuse are able to connect to their

186. See Aiken & Goldwasser, supra note 156, at 161–62.
187. Id. at 162.
188. See Johnson, Home, supra note 15, at 16–17 (discussing importance of supporting agency and dignity to appropriately address domestic violence).
191. Id. at 10 (quoting Dworkin, supra note 4, at 426).
192. Id. at 11 (citing Martha Nussbaum, Human Dignity and Political Entitlements, in Human Dignity and Bioethics: Essays Commissioned by the President’s Council on Bioethics (2008)).
193. See id. at 16–17.
community, control their physical environment by leasing or owning their own home, and build assets, such as home ownership or a small business.\textsuperscript{194} And when persons subjected to abuse make their own informed choices about physical separation or what is safe, they are more satisfied with their life and less at risk of physical violence.\textsuperscript{195} Moreover, agency fosters resilience, a necessary feature for addressing abuse in one’s life.

Researchers have shown that women who experience intimate partner abuse often decrease their exposure to violence when they exercise their own agency.\textsuperscript{196} As domestic violence is the systemic operation of power and control, “‘[a]n important element of responding to the problem [of domestic violence] is to restore a victim’s fundamental rights of freedom, choice and autonomy.’”\textsuperscript{197} For example, a woman exercises her agency when she decides to obtain a civil protection order that will provide the specific relief she is seeking, perhaps in a form as simple as an injunction against further abuse or as complicated as orders for child custody, child support, eviction of the abusive partner from the home, or an injunction against any contact.\textsuperscript{198}

On the other hand, when a landlord or employer requires a woman subjected to abuse to obtain a CPO in order to avoid eviction or termination, the woman’s agency is constrained by seeking a CPO she would not have otherwise sought. In addition, a woman is unable to exercise her complete agency when the court refuses to grant a civil protection order unless the woman reluctantly requests an injunction against physical contact.\textsuperscript{199} Such constraints on agency are not uncommon. All agency is circumscribed by the context in which we live and the systemic operation of power. For a security paradigm focused on supporting women’s agency, we should focus on supporting women’s exercise of agency that is most linked to a decrease of intimate partner abuse: her complete agency that is decided upon by her and that is countervailing to the systemic operation

\textsuperscript{194.} See id.
\textsuperscript{195.} Cf. supra notes 52–119 and accompanying text.
\textsuperscript{197.} Johnson, \textit{Redefining Harm}, supra note 5, at 1151 (second alteration in original) (quoting Tamara L. Kuennen, \textit{Analyzing the Impact of Coercion on Domestic Violence Victims: How Much Is Too Much?}, 22 \textit{Berkeley J. Gender L. & Just.} 2, 30 (2007)).
\textsuperscript{198.} See id. at 1131 (describing various remedies available under civil protection order statutes).
\textsuperscript{199.} See id. at 1150. Research shows that leaving an abusive relationship can be more dangerous for women because it “may increase stalking, harassment, and may decrease the woman’s ability to influence him.” \textit{Id.} at 1127 (citing LENOIRE E. A. WALKER, \textit{ABUSED WOMEN AND SURVIVOR THERAPY: A PRACTICAL GUIDE FOR THE PSYCHOTHERAPIST} 55 (1994)). Separation Assault, as it is known, results from the abuser losing control over his victim and lashing out with increased violence. See Goodman & Epstein, supra note 8, at 97–98; Aiken & Goldwasser, supra note 156, at 162. Because separation assault is common, it is critical that the woman is the one making the decision to exercise her agency in light of the abusive relationship and her objectives, as she is in the best position to predict her risk of danger. See Johnson, \textit{Balancing Liberty}, supra note 4, at 558–61.
of power, especially that of the domestic abuse. Therefore, outsiders forcing action by the woman, rather than permitting her to make the informed choice about which action to take, can be counterproductive.

Institutional actors can promote and support women’s agency by expanding options for women to choose in their efforts to address the intimate partner abuse.\textsuperscript{200} One study explored the agency of low-income Puerto Rican women who were living with abusive partners.\textsuperscript{201} The study found that when the women decided to rent and control their own apartments, they were able to decrease their exposure to violence. The study found that this decrease in violence occurred even when the women continued their relationships with the partners who had been abusive.\textsuperscript{202} Another study shows that women who received autonomy-respecting “support and assistance” during a CPO proceeding “thought more positively about the proceedings” and “reported having good social support networks, fewer feelings of isolation, and better access to child care after receiving assistance.”\textsuperscript{203} The options can—and should—include criminal justice and civil legal system responses that provide physical separation for either safety or punitive purposes. But, to effectuate agency and thereby decrease future violence, the woman should be provided the opportunity to make these decisions based upon her experience of the violence, her prediction of the risk, and her determination of options that would increase her life satisfaction.

The person subjected to abuse may not be able to achieve her goals if she does not have real options—either because she has limited economic resources or because our system has not chosen to fund other options because they do not fit the “safety” paradigm.

Accordingly, a focus on security more effectively addresses intimate partner abuse because it offers more avenues to address intimate partner abuse, ensuring the economic, housing, health, and relationship security of the woman subjected to abuse. Further, I define security in this Article as requiring that the person subjected to abuse, not the state or other institution, is the decision maker as to which options to choose to best address the abuse.\textsuperscript{204} Built into the security paradigm is a recognition of the agency and dignity of the person subjected to abuse.

\textsuperscript{200} See Johnson, \textit{Redefining Harm}, supra note 5, at 1151.


\textsuperscript{202} See id. at 385 (“Mothers […] interpreted housing as a valued resource in intimate partner relationships in divergent ways with independent housing being seen as a bargaining tool to maintain or initiate relationships as well as a refuge for terminating relationships that experienced conflict.”).

\textsuperscript{203} Wan, supra note 169, at 611.

\textsuperscript{204} See supra Part III.A.
D. Focusing on Security Will Increase Access to Resources

A security-focused approach to address domestic violence emphasizes increased access to economic, housing, health, relationship, and other resources for persons subjected to intimate partner abuse. As Dana Harrington Conner discusses, “[f]inancial independence [for the person subjected to abuse] shifts power within the intimate relationship.”205 Without such independence, one is vulnerable to a controlling or abusive partner whether the person wants to remain in the relationship or leave it.206 Jody Raphael documented that women on welfare experience intimate partner abuse at four to five times the rate of all women.207 Donna Coker analyzed a body of research showing the inextricable link between access to economic resources and intimate partner abuse. For instance, Coker analyzed research conducted by Cris Sullivan, showing the connection between financial resources, future violence, and overall well-being.208 The study evaluated two groups of women who had been subjected to abuse and were set to leave a shelter. The control group received no services upon leaving the shelter, while the other group of women was provided with an advocate who assisted the women in accessing “educational resources, legal assistance, employment, services for their children, housing, child care, transportation, financial assistance, health care, and social supports.”209 These resources were provided in a woman-centered manner based on the women’s individual goals, not state or institutional goals. As a result of this access to resources, the women experienced significantly reduced psychological abuse, fewer mental health issues, greater social capital, higher quality of life, and no physical abuse.210 These results were in stark contrast to the control group.211 Similarly, JoAnn Miller and Amy Krull found that persons subjected to abuse who were unemployed experienced greater rates of recurring abuse than those who were employed. Two other studies showed that the length of unemployment correlated with the rate at which abuse reoccurred.212 In addition, a study by Amy Farmer and Jill Tiefenthaler showed that, as women’s in-

205. Conner, supra note 36, at 374.
206. See Coker, Economic Rights, supra note 9, at 188 (“Inadequate material resources render women more vulnerable to violence. Inadequate material resources increase the batterers’ access to women who do try to separate. Inadequate material resources are a primary reason why women do not try to separate.”).
209. Id. at 1022.
210. See id. at 1022–23.
211. See id.
212. See id.
come increases, the rate of domestic violence decreases, except for women in the highest income bracket.\textsuperscript{213}

Accordingly, the relative economic power within a relationship is correlated to occurrence of abuse. Yet it is important to understand that neither the woman nor the couple is an economically autonomous entity. Both need to be viewed within the broader context of their friends, family, neighbors, community, and society and the way in which they interact with the larger system of economic dependencies and resources.\textsuperscript{214} In her examination of economic security of women subjected to abuse, Kameri Christy-McMullin’s research shows that there are “six overlapping components of economic security: public assistance, education, employment, income, assets, and divorce and child custody settlements.”\textsuperscript{215} Accordingly, this section examines the research regarding the relative economic power within the relationship, including access to public benefits, education, employment, income, assets, and family law support, including civil protective orders.

As discussed by Christy-McMullin, the first area of economic security is public assistance. There are multiple government assistance programs that are intended to provide those in poverty with a safety net and security. Unfortunately, as many have documented, the public safety net has shrunk since the welfare reforms of 1996. For instance, Temporary Assistance for Needy Families (TANF) replaced Aid to Families with Dependent Children (AFDC) and placed a five-year cap on cash assistance. TANF also instituted work requirements, narrowed educational waivers, and narrowed other waivers of the cap.\textsuperscript{216} As Michele Gilman documents, TANF has not decreased the numbers of those experiencing poverty but rather has “pushed many poor mothers into the low-wage workforce, where they struggle to survive on meager wages.”\textsuperscript{217} Yet, even those who might benefit from the minimal monthly average assistance of $429 do not receive the assistance because of the bureaucratic and arbitrary gatekeeping of the

\textsuperscript{213.} See Christy-McMullin, Designing Policies, supra note 9, at 113 (citing Amy Farmer & Jill Tiefenthaler, An Economic Analysis of Domestic Violence, 55 Rev. Soc. Econ. 337 (1997)). Farmer and Tiefenthaler’s study also showed that violence increased with an increase in the man’s income level. See id.; see also Kameri Christy-McMullin, An Evidenced-Based Approach to a Theoretical Understanding of the Relationship Between Economic Resources, Race/Ethnicity, and Woman Abuse, 3 J. Evidence-Based Soc. Work 1, 23 (2006) [hereinafter Christy-McMullin, An Evidence-Based Approach] (explaining that studies examining relationship between economic resources and woman abuse provide mixed results, showing three recent studies demonstrate significant relationship between economic resources and abuse, while two older studies show no significant relationship).

\textsuperscript{214.} See Coker, Shifting Power, supra note 24, at 1024.

\textsuperscript{215.} Christy-McMullin, Designing Policies, supra note 9, at 111.

\textsuperscript{216.} See Michele Estrin Gilman, The Return of the Welfare Queen, 22 Am. U. J. Gender Soc. Pol’y & L. 247, 247, 249 (2014); see also id. at 254–56 (providing helpful summary of TANF work requirements).

\textsuperscript{217.} Id. at 249.
TANF program and its workers. The net result is “4.5 million people receive cash assistance through TANF (amounting to 0.47% of the federal 2012 budget) . . . .” As such, TANF is a very small public assistance program, with capped spending at $16.5 billion.

Moreover, although the numbers of people receiving TANF cash assistance have decreased, “poverty and unemployment rates have increased.” The unemployment rates for TANF recipients are explained in part due to the higher number of barriers to work that TANF recipients face, including “lack of education, mental or physical disabilities, substance abuse or alcoholism, limited work experience, and caregiving responsibilities for disabled children.” The poverty rates are explained in part because “income gains from employment are often reduced by the loss of public benefits and are eaten up by the very costs of working—child care, transportation, uniforms, and other expenses.” Thus, even healthy individuals experiencing poverty with some degree of higher education and a work history experience difficulty gaining economic stability. Additionally, the Supplemental Nutritional Aid to Families Program (SNAP), otherwise known as food stamps, benefits 46 million people. Even when families receive TANF and SNAP benefits, however, most of those families remain below the poverty line. Further, single mothers have a higher unemployment rate than in the past.

The largest public assistance program is the Earned Income Tax Credit (EITC), which provides a tax credit for low-income working persons. In 2011, the EITC provided $58.6 billion to 26.2 million families, helping millions rise out of poverty. To receive this assistance, however, one must be employed. There are other public assistance programs as well, which are more fully-discussed below in Part IV.E.

The second area of economic security is education, as discussed by Christy-McMullin. Education serves as an important tool for increasing opportunities for employment. Access to education can increase one’s employment opportunities and thereby increase one’s economic resources. As part of the isolation that may result from the abuse, a woman may not have been permitted to complete or pursue education that would have enhanced her employment marketability. For low-income women,
another barrier to education has been TANF regulations that permit only 30% of a state’s recipients to be exempted from the work requirements to pursue education, and even that exemption is limited to vocational training or completion of high school. Christy-McMullin’s research shows that having at least a four-year college degree has a positive impact on physical abuse, while having only a high school diploma or some college does not. It is important to note, however, that her research also shows that education is not related to emotional abuse.

The third and fourth areas documented by Christy-McMullin for economic security are employment and income. Not all employment leads to economic security. Rather, employment is integral to security when it provides a decent wage that raises the employee and her family above the poverty line at a minimum, and provides the employee with job flexibility, paid leave, and/or a chance for promotion. A broad range of long-standing barriers to such employment exist, such as discrimination against persons subjected to abuse, unequal pay between men and women, lack of a broad-scale living wage for minimum wage workers, and gender segregation relegating women into part-time employment with few or no benefits. Research also shows that employment is beneficial because physical abuse is less prevalent in households with higher income.

The fifth area of economic security is assets. Studies show that assets such as ownership of a home, real property, or a small business are important aspects of a woman’s perception of economic security. Studies by Deborah Page-Adams and Roger Peterson show that women who own their own homes are less likely to experience intimate partner abuse than women who rent. Moreover, when women controlled their assets like a home, even if it was rented, they were able to have more satisfying intimate relationships and decrease the rate of intimate partner abuse. Studies also show that there is a significant relationship between asset ownership and intimate partner abuse, and that the two are negatively correlated.

230. See 45 C.F.R. § 261.33 (2014); Christy-McMullin, Designing Policies, supra note 9, at 112.
232. See id. at 18 (noting research is limited and that women’s level of income may play large factor in issue as well).
234. See Conner, supra note 36, at 383–86.
235. See Christy-McMullin & Shobe, supra note 231, at 17–18 (providing this conclusion, but with caveat that significant relationship may not exist if race, ethnicity, marital status, and age are accounted for).
236. See Christy-McMullin, Designing Policies, supra note 9, at 113–14.
237. See id. at 114. But see Christy-McMullin & Shobe, supra note 231, at 18 (noting that, while their study supported this conclusion, it may be that female-only asset owners may be at higher risk of physical abuse because male perpetrators of abuse are not at risk of losing assets).
Accordingly, women’s ownership of assets is correlated with reduced abuse, whether the assets are homes, small businesses, savings accounts, retirement accounts, pensions, or other personal property.239

The sixth area of economic security is family law support orders. Many, though not all, civil protective order laws permit the petitioners to request economic support for children and themselves, reimbursement for expenses incurred due to the violence, control over the home, support for the home, such as rent and utilities, and control over a car for transportation.240 Divorce and child custody orders also include support orders for children and spouses. However, the current family and domestic violence laws, as written and as applied, create obstacles to women’s economic security. For instance, although laws may permit emergency and temporary custody and support orders, courts may not in fact schedule hearings in a timely manner.241 Moreover, although divorce laws often require the division and distribution of marital property in an equitable manner, such equitable distribution may not in fact consider the economic effects of domestic violence, including the effects of economic abuse such as sabotaging education and employment opportunities and the isolation that reduces access to employment networks and social support.242 Further, if the woman remains the primary caretaker, a 50/50 split of assets, coupled with child support, calculated pursuant to the guidelines, does not adequately compensate the woman for the burden that caretaking responsibilities place on the woman’s finances and employment opportunities.243

E. Supporting Relationship Security Enhances the Social Capital of Persons Subjected to Abuse

Focusing on security reorients the framework for addressing domestic violence from the danger/haven false dichotomy to an understanding that the person subjected to abuse is someone who is at the center of a broad web of relationships, connections, and support, composed of intimate and familial relationships, community, employment, home and health; and this web contributes to necessary security. As discussed above, respect for a woman’s agency and dignity, including support for the woman’s connection to her community, is critical for a successful response to domestic violence. Connections to the community help to develop and sustain her resilience in addressing the intimate partner abuse in her life. Because abuse is systemic in nature, a person subjected to abuse will be more satisfied with her response to abuse if she can also employ a systemic response.

240. See Klein & Orloff, supra note 69, at 931–41, 990–1004.
241. This has been the experience of the student attorneys in my Family Law Clinic.
242. See, e.g., MD. CODE ANN., FAM. LAW § 8-205 (West 2014).
243. See Christy-McMullin, Designing Policies, supra note 9, at 114.
As seen above in Part II.A., the coercive tactics of abuse include tactics of isolation and economic abuse, which create greater dependence upon the abuser and can hamper the ability of a woman subjected to abuse to counter the detrimental financial and health effects of the abuse. Research has shown that for women experiencing intimate partner abuse, their connection to their community is critical for their physical and psychological well-being. The likelihood of future abuse diminishes with a stronger network of supportive community for the woman subjected to abuse. Critically, though, the community will best serve the woman who has experienced intimate partner abuse if it is attuned to the woman’s experience of abuse and her judgments regarding how to best respond to the abuse.

Research shows that community connections are critical to family success. For instance, Carol Stack conducted an early study of survival strategies in an urban African-American community. Stack found that “domestic functions are carried out for urban Blacks by clusters of kin who do not necessarily live together . . . .” Rather, “the basis of these units is the domestic cooperation of close adult females and the exchange of goods and services between male and female kin.” The families observed by Stack consisted of kin and non-kin who made up a small, organized, and stable network of persons who “interact[ed] daily,” “share[d] reciprocal obligations,” and provided for the “domestic needs of children and assuring their survival.” Stack’s study showed the “adaptive strategies, resourcefulness, and resilience of urban families under conditions of perpetual poverty or the stability of their kin networks.” These networks also engaged in “swapping” of items such as “food stamps, rent money, a TV, hats, dice, a car, a nickel here, a cigarette there, food, milk, grits, and children.”

The reality that parenting and families exist within a web of support has been shown to be vital to family success in current times as well.

244. See Goodman & Epstein, supra note 8, at 99.
245. See id. at 101.
247. Id. (citing Carol Stack, The Kindred of Viola Jackson: Residence and Family Organization of an Urban Black American Family, in Afro-American Anthropology: Contemporary Perspectives (N.E. Whitten & John F. Szwed eds., 1970)).
248. Id. at 30–31.
249. Id. at 22.
250. Id. at 32.
Some options supporting the connection between women experiencing abuse and their community currently receive some federal funding. For instance, open shelters—shelters which are situated in a residential community and are known to be shelters by the community—permit women subjected to abuse to be in a neighborhood and be known by their neighbors.\textsuperscript{252} Other shelters are located in confidential locations. They do not identify their purpose to people and entities in their proximity. The women residents are not permitted to share the location of the shelter or invite their family, friends, lawyers, or other service providers from outside the shelter to the shelter. Moreover, confidential shelters may be located a great distance from the woman’s actual neighborhood and community.\textsuperscript{253} Therefore, the shelter serves as an actual physical barrier between the woman and her support network and community. Research shows that open shelters are just as safe as confidential shelters because, although the location is known to the abusive person, the community acts vigilantly on behalf of the woman.\textsuperscript{254} She is not as isolated as she would be in the confidential shelter, and so her life satisfaction is increased. By focusing on security, we can develop and maintain other options to support the choice of the person subjected to abuse to remain connected with her community. For example, policies should respect a person’s decision to avoid shelters and remain at her home in close proximity to her community support network of family, friends, and neighbors. This network could be critical to maintaining her employment, child care, and general support.

