Texas Law Review

THE INCENTIVES MATRIX: THE COMPARATIVE EFFECTIVENESS OF REWARDS, LIABILITIES, DUTIES, AND PROTECTIONS FOR REPORTING ILLEGALITY
Yuval Feldman & Orly Lobel

KEEP CHARITY CHARITABLE
Brian Galle

THE TAKING/TAXING TAXONOMY
Amnon Lehavi

BOOK REVIEW—SOME MOVING PARTS OF JURISPRUDENCE
Kevin Toh

POST-9/11 ANTI-TERRORISM POLICY REGARDING NONCITIZENS AND THE CONSTITUTIONAL IDEA OF EQUAL PROTECTION UNDER THE LAWS

THE CONMAN AND THE SHERIFF: SEC JURISDICTION AND THE ROLE OF OFFSHORE FINANCIAL CENTERS IN MODERN SECURITIES FRAUD

A COMPARATIVE "HARD LOOK" AT *CHEVRON*: WHAT THE UNITED KINGDOM AND AUSTRALIA REVEAL ABOUT AMERICAN ADMINISTRATIVE LAW

MAY 2010 VOL. 88 NO. 6 PAGES 1151 TO 1399

Texas Law Review (ISSN 0040-4411) is published seven times a year—November, December, February, March, April, May, and June. The annual subscription price is $47.00 except as follows: Texas residents pay $50.88 and foreign subscribers pay $55.00. All publication rights are owned by the Texas Law Review Association. *Texas Law Review* is published under license by The University of Texas at Austin School of Law, P.O. Box 8670, Austin, Texas 78713. Periodicals Postage Paid at Austin, Texas, and at additional mailing offices.

POSTMASTER: Send address changes to The University of Texas at Austin School of Law, P.O. Box 8670, Austin, Texas 78713.

Complete sets and single issues are available from WILLIAM S. HEIN & CO., INC., 1285 Main St., Buffalo, NY 14209-1987. Phone: 1-800-828-7571.

Single issues in the current volume may be purchased from the *Texas Law Review* Publications Office for $15.00 per copy plus shipping. Texas residents, please add applicable sales tax.

The *Texas Law Review* is pleased to consider unsolicited manuscripts for publication but regrets that it cannot return them. Please submit a single-spaced manuscript, printed on one side only, with footnotes rather than endnotes. Citations should conform with the *Texas Rules of Form* (11th ed. 2006) and *The Bluebook: A Uniform System of Citation* (18th ed. 2005). Except when content suggests otherwise, the *Texas Law Review* follows the guidelines set forth in the *Texas Law Review Manual on Usage & Style* (11th ed. 2008), *The Chicago Manual of Style* (15th ed. 2003), and Bryan A. Garner, *A Dictionary of Modern Legal Usage* (2d ed. 1995).

Editorial Offices: *Texas Law Review*
727 East Dean Keeton Street, Austin, Texas 78705
(512) 232-1280 Fax (512) 471-3282
tlr@law.utexas.edu
http://www.texaslrev.com

THE UNIVERSITY OF TEXAS SCHOOL OF LAW

ADMINISTRATIVE OFFICERS

LAWRENCE G. SAGER, B.A., LL.B.; *Dean, John Jeffers Research Chair in Law, Alice Jane Drysdale Sheffield Regents Chair.*
MECHELE DICKERSON, B.A., J.D.; *Associate Dean for Academic Affairs, Arthur L. Moller Chair in Bankruptcy Law and Practice.*
ALEXANDRA W. ALBRIGHT, B.A., J.D.; *Associate Dean for Administrative Services, Senior Lecturer.*
EDEN E. HARRINGTON, B.A., J.D.; *Assistant Dean for Clinical Education, Director of William Wayne Justice Center for Public Interest Law.*
KIMBERLY L. BIAR, B.B.A.; *Assistant Dean for Financial Affairs, Certified Public Accountant.*
CARLA COOPER, B.A., M.A., Ph.D.; *Assistant Dean for Alumni Relations and Development.*
MICHAEL J. ESPOSITO, B.A., J.D., M.B.A.; *Assistant Dean for Continuing Legal Education.*
KIRSTON FORTUNE, B.F.A.; *Assistant Dean for Communications.*
MONICA K. INGRAM, B.A., J.D.; *Assistant Dean for Admissions and Financial Aid.*
DAVID A. MONTOYA, B.A., J.D.; *Assistant Dean for Career Services.*
LESLIE OSTER, B.A., J.D.; *Assistant Dean for Strategic Planning.*
REYMUNDO RAMOS, B.A.; *Assistant Dean for Student Affairs.*