In addition, the development of a woman’s social capital, “social relationships based on trust that have value or can be used productively,”\textsuperscript{255} can be instrumental in addressing domestic violence as well.\textsuperscript{256} While social capital is important for persons in all economic classes,\textsuperscript{257} it is essential for those experiencing poverty in order to access employment with a living wage. Moreover, persons subjected to abuse must build social capital across various institutions, communities, and individuals in order to rise from poverty and address persistent abuse. For welfare recipients, “social capital generated within closed networks of sharing relationships as well as bridging capital that crosses class, race, and social boundaries”

\textsuperscript{252} See Goodman & Epstein, supra note 8, at 102.
\textsuperscript{253} See id. at 102, 121–22.
\textsuperscript{254} See Johnson, Home, supra note 15, at 38 (citing Goodman & Epstein, supra note 8, at 102).
\textsuperscript{255} Jo Anne Schneider, Social Capital and Welfare Reform: Organizations, Congregations, Communities 9 (2006).
\textsuperscript{256} See Goodman & Epstein, supra note 8, at 101; Conner, supra note 36, at 366–69.
helps “to create relationships of trust” that are instrumental for accessing and maintaining critical resources.\textsuperscript{258}

IV. Proposals Toward Security—The Legal and Funding Landscape

In this Part, I set forth a variety of legal, funding, and policy proposals intended to promote women's security.

A. Expand Actionable Harm in Civil Protection Order Laws

Focusing on security requires institutions to identify all forms of abuse, including non-physical abuse and all forms of harm from the broad range of abuse. Civil protection order laws should be expanded to address all forms of abuse. Relief should be available to persons subjected to emotional, psychological, and economic abuse, in addition to physical abuse. Immigration law, divorce law, and public benefits law provide useful models, as each identifies domestic violence as physical and emotional cruelty.\textsuperscript{259} By expanding the forms of abuse that are actionable, the law will address the actual experience of petitioners who have been subjected to abuse and create further options for them to seek a court-ordered remedy to address the abuse.

B. End Mandatory Responses

To effectuate security, jurisdictions should end all mandatory responses to domestic violence, such as laws and policies that mandate arrest, “no-drop” prosecution policies, and criminal no contact orders against the will of the person subjected to abuse. As discussed above, the research shows that women subjected to abuse are the best predictors of their risk of future abuse.\textsuperscript{260} In addition, when women subjected to abuse are able to make their own decisions about how to respond to the abuse, they are more satisfied with the results—and the violence to which they are subjected tends to decrease.\textsuperscript{261}

C. Address Barriers to Employment

Economic security can be promoted for persons subjected to intimate partner abuse by legislation that addresses barriers to employment and provides exceptions to unemployment insurance for persons subjected to abuse. These laws need to define domestic violence broadly to include the full range of abuse, giving women subjected to abuse full access to economic security. As discussed above, economic abuse can include sabotaging one’s employment through a variety of tactics, such as constraining the


\textsuperscript{259} See Johnson, Redefining Harm, supra note 5, at 1156–61.

\textsuperscript{260} See supra notes 188–204 and accompanying text.

\textsuperscript{261} See id.
employee’s ability to leave for work in a timely fashion, sabotaging one’s transportation to and from work, showing up at one’s place of work, and contacting the employee at work to render her unable or seemingly unable to perform her duties. Moreover, the person subjected to abuse may need to take leave in order to address the domestic violence, for matters such as criminal or civil litigation that require multiple court appearances. Similarly, up to one half of persons subjected to abuse report that they lost their employment due to the abuse.262 Accordingly, legislation protecting the employment of persons subjected to abuse is necessary.

Some states have passed such legislation, but all states need to do so. As of 2013, fifteen states and the District of Columbia had laws in place to protect employment rights of persons subjected to abuse.263 Nine states prohibit discrimination or retaliation against an employee who was subjected to domestic violence either because she was a subjected to abuse, because there was disruption at the workplace related to the abuse, or because she took time off from work to address the domestic violence.264 Three states have laws requiring the employer to offer workplace accommodations to the employee because of the abuse, such as relocation, flexible work schedule, and screening of incoming telephone calls.265 Further, ten states and the District of Columbia require employers to offer some form of leave to employees subjected to abuse when necessary to address the abuse, such as preparing to participate in court proceedings, to receive medical treatment or counseling, to obtain safety planning resources, and to secure housing assistance, relocation assistance, and legal assistance.266 The states vary on the amount of leave, whether it is paid or unpaid, and on the types of actions that will be eligible for the leave.267 Economic security would be greatly increased if each jurisdiction passed similar legislation recognizing the employment rights of persons subjected to abuse.


264. See id. at 1–9 (noting, as of 2013, states with some form of such legislation are California, Connecticut, Hawaii, Illinois, Kansas, New York, North Carolina, Oregon, and Rhode Island).

265. See id. (showing Hawaii, Illinois, and Oregon have some such form of legislation).


267. See id.
In addition, federal legislation in the area of employment rights for persons subjected to abuse needs to be passed. There have been major legislative initiatives that have been introduced in Congress, but as of yet, not passed. For instance, Congress proposed the Security and Financial Empowerment Act, which provides for workplace emergency leave, anti-discrimination provisions, and reasonable accommodation provisions for persons subjected to abuse. Congress also proposed the Healthy Families Act, which provides for paid sick leave that can be used by persons subjected to domestic abuse to pursue legal action and to obtain medical attention, services, counseling, and relocation. In addition, the Healthy Families Act precludes retaliation against the employee for taking such leave. Such federal legislation needs to be passed to ensure economic security.

To assist persons subjected to abuse in finding employment, programs should be offered that provide effective career counseling for such persons. One effective model of career counseling is based upon an empowerment model and is called Social Cognitive Career Theory (SCCT). SCCT requires career counselors to “facilitate critical reflection and awareness of the power dynamics at work in battered women’s lives; facilitate the recognition, enhancement, and use of the skills and resources these women have; and, ultimately, facilitate the ability of these women to contribute to the empowerment of others.” This model uses Ellen Harley McWhirter’s empowerment theory, which defines empowerment as:

“[T]he process by which people, organizations, or groups who are powerless or marginalized (a) become aware of the power dynamics at work in their life context, (b) develop the skills and capacity for gaining some reasonable control over their lives, (c) which they exercise, (d) without infringing on the rights of others, and (e) which coincides with actively supporting the empowerment of others.”


269. See LEGAL MOMENTUM, supra note 263, at 14 (citing H.R. 1229, 113th Cong. (2013)).

270. See id. at 14–15 (citing H.R. 1286/S. 631, 113th Cong. (2013)).

271. See id. at 15.


273. Id.

274. Id. (quoting ELLEN HARLEY McWHIRTER, COUNSELING FOR EMPOWERMENT 12 (1994)).
Accordingly, SCCT uses the “five Cs of empowerment,”275 which include: collaboration,276 context,277 competence,278 critical consciousness,279 and community,280 in order to effectively assist women subjected to abuse in identifying and obtaining their career goals.

D. Expand Eligibility for Unemployment Insurance Benefits

Reformers can promote the economic security of persons subjected to abuse who lose their employment due to abuse by expanding eligibility for unemployment insurance.281 At present, individuals cannot qualify for unemployment insurance when they leave their job voluntarily, unless they have “good cause” for leaving.282 In many states, personal reasons do

275. Id. (“Recommendations for empowering battered women and for developing interventions that address the variables and relationships defined by SCCT are aligned with the five Cs of empowerment: collaboration, context, competence, critical consciousness, and community.”).

276. See id. Collaboration is essential for effective career counseling because it frames the process as the “mutual definition of the problems and construction of goals as well as collaborative and flexible strategies for change.” Id. As such, the person subjected to abuse is identified as central to the counseling as it is her life and goals at issue, while the counselor has an integral role in facilitating the counseling. See id.

277. See id. To be effective in the counseling, assumptions or generalizations cannot guide the process. Rather, it is “essential that the battered woman’s life situation, including her educational and career concerns, be understood in context.” Id. Specifically, effective career counseling needs to account for her culture; family structure; religious, community, economic and support network; law enforcement sensitivity; and local opportunities for employment and school. Id.

278. See id. Chronister and McWhirter explain that effective career counseling requires the counselor and the woman to recognize “the skills, resources, and experiences that women possess and that may contribute to achieving their counseling goals as well as to developing new skills.” Id. (citation omitted).

279. See id. Critical consciousness emphasizes the individual’s ability to examine herself within her life context and her ability. See id. Specifically, this component engages a power analysis for the woman to identify how she is and can be transforming the identified power operating. See id. It also serves to identify support for the woman in her community as well as obstacles. The career counselor also needs to engage in this reflection to ensure that her privilege does not undermine the woman’s work. Id. at 423.

280. See id. Community serves many goals for career counseling. Community can be created through support groups or facilitated through similar interests or existing connections. This community can then “provide validation of roles and identity; physical, emotional, and social support; and opportunities for belonging and contribution.” Id. In addition, the community can serve as a network for school and job opportunities, development of specific skills, and development of general skills required to be persuasive and equipped in oral advocacy. See id.

281. See generally Runge, supra note 268 (noting some states have amended unemployment insurance statutes to include or exclude persons subjected to domestic violence).

not qualify as good cause.\textsuperscript{283} Some states have established a replicable model, in which experiencing domestic violence can qualify as “good cause” for leaving employment.\textsuperscript{284} Other states have recognized that people who experience domestic violence may be unable to satisfy a requirement of being “able and available” to work.\textsuperscript{285} More than thirty jurisdictions have passed laws that explicitly include domestic violence within the definition of “good cause” or otherwise provide unemployment insurance to domestic violence survivors.\textsuperscript{286} While the details of each law vary, in most cases, the applicant must fulfill all other eligibility requirements for unemployment insurance, and often the applicant must provide documentation or certification of the violence.\textsuperscript{287} In addition, under the American Recovery and Reinvestment Act of 2009 and the Worker, Homeownership, and Business Assistance Act of 2009, grants were provided to the states to extend unemployment insurance benefits to workers who leave their jobs due to domestic violence, but those extended benefits ended in the beginning of 2014.\textsuperscript{288}
E. Increase Access to Public Benefits

Another way to increase income for women subjected to abuse is by increasing their access to public benefits if they are income eligible. For instance, the TANF program provisions that permit persons subjected to abuse to obtain TANF without increasing their risk of abuse should be available to persons subjected to all forms of abuse, not simply physical abuse. Ordinarily, TANF has strict requirements for eligibility: the individual must be employed and must be seeking child support where applicable. Additionally, TANF has a five-year (or sixty-month) limit, meaning that an individual may not receive TANF for more than sixty months (whether consecutive or non-consecutive) total in her life. But because domestic violence could increase if the TANF applicant were to seek child support from her abuser, or if she reported to her worksite, Congress passed the Family Violence Option (FVO), wherein the states are allowed to waive any of the TANF requirements for persons subjected to abuse. By adopting the FVO, a state certifies that it will screen to identify persons subjected to abuse while maintaining their confidentiality; will refer them to supportive services; and will waive program requirements such as time limits on the receipt of benefits, work requirements, or cooperation with child support enforcement if those requirements make it more difficult to escape the violence or would unfairly penalize the person. Notably, “[a]ll states have either formally certified adoption of the FVO [forty-one states and the District of Columbia] or reported to the federal government adoption of a comparable policy.” Additionally, states are allowed to exempt 20% of their TANF caseload from the sixty-month time limit for cases of “hardship,” including women who have been

289. Cf. Richard C. Fording & William D. Berry, The Historical Impact of Welfare Programs on Poverty: Evidence from the American States, 35 Pol. Stud. J. 37, 37 (2007) (“Many analysts have maintained that public assistance expansion during this period decreased poverty by raising the incomes of the poor (an income enhancement effect), while others have contended that welfare expansion increased poverty by discouraging the poor from working (a work disincentive effect.”).


291. See id. at 13.

292. See id.


295. Id.
battered or subjected to extreme cruelty.296 Unfortunately, despite the fact that the TANF waiver defines domestic violence broadly as “battered or subjected to extreme cruelty,”297 TANF case workers interpret domestic violence as physical violence, thus limiting who can gain access to the TANF economic and other resources.298 For instance, in New York State, waivers of TANF requirements are granted only where applicants can show that they are currently in danger.299 This excludes persons who are subjected to nonphysical abuse or who may not be in current danger because the person committing the acts of abuse is incarcerated, for example. Persons subjected to abuse are also denied the waiver if they are unable to comply with the monthly reevaluation requirement to remain qualified for the FVO because of the other demands on their time from addressing the intimate partner abuse.300 TANF has been shrinking and its effectiveness in lifting persons out of poverty and providing economic security has decreased despite calls from advocates who recommend that TANF benefits be increased.301

Additional provisions should be passed so that persons subjected to abuse may access public benefits. The EITC is the largest and most effective public assistance benefit available.302 Currently, under the EITC, mar-

---


297. 42 U.S.C. § 608(a)(7)(C)(iii) (2012). The definition of “battered or subjected to extreme cruelty” is: (I) physical acts that resulted in, or threatened to result in, physical injury to the individual; (II) sexual abuse; (III) sexual activity involving a dependent child; (IV) being forced as the caretaker relative of a dependent child to engage in nonconsensual sexual acts or activities; (V) threats of, or attempts at, physical or sexual abuse; (VI) mental abuse; or (VII) neglect or deprivation of medical care.

Id.

298. See CASEY ET AL., supra note 294, at 11–12; Gallagher, supra note 296, at 1002 (“There is, however, substantial evidence that TANF case workers often fail to effectively screen for domestic violence and/or to offer waivers and service referrals when appropriate.”); SHORTCHANGING SURVIVORS, supra note 290, at 17 (describing screening process); see also Shelly Kintzel, Comment, The Effects of Domestic Violence on Welfare Reform: An Assessment of the Personal Responsibility and Work Opportunity Reconciliation Act as Applied to Battered Women, 50 U. KAN. L. REV. 591, 592–94 (2002) (discussing how domestic violence cannot be limited to physical abuse).

299. See SHORTCHANGING SURVIVORS, supra note 290, at 6–7.

300. See CASEY ET AL., supra note 294, at 12.

301. See id., at 16.

ried couples who have lived together for at least six months of the last taxable year must sign a joint return in order to be eligible for the EITC.303 They cannot access the credit if they file “married filing separately” (MFS).304 Accordingly, the EITC would be unavailable in a situation in which the person subjected to abuse is married and living with her spouse but wishes to file separately to avoid further or future economic abuse or coercion. The EITC requirements should be modified to permit the MFS status to receive the EITC benefit under these circumstances.

There are a few ways in which this could be accomplished.305 First, as it did for a credit under the Affordable Care Act (ACA), the IRS could issue a Notice to permit the EITC to be available to persons subjected to abuse who are unable or unwilling to file jointly because of the abuse.306 Under the ACA Notice, a married taxpayer may file MFS if she is living apart at the time she files her tax return, she cannot file jointly because of the abuse, and she describes the situation in her tax return.307 A similar Notice could be issued for the EITC. Further, towards more security, the credit could be available even if the person subjected to abuse does not live separate and apart from her spouse who is abusive. One proposal for this is to permit her to file MFS and indicate that she is filing separately due to the abuse, is economically independent of her spouse, and requests an independent credit.308 Another option is to permit her to file MFS because of the abuse and then require the IRS to calculate the amount of the credit based on each spouse’s separate filings, with the refund being split between the spouses.

303. See SUSAN MORGENSTERN & MARY M. GILLUM, CONSUMER RTS. FOR DOMESTIC VIOLENCE SURVIVORS INITIATIVE, FEDERAL TAX ADVOCACY FOR DOMESTIC VIOLENCE SURVIVORS 5 (2009), available at http://www.nclc.org/images/pdf/domestic_violence/intersection-tax-law02-24-2010.pdf. There is an exception if one’s spouse did not live in the home for the last six months of the year. See IRS, PUBLICATION 596: EARNED INCOME CREDIT 5 (2014), available at http://www.irs.gov/pub/irs-pdf/p596.pdf. If the spouse has not lived with the other spouse for six months in the year, she may be able to file as head of household instead. See id.

304. See IRS, supra note 302, at 5.

305. Thanks to Fred Brown and Leslie Book for thinking through this problem with me and providing sound suggestions.


307. See id.

308. Thanks to Fred Brown for this idea, which I have simplified for purposes of this discussion.
F. Block Coerced Debt from Credit Reporting Agencies

As discussed in Part III above, access to assets is critical for security because such access has a negative correlation to intimate partner abuse. Accordingly, proposals that enhance the assets of a person subjected to abuse will increase the likelihood of reduced exposure to future violence. One way to increase assets is to increase one’s credit rating or address negative credit ratings. As Angela Littwin has documented, financial abuse can be perpetrated by “coerced debt,” defined as “nonconsensual, credit-related transactions . . . .” Debt can be coerced in multiple ways, including taking credit cards out in a partner’s name, forcing a partner to take loans, tricking a partner into signing a quitclaim deed to a home, and excessively charging a partner’s credit cards. Coerced debt can destroy a person’s credit rating, which, in turn, can limit the ability to rent an apartment or take out a car loan, and can increase the interest rate on a loan. By creating obstacles to obtaining housing, coerced debt impinges on the security of a person subjected to abuse. Littwin proposes blocking coerced debt from credit reporting agencies and having family courts decide the parties’ responsibility for coerced debt. Littwin’s proposal provides an initial step towards addressing the security of persons subjected to abuse. To further address these issues, community-based educational programs could be funded to address economic abuse and coerced debt, as well as various other ways to address them.

G. Allow Monetary Damages in Civil Protection Order Cases

All United States jurisdictions could include monetary damages in their CPO laws for injuries to persons subjected to abuse, whether they be physical, sexual, emotional, psychological, or economic. Only twenty-two jurisdictions specifically permit restitution as a form of relief from domestic violence. Compensatory damages are another area that should

310. See id. at 986–91. Misuse of one’s partner’s social security number to take out loans or credit cards in that person’s name is another form of coerced debt. See id. at 987. Fortunately, a person subjected to abuse may request a new social security number if they are being harassed, abused, or if their life is in danger. See Soc. Sec. Admin., New Numbers for Domestic Violence Victims (Sept. 2013), available at http://www.ssa.gov/pubs/EN-05-10093.pdf. This provision is a good example of a security provision, as it defines domestic violence broadly to include all forms of abuse.
311. See Littwin, supra note 309, at 1000–03.
312. See id. at 992–93, 1002.
313. See id. at 1002–03 (noting coerced debt often prevents women from leaving abusive relationships).
315. See Johnson, Redefining Harm, supra note 5, at 1155.
316. See ABA Comm’n on Domestic Violence, supra note 151 (noting states with CPO remedy of restitution are Alaska, California, Connecticut, Delaware, Illi-
be expanded in CPO laws. One effective approach to granting this remedy is seen in those jurisdictions that permit “‘any other relief that would address the domestic violence’...”317 Such a “catch-all” remedy permits persons subjected to abuse to identify how they have been harmed by the abuse and to seek a remedy for their injuries.318 While forty jurisdictions in the United States include a catch-all remedy in their CPO laws, they are rarely, if ever, used for compensatory damages and full restitution. To ensure security for persons subjected to abuse, all jurisdictions should both include a robust catch-all remedy in their CPO laws and use them to award compensatory damages.

H. Increase Housing Options319

Further, all civil protective order laws should create viable short- and long-term housing remedies that provide shelter and are sustained by government subsidy or by shifting the cost of the housing to the person who perpetrated the abuse.320 Similarly, all jurisdictions should pass laws that permit renters subjected to intimate partner abuse to maintain secure housing options by evicting the person perpetrating the abuse, terminating her lease early if she wants to leave the shared rented home, and changing the locks of a rented apartment.321

To increase housing security, certain steps could be taken. First, as discussed above, government money allocated to less effective safety paradigm responses to domestic violence could be shifted to fund security paradigm options, such as low-barrier open shelters that have few treatment or other requirements to maintain residency, transitional housing, and long-term housing.322 CPO laws should not only allow the petitioner to exclude the person committing the abuse, they should also allow the petitioner to request that the respondent contribute to alternative housing for the person subjected to abuse if she chooses to leave the home.323 Jurisdictions could also provide rental assistance for an extended period of time to support a person subjected to abuse who lives in alternative housing, Indiana, Maine, Massachusetts, Minnesota, Mississippi, Missouri, Montana (though it is unclear), Nevada, New Hampshire, New Jersey, New Mexico, New York, Oregon (limited to emergency monetary award), Pennsylvania, Tennessee, West Virginia, Wyoming (only medical)).

317. See Johnson, Redefining Harm, supra note 5, at 1155 (quoting Klein & Orloff, supra note 69, at 912–14).
318. See id.
320. See id. at 26.
321. See id. at 18–26.
322. See id. at 50–51.
323. See id. at 53 (discussing CPO laws in New Jersey and West Virginia that permit petitioners to obtain funds for alternative housing from respondents, as well as CPO laws in ten jurisdictions requiring respondents to provide alternative housing to petitioners).
ing.\textsuperscript{324} If, on the other hand, the CPO petitioner is able to vacate the respondent from the home, CPO laws should afford petitioners the ability to obtain a share of the household maintenance expenses from the respondent to ensure that she can maintain the home.\textsuperscript{325} All jurisdictions should have laws that permit persons subjected to abuse to terminate their leases early if necessary to address the violence.\textsuperscript{326} Also, all jurisdictions should have laws that prohibit discrimination in housing, rental or owned, against persons subjected to abuse.\textsuperscript{327} Such laws will permit persons to access and maintain stable housing. All jurisdictions should have laws that permit the person subjected to abuse to defend against a landlord’s eviction action based on the domestic violence.\textsuperscript{328} Further, all jurisdictions should have laws that permit renters subjected to abuse to require their landlords to change the locks to protect against future domestic violence.\textsuperscript{329} Finally, all jurisdictions should permit a petitioner to seek a CPO injunction against future abuse without requiring she stay away from or not contact the respondent.\textsuperscript{330} When the person subjected to abuse chooses to stay in the relationship that had been abusive, her decision should be respected and accepted, to enhance her relationship, health, and, here, her housing security.

I. Time Banking

Another way to increase the relationship security of a person subjected to abuse is to create or recreate her connection to the community and build her social capital. One way to effectively build social capital and connect a person subjected to abuse to her community is Time Banking

\textsuperscript{324} See id. (citing Alaska’s law providing rental assistance for thirty-six months).
\textsuperscript{325} See id. at 26 (discussing CPO vacate provisions in New Jersey and Missouri that require respondent to pay portion of rent or mortgage if excluded from home).
\textsuperscript{326} Id. at 33 n.141 (citing rental laws in Arizona, Colorado, Delaware, Washington, D.C., Illinois, Indiana, Maryland, Minnesota, New York, North Carolina, Oregon, Texas, Washington, and Wisconsin that permit persons subjected to abuse to terminate their leases early).
\textsuperscript{327} See id. at 37 n.153 (citing anti-discrimination laws in Washington, D.C., Indiana, North Carolina, Oregon, Rhode Island, Washington, Wisconsin, and Westchester County, New York that prohibit housing discrimination against persons subjected to abuse).
\textsuperscript{328} See id. at 19–20. Currently, only eleven jurisdictions have laws protecting persons subjected to abuse from this variety of landlord action. See id. at 20 n.81 (noting Colorado, District of Columbia, Iowa, Louisiana, Maryland, Minnesota, New Mexico, Oregon, Virginia, Washington, and Wisconsin have such laws).
\textsuperscript{329} See id. at 19–20. Currently, only ten states have laws that allow tenants to change the locks on their apartments because of abuse. See id. at 20 n.88 (noting that Arizona, Arkansas, Illinois, Indiana, Maryland, North Carolina, Oregon, Utah, Virginia, Washington, and Washington, D.C. currently have laws to this effect).
\textsuperscript{330} See id. at 45 (citing Maryland’s law allowing petitioner to seek injunction against future abuse without also requiring abuser to leave home or stop contacting petitioner).
and Co-Production.\textsuperscript{331} Time Banking is an economy that relies on people’s work and giving time rather than money. It is being used in over twenty countries to meet people’s needs and deal with social problems—such as childcare, mental illness, job training, social justice movements, transportation, and drug treatment.\textsuperscript{332} Unlike the monetary economy, as Professor Edgar Cahn explains, the invisible core economy that Time Banking relies upon is made up of “family, neighborhood, community, and civil society.”\textsuperscript{333}

Time Banking uses human resources, where a person can give an hour to someone else to perform such assistance as caring for a child or tutoring someone, and that hour spent also creates a relationship. As Cahn explains, “[w]hen we give an hour to rebuild community, we are building something we cannot buy. Time Banking enables us to value that and in doing so, we value ourselves, our time, our being.”\textsuperscript{334} When a person gives an hour to another, the person earns one Time Dollar.\textsuperscript{335} The person who is helped owes one Time Dollar and can pay it by helping anyone.\textsuperscript{336} There is a database that identifies what persons in the community are able and willing to do and when they can do it.\textsuperscript{337} A need for assistance is matched using the database.\textsuperscript{338} This same system keeps track of what a person is owed or has earned.\textsuperscript{339}

Time Banking is beneficial not just because it provides the person subjected to abuse with assistance but also because its reciprocal nature allows her to feel as though she is connected and contributing to her community. Research shows that acting more like friends is a healing and restorative venture.\textsuperscript{340} Co-Production is a construct that helps explain the value and mechanics of Time Banking and the shared social justice goals of respecting human beings.\textsuperscript{341}

According to Cahn, social capital, “composed of trust, reciprocity, and civic engagement,” is generated through Time Banking and Co-Produc-
These mechanisms could be a useful vehicle for persons subjected to abuse who want to be connected to their community and also have needs that their neighbors could help fulfill. In return, the person subjected to abuse would also help others, in order to earn Time Dollars for those owed. The benefits would be the ability to address child care, job training, and health care needs without money; identifying one’s own capacity and abilities; feeling empowered by helping another and making a difference; and gaining friends and a stronger, larger network. Moreover, Time Banking is built around the idea that all people are agents and have vast capabilities. They are not problems or victims, but rather solutions. People are not in need of charity or a hand out, but rather can make their own change. Such ideas are integral to a security paradigm as well, making Time Banking a useful model for expanding security options.

J. Increase Utilization of Tort Law

Tort law is an underutilized potential vehicle for recovery of economic damages for people subjected to abuse. People subjected to threatened or actual physical violence may recover damages for battery and assault. A claim of intentional infliction of emotional distress may provide damages for emotional distress resulting from "extreme and outrageous conduct." In addition to common law tort claims, several states and municipalities have laws that are specific to domestic and gender-based violence. For instance, in New Jersey and Washington, the “battered women's syndrome” tort allows consideration of a pattern of coercive behavior that constitutes intimate partner abuse.

342. Id. at 169.
343. Similar to the ideas of Time Banking and Co-Production, Carol Stack studied and then documented how Urban Blacks in the 1960s survived despite limited economic resources. See supra notes 246–50 and accompanying text.
345. Time Banking has been used by a project working with 5,000 women, as discussed by Elvira Méndez and her co-authors. See id. at 2. If the project’s screening showed domestic violence, then a group of men and women worked together to respond to the situation in a supportive group environment focused on empowerment. See id.
347. See Carey, supra note 346, at 696, 702–03 (discussing intentional infliction of emotional distress).
348. RESTATEMENT (THIRD) OF TORTS: LIABILITY FOR PHYSICAL AND EMOTIONAL HARM § 46 (2011); see also Johnson, Redefining Harm, supra note 5, at 1158.
349. See Carey, supra note 346, at 709–12.
based crime committed against her, while in Illinois and California, women can bring a tort action for gender-based violence. This tort is based on an actual or threatened battery or sexual assault. In addition, California law recognizes a domestic violence tort that defines abuse as “intentionally or recklessly causing or attempting to cause bodily injury, or placing another person in reasonable apprehension of imminent serious bodily injury to himself or herself, or another.”

Thus, tort law provides a mechanism through which a woman subjected to abuse may receive monetary damages for emotional and physical harm resulting from physical abuse and may do so without engaging the criminal justice system. Like criminal law, however, most cognizable tort law claims prioritize physical violence, failing to provide a remedy for other forms of abuse, such as emotional or economic abuse. Expanding tort law remedies to permit damages for emotional or economic abuse would be more effective in promoting women’s economic security.

V. CONCLUSION

This Article began with a quotation from scholars in the area of cybersecurity. They state that cybersecurity requires “organizational, legal, economic, and social” aspects. In sum, because our national computer network will never be “safe” from hackers, they argue, the goal of the system should be one that is secure because it is resilient and can withstand, recover, and be strong in the face of breaches of security. Similarly, our legal system and funding schemes could benefit from focusing on supporting the resiliency of persons subjected to abuse—what I call security—which encompasses organizational, legal, economic, and social aspects that help the person subjected to abuse fortify her liberty, dignity, and agency.

Current domestic violence policy, with its limited goal of safety—primarily focused on short-term physical separation to decrease physical intimate partner violence—is both too broad and too narrow a goal. It is too broad a goal because actions taken against the will of the woman subjected to abuse may actually increase her risk of abuse rather than make her safe. In addition, physical separation can also increase abuse. On the other hand, it is too narrow a goal, because the safety paradigm fails to recognize and respond to all forms of domestic abuse, all forms of harm from such abuse, other responses that are not focused merely on short-term physical

350. See id. at 715–16.
351. Id. at 716 (quoting CAL. PENAL CODE § 13700(a) (West 2005)).
352. See id. at 735–53.
353. Carey states that most successful intentional infliction of emotional distress claims were based on actual or threatened physical abuse. See id. at 702.
354. See id. at 754–55.
355. SINGER & FRIEDMAN, supra note 1, at 36.
356. See id.
separation, and the decision making capabilities of the woman subjected to abuse.

Instead, this Article argues for domestic violence policy, which shapes the civil legal system and government funding criteria, to focus on security. Security incorporates all forms of abuse, including physical, sexual, emotional, psychological, and economic; focuses on all harms that result from these varying forms of abuse; identifies short- and long-term responses that enhance the economic, housing, health, and relationship security of the person subjected to abuse because of the connection between such security and the decrease of domestic violence; and ensures that the person subjected to the abuse is the decision maker as to the appropriate responses to the abuse, because her agency and dignity have a positive relationship to her satisfaction with her life and the recurrence of domestic violence.

As stated earlier, the domestic violence movement is at a choice moment where changing from a safety paradigm to a security paradigm could be beneficial to persons subjected to abuse. A security paradigm could be more effective in identifying and remedying all forms of abuse and the harms from abuse, as well as achieving the goals of addressing intimate partner violence, because the domestic violence legal, policy, and funding systems would address domestic violence in the way that the person subjected to domestic violence would like to address it. Such goals, weighed in the context of her other goals—such as continuing her relationship, achieving economic stability, evaluating her safety, and protecting her children—may result in her choosing to eliminate, to decrease, or to tolerate a measure of the domestic violence. A shift from safety to security would provide state and institutional short- and long-term options to better address all forms of intimate partner abuse beyond physical violence; better address all harms of intimate partner abuse; and enhance effective responses to intimate partner abuse by supporting the agency of persons subjected to abuse, by supporting their community connections and increasing their social capital, and by increasing their access to economic, housing, and health resources.
WHEN the Supreme Court recently handed down its decision in Halliburton Co. v. Erica P. John Fund, Inc. ("Halliburton II"), the decision was not greeted with enthusiasm. For liberals, it was adding yet another hurdle to the class certification process. On the other hand, conservatives had hoped that the Court would overturn its earlier decision in Basic Inc. v. Levinson, which had adopted the “fraud on the market” theory of reliance. Justice Thomas, in his concurring opinion, argued for just such a result.

Halliburton II is in some respects a strange opinion. Both the majority opinion and Justice Thomas’s concurring opinion, which was in effect a dissenting opinion, evidence both insight and tunnel vision. Justice Roberts preserved the opportunity to establish reliance in a class action but continued to focus on market efficiency, a questionable, troublesome, and, as applied, dubious concept. Justice Thomas correctly pointed out some of the flaws in the fraud on the market theory, but he would require individualized proof of reliance, thereby eliminating the class action as a remedy for misrepresentations by management of publicly traded corporations. While the Basic theory is flawed, there is another, more straightforward way to establish reliance by members of a class: namely, that investors rely, not on the integrity of market price, but rather on the integrity of the information that corporate management inserts into the market. Arguing for this approach is the purpose of this Article.

In evaluating Supreme Court jurisprudence in the securities area, it is essential to ask the following question: for whose benefit were the securities laws enacted? The Securities Exchange Act of 1934 (“1934 Act”) early on provides that the purpose of the Act is “to insure the maintenance of fair and honest markets,” while section 10, pursuant to which Rule 10b-5 was promulgated, gives the Securities and Exchange Commission (SEC)
authority to promulgate anti-fraud regulations “as necessary or appropriate in the public interest or for the protection of investors.”\footnote{Id. § 78m(d)(1).}

The Supreme Court itself has said that a “fundamental purpose [of the securities laws is] to substitute a philosophy of full disclosure for the philosophy of caveat emptor and thus to achieve a high standard of business ethics in the securities industry.”\footnote{SEC v. Capital Gains Research Bureau, 375 U.S. 180, 186 (1963).} The Court added that Congress intended securities laws enacted for the purpose of avoiding frauds to be construed “not technically and restrictively, but flexibly to effectuate its remedial purposes.”\footnote{Id. at 195.}

Contrariwise, the recent stance of the Supreme Court and other federal courts often is to protect corrupt management from innocent investors. Justice Thomas, in the conservative concurrence in \textit{Halliburton II}, asserted that “[l]ogic, economic realities, and our subsequent jurisprudence” dictate that “\textit{Basic} should be overruled.”\footnote{Halliburton II, 134 S. Ct. 2398, 2418 (2014) (Thomas, J., concurring).} But the Court’s prior jurisprudence in the Rule 10b-5 area is substantially flawed,\footnote{See generally Charles W. Murdock, \textit{Janus Capital Group, Inc. v. First Derivative Traders: The Culmination of the Supreme Court’s Evolution from Liberal to Reactionary in Rule 10b-5 Actions}, 91 DENV. U. L. R EV. 369 (2014).} and, contrary to Justice Thomas’s assertion, logic and economic reality demonstrate that the conservative jurisprudence would emasculate investor protection and provide a shield for unscrupulous management.

Protecting management at the expense of investors is getting the statutory scheme upside down. It would appear that the federal courts often are living in an alternate universe.\footnote{See generally Charles W. Murdock, \textit{The Private Securities Litigation Reform Act and Particularity: Why Are Some Courts in an Alternate Universe?}, 45 LOY. U. CHI. L.J. 615 (2014).} For example, the Fifth Circuit Court of Appeals believed that it had the authority under \textit{Basic}, a decision which is universally recognized as favorable to plaintiffs, “to tighten the requirements for plaintiffs seeking a presumption of reliance”\footnote{Oscar Private Equity Invs. v. Allegiance Telecom, Inc., 487 F.3d 261, 264–65 (5th Cir. 2007).} in order to protect defendants from the “\textit{in terrorem} power of certification.”\footnote{Id. at 267.} It was the dissent in \textit{Basic} that was worried about the \textit{in terrorem} effect of class certification;\footnote{See \textit{Basic Inc. v. Levinson}, 485 U.S. 224, 262 (1988) (White, J., dissenting) (“I suspect that all too often the majority’s rule will ‘lead to large judgments, payable in the last analysis by innocent investors, for the benefit of speculators and their lawyers.’” (quoting SEC v. Tex. Gulf Sulphur Co., 401 F.2d 833, 867 (2d Cir. 1968) (en banc) (Friendly, J., concurring))).} The majority stated:

\begin{quote}
The modern securities markets, literally involving millions of shares changing hands daily, differ from the face-to-face transac-
tions contemplated by early fraud cases, and our understanding of Rule 10b-5’s reliance requirement must encompass these differences.15

Federal courts often seem more concerned with protecting management from strike suits than in protecting investors from fraud.16 In many instances, either the courts’ understanding of business is woeful or the opinions evidence a bias in favor of management.17 The Supreme Court could hardly be unaware of the extensive fraud by many corporations in the cases that have come before it. Recent Supreme Court decisions have dealt with channel stuffing and lying about whether a new product was being shipped,18 conspiring with a supplier to inflate earnings,19 misrepresenting that a product did not cause patients to lose their sense of smell when management was aware of injured users,20 changing accounting methodology to inflate earnings,21 hiding wrongful conduct in connection with market timing,22 and asserting that an FDA meeting was not focused

15. Id. at 243–44 (majority opinion) (footnote omitted).


17. See Murdock, supra note 11, at 624–30 (discussing court’s characterization of plaintiff’s channel stuffing allegation—that K-Mart stores had fifty to one hundred weeks of battery inventory, and that Walmart had thirty to fifty weeks of battery inventory—as being conclusory and lacking particularity). Walmart in particular is known as a leader in just-in-time inventory control. See id. at 635–39 (discussing court’s characterization of plaintiff’s extensive allegations of budgets, forecasts, daily reports, monthly reports, and “Stop Ship” reports, which notified management of manufacturing problems and their impact on volume shipments, as boilerplate because other cases had also referenced internal reports as being basis to hold senior management liable).


upon the company’s products when in fact such products were the focus of the meeting. 23

From the foregoing litany, it is inconceivable that the Supreme Court is not aware of the many circumstances in which ethically indifferent or corrupt management has misled the investing public. Yet, conservative members of the Court have either turned a blind eye to management fraud when issuing majority opinions 24 or, in dissents or concurrences, have advocated positions that would undercut investor protection in favor of insulating corporations and management from accountability. 25 There is considerable antipathy by the conservative members of the Court to a private civil remedy in the first place. Justice Thomas, in his concurrence in Halliburton II, began his opinion by asserting that “[t]he implied Rule 10b-5 private cause of action is ‘a relic of the heady days in which this Court assumed common-law powers to create causes of action.’” 26

However, former Justice Stevens has pointed out that these heady days persisted for 200 years: "Fashioning appropriate remedies for the violation of rules of law designed to protect a class of citizens was the routine business of judges.” 27 Moreover, highly regarded conservative Supreme Court Justices, such as Tom Clark 28 and John Harlan, 29 have recognized the important role that implied private causes of action play in the enforcement of the securities laws. Justice Clark, in recognizing a private cause of action under the proxy provisions of the 1934 Act, stated that “[w]hile this language [of the Act] makes no specific reference to a private right of action, among its chief purposes is ‘the protection of investors,’ which certainly implies the availability of judicial relief where necessary to achieve that result.” 30 Expanding upon this statement in a later decision,

25. See, e.g., Halliburton II, 134 S. Ct. 2398, 2417–18 (2014) (Thomas, J., concurring) (arguing that Basic should be overruled); Tellabs, Inc. v. Makor Issues & Rights, Ltd., 551 U.S. 308, 329–30 (2007) (Scalia, J., concurring) (arguing that inference of scienter is established only when plaintiff’s inference is “more plausible than the inference of innocence”).
27. Stoneridge, 552 U.S. at 177 (Stevens, J., dissenting). Justice Stevens referenced Marbury v. Madison, 5 U.S. 137 (1803), and stated, “[w]hile this language is true that in the early days state law was the source of most of those rules, throughout our history—until 1975—the same practice prevailed in federal courts with regard to federal statutes that left questions of remedy open for judges to answer.” Id.
Justice Harlan expressed his concern for the plight of small shareholders and declined to take a position that "would be bound to discourage such shareholders from the private enforcement of the proxy rules that 'provides a necessary supplement to Commission action.'" 31

While courts have been concerned about the *in terrorem* impact upon corporations and management of permitting class actions to proceed, there actually is a simple way for defendants to avoid such class action certification: let management tell the truth in the first place. Now, many will regard such a solution as not only simplistic but also naïve: cannot plaintiff lawyers always twist some statement to make it appear misleading? This concern is greatly overstated. The particularity demanded by the Private Securities Litigation Reform Act of 1995 (PSLRA) dampens this concern. Under that Act, plaintiffs must allege the misleading statements with particularity, without the benefit of discovery, and with their attorneys exposed to sanctions if they do not conduct due diligence before filing suit. 35

On the other hand, under the present regime, there are many meritorious cases that are not certified as class actions, and thereby generally die, because of the unrealistic requirements—as developed by the federal courts—for class action certification. Much of this revolves around whether the securities are traded in an efficient market. This determination, in turn, is required because of the “fantasy”—in the view of Justice Thomas—that investors rely upon the integrity of the market price. 37 This Article asserts that investors rely not upon price, but upon the total mix of information in the marketplace. If courts do not recognize this reality, they should at least recognize that what is necessary to establish reliance is not that the stock is traded in an efficient market, but rather that the misrepresentations in question had “price impact.” 38

---


34. *See id.*

35. *See id.*

36. For example, see the First Circuit’s confounding analysis of whether the market is efficient in *In re PolyMedica Corp. Sec. Litig.*, 432 F.3d 1 (1st Cir. 2005), followed by an absurd opinion on remand, 453 F. Supp. 2d 260 (D. Mass. 2006). For an in-depth discussion of the opinions in the *PolyMedica* litigation, see *infra* notes 121–82 and accompanying text.


38. *See generally* Lucian A. Bebchuk & Allen Ferrell, *Rethinking Basic*, 69 Bus. Law. 671 (2014) (arguing that what plaintiffs should need to establish reliance is not that there is efficient market, but rather that defendants’ representations have caused fraudulent distortion of market). For a further discussion of the arguments presented by Professors Bebchuk and Ferrell, see *infra* notes 243–52 and accompanying text.
Since the *Halliburton* decisions build upon *Basic*, *Basic* and its weaknesses will first be considered in Part II. The basic flaw in *Basic* was in adopting an approach to reliance that focused upon price, which is the result of information, rather than the information itself. It is information that investors rely upon, not price.

In Part III, the flaw in *Halliburton II* is analyzed, namely, the continued reliance upon the notion of an “efficient” market as a predicate for presuming reliance. Paradoxically, this Article uses the Court’s own hypothetical to illustrate that the success of plaintiffs’ litigation should depend, not upon market efficiency, but rather upon price impact. Part IV continues the analysis of why market efficiency, particularly as interpreted by the federal courts, is an outmoded and dubious concept. The Article uses the various opinions in the *PolyMedica* litigation to illustrate the unreasonable approach federal courts have taken in determining whether the market is efficient.

Part V looks at the internal inconsistencies between *Halliburton Co. v. Erica P. John Fund, Inc.* (“*Halliburton I*”),39 *Amgen Inc. v. Connecticut Retirement Plans and Trust Funds,*40 and *Halliburton II*, and argues that price impact, materiality, and loss causation are all interrelated concepts that should be determined, not at the certification stage, but in a trial on the merits. While this Article argues that all that should need to be proven at the certification stage is information distortion, not price distortion, it also considers an alternative approach using price distortion, as opposed to market efficiency, as a basis for presuming reliance.

The Article’s conclusion asserts that the normative position of federal courts should recognize that investors rely upon informational integrity, rather than price integrity, and that the interrelated, intensely factual issues of price impact, materiality, and loss causation should be reserved for trial.

II. THE BASIC FLAW: FOCUSING ON EFFECT RATHER THAN CAUSE

A. The Problematic Focus on Price Integrity, Contrasted with the Underlying Policies

*Basic Inc. v. Levinson*41 was an unusual case from a factual perspective. Normally, when investors claim that they were defrauded, the fraud arises from the company putting out positive information, thereby inducing the investors to buy, when the situation was actually negative. Consider *Makor Issues & Rights, Ltd. v. Tellabs, Inc.*,42 where the CEO made a series of positive and specific statements that one key product (the Titan 5500) was experiencing continued growth and that a newer product (the Titan

40. 133 S. Ct. 1184 (2013).
42. 437 F.3d 588 (7th Cir. 2006).
6500) was ready to ship. Contrariwise, demand for the 5500 was precipitously declining, and the 6500 was not yet being produced. These misrepresentations induced investors to buy. On the other hand, in Basic, the company put out negative information, namely, that no merger negotiations were taking place, when the situation was actually positive. Consequently, the plaintiff investors sold shares.

In Basic, the Court considered two issues: whether the merger negotiations were material and whether the investors could establish reliance based upon the fraud on the market theory. According to the Court, the fraud on the market theory, “[s]uccinctly put,” is as follows:

The fraud on the market theory is based on the hypothesis that, in an open and developed securities market, the price of a company’s stock is determined by the available material information regarding the company and its business . . . . Misleading statements will therefore defraud purchasers of stock even if the purchasers do not directly rely on the misstatements . . . . The causal connection between the defendants’ fraud and the plaintiffs’ purchase of stock in such a case is no less significant than in a case of direct reliance on misrepresentations.

When the majority in Basic adopted the fraud on the market theory, the Court stated that, in accepting the presumption of reliance, “we need only believe that market professionals generally consider most publicly announced material statements about companies, thereby affecting stock market prices.” The Court also stated that, by accepting this rebuttable presumption, it did “not intend conclusively to adopt any particular theory of how quickly and completely publicly available information is reflected in market price.”

The Court, while recognizing that “reliance is an element of a Rule 10b-5 cause of action” and that reliance “provides the requisite causal connection between a defendant’s misrepresentation and a plaintiff’s injury,” pointed out that there is “more than one way to demonstrate the causal connection.” It acknowledged that Affiliated Ute Citizens of Utah v. United States “dispensed with a requirement of positive proof of reliance, where a duty to disclose material information had been breached, concluding that the necessary nexus between the plaintiffs’ injury and the defendant’s

43. Id. at 593.
45. Id. at 226.
46. Id. at 241–42 (alterations in original) (quoting Peil v. Speiser, 806 F.2d 1154, 1160–61 (3d Cir. 1986)) (internal quotation marks omitted).
47. Id. at 246 n.24.
48. Id. at 248 n.28.
49. Id. at 243.
wrongful conduct had been established [by the materiality of the omitted information].”

In Basic, the Court should simply have built upon its analysis from Affiliated Ute Citizens. In this latter case, the Court stated that where the facts “involve[ed] primarily a failure to disclose, positive proof of reliance is not a prerequisite to recovery. All that is necessary is that the facts withheld be material in the sense that a reasonable investor might have considered them important in the making of this decision.” Every misrepresentation also includes a failure to disclose.

For example, again focusing on Tellabs, management stated that the Titan 5500 was experiencing continued growth and that the Titan 6500 was ready to ship. These were the misrepresentations. But management also failed to disclose that demand for the 5500 was precipitously declining and that the 6500 was not yet ready to ship. These were material facts, and the failure to disclose these material facts was fraudulent. Courts today often fail to recognize that the antifraud provision embodied in Rule 10b-5 forbids not only misrepresentations (lies), but also failures to disclose facts to ensure that facts stated are not misleading (half-truths).

The Basic Court clearly recognized that there needed to be a different approach with regard to reliance in transactions over “modern securities markets, literally involving millions of shares changing hands daily,” as contrasted with face-to-face transactions. In this regard, the Court quoted the Ninth Circuit’s decision in Blackie v. Barrack, to the effect that the “causal nexus can be adequately established indirectly, by proof of materiality coupled with the common sense that a stock purchaser does not ordinarily seek to purchase a loss in the form of artificially inflated stock.” The Court in effect recognized that requiring proof of individualized reliance would effectively destroy the class action as a vehicle to remedy securities fraud.

The Basic Court also cited Blackie in support of its assertion that the market informs the investor by impounding information into the value of the stock. There is language in Blackie to support that position: “Nevertheless, [the investor] relies generally on the supposition that the market price is validly set and that no unsuspected manipulation has artificially inflated the price, and thus indirectly on the truth of the representations underlying the stock price whether he is aware of it or not . . . .” But

54. 524 F.2d 891 (9th Cir. 1975).
55. Basic, 485 U.S. at 245 (quoting Blackie, 524 F.2d at 908) (internal quotation marks omitted).
56. Blackie, 524 F.2d at 907.
both before and after the above quoted language, the Blackie court used the Affiliated Ute Citizens approach to causation. The court first stated: “We think causation is adequately established in the impersonal stock exchange context by proof of purchase and of the materiality of misrepresentations, without direct proof of reliance.”57 Later, the Blackie court stated:

Here, the requirement [the burden of individualized reliance] is redundant[:] the same causal nexus can be adequately established indirectly, by proof of materiality coupled with the common sense that a stock purchaser does not ordinarily seek to purchase a loss in the form of artificially inflated stock. Under those circumstances we think it appropriate to eliminate the burden [of individualized reliance].58

Consequently, the Basic Court could have determined commonality under the Affiliated Ute Citizens approach by determining that investors rely upon the materiality of the misdisclosed facts. When management speaks, investors are entitled to rely on the truthfulness of the representations. When those representations are material and false, the class action remedy should be appropriate.

When the Court in Basic spoke of “integrity” in the market, it was referring to the integrity of information, rather than of the stock price: “Congress expressly relied on the premise that securities markets are affected by information, and enacted legislation to facilitate an investor’s reliance on the integrity of those markets . . . .”59 While the legislative history upon which the Court relied also referred to market price, the key phrase in the cited reports is that “the hiding and secreting of important information obstructs the operation of the markets as indices of real value.”60 As the Court observed, “[w]ho would knowingly roll the dice in a crooked crap game?”61

From a policy standpoint, what does the foregoing reflect? The Court was clearly positively inclined toward the use of the class action as a vehicle to remedy securities fraud in the public markets when it adopted the fraud on the market theory. Contrariwise, the dissent thought that the majority opinion was a bit of a reach.62 The Court also recognized the importance

57. Id. at 906.
58. Id. at 908 (footnote omitted).
59. Basic, 485 U.S. at 246.
60. Id. (quoting H.R. REP. NO. 1383, at 11 (1934)) (internal quotation marks omitted).
62. See id. at 250–53 (White, J., concurring in part and dissenting in part). The dissent stated, “[y]et today, the Court embraces this theory with the sweeping confidence usually reserved for more mature legal doctrines.” Id. at 250–51. “But with no staff economists, no experts schooled in the ‘efficient-capital-market hypothesis,’ no ability to test the validity of empirical market studies, we are not well
of accurate information to inform the market. Unfortunately, the Court got the cart before the horse. What investors rely upon is the integrity of the information in the market. The information then impacts the price of the shares. But the information is the cause, and the price is the effect. In focusing upon price, the fraud on the market theory is subject to the criticism that investors do not believe that the stock price is "true": investors buying stock believe that the price is too low, whereas investors selling stock believe that the price is too high.

Justice White, writing for the dissenters, challenged the assertion from the cases relied upon by the majority that investors "rely on the price of a stock as a reflection of its value." According to the dissent, "many investors purchase or sell stock because they believe the price inaccurately reflects the corporation’s worth." Moreover, according to the dissent, "[i]f investors really believed that stock prices reflected a stock’s ‘value,’ many sellers would never sell, and many buyers never buy (given the time and cost associated with executing a stock transaction)." It is hard to quibble with the foregoing assertions since, if you asked a buyer why she purchased, she might well respond “because the stock was undervalued;” whereas, if you asked a seller of the same stock why he sold, he might well respond that “the stock was overvalued.”

However, the Basic dissent did recognize that the congressional policy reflected in the securities laws called for "widespread public disclosure and distribution to investors of material information concerning securities." This policy “is expressed in the numerous and varied disclosure requirements found in the federal securities law scheme." In addition to the affirmative disclosure obligations under the securities laws, Congress also equipped to embrace novel constructions of a statute based on contemporary microeconomic theory.”

63. Id. at 255 (quoting id. at 244 (majority opinion); Peil v. Speiser, 806 F.2d 1154, 1161 (3d Cir. 1986)) (internal quotation marks omitted).

64. Id. at 256 (quoting Barbara Black, Fraud on the Market: A Criticism of Dispensing with Reliance Requirements in Certain Open Market Transactions, 62 N.C. L. Rev. 435, 455 (1984)) (internal quotation marks omitted).

65. Id.


67. Basic, 485 U.S. at 258 (White, J., concurring in part and dissenting in part).

68. Id. at 258–59 (citing 15 U.S.C. §§ 78m, 78o(d)).
enacted a catchall anti-fraud prohibition, pursuant to which Rule 10b-5 was promulgated.

B. Upon What Do Investors Actually Rely?

If investors do not rely upon price, then upon what do they actually rely? As Justice White recognized, Congress has created an elaborate system of disclosure in order that investors be informed. Are not investors entitled to rely on the presumption that corporations and their management comply with the securities laws and make honest, accurate, and truthful disclosures? When people buy drugs, are they not entitled to rely on the fact that the drugs meet FDA requirements? When they purchase meat, are they not entitled to rely on the fact that the meat meets U.S. Department of Agriculture labeling standards? As noted earlier, “[w]ho would knowingly roll the dice in a crooked crap game?”69 Congress clearly was concerned with “fair and honest markets,”70 and with the “protection of investors.”71

Unfortunately, the Rehnquist and Roberts Courts have become not just conservative, but reactionary in their approach to enforcement of the securities laws.72 According to the dissent in Basic, “[i]nvestors act on inevitably incomplete or inaccurate information, [consequently] there are always winners and losers; but those who have ‘lost’ have not necessarily been defrauded.”73 This cavalier approach to disclosure has had untoward consequences. The Dirks v. SEC74 decision, which was the source of this quote, was a disaster, precipitating a dramatic increase in insider trading.75

It is true that, in the securities world, there are winners and losers. Supposedly, the market reflects the present worth of a company’s stream of income out into the future. Some people are more adroit than others at analyzing the business plan of the company and the probability and magnitude of its stream of income. Thus, there are winners and losers. In

69. Id. at 247 (majority opinion) (quoting Schlanger v. Four-Phase Sys. Inc., 555 F. Supp. 535, 538 (S.D.N.Y. 1982)) (internal quotation marks omitted).
71. Id. § 78l.
72. See generally Murdock, supra note 10.
73. Basic, 485 U.S. at 256 (White, J., concurring in part and dissenting in part) (second alteration in original) (quoting Dirks v. SEC, 463 U.S. 646, 667 n.27 (1983)) (internal quotation marks omitted).
75. See H.R. R EP. NO. 98-355, at 5 (1984) [hereinafter 1984 ITSA Report], reprinted in 1984 U.S.C.C.A.N. 2274, 2278 (“The current commission has made the prosecution of insider trading a priority, and has brought more such cases during the past four years than in all previous years combined.”); see also JAMES B. STEW-ART, DEN OF THIEVES (1992). Dirks is discussed at length in Murdock, supra note 10, at 378 n.50. The Dirks majority determined that a tippee was not liable for his insider trading unless the tipper had sinned—breached a common-law fiduciary duty—and that the tipper did not sin unless he received a pecuniary or a reputational benefit. See Dirks, 463 U.S. at 653–64.
addition, the price of a company’s stock may be affected by unexpected and unanticipated events, such as a typhoon or tsunami affecting the available supply of materials that the company needs, or a financial meltdown caused by a real estate bubble. Or, a company’s management may turn out to be incompetent. These are all risks to which all investors are subject. But investors should not expect to bear the risk of securities fraud, which is why the securities laws were enacted to prevent it.

Thus, investors do rely upon the integrity of the information in the securities markets about a company’s operations. When management misrepresents the company’s operations and potential earnings, and investors buy at an inflated price, they should be able to bring a class action lawsuit to recover the difference between the amount they paid—which was inflated by management’s misrepresentations—and the value at the time of purchase of the securities—which often may be calculated by reference to the drop in price when the true state of facts becomes public.

Consider once again the Tellabs situation. In 2000, Tellabs was a highflying tech company with a major product, the Titan 5500, and an upgraded version, the Titan 6500, which was supposedly ready to ship. The stock was a growth stock, reaching a high of $67 per share and selling at almost forty times earnings. To maintain such a price, continued growth was essential.


77. After the financial meltdown, real estate lending dried up, thereby adversely affecting many businesses.

78. Consider the bailout of General Motors resulting from decades of mismanagement; also consider the necessity to bail out major banks in the United States as a result of their improvident investing in toxic securities that led to the financial meltdown. More recently, after the London Whale debacle, in which J.P. Morgan lost about $6 billion, the stock of J.P. Morgan dropped as much as 24% in a month.

79. See Anna Marie Kukec, Naperville’s Tellabs to Be Acquired for $891 Million, Daily Herald (Oct. 22, 2013, 7:36 AM), http://www.dailyherald.com/article/20131021/business/710219902. The stock also reached a high of $67. See Makor Issues & Rights, Ltd. v. Tellabs, Inc., 437 F.3d 588, 593 (7th Cir. 2006). Thus, the price-earnings ratio was almost forty.


81. When stock sells at a high price/earnings ratio, the value of the stock is heavily dependent upon the anticipated growth rate:

\[
CR = DR - G
\]
From December 2000 through April 2001, the CEO made, or caused to be made, a series of positive statements about the company that were designed to assure investors that the existing growth would continue. On December 11, 2000, a press release was issued to the effect that the 5500 was experiencing continuing growth and that the 6500 was now available.82 On January 23, 2001, there was another press release and an investor conference call in which “robust growth” was touted and the fact that “customers are buying more and more Tellabs equipment” was asserted.83 On February 14, 2001, a letter was sent to Tellabs’ shareholders toting “robust growth,” asserting that sales of the 5500 had soared 56% in 2000, and stating that customers were embracing the 6500.84 The accompanying annual report stated that the 5500 was going strong. In March, there was another investor telephone call at which it was stated that the 5500 was maintaining its growth rate85 and that demand for the 6500 continued to grow.86

In March, the first quarter revenue projection was reduced from a range of $865–$890 million to $830–$865 million.87 On April 6, the first quarter revenue projection was reduced to $772 million, and the price of the stock dropped to about $30 a share.88 It gradually moved up to $38 per share, but, on June 19, 2001, Tellabs announced a substantial reduction in its second quarter projections—from $780–$820 million to only $500 million—and, at another investor conference call, the CEO announced that the reduction “was almost entirely because of an enormous reduction in TITAN 5500 sales;” the price plummeted to about $16 per share.89 Tellabs had been engaged in channel stuffing,90 and the channel stuffing was so extensive that, “[a]ccording to the plaintiffs’ confidential sources, Tellabs had to lease extra storage space in January and February 2001 to accommodate the large number of returns.”91 Tellabs manage-

Therefore, the larger the growth rate, the lower the capitalization rate. The capitalization rate is inversely proportional to the price-earnings multiple. Thus, the larger the growth rate, the higher the price-earnings multiple. Since a rudimentary approach to the value of a company is earnings times the price-earnings multiple, a decrease in earnings (and, consequently, growth) is a double whammy: it reduces the earnings, and the reduced earnings lower growth and thereby increase the capitalization rate, consequently reducing the price-earnings ratio.

82. *Tellabs*, 437 F.3d at 592.
83. *Id.*
84. *Id.*
85. *Id.* at 592–93.
86. *Id.* at 592.
87. *Id.*
88. *Id.* at 593.
89. *Id.*
90. The practice of channel stuffing, and its impact upon earnings and stock price, is discussed extensively in Murdock, *supra* note 11, at 625–32.
91. *Tellabs*, 437 F.3d at 598.
ment clearly knew that sales of the 5500 were in the process of dropping dramatically.\textsuperscript{92}

Justice Thomas, in his concurring opinion in \textit{Halliburton II}, asserted that courts should employ “logic” and observe “economic realities.”\textsuperscript{93} Then let us look at the logic and economic realities of the Tellabs situation. First of all, the information in the marketplace in the first quarter of 2001 was that Tellabs was continuing as a high-tech growth company. Someone buying stock in Tellabs would have that view of the company, a view that management sought to reinforce by its comments from December 2000 to March 2001.

That is the factual milieu upon which anyone purchasing Tellabs’ stock would operate. That is the factual milieu which would determine the price of the Tellabs stock. Investors\textsuperscript{94} universally rely upon the business plan of the company and the cash flow generated by its operations pursuant to that business plan; the stock price then follows. Information about a company’s business plan and the results of its operations are disclosed to the market through the securities laws’ mandatory disclosure system and management’s discussions and releases complementing this disclosure system. Thus, investors rely upon the total mix of information in the marketplace. This mix of information presumptively is the same for all investors. Some investors may bring more skilled analysis to this information than other investors. But the mix of information upon which they are operating is the same, thus the reliance element in Rule 10b-5 is one that is common to all investors.

Some investors may have been directly aware of some of this information, for example, those who sat in on the investor conference call. Others may have been aware of this mix of information indirectly, through re-

\textsuperscript{92} The Seventh Circuit described the situation as follows: By January 2001, the complaint asserts, demand for Tellabs’s “best seller”—the TITAN 5500—was drying up. Verizon, Tellabs’s largest customer, reduced its orders for the TITAN 5500 by roughly 25% in late 2000 and by roughly 50% in January 2001. Customers in Latin America and Central America were no longer buying the product. By late 2000, according to a couple of the confidential sources, Tellabs had excess TITAN 5500s on hand because of the lack of demand. One confidential source informed the plaintiffs that Tellabs paid Probe Research, an outside company, $100,000 to forecast demand for the TITAN 5500. Completed “in or about early 2001,” the report showed that the market need for the TITAN 5500 was evaporating. Based on this research, Tellabs’s marketing strategy department distributed an internal memorandum that concluded that revenue from the TITAN 5500 would decline by about $400 million. \textit{Id.} at 597.

\textsuperscript{93} \textit{Halliburton II}, 134 S. Ct. 2398, 2418 (2014) (Thomas, J., concurring).

\textsuperscript{94} It is important to draw a distinction between investors and traders. Traders, particularly high-speed traders, make their money on market volatility. Thus, they are only indirectly concerned with the company’s business plan and operations. \textit{See generally }MICHAEL LEWIS, \textit{Flash Boys} (2014); SCOTT PATTERSON, \textit{Dark Pools} (2013).
ports in the media or recommendations of their stockbrokers. But all are operating on the same mix of information. If we were to require individual reliance, some institutional investors who were following Tellabs stock could bring suit directly and might have a sufficient economic stake to bear the risk of litigation.

But no small investor would have a sufficient economic stake to bring an individual cause of action. I personally bought 1000 shares of Tellabs stock around this time. But even if I were to lose $30 a share, from the standpoint of time, it would not be worth it to me to file suit on my own behalf. A non-attorney investor is in an even worse place. She must engage a lawyer sufficiently sophisticated in securities litigation to represent her at $500 to $400 per hour or even higher. The attorney must conduct a substantial due diligence investigation prior to bringing suit or be subject to sanctions pursuant to the PSLRA. The complaint must be drafted with great detail (particularity) pursuant to the PSLRA, in order to defeat a motion to dismiss, and there is no discovery prior to the hearing on the motion to dismiss. Thus, an individual investor could expect the need to front a substantial amount of money to bring suit, which could end up exceeding any possible recovery. And no attorney would take such a suit on a contingency basis. There simply is not a sufficient economic stake to warrant the time, cost, and risk of litigation. Thus, were the courts to rule out class action status, the net result would be to leave the small investor without a remedy. Should the securities laws be premised on the basis that only some institutional investors are entitled to the protection of the law?

Consequently, Basic appropriately recognized that, in these situations, reliance must be based on that which provides commonality to all investors. The error in Basic was in adopting a doctrine based on an economic theory that an individual investor could not beat the market. This transferred the reliance element from a focus on the information to a focus on the price of the stock that flows from the information. Federal courts should leave economic theory to economists and instead deal with that with which the federal courts have dealt for decades: the disclosure obligations of a company and the fraudulent misrepresentations by those in control of a company that are aimed at defrauding the market in general. Therefore, courts should recognize that all investors rely upon the total mix of information in the market, most of which—practically all the misleading information—is introduced to the market by corporate management.

96. See In re PolyMedica Corp. Sec. Lit., 432 F.3d 1, 8 (1st Cir. 2005) (“The efficient market hypothesis began as an academic attempt to answer the following question: Can an ordinary investor beat the stock market, that is, can such an investor make trading profits on the basis of new information? In an efficient market, the answer is ‘no,’ because the information that would have given the investor a competitive edge and allowed the investor to ‘beat’ the market is already reflected in the market price.” (citing Lynn A. Stout, The Mechanisms of Market Inefficiency: An Introduction to the New Finance, 28 J. Corp. L. 635, 639 (2005))).
As a caveat to this approach, it should be recognized that a particular statute, section 18 of the 1934 Act, provides a cause of action for a misleading statement in any document filed with the SEC. With respect to this cause of action, courts have held that individualized reliance must be alleged and proven. Thus, supposedly by analogy, since the mix of information in the marketplace includes documents filed with the SEC, it could be argued that individualized reliance, and not a presumption of common reliance, is necessary when dealing with misleading corporate filings.

However, two factors militate against importing the section 18 approach to reliance into a Rule 10b-5 class action case. First of all, the mix of information in the marketplace includes not just documents filed with the SEC but also much additional material. As illustrated by the previous discussion with respect to Tellabs, misleading information produced by corporate management may be promulgated through press releases, letters to shareholders, or investor conference calls. Thus, a cause of action under Rule 10b-5 conceivably could rely exclusively upon statements made outside of documentation filed with the SEC.

More importantly, even if some of the misleading statements are included in documents filed with the SEC, a cause of action under Rule 10b-5 differs substantially from a cause of action under section 18. Under section 18, courts have held that plaintiffs need not allege scienter, whereas under Rule 10b-5, scienter must be pleaded with particularity.

III. The Halliburton Flaw: Retaining the Focus on Market Efficiency

A. The Halliburton II Decision

Halliburton II differs markedly from Halliburton I. The Halliburton I decision was short, was adopted unanimously, and focused upon only one issue: whether plaintiffs must establish loss causation at the certification stage. The Halliburton II decision was lengthy, produced various

98. See, e.g., In re Suprema Specialties, Inc. Sec. Litig., 438 F.3d 256, 283 (3d Cir. 2006) (“SSF Plaintiffs concede that to state their claim under Section 18, they were required to plead actual, as opposed to presumed, reliance upon a false or misleading statement.”); Howard v. Everex Sys., Inc., 228 F.3d 1057, 1063 (9th Cir. 2000) (“[C]ourts have required a purchaser’s actual reliance on the fraudulent statement under § 18(a), as opposed to the constructive reliance, or fraud-on-the-market, theory available under § 10(b).”); see also Brief on Behalf of Appellants at 44, In re Suprema Specialties., 438 F. 3d 256 (No. 04-3755), 2004 WL 5327687 (citing Heit v. Weitzen, 402 F.2d 909, 916 (2d Cir. 1968)).
99. See, e.g., In re Suprema Specialties, 438 F.3d at 283 (“A Section 18 plaintiff, however, bears no burden of proving that the defendant acted with scienter or any particular state of mind.”).
103. The Court determined that plaintiffs need not. See id. at 2184–87.
opinions, though only one judgment, and dealt with three issues: (i) whether Basic should be overruled, (ii) whether plaintiffs should be required to establish price impact at the certification stage, and (iii) whether defendants should have the opportunity to establish lack of price impact at the certification stage.\textsuperscript{104}

With respect to the first issue, the Court in \textit{Halliburton II} found no policy reason to overrule Basic. However, the Court did engage in some convoluted analysis to justify the “reliance on price” aspect of Basic and, unfortunately, retained “market efficiency” as one of the four prerequisites for invoking the presumption of reliance.\textsuperscript{105} The other three prerequisites are publicity, materiality, and market timing.\textsuperscript{106} Plaintiffs have the burden of proving these prerequisites, and, with the exception of materiality,\textsuperscript{107} they must be satisfied before class certification.\textsuperscript{108} As will be addressed below, the Court’s own logic supports the argument that what plaintiffs must establish is not market efficiency, but rather price impact.\textsuperscript{109}

The Court also rejected Halliburton’s argument that plaintiffs should be required to prove that “defendant’s misrepresentation actually affected the stock price—so-called ‘price impact’—in order to invoke the Basic presumption.”\textsuperscript{110} According to the Court:

What is called the Basic presumption actually incorporates two constituent presumptions: First, if a plaintiff shows that the defendant’s misrepresentation was public and material and that the stock traded in a generally efficient market, he is entitled to a presumption that the misrepresentation affected the stock price. Second, if the plaintiff also shows that he purchased the stock at the market price during the relevant period, he is entitled to a further presumption that he purchased the stock in reliance on the defendant’s misrepresentation.\textsuperscript{111}

According to the Court, “[b]y requiring plaintiffs to prove price impact directly, Halliburton’s proposal would take away the first constituent presumption.”\textsuperscript{112} The Court concluded that “[f]or the same reasons we declined to completely jettison the Basic presumption, we decline to effectively jettison half of it by revising the prerequisites for invoking it.”\textsuperscript{113}

\begin{itemize}
  \item \textsuperscript{104} See \textit{Halliburton II}, 134 S. Ct. at 2405.
  \item \textsuperscript{105} See \textit{id.} at 2416, 2424.
  \item \textsuperscript{106} See \textit{id.} at 2412.
  \item \textsuperscript{107} In \textit{Amgen}, discussed \textit{infra} at notes 222–31 and accompanying text, the Court, in a conflictual opinion, determined that plaintiffs need not establish materiality at the certification stage.
  \item \textsuperscript{108} See \textit{Halliburton II}, 134 S. Ct. at 2412.
  \item \textsuperscript{109} See \textit{infra} notes 110–14 and accompanying text.
  \item \textsuperscript{110} \textit{Halliburton II}, 134 S. Ct. at 2413
  \item \textsuperscript{111} \textit{id.} at 2414.
  \item \textsuperscript{112} \textit{id.}
  \item \textsuperscript{113} \textit{id.}
\end{itemize}
What the Court did determine was that the defendant could introduce evidence to establish a lack of price impact in the particular case so as to rebut the presumption of reliance. The Court held that “defendants must be afforded an opportunity before class certification to defeat the presumption through evidence that an alleged misrepresentation did not actually affect the market price of the stock.”

This is basically gobbledygook. The Court acts as though it is doing plaintiffs a favor by rejecting the requirement for plaintiffs to show “price impact” specifically, but the Court retains the requirement to show that the market was efficient—a requirement which necessitates plaintiffs establishing price impact generally, rather than specifically. Additionally, by permitting defendants to prove the lack of price impact, the Court indirectly requires plaintiffs to litigate price impact at the certification stage. Plaintiffs would have been better off if the Court had required plaintiffs to prove price impact at the certification stage but had jettisoned the requirement to establish “market efficiency,” a concept which is not just amorphous and disputed, but which also adds substantially to the cost of litigation by generating a battle of the experts on a generalized proposition.

Litigation would be much more efficient if the role of experts were limited to determining whether, in this particular case, there was a price impact from the misrepresentation. And, determining price impact is basically the same concept as loss causation, which the Court determined plaintiffs need not establish at the certification stage. Thus, as discussed in the analysis of Basic, all plaintiffs should need to show at the certification stage is that the defendants introduced misrepresentations into the marketplace which altered the mix of reliable information upon which investors rely. But this is basically the definition of materiality, which the Court has determined need not be established until the trial stage. Thus, logically, price impact, materiality, and loss causation are related concepts that should be litigated at a trial on the merits.

B. *The Court’s Hypothetical—A Two-Way Street*

In holding that defendants must be afforded the opportunity to rebut the presumption of reliance by establishing that the misrepresentation did not affect the market price, the Court used this hypothetical:

Suppose a defendant at the certification stage submits an event study looking at the impact on the price of its stock from six discrete events, in an effort to refute the plaintiffs’ claim of general market efficiency. All agree the defendant may do this. Suppose one of the six events is the specific misrepresentation asserted by the plaintiffs. All agree that this too is perfectly acceptable. Now suppose the district court determines that, despite the defen-
In this hypothetical situation, the market for the stock is generally efficient, which would normally suffice to establish the fraud on the market presumption. But the particular misrepresentation that is being litigated did not affect the price of the stock. The Court thought that it would be “bizarre”\(^{116}\) to enable the plaintiff to get past the certification stage by virtue of indirect evidence when direct evidence was available to show the plaintiff was not injured.\(^{117}\) The Court concluded: “While \textit{Basic} allows plaintiffs to establish that precondition indirectly, it does not require courts to ignore a defendant’s direct, more salient evidence showing that the alleged misrepresentation did not actually affect the stock’s market price and, consequently, that the \textit{Basic} presumption does not apply.”\(^{118}\)

The Court’s position would seem reasonable: if the price were not affected by the alleged misrepresentation, then either the misrepresentation was not material or the plaintiff had not suffered any loss. Why proceed to trial?

On the other hand, the converse of the Court’s hypothetical can demonstrate why the fraud on the market theory is flawed as a result of its focus upon market efficiency. Suppose, as the Supreme Court did, that there was a situation in which there were six event studies, one of which was the specific representation in the instant suit. Assume further that five of the event studies showed no price impact, but that the specific representation in the instant suit did show a marked impact on price. A defendant could make the case that the market was not efficient, and, therefore, the fraud on the market presumption was not applicable. Consequently, class certification would fail, and, most likely, the suit would die, particularly if plaintiffs were only small investors who could not afford to bring a direct action. This would be so even though the plaintiffs established that the corrective disclosure caused the market price to drop dramatically, thereby leaving plaintiffs with worthless stock. Does this make sense?

A much more sensible regime would be to have a presumption—which is not really a presumption but rather a reality—that investors rely upon the total mix of information in the market; when this mix is fraudulently created by the misrepresentations of management, affected inves-
tors can bring a class action lawsuit. This should suffice for class certification. This is what should be the normative position for federal courts to take.

But, if more be needed, then require plaintiffs to establish “price impact.” Under this variation, plaintiffs’ burden would be less than exists at present to establish that the misrepresentation took place in an efficient market, since plaintiffs need deal only with one event-study—the current misrepresentation—rather than a multitude of event studies. Moreover, plaintiffs would not be assaulted by the various views as to whether all information must be incorporated into the price of the stock, or whether merely most publicly available information must be so incorporated, or how quickly the information must be incorporated into the price.

But evidence establishing price impact would probably also establish materiality and loss causation. In effect, we would be moving the trial on the merits stage to the class certification stage. It is better to presume reliance based upon the defendant’s insertion of misleading information into the market, grant class certification, and leave any factual issues for trial. Otherwise, as Professor Langevoort has stated:

A detailed resolution of market-impact and loss-causation issues at the class-certification stage substitutes fact finding by the judge on an extraordinarily complex issue at an early stage of the proceeding for fact finding by the jury at trial and gives appellate courts far greater authority to review those findings.119

IV. RETAINING A FOCUS ON MARKET EFFICIENCY IS UNREASONABLE, OUTDATED, AND IS MERELY A DEVICE TO INCREASE LITIGATION COSTS AND THWART MERITORIOUS LITIGATION

The converse of Justice Roberts’s hypothetical, discussed above, illustrates one reason why requiring market efficiency as a condition for reliance is irrational. What difference does it make whether or not other information is integrated into the market price if the plaintiff can demonstrate that the misrepresentation of which the plaintiff complained did affect market price? A market could be fairly inefficient but, if the misrepresentation were significant enough, it could affect market price and adversely impact plaintiffs, thereby causing their loss.

Moreover, as Professors Prichard and Henderson assert in their brief to the Supreme Court in the Halliburton II case:

[P]roving the efficiency of the market as a whole is only an indirect means of proving that the market relied on a particular statement. . . . [D]etermining whether a misstatement distorted the market is typically easier than demonstrating efficiency of the market as a whole. It is also a more direct means of inquiring

into reliance, and a more reliable method of showing whether the complained-of fraud was, in fact, a “fraud on the market.”

A. PolyMedica: A Leading Case Exemplifying the Unreasonable Application of Market Efficiency

1. The PolyMedica Facts

The stock of PolyMedica, a distributor of diabetes testing supplies, traded as high as $58.25 during the quarter ending December 31, 2000.121 Apparently because of short selling pressure,122 the high in the following quarter was $44.123 Toward the end of that quarter, CIBC World Markets announced that there was an investigation into PolyMedica for Medicare fraud. The stock plummeted from $33 to $17.124 The next business day, March 26, 2001, PolyMedica held an investor conference call, reassuring investors and extolling its efficient systems and highly trained and dedicated personnel.125 A company official referenced a two-year-old complaint to Health and Human Services and asserted that, in view of the 60,000 to 90,000 shipments PolyMedica made per month, there will always be some customer complaints.126 Thereafter, the stock rose to $48 in the quarter ending September 30, 2001.127 However, in late July of that quarter, the New York Stock Exchange declined to list PolyMedica stock, and, on August 6th, Barron’s disclosed that a federal grand jury was looking into possible Medicare and investor fraud.128 The stock dropped to $22 after the Barron’s article, and it further dropped to $14 when, a couple of days later, PolyMedica finally admitted that a criminal investigation was underway.129 The stock eventually dropped to $11.25.130

123. See PolyMedica Annual Report, supra note 121, at 9.
124. See PolyMedica Plunges, supra note 122; see also PolyMedica Annual Report, supra note 121, at 9.
125. See PolicyMedica Plunges, supra note 122.
126. See id.
130. See PolyMedica Annual Report, supra note 121.
According to PolyMedica’s employees, the company often shipped products to customers that the customers did not order, and, even though these products were often returned, “[o]ver 90% of returns are not refunded to Medicare.” At the end of one month on a Friday, the company was 10,000 orders short of their goal; magically, the 10,000 orders appeared by the close of the day on Saturday. The company would also load up trucks from its warehouses and then have the trucks sit in the warehouse parking lot because there were no orders. When employees tried to inform management of improper activity, they were “met with silence or reprimand.” PolyMedica eventually settled the criminal charges with the Justice Department by paying $35 million.

One would think that the foregoing situation would be a matter of concern for the federal courts. Medicare fraud is a serious matter. Consequently, it will be instructive to contrast the approach of the district court in its initial opinion granting certification, with the circuit court opinion reversing on the basis of its definition of an efficient market, and the district court opinion on remand, applying the circuit court’s view on market efficiency.

2. The Initial District Court Opinion: A Sensible Approach Applying Basic

Judge Keeton’s opinion in In re PolyMedica Corp. Securities Litigation, which granted class certification, is a joy to read: it is clear, concise, and cogent. In applying the fraud on the market theory, he relied upon the Supreme Court’s opinion in Basic, rather than upon academics arguing about the concept of an efficient market and courts relying upon the academics. He essentially adopted the plaintiffs’ assertion that “the essential factor in determining market efficiency is whether the stock price was affected by the material information available in the market.”

Looking to Basic, Judge Keeton relied upon three statements by the Basic Court:

(i) The presumption is support [sic] by common sense and probability . . . . Because most publicly available information is reflected in market price, an investor’s reliance . . . may be presumed . . . .

131. Einhorn, supra note 128.
132. Id.
133. Id.
134. Id.
137. See id. at 39.
138. Id. at 40.
139. Id. (first alteration in original) (quoting Basic Inc. v. Levinson, 485 U.S. 224, 247 (1988)).
The Need for a New Paradigm

(ii) We need not determine by adjudication what economists and social scientists have debated. . . . We need only believe that market professionals generally consider most publicly announced material statements about companies, thereby affecting stock market prices.140

(iii) We do not intend conclusively to adopt any particular theory of how quickly and completely publicly available information is reflected in market price.141

Based upon the foregoing, he concluded:

Considering the three statements together, I conclude that an “efficient” market in the context of the “fraud on the market” theory is not one in which a stock price rapidly reflects all publicly available material information. Rather, the “efficient” market required for “fraud on the market” presumption of reliance is simply one in which “market professionals generally consider most publicly announced material statements about companies, thereby affecting stock market prices.”142

Applying the above standard to the facts presented to him, Judge Keeton noted the following:

(i) There was a cause-and-effect relationship between unexpected disclosures regarding the corporation and an immediate response in the stock price: “In particular, at five points in 2001 . . ., PolyMedica stock price rose [dramatically] on news of greater than expected growth and fell [dramatically] on negative news about the Company.”143

(ii) The average weekly trading volume exceeded one million shares, which was around 10% of the shares outstanding.144

(iii) “[E]ighteen different securities analysts followed PolyMedica” and “at least four securities analysts issued at least seventeen analyst reports.”145

(iv) There were “at least 283 market makers [ ] in the market for PolyMedica common stock . . . .”146

140. Id. (quoting Basic, 485 U.S. at 246 n.24).
141. Id. at 40–41 (quoting Basic, 485 U.S. at 249 n.9).
142. Id. at 41 (quoting Basic, 485 U.S. at 246 n.24).
143. Id. at 43 (second and third alterations in original) (internal quotation marks omitted).
144. See id.
145. Id.
146. Id.
(v) The market capitalization exceeded $200 million.147

From the foregoing, Judge Keeton, in the opinion of the Author, reasonably concluded that the market was efficient and that the fraud on the market presumption applied.148

3. *The Circuit Court and the District Court on Remand: An Example of Courts in an Alternate Universe*

a. The Circuit Court: A Stringent Definition of Market Efficiency

At the outset of the First Circuit’s opinion in *PolyMedica*,149 it acknowledged that, according to the plaintiff, PolyMedica’s stock had lost more than 80% of its value, and PolyMedica had engaged in welfare fraud.150 At the time of its opinion, in December 2005, these facts were easily verifiable. While judges are supposed to be bias free (an illusion since everyone’s decision-making is affected by bias151), would not the instinct of someone who was concerned about fraud in the securities markets be to certify the class to determine what the facts really were? This would be particularly so when presented with the clear, concise, and cogent analysis of Judge Keeton.

The circuit court recognized the *Basic* pronouncements relied upon by the district court but also focused upon other statements, such as:

“[T]he market is acting as the unpaid agent of the investor, informing him that given all the information available to it, the value of the stock is worth the market price,” and that “the market price of shares traded on well-developed markets reflects all publicly available information, and hence, any material misrepresentations.”152

The circuit court concluded that the Supreme Court’s language in *Basic* could support either the lower court’s interpretation of efficiency or the view adopted by most lower courts that the stock price must “fully reflect all publicly available information.”153

147. See id.

148. See id.

149. *In re PolyMedica Corp. Sec. Litig.*, 432 F.3d 1 (1st Cir. 2005).

150. See id. at 3.

151. See generally Daniel Kahneman, *Thinking, Fast and Slow* (2011). Kahneman is a Nobel Prize winner for his work in behavioral economics. His recent book discusses two systems of thinking: “thinking fast,” which operates automatically and quickly with little effort, and “thinking slow,” which is deliberative and analytical. The first system initially produces the information upon which the second system thoughtfully operates. That is how bias can creep into decisions which we believe are thoughtfully made. For a more extensive discussion, see Charles W. Murdock & Barry Sullivan, *What Kahneman Means for Lawyers: Some Reflections on Thinking, Fast and Slow*, 44 Loy. U. Chi. L.J. 1377 (2013).

152. *PolyMedica*, 432 F.3d at 11 (quoting Basic Inc. v. Levinson, 485 U.S. 244, 246 (1988)).

153. Id. at 12–13.
The problem with this formulation, as discussed in the next section, is that the “fully reflect all publicly available information” standard is impossible to prove, relies upon the conjecture of so-called experts, and is not realistic, particularly when coupled with the circuit court’s add-on that the information must be “incorporate[d] rapidly or promptly.”\textsuperscript{154} The court stated that, “[b]y ‘fully reflect,’ we mean that market price responds so quickly to new information that ordinary investors cannot make trading profits on the basis of such information. This is known as ‘informational efficiency.’”\textsuperscript{155} Although the court did not require “fundamental value efficiency,” which means that the market price is also accurate, it did recognize that fundamental value efficiency could be relevant to informational efficiency.\textsuperscript{156}

Again, as developed in the next section, it is clear that markets incorporate different types of information at different rates, depending upon the complexity and mode of distribution of the information.\textsuperscript{157}

b. The District Court on Remand: An Unreasonable Application of a Stringent Standard

(i) Plaintiffs’ Strong Evidence of Market Efficiency

On remand, the district court\textsuperscript{158} looked at the plaintiffs’ evidence, primarily focusing upon the five \textit{Cammer} factors, which many courts have used to determine whether the market for a particular stock is efficient.\textsuperscript{159} The court analyzed them as follows:

(1) Average trading volume: the average weekly trading volume was 4,140,232 shares, which accounted for 31\% of the 13,280,000 total shares outstanding. \textit{Cammer} suggested that a 2\% weekly trading volume “warranted a strong presumption of market efficiency.”\textsuperscript{160}

(2) Number of securities analysts: seven analysts followed the stock, and there were 348 articles mentioning

\textsuperscript{154. Id. at 12.}
\textsuperscript{155. Id. at 19.}
\textsuperscript{156. See id.}
\textsuperscript{157. See id.}
\textsuperscript{158. In re Polymedica Corp. Sec. Litig., 453 F. Supp. 2d 260 (D. Mass. 2006). In the interim, Judge Keeton had retired, and the case was reassigned to Judge Young. See id. at 264.}
\textsuperscript{159. See Cammer v. Bloom, 711 F. Supp. 1264, 1286–87 (D.N.J. 1989). The factors are: (1) the stock’s average trading volume; (2) the number of securities analysts that followed and reported on the stock; (3) the presence of market makers and arbitrageurs; (4) the company’s eligibility to file a Form S–3 Registration Statement; and (5) a cause-and-effect relationship, over time, between unexpected corporate events or financial releases and an immediate response in stock price. See id.}
\textsuperscript{160. See PolyMedica, 453 F. Supp. 2d at 267.}
PolyMedica.161 A previous decision in the First Circuit had approved efficiency even though there was only one analyst.162 The court said that the strength of this factor “is uncertain because there exists no coherent yardstick against which to measure it,” even though the court also recognized that, according to one study, this factor “is one of only two which actually have statistically significant, empirical support.”163

(3) Presence of market makers: there were 193 market makers, of whom 27 traded over one million shares each.164 Previous First Circuit decisions had approved class certification with substantially less market-making activity.165 However, the court stated that, since there is “no accepted standard” to determine the minimum number, the court would place little weight on this factor.166

(4) Eligibility to file an S-3 registration statement: in order to use this form, the SEC requires a market capitalization of $75 million, not counting stock held by management.167 The rationale behind this requirement is that, at this level of market capitalization, there will be sufficient analyst and investor interest to ensure that information is widely disseminated and integrated. The court acknowledged that “[c]ourts have found that the SEC permits an S–3 Registration statement ‘only on the premise that the stock is already traded on an open and efficient market, such that further disclosure is unnecessary.’”168 The market capitalization of PolyMedica during the class period was about $500 million.

(5) Cause and effect relationship: plaintiffs’ expert, Miller, presented a chart which listed “the price change in PolyMedica stock on [ ] five days, each of which had significant news events:

- Reports of consumer complaints to government investigators (Mar. 23, 2001: 49.54% decline);

161. See id.
162. See In re Xcelera.com Sec. Litig., 430 F.3d 503, 514–15 (1st Cir. 2005).
164. See id. at 268.
165. See, e.g., Xcelera.com, 430 F.3d at 516 (approving class certification where there were only twenty market makers and seven that traded over one million shares).
166. PolyMedica, 453 F. Supp. 2d at 268.
167. See id.
PolyMedica’s response that those reports were rumors and that it had not been contacted by any government agency (Mar. 26, 2001: 42.65% rise);

- Announcement that shares would no longer be listed on the NYSE (July 23, 2001: 29.52% decline);

- Report that PolyMedica may be indicted for Medicare and investor fraud (Aug. 6, 2001: 32.17% decline);

- PolyMedica announcement that the U.S. attorney for the Southern District of Florida was conducting an investigation into one of its units (Aug. 8, 2001: 17.65% decline).“169

Miller also presented “a side-by-side comparison of movements in the PolyMedica stock price, ‘peer group’ stock prices, and the NASDAQ index and [stated] that ‘the dramatic price increases and declines in the price of PolyMedica stock during the disputed period in response to new company-specific information were not mirrored in price movements of the NASDAQ Composite Index or the comparable company index.’”170

I would expect that most persons, reviewing the foregoing facts, would conclude that PolyMedica traded in an efficient market. To disregard or fail to see the significance of facts such as an average weekly trading volume of 31% of the outstanding shares, seven analysts following the stock, 193 market makers of whom 27 traded over one million shares each, and a market capitalization of about $500 million (whereas a market capitalization of only $75 million would be sufficient to be eligible to file an S-3 (short form) registration statement), suggests a lack of objectivity and an outcome determinative mentality from the court.

(ii) The District Court’s Unrealistic Reliance upon Defendant’s Expert

According to the district court, the foregoing analysis by plaintiffs’ expert “leaves much to be desired.”171 To support this conclusion, the court relied upon the following testimony of Dunbar, PolyMedica’s expert:

[Y]ou went and searched for the largest price drops. That’s not a scientific study. A scientific study is one where you draw a sample and then you compare a test statistic from that sample to another sample . . . . All you did was went and picked the largest stock price drops and said, oh, gee, that just shows that it’s informationally efficient. You picked five days out of about 160 trading days. What you should do is look at all 160 trading days and do a scientific study to see if there’s a difference between the news days and the non-news days. And if you would have done that

169. Id. at 269 (bulleted list as appears in original).
170. Id.
171. Id.
you would have found that there wasn’t any difference between them.

If you picked news days as a sample, all news days, not just the ones you self selected. I mean, you selected the few news days that would prove your point. But there’s many other news days in that contested period, anywhere from 23 to 59, versus [sic] on how you want to count them. If you want to look at what the stock price reaction is on those news days versus the non-news days, you’ll find that you can’t say that the news days were drawn from a different sample than the non-news days. In other words, they were providing as much information to the market as the non-news days.172

From the standpoint of rebutting plaintiffs’ evidence, it is the above opinion of defendant’s expert that “leaves much to be desired.” Dunbar indicated that there were between twenty-three and fifty-nine news days. This is a wide range of uncertainty, which also raises questions about the materiality of the information on these various days, and whether any particular news item was within the expectations of the market or, on the other hand, was a surprise. Moreover, was any of the information on the so-called news days as dramatic as the allegations of Medicare fraud, denial of listing by the New York Stock Exchange, or possible criminal indictment?

The following is the only stock graph during the relevant period that I have been able to access.173

![Stock Graph](image)

It shows a precipitous drop on March 26th, upon news of the government’s investigation into Medicare fraud. It shows a subsequent rise the

---

172. *Id.* at 269–70 (alterations and error in original).
following Monday, when PolyMedica’s investor conference call assured investors there were no problems. The graph further shows a sharp decline when it was announced that the New York Stock Exchange would not be listing PolyMedica shares. The graph ends before the announcement that PolyMedica might be indicted for Medicare fraud, when another substantial drop in price occurred. The extent to which the market moved on the 150 or so days when there was either no news or “less significant” news is irrelevant to the efficiency of the market with respect to the dramatic news reports about PolyMedica. As discussed in the next section, PolyMedica’s expert fell into the trap of assuming that all news is digested immediately by the market, without recognizing that the significance, complexity, and mode of disclosure all affect the pace at which the market processes news.

Nevertheless, the court endorsed the above criticism of Miller and stated with respect to the plaintiffs’ expert: “Miller’s mere listing of five days on which news was released and which exhibited large price fluctuations proves nothing. Miller’s only marginally useful analysis—which he oddly labels ‘not a significant factor’—is his unscientific comparison of PolyMedica Stock to the NASDAQ index.”

It is incomprehensible that a court could look at the price movements reflected in the above graph, and as set forth by plaintiffs’ expert, and conclude that they prove nothing. This is particularly true when the testimony is not in connection with a trial on the merits but rather relates to the issue of whether the plaintiff has presented sufficient evidence of commonality for class certification. Moreover, in point of fact, the analysis by plaintiffs’ expert is considered scientific. Comparing the price movements in the defendant company’s stock with price movements of peer stocks and an index such as NASDAQ is a scientific way of establishing that the price movement with respect to the defendant company was caused by factors over and above whatever factors impacted the movement of peer stocks or the market as a whole.

(iii) The District Court’s Unrealistic Reliance on the Lack of Shares Available for Short Selling

The district court was also influenced, in determining that the market for PolyMedica stock was not efficient, by the situation with regard to short selling. By April 2001, there was a huge amount of short interest; this made obtaining additional shares to short difficult. In addition, a typical loan fee for lending stocks was 1%; during the summer of 2001, however, the loan fee with regard to PolyMedica was between 15% and 35%. According to Dunbar, this fact meant the market was inefficient because persons who would have liked to have shorted the stock were not

175. See id. at 273–75.
176. See id. at 273.
177. See id. at 274.
able to do so; thereby the market was denied the information that additional short selling would have brought to the market. As Dunbar stated, “the constraints on shorts will prevent people who don’t own the stock from providing their viewpoints to the market.”\(^\text{178}\) The court agreed that this was another significant factor supporting a lack of market efficiency.\(^\text{179}\)

What is paradoxical about this analysis is that the extent of short selling shows that many sophisticated investors believed that PolyMedica stock was overpriced. The large premium charged for lending the stock also reflected the fact that the holder was worried that the stock might drop precipitously before the holder received the stock back from the short seller. Thus, the short selling, and the impediments to additional short selling, confirms the fact that PolyMedica was overpriced because of the misrepresentations of the defendant. Instead of focusing on this, the court rejected efficiency because the market did not have all the short selling information that the court believed it should have had. Contrary to the court’s opinion, the extent of short selling, the loan fee charged, and the difficulty in accessing shares to short are all facts that are in the public domain.

In addition, if stock is not available to loan to bearish investors to enable them to sell it short, there is another vehicle for “bears” to show the market that they believe the stock is overvalued—they could purchase “put” options.\(^\text{180}\) But this possibility was never considered by the court.

(iv) The Unreasonable “One Day” Standard for “Informational Efficiency”

Finally, the court determined that the time within which the market must react to new information in order to be efficient was one day:

This Court, however, holds that the First Circuit’s definition and relevant explanation of efficiency in \textit{PolyMedica}, which stated that stock price must quickly and fully reflect the release of public information such that ordinary investors cannot profitably trade on the basis of it, requires that the reaction to news be fully completed on the same trading day as its release—and perhaps even within hours or minutes.\(^\text{181}\)

Once again, the court fell into the trap of assuming that all news is digested immediately by the market, without recognizing that the significance, complexity, and mode of disclosure all affect the pace at which the market processes news. Moreover, paradoxically, the negative disclosures

\(^{178}\) Id. at 276.

\(^{179}\) See id.


\(^{181}\) \textit{PolyMedica}, 453 F. Supp. 2d at 278 (footnote omitted).
about PolyMedica relied upon by Miller, plaintiffs’ expert, all had a dramatic negative impact on the price of the stock within one day. Consequently, these disclosures thereby met the district court’s standard that the information be impounded into the market within one day. On the other hand, the positive information injected by PolyMedica’s investor conference call probably took longer for the market to fully incorporate because many potential buyers were probably waiting to see if confirmation of the negative disclosures was in the offing.

The opinion on remand was also not true to the circuit court’s reading of Basic. To counter Judge Keeton’s analysis of Basic, the circuit court quoted other provisions in Basic in which the Supreme Court referred to all information: “[t]he market is acting as the unpaid agent of the investor, informing him that given all the information available to it, the value of the stock is worth the market price,” and that “the market price of shares traded on well-developed markets reflects all publicly available information.” Note that the language of Basic that the circuit court quoted did not unqualifiedly refer to “all information.” Rather it referred to all the “available” information. Arguably, then, if short selling information was not available to the market, it was not relevant under Basic. Again, paradoxically, much of the information the court discussed in determining that the market lacked the information of putative short sellers was actually available to the market.¹⁸²

The opinions of the circuit court and the district court on remand are classic examples which illustrate that a little knowledge can be a dangerous thing. These opinions also support the conclusion that a legal liability structure predicated upon a determination of whether or not the market for a particular stock is efficient is based on whimsy, rather than reality.

B. Scholarly Rejection of Using Market Efficiency to Determine Reliance

Early on, Basic was criticized for arguably requiring an efficient market in order to create a presumption of reliance. A joint effort of law and business professors asserted “substantial disagreement exists about to what degree markets are efficient, how to test for efficiency, and even the definition of efficiency.”¹⁸³ The same criticism holds true today.¹⁸⁴ Moreover, two Nobel Prize winners in economics, Eugene Fama¹⁸⁵ and Robert


¹⁸⁴. See generally Langevoort, supra note 119.

Shiller,186 have taken opposing positions on whether the market is efficient, Fama arguing that it is efficient and Shiller arguing that it is not.

The logic employed by many courts is that reliance depends upon market efficiency, and market efficiency depends upon speed of adjustment. Thus, if market adjustment is “slow,” then the investor does not rely for purposes of achieving class certification. Professor Langevoort, in his seminal article on the history of Basic, raises the fundamental question about market efficiency and its reliance upon speed of market adjustment: “what does speed of adjustment have to do with reliance on stock-price integrity?”187 He asserts that “fraud can and does distort prevailing prices even when adjustment is delayed or incomplete.”188

By way of illustration, Professor Langevoort used In re Merck & Co. Securities Litigation.189 Merck, in 2002, ranked 24th on the Fortune 500 list;190 you would expect a market for its shares to be efficient. However, the question in Merck was not whether the market was efficient, but rather at what point in time would information be impounded into the market and affect the price. Medco, a wholly-owned subsidiary of Merck, recognized income attributable to a copayment made by a pharmacy customer to the pharmacy, even though the subsidiary did not handle these copayments.191 In a registration statement filed April 17, 2002, in connection with a spinoff of Medco, Merck disclosed for the first time that Medco had been recognizing revenue on this basis, but did not disclose the total amount of revenue that had been improperly recognized.192 On the day the registration statement was filed, Merck stock price went up from $55.02 to $55.05.193

Almost two months later, “[o]n June 21, 2002, The Wall Street Journal reported that Medco had been recognizing co-payments as revenue and estimated that in 2001 $4.6 billion in co-payments had been [thereby] recognized.”194 That day Merck stock dropped $2.22, from $52.20 to $49.98.195 Merck filed an amended registration statement on July 5, 2002, and disclosed that Medco had recognized over $12.4 billion in “copay-

187. Langevoort, supra note 119, at 169.
188. Id. at 161.
189. See id. at 173–78 (discussing In re Merck & Co. Sec. Litig., 432 F.3d 261 (3d Cir. 2005)).
191. See Merck, 432 F.3d at 264.
192. See id.
193. See id.
194. Id. at 265.
195. Id.
ment” revenue, of which $5.537 billion was recognized in 2001. The Merck stock declined to $43.57 on July 10th.

The court noted that the Wall Street Journal reporter had derived her figure by subtracting home delivery prescriptions from total prescriptions filled in order to determine the number of retail prescriptions and then assumed an average $10 copayment for each retail prescription. The reporter then multiplied the two numbers together to arrive at the conclusion that $4.6 billion had been improperly recognized. However, according to the circuit court, the market made this determination on April 17th when the price of Merck stock rose $.03, thereby demonstrating that the information was not material: “The Journal reporter simply did the math on June 21; the efficient market hypothesis suggests that the market made these basic calculations months earlier.”

In a very telling quote, the court stated: “If these analysts [who followed Merck and focused on revenue growth] were unable for two months to make a handful of calculations, how can we presume an efficient market at all?” My conclusion from this data is that the analysts were capable of making these computations but did not, which in turn suggests that the market is not necessarily all that efficient, even for a stock as significant and widely traded as Merck.

Also consider Enron. At the beginning of 2001, Enron stock was trading at about $84. On March 5, 2001, it was trading at about $70, when Bethany McLean wrote her famous article in Fortune, Is Enron Overpriced? That week, Enron trading was essentially flat; in fact the high on Friday was higher than the high on Monday. The following week, Enron dropped roughly 10 points to about $60. For the next month, there was a 10% dip and then a rise back to $60 in early May. So much for the diligence and perspicaciousness of analysts!

By September, Enron stock had steadily dropped to $35, and by the end of the year, it was worthless. Granted, the Enron business model was more complicated than that of Medco; however, that complexity should have led an “efficient” market, and the many analysts who covered Enron, to be suspicious and inquisitive. As McLean pointed out, Enron

196. Id.
197. Id.
198. See id. at 270.
199. Id. at 271.
200. Id. at 270.
203. See Ellison, supra note 201.
204. See id.
205. See id.
206. See id.
was the “It Girl” of Wall Street. While Enron was secretive about its business plan, warning signs were out there, including the disparity between cash flow and income and the risky business model that it was employing.207

Apropos to the subject of analyst diligence is the multi-analyst scandal in connection with the dot-com bubble,208 which came to light in spring 2002. New York State Attorney General Eliot Spitzer uncovered a number of e-mails written by Merrill Lynch investment analysts, describing the stock of various companies as “junk,” “crap,” and “a disaster,” unbelievably, they publicly rated these stocks as a “buy.”209 Analysts hyped their ratings, not just to obtain investment banking work for their firms, but sometimes for personal profit, such as getting their daughters into nursery school.210

Merrill Lynch settled for $100 million and agreed to revise its practices with respect to analysts.211 Within a year of Merrill Lynch being sued, ten top United States investment banking firms agreed to pay a total of $875 million in penalties and disgorgement for similar practices.212 Pursuant to the settlement agreements, Wall Street agreed not to pressure

207. See McLean, supra note 202.
   a. Internet Capital Group (ICGE):
      1. E-mail: October 5, 2000—“Going to 5?­” (strong sell); October 6, 2000—“No helpful news to relate, I’m afraid. This has been a disaster—there really is no floor to the stock.”
      2. Investor advice: October 5, 2000—2-1 rating (buy to strong buy);
   b. excite@home (ATHM):
      1. E-mail: June 3, 2000—“ATHM is such a piece of crap!”
      2. Investor advice: June 3, 2000—2-1 rating (buy to strong buy);
   c. Lifeminders (LFMN):
      1. E-mail: December 4, 2000—“I can’t believe what a POS that thing is.”
      2. Investor advice: December 4, 2000—2-1 rating (buy to strong buy).
Id.

210. Jack Grubman, one of the leading analysts on Wall Street, sent an e-mail stating that his boss, Sanford Weill, the chairman of Citigroup and a member of the Board of Directors of AT&T, helped Grubman to get his twin daughters enrolled in an exclusive nursery school after Grubman began recommending AT&T stock. Mr. Weill has acknowledged that he asked Grubman to “take a fresh look at AT&T,” which was code on Wall Street for changing your opinion. See Gretchen Morgenson & Patrick McGeehan, Wall Street and the Nursery School: A New York Story, N.Y. Times (Nov. 14, 2002), http://query.nytimes.com/gst/fullpage.html?res=9C03E3D163F937A25752C1A9649C8B63&f######
211. See The Merrill Lynch Settlement: Good for Merrill, Not for Investors, WHARTON (June 5, 2002), http://knowledge.wharton.upenn.edu/article/the-merrill-lynch-settlement-good-for-merrill-not-for-investors/.
212. See Joint Press Release, SEC, Ten of Nation’s Top Investment Firms Settle Enforcement Actions Involving Conflicts of Interest Between Research and Invest-
analysts to give unwarranted ratings. However, a couple of years later, Rodman and Renshaw fired an analyst because he would not raise his price target for a particular stock.\footnote{213} So much for assuming the integrity of analysts!

What the foregoing demonstrates is the folly of creating a legal structure predicated upon the efficiency of the market and the competence and integrity of analysts.

There is a classic joke about an imprisoned economist who said that he planned to escape by assuming a ladder. This, in tongue-in-cheek fashion, illustrates the extent to which economic theory is predicated upon assumptions. Unfortunately, some of the assumptions in the efficient market theory are questionable. The basic idea is that arbitrageurs will respond to information by driving the price in the right direction. If they believe the price is too low, they will buy; if they believe the price is too high, they will sell short. This requires “that these arbitrageurs should face no capital constraints and have infinite horizons. More precisely, they must have sufficient capital and long enough horizons to await patiently the revelation of information or the reversal of sentiment underlying the mispricing.”\footnote{214} Particularly with respect to short selling, where the risk is unlimited if the stock moves in the wrong direction, there are substantial impediments to the idealized theory as to what the bearish arbitrageur will commit to do.\footnote{215} In addition, both momentum traders\footnote{216} and the existence of herding\footnote{217} can lead to inefficiency. Again, is this the foundation upon which to build a legal theory?

Another problem with the efficient market theory that is raised by Professor Langevoort is that it is binary—either a market for a company’s security is efficient or it is not efficient.\footnote{218} “But there is no such thing as a perfectly efficient market, and so it becomes easy for a court to miss the forest for the trees by accepting too readily the defendants’ statistical evidence of imperfection as reason not to certify.”\footnote{219} One critique of the PolyMedica decision is that the court never considered that “different kinds

\begin{itemize}
\item 213. The analyst originally had an “outperform” recommendation, with a price target of $2.88. As the stock approached this price, he sought to downgrade his recommendation to “market perform,” but instead, his director of research recommended that he increase his price target. See Gretchen Morgenson, \textit{Did Wall Street Really Learn Its Lesson?}, N.Y. Times (Apr. 9, 2006), http://query.nytimes.com/gst/fullpage.html?res=9905E6D91130F93AA35757C0A9609C8B63.
\item 215. See id. at 480–83.
\item 216. See id. at 496–97.
\item 217. See id. at 494–95.
\item 218. See Langevoort, supra note 119, at 167.
\item 219. Id. at 173.
\end{itemize}
of information are likely impounded at different rates of speed, even for the same issuer."\textsuperscript{220} The same is true of the court in \textit{Merck}.

It is \textit{Alice in Wonderland} type thinking to assume that all information—irrespective of its significance, complexity, mode of disclosure, and context of disclosure—would be assimilated at the same rate, particularly within one day. The pattern of the Merck stock movement actually illustrates this fact. Is the market as likely to respond to an arguably buried fact in a multipage registration statement in as quick and decisive a manner as it would to a thoughtful and clearly laid out article in the \textit{Wall Street Journal}, which in turn must be contrasted with an affirmative, clear, detailed acknowledgment by the company? To hold that all information is rapidly (i.e., within one day) assimilated by the market, irrespective of its complexity and mode of distribution, is to endow the market with omniscience.

At a minimum, even if reliance is related to stock price integrity per \textit{Basic}, rather than informational integrity, "rigorous insistence on a showing of market impact would seem to obviate the need to also show market efficiency . . . ."\textsuperscript{221}

\section*{V. The Interrelationship of Reliance, Materiality, and Loss Causation}

\subsection*{A. Fraud on the Market Reliance and Materiality: The Amgen Confusion}

Justice Ginsburg’s opinion in \textit{Amgen}\textsuperscript{222}—in rejecting the requirement that plaintiffs must prove materiality in order for a class to be certified—was characterized by the dissenters as simplistic and illogical\textsuperscript{223} and further muddied the waters as to what must be proved and when.

Justice Ginsburg acknowledged that materiality “indisputably” “is an essential predicate of the fraud-on-the-market theory . . . .”\textsuperscript{224} She also acknowledged that materiality is an essential element of a Rule 10b-5 cause of action.\textsuperscript{225} However, since materiality is an objective standard, the issue of materiality is common to all class members and thus need not be proven at the class certification stage.\textsuperscript{226} Rather, it is an element of a Rule 10b-5 action that will be determined on the merits at trial.\textsuperscript{227} Thus far the opinion makes sense.

However, Justice Thomas, speaking for the dissenters, noted that materiality is an element of the fraud on the market theory and critiqued the majority opinion on the basis that “the Court transforms the predicate

\begin{footnotesize}
\begin{enumerate}
\item \textsuperscript{220} \textit{Id.} at 170.
\item \textsuperscript{221} \textit{Id.} at 179 (citing Macey et al., \textit{supra} note 183, at 1018).
\item \textsuperscript{222} 133 S. Ct. 1184 (2013).
\item \textsuperscript{223} \textit{See id.} at 1206–07 (Thomas, J., dissenting).
\item \textsuperscript{224} \textit{Id.} at 1195 (majority opinion).
\item \textsuperscript{225} \textit{See id.} at 1191–92.
\item \textsuperscript{226} \textit{See id.} at 1197.
\item \textsuperscript{227} \textit{See id.}
\end{enumerate}
\end{footnotesize}
certification inquiry into a novel either-or inquiry occurring much later on the merits." He argued:

According to the Court, either (1) plaintiffs will prove materiality on the merits, thus demonstrating *ex post* that common questions predominated at certification, or (2) they will fail to prove materiality, at which point we learn *ex post* that certification was inappropriate because reliance was not, in fact, a common question. In the Court’s second scenario, fraud on the market was never established, reliance for each class member was inherently individualized, and Rule 23(b)(3) in fact should have barred certification long ago. The Court suggests that the problem created by the second scenario is excusable because the plaintiffs will lose anyway on alternative merits grounds, and the case will be over.

Justice Ginsburg’s holding that materiality need not be proven at the class certification stage would seem to be sophistry if, in establishing reliance through fraud on the market, materiality must be proven at the class certification stage as an element of the fraud on the market doctrine. On the other hand, her opinion would make sense if either there are two different aspects of materiality, or if materiality is not an element of the fraud on the market hypothesis.

Justice Ginsburg’s holding could be justified on the basis that there is materiality in general and materiality specific to plaintiffs’ claims. Consider again Justice Roberts’s hypothetical in *Halliburton II*, in which there were six event studies, one of which involved the specific misrepresentation alleged by plaintiffs as the basis for their suit; even though there was no price impact with respect to the specific misrepresentation challenged in the suit, the evidence showed an efficient market. To Justice Roberts, it would make no sense to certify the class action in such a situation.

But what if the defendant and/or plaintiffs had provided price impact with respect to five events, none of which involved the current alleged misrepresentation, but all of which showed price impact, consequently

---

228. Id. at 1211 (Thomas, J., dissenting).
229. Id. (footnote omitted).
231. His discussion of the consequences of the foregoing is confusing, as he posits that there is an efficient market, but in summarizing the plaintiffs’ view, states that the “action should be certified and proceed as a class action (with all that entails), even though the fraud-on-the-market theory does not apply and common reliance thus cannot be presumed.” *Id.* But, if the class were to be certified, this would mean that the fraud on the market theory did apply. The paradox is that the class action could proceed even though the ultimate result is already known, namely, the plaintiffs must fail because the plaintiffs cannot establish price impact, which means either that the misrepresentation was not material or that there were no damages.
compelling the conclusion that the market was efficient? Here we have an alleged misrepresentation made in an efficient market, which, under Basic, should support a presumption of reliance—and the materiality of the misrepresentation which led to the present litigation would be determined at trial. This is an example of materiality in general but not with respect to the particular case. While this is a possibility, it is not a very likely possibility.

The materiality of the litigated misrepresentation is not necessarily relevant to the question of whether the market is efficient. Arguably, the market impounds all information, both material and immaterial, since the price of the stock impounds many pieces of information, any one of which might not be significant but could become significant in the aggregate. Consequently, whether the alleged misrepresentations were material does not necessarily impact reliance. The investor would rely upon the integrity of the market price, according to Basic, or on the total mix of information in the market, according to this Article, in purchasing the stock—irrespective of whether the investor was defrauded through misrepresentations or otherwise. Whether the investor relied and whether the investor was defrauded are two separate questions.

This conclusion is supported by the facts of the Basic case. The district court certified the class on the fraud on the market theory, but it granted a motion for summary judgment on the ground that the misrepresentations were not material. Thus, at the district court level, the court certified a class irrespective of the materiality of the representations.

Notwithstanding the foregoing defense of Justice Ginsburg’s opinion, if materiality is an element of the fraud on the market theory, then it makes no sense to litigate the issue twice. Arguably, if lack of materiality is established in the certification process, then issue preclusion should prevent re-litigation of this issue. But, according to Justice Thomas, what is the point of continuing the litigation if, at the certification stage, we know the plaintiff cannot prevail?

B. Materiality, Price Impact, and Loss Causation

Price impact, materiality, and loss causation are all intensely factual issues which should be determined in a trial on the merits, before a trier of fact—which could well be a jury—not in a mini-trial at the class certification stage, with a judge determining the factual issues. Previous Supreme Court decisions have determined that neither materiality nor loss causation need to be proven by plaintiffs at the class certification stage. But Justice Roberts, in Halliburton II, determined that “defendants may introduce price impact evidence at the class certification stage, so long as it is for the purpose of countering a plaintiff’s showing of market

232. See Amgen, 133 S. Ct. at 1197.
efficiency, rather than directly rebutting the presumption [of reliance]."234

But, price impact, materiality, and loss causation are all interrelated. Thus, evidence related to all three will be adduced at the class certification stage, and, if plaintiffs are successful in achieving certification, again at the trial on the merits. For example, the expert report of the plaintiffs in *Halliburton* I235 dealt with market efficiency (of which price impact is a factor),236 materiality,237 and loss causation.238 With respect to materiality, Jane D. Nettesheim, plaintiffs’ expert, stated:

> It must be noted that a statistically significant change in a company’s security’s price (net of market and industry effects) is an indicator that new company-specific information has dramatically changed the total mix of information about a company . . . . Information “important enough to affect security prices when publicly released provides compelling evidence that a reasonable investor would consider the information important in making an investment decision.”239

Thus, the correlation of price impact upon materiality; further, without price impact, there would be no loss causation.

If investors are rational—and there is thinking that questions that assumption240—and if there is no market manipulation,241 then immaterial

236. *See id.* at 167a–208a.
237. *See id.* at 208a–13a.
238. *See id.* at 213a–62a.
241. When the market goes up and the economic and political news is negative, it could be asserted that either the market is not rational or that the market is being manipulated. From February to August 2014, the Dow Jones Industrial Average moved from under 16,000 to over 17,000 before—in the first part of August 2014—falling back to the middle of that range. The only positive domestic economic news had been in job gains, but the wages in the new jobs were 23% lower than the jobs that were lost in the recession. *See U.S. Conference of Mayors, U.S. Metro Economies: Income and Wage Gaps Across the U.S.* 2 fig.1 (Aug. 2014), *available at* http://www.usmayors.org/metroeconomies/2014/08/report.pdf. Thus, purchasing power was lagging. Capital expenditures were 15% less than the pattern prior to the recession; businesses were holding onto $1.8 trillion in cash, rather than investing to create jobs and increase productivity. *See Jamie McGeever, “Capex” Mystery Confounds Experts, Chi. Trib.* (Aug. 12, 2014), *http://ircreader.olivesoft
information, by itself, should have no impact on price and material information should affect the price. Much of the alleged fraud in publicly traded securities arises from management’s attempt to maintain the price of the stock when it knows that the position of the company is deteriorating. In such a situation, the arguably material misrepresentation takes place at one point in time, and the price impact is reflected at another point in time, when the truth comes out. But the materiality of the information will be evidenced by the price impact when the truth comes out.

Again, consider the Tellabs situation. In December 2000 through March 2001, the CEO made misleading statements with respect to the continued growth of the company. In June, when the truth came out—namely, that sales of a major product had markedly decreased—the price of the stock plummeted. This is a pattern that repeats itself over and over again. Consequently, there is a clear correlation between the materiality of an alleged misrepresentation, the price impact of such misrepresentation, and the plaintiffs' damages. It would seem logical and efficient that these issues would then be litigated at the same time.

The market for individual stocks is also of questionable rationality. When you compare the price of Amazon on August 1, 2013 with that of August 1, 2014, the price would appear to be essentially flat. Amazon closed at $305 on August 1, 2013, and at $307 on August 1, 2014. But the price of Amazon dropped to $279 on August 28, 2013, and then rose to $309 on December 3, 2013. This is a 43% increase in a little over three months. It rose to $408 on January 22, 2014, but then dropped to $337 on February 5, 2014. This is a 17% drop in two weeks. It then rose to $383 on March 13, 2014, or a 13% increase in a little over a month, before dropping to $284 on May 9, 2014. This is a 25% drop in less than two months. It then rose to $364 on July 24, 2014, a 28% increase in a little over two months, before dropping to $337 on August 1, 2014, a 16% drop in a week. Amazon’s P/E ratio is almost 500. See Amazon.com Inc. (AMZN) Historical Prices, YAHOO FIN., http://finance.yahoo.com/q/hp?s=AMZN+Historical+Prices (last visited Jan. 12, 2015).

The volatility of Netflix was even greater. Its price rose from $319 on January 15, 2014, to $457 on February 25, 2014, a 42% increase in less than a month and a half. It then dropped to $299 on April 28, 2014, a 34% drop in two months. Finally, it rose to $475 on July 2, 2014, a 58% increase in a little over two months. Netflix’s P/E ratio is around 160. See Netflix, Inc. (NFLX) Historical Prices, YAHOO FIN., http://finance.yahoo.com/q/hp?s=NFLX+Historical+Prices (last visited Jan. 12, 2015).

At best, this is a pattern of trading, not investing, and, at worst, raises a question of manipulation. The market for the stocks would certainly seem to be “efficient.” Amazon had a market capitalization of over $140 billion and an average daily volume of about 3.8 million shares. Netflix had a market capitalization of about $26 billion and an average daily volume of about 2.7 million shares. There appears to be little in the way of “news” that would justify price swings of such magnitude.

242. See generally Murdock, supra note 11 (discussing cases that illustrate this pattern).
Justice Roberts, in *Halliburton I*, determined that loss causation need not be established at the class certification stage since reliance relates to the transaction causation and not loss causation. Thereafter, Justice Ginsburg, in *Amgen*, determined that materiality need not be proven at the class certification stage, because materiality is common to all plaintiffs as it is an objective standard. But if price impact is litigated at the class certification stage, according to Justice Roberts in *Halliburton II*, then we are once again bifurcating issues that ought to be tried together. From the standpoint of judicial economy and cost of litigation, these intensely factual issues should be determined in a trial on the merits. This would be possible if we backed away from the current interpretation of fraud on the market as indicating reliance on price integrity and recognized that what the market relies upon is information integrity.

VI. FRAUDULENT DISTORTION

As this Article was nearing completion, a very thoughtful article by Professors Bebchuk and Ferrell was published in *The Business Lawyer*. While the article was published after the Supreme Court decided *Halliburton II*, an earlier draft of the article had been circulated before the briefing of the case and was cited by the main briefs of both sides. Briefly stated, the authors would replace reliance on the efficiency of the market with “reliance on the market price not being impacted by (and thus reflecting) misstatements and omissions” or, in other words, not being impacted by fraudulent distortion.

The foregoing authors do an excellent job of arguing against the continued reliance on market efficiency as a basis for presuming reliance. While the courts, in dealing with the fraud on the market presumption, have treated efficiency as a binary concept—the market is either efficient or not efficient—Professors Bebchuk and Ferrell also argue that efficiency is a “continuum.”

While the approach of Professors Bebchuk and Ferrell is superior to the present focus on market efficiency, it poses two difficulties. First of all, it is phrased in the negative, namely, that investors rely upon an absence of fraudulent distortion; secondly, it posits that investors rely upon a result rather than a cause. Once again, consider *Tellabs*. When management states that one major product is experiencing high demand, whereas in fact demand is drying up, and that a successor product is ready to ship, when in fact it is not, the market price is maintained at an artificially high

243. See Bebchuk & Ferrell, supra note 38.
244. The article was published in the May issue of *The Business Lawyer*, but it was not distributed until late July or early August. The decision of the Supreme Court in *Halliburton II* came down on June 23, 2014.
245. See Bebchuk & Ferrell, supra note 38, at 675.
246. Id. at 686.
247. See id. at 689–90.
level. This situation is in fact fraudulent distortion. But what investors rely upon is that the information introduced into the market by management is in fact truthful and accurate. If management had said that Tel-labs’ business was substantially declining, investors would not have bought at the then existing market price. What investors rely upon is the truthfulness of the information in the market.

Moreover, any liability regime which posits reliance on the price of a security remains open to those who argue that investors do not rely on the price of the security. Professors Bebchuk and Ferrell assert that “[u]nder our approach, market prices need not be relied on or assumed by investors to reflect true value. Fraudulent distortion merely turns on whether the market price is different from what it otherwise would have been absent the fraud.”248 What then do investors rely upon? If you assert that they rely upon a price not affected by fraudulent misrepresentations, are you not really asserting that they rely upon the truthfulness of the mix of information in the marketplace?

The professors phrased the fraud on the market rule as consisting of three propositions:

(A1) The price of a security traded in an efficient market will reflect all publicly available information about a company;

(A2) Accordingly, a buyer of the security in an efficient market may be presumed to have relied on public information in purchasing the security; and

(A3) Where the market for a security is inefficient, a plaintiff cannot invoke the fraud on the market presumption.249

They then recast the foregoing propositions as follows:

(B1) The price of a security traded in an efficient a public market will reflect all some publicly available information about a company;

(B2) Accordingly, a buyer of the security in an efficient a public market may be presumed to have relied on public information in purchasing the security on the market price not being fraudulently distorted, i.e., not being different from what it would have been absent the disclosure deficiency; and

(B3) Where the market price for a security is inefficient not fraudulently distorted, a plaintiff cannot invoke the fraud-on-the-market classwide reliance presumption.250

248. Id. at 688.
249. Id. at 685.
250. Id. at 686 (strikethroughs indicating deletions and italics indicating additions).
Let’s examine this new proposition. First of all, it recognizes that a public market need not impound all publicly available information before investors can be defrauded by misleading statements from corporate management. This is sound. Secondly, it recognizes that misleading information may have the effect of distorting price. This also is sound. But the third proposition then uses the presence or absence of price distortion to determine reliance, whereas the existence of price distortion, or the lack thereof, much more logically relates to both materiality and loss causation. In addition, this approach arguably is also subject to the issues raised by Justice White in his dissent in *Basic*.

As stated earlier, what investors really rely upon is the truthfulness of the information that is in the public markets. In the *Tellabs* example, investors relied upon public information that the sales of the 5500 continued apace and that the 6500 was ready for shipment and customers were eagerly awaiting it. If management had told the truth, investors would not have bought, and the price of Tellabs’ stock would have plummeted, as it did when the truth was revealed.

Let me paraphrase what I have been asserting in this Article as a modification of Professors Bebchuk and Ferrell’s modification of the three *Basic* propositions. Thus:

(i) The price of a security traded in a public market will reflect some publicly available information about a company (thus, my first proposition would be identical to their’s, although I would accept substituting “most” for “some”);

(ii) Accordingly, a buyer of the security in a public market may be presumed to have relied on public information in purchasing the security; and

(iii) A defendant may rebut the presumption of reliance by proving that the misrepresentation did not affect the decision of a typical investor to purchase, i.e., the misrepresentation was not objectively material.251

This is clearly a simpler approach to the issue of reliance. Arguably, it is also more straightforward. Adopting this approach would undoubtedly enrage defense counsel and the Business Roundtable, because it would simplify the certification process and eliminate all of the litigation over market efficiency. It would then leave the issues of materiality and loss causation to trial since they are factual issues common to all the plaintiffs. On the other hand, if an alleged material misstatement is so insubstantial

251. All of these three-proposition formulations are assuming that the plaintiff is a purchaser. This assumption is because most of the time management is trying to support the price of the shares and the purchaser overpays. In other words, management is putting good news into the marketplace. However, *Basic* itself dealt with a “bad news” situation, i.e., that no merger negotiations were taking place. In such a situation, the focus would be upon the seller being defrauded.
as to be determined immaterial as a matter of law, this could be dealt with at the class certification stage, or on a motion to dismiss or for summary judgment. All that would need to be established at the class certification stage is that the defendant company, and its management, inserted misleading material into the public marketplace. I would also accept a limited determination of materiality at the class certification stage: namely, whether the alleged misrepresentations were of the kind that a reasonable investor would normally rely upon.

Another way of comparing this approach to that of Professors Bebchuk and Ferrell is that their approach would permit, at the class certification stage, what Justice Roberts ultimately held in *Halliburton II*: that “defendants must be afforded an opportunity before class certification to defeat the presumption through evidence that an alleged misrepresentation did not actually affect the market price of the stock.”252 While Professors Bebchuk and Ferrell did not take a definitive position on whether the plaintiff or the defendant, at the class certification stage, should have the burden of proof on price impact, Justice Roberts appears to make the lack of price distortion an affirmative defense, with respect to which the defendant should have the burden. Putting that issue to one side, their approach clearly presents this issue of price impact for resolution at the certification stage, with the bifurcation of issues that ought to be tried together.

On the other hand, under the approach advocated by this Article, reliance flows from “information distortion,” not price distortion. Accordingly, at the certification stage, the only issue for resolution is whether the defendant corporation and its management introduced misleading information into the marketplace. If they have, reliance should be presumed. It could be argued that the effect of this approach is to make class certification a routine matter. And that is correct. It should be a routine matter, as it was before courts started using the class certification process to hold a mini-trial. As has been developed in this Article, factual matters, such as materiality, price distortion, and loss causation, together with scienter, should be developed in one proceeding at trial.

### VII. Conclusion

Where are we after the Sturm und Drang of three United States Supreme Court decisions in three years253 grappling with the contours of the fraud on the market theory? I would argue that the present legal regime for private securities litigation is irrational, wasteful and inefficient, biased in favor of management as opposed to investors, and built upon illusion, not reality. The Supreme Court is wandering in an illogical thicket from which it cannot extricate itself. It is ideologically divided: the conserva-

---

tives are on a mission to protect business and management, sometimes corrupt management, from the onslaught of consumers, employees, and investors, while the liberals are fighting a rearguard action to prevent as much damage as possible.

Consider the contours of a private securities case. The plaintiff files a complaint which must set forth the misrepresentations with specificity—sufficient particularity to satisfy the sometimes unreasonable expectations of the federal courts—and must also set forth facts giving rise to a strong inference of scienter. The plaintiff is then met with a motion to dismiss which is to be heard before the plaintiff has an opportunity for discovery. The detail expected by federal courts often can only be obtained through confidential informants. However, confidential informants are viewed with a high degree of suspicion by federal courts.\footnote{See, e.g., Higginbotham v. Baxter Int'l Inc., 495 F.3d 753, 757 (7th Cir. 2007) (Easterbrook, J.) (asserting that information supplied by confidential sources must be “discounted,” usually “steep[ly]”).}

If the plaintiff surmounts the motion to dismiss and the expectations of pleading particularity, the next step is class certification. Some courts bifurcate discovery between that necessary for certification and that necessary for the trial on the merits. Supposedly, the discovery for class certification would be limited but, as a result of \emph{Halliburton II}, the defendant can now raise price impact, which is intimately related to materiality and loss causation. However, the plaintiff is entitled to a presumption of reliance, or transaction causation, as a result of the fraud on the market theory. But one of the elements of the fraud on the market theory, according to the courts, is materiality, which, according to \emph{Amgen}, is a matter to be addressed at a trial on the merits. Is this the clarity we would expect when billions of dollars are at stake?

According to most courts, the benefit of the fraud on the market presumption is only available when the market is efficient. Market efficiency is a concept developed by economics and finance professors to support the arguable proposition that the individual investor cannot beat the market. Its origins had nothing to do with securities fraud. The Nobel Prize in economics has been awarded to two professors who have opposing viewpoints on whether the market is efficient. Courts seem to regard market efficiency as a binary proposition and expect that, in an efficient market, material information will be quickly—in as little time as one day—impounded into the price of the company's stock. They give no credence to the idea that there are degrees of market efficiency, degrees of materiality, varying modes of communicating information to the market, varying probability of the sources of the information, and different degrees of complexity of the information so transmitted—all of which logically would affect the time necessary for information to be impounded into the market. This is a legal structure built upon an illusion.
Since the reality is that there is no such thing as a perfectly efficient market, as one scholar has stated, “it becomes easy for a court to miss the forest for the trees by accepting too readily the defendants’ statistical evidence of imperfection as reason not to certify.”

If the plaintiff does not succeed in obtaining class certification, that is usually the end of the litigation because most individual plaintiffs do not have a sufficient economic stake to carry on the cost of litigation on their own. If the plaintiff does succeed in certification, we then move on to discovery on the merits and eventually to trial.

Materiality, price distortion, and loss causation are all intertwined and highly factual issues. But the effect of these three Supreme Court decisions is to bifurcate these issues. This bifurcation only adds to the expense and confusion of litigation.

After almost seventy years of private securities litigation, it is time to step back and examine whether there is a more sensible approach. The starting place would be the securities laws themselves and the purposes for which they were enacted. The prevalence of fraud in the securities markets was clearly an impetus for the enactment of the securities laws; it is further clear that the securities laws were enacted for the protection of investors and to insure the maintenance of fair and honest markets. Moreover, as former Justice Stevens has stated, “[f]ashioning appropriate remedies for the violation of rules of law designed to protect a class of citizens was the routine business of judges,” or at least it was until a supposedly conservative, but in reality reactionary, Supreme Court began cutting away at investor protection.

As other scholars have recognized, “proving the efficiency of the market as a whole is only an indirect means of proving that the market relied on a particular statement.” Basic, in discussing materiality, focused on the fact that investors would be guided by the total mix of information available. It is this total mix of information upon which the investment community relies and which determines the price of the stock. Mandatory reporting under the securities laws provides much of this mix of information, while voluntary corporate disclosures, such as press releases and investor conference calls, provide additional information. These sources of information are supplemented by analyst reports and other third-party-generated information.

Since a major purpose of the securities laws was to provide investors with information, and since the purpose of the antifraud provisions is to ensure truthfulness, it seems far more logical to presume that investors rely upon this total mix of information, rather than believing the fantasy.

255. Langevoort, supra note 119, at 173.
257. Brief of Law Professors, supra note 120, at 24.
that they rely upon the integrity of market prices. What they rely upon is the integrity of the information that corporate management has introduced into the market. The integrity, or lack thereof, of market prices is a matter for a damage determination, namely, loss causation, not for reliance, transaction causation.

If this view of reliance is adopted, then the meaty factual issues of materiality, price distortion, and loss causation can be addressed at trial—which is where they belong. Rejecting the focus on market efficiency will, first, eliminate the present practice of ruminating about some idealized notion of an efficient market. Second, it will eliminate the use of minor imperfections in this idealized model to justify dismissal of class certification, and thus effectively avoid the termination of lawsuits that may well be meritorious. The focus of courts should be upon protecting investors and holding management accountable; too often the converse is true.
250

Villanova Law Review

[Vol. 60: p. 203]