FACULTY EMERITI

HANS W. BAADE, A.B., J.D., LL.B., LL.M.; *Hugh Lamar Stone Chair Emeritus in Civil Law.*
RICHARD V. BARNDT, B.S.L., LL.B.; *Professor Emeritus.*
WILLIAM W. GIBSON, JR., B.A., LL.B.; *Sylvan Lang Professor Emeritus in Law of Trusts.*
ROBERT W. HAMILTON, A.B., J.D.; *Minerva House Drysdale Regents Chair Emeritus.*
DOUGLAS LAYCOCK, B.A., J.D.; *Alice McKean Young Regents Chair Emeritus.*
J. L. LEBOWITZ, A.B., J.D., LL.M.; *Joseph C. Hutcheson Professor Emeritus.*
JOHN T. RATLIFF, JR., B.A., LL.B.; *Ben Gardner Sewell Professor Emeritus in Civil Trial Advocacy.*
JOHN F. SUTTON, JR., J.D.; *A.W. Walker Centennial Chair Emeritus.*
JAMES M. TREECE, B.A., J.D., M.A.; *Charles I. Francis Professor Emeritus in Law.*
RUSSELL J. WEINTRAUB, B.A., J.D.; *Ben H. & Kitty King Powell Chair Emeritus in Business and Commercial Law.*

PROFESSORS

JEFFREY B. ABRAMSON, B.A., J.D., Ph.D.; *Professor of Law and Government.*
DAVID E. ADELMAN, B.A., Ph.D., J.D.; *Harry Reasoner Regents Chair in Law.*
DAVID A. ANDERSON, A.B., J.D.; *Fred and Emily Marshall Wulff Centennial Chair in Law.*
MARK L. ASCHER, B.A., M.A., J.D., LL.M.; *Joseph D. Jamail Centennial Chair in Law.*
RONEN AVRAHAM, M.A., LL.M., J.S.D.; *Thomas Shelton Maxey Professor in Law.*
LYNN A. BAKER, B.A., Honours B.A., J.D.; *Frederick M. Baron Chair in Law, Co-Director of Center on Lawyers, Civil Justice, and the Media.*
MITCHELL N. BERMAN, A.B., M.A., J.D.; *Richard Dale Endowed Chair in Law.*
BERNARD S. BLACK, A.B., M.A., J.D.; *Hayden W. Head Regents Chair for Faculty Excellence, Director of Center for Law, Business, and Economics.*
LYNN E. BLAIS, A.B., J.D.; *Leroy G. Denman, Jr. Regents Professor in Real Property Law.*
ROBERT G. BONE, B.A., J.D.; *Professor.*
NORMA V. CANTU, B.A., J.D.; *Professor of Law and Education.*
LOFTUS C. CARSON, II, B.S., M. Pub. Aff., M.B.A., J.D.; *Ronald D. Krist Professor.*
ROBERT M. CHESNEY, B.S., J.D.; *Charles I. Francis Professor in Law.*
MICHAEL J. CHURGIN, A.B., J.D.; *Raybourne Thompson Centennial Professor.*
JANE M. COHEN, B.A., J.D.; *Edward Clark Centennial Professor.*
FRANK B. CROSS, B.A., J.D.; *Herbert D. Kelleher Centennial Professor of Business Law.*
JOHN DEIGH, B.A., M.A., Ph.D.; *Professor of Law and Philosophy.*
GEORGE E. DIX, B.A., J.D.; *George R. Killam, Jr. Chair of Criminal Law.*
JOHN S. DZIENKOWSKI, B.B.A., J.D.; *Dean John F. Sutton, Jr. Chair in Lawyering and the Legal Process.*
KAREN L. ENGLE, B.A., J.D.; *Cecil D. Redford Professor in Law, Director of Bernard and Audre Rapoport Center for Human Rights and Justice.*
WILLIAM E. FORBATH, A.B., B.A., Ph.D., J.D.; *Lloyd M. Bentsen Chair in Law.*
JULIUS G. GETMAN, B.A., LL.B., LL.M.; *Earl E. Sheffield Regents Chair.*
STEVEN GOODE, B.A., J.D.; *W. James Kronzer Chair in Trial and Appellate Advocacy, University Distinguished Teaching Professor.*
LINO A. GRAGLIA, B.A., LL.B.; *A. Dalton Cross Professor.*
PATRICIA I. HANSEN, A.B., M.P.A., J.D.; *J. Waddy Bullion Professor.*
HENRY T. HU, B.S., M.A., J.D.; *Allan Shivers Chair in the Law of Banking and Finance.*
DEREK P. JINKS, B.A., M.A., J.D.; *The Marrs McLean Professor in Law.*
STANLEY M. JOHANSON, B.S., LL.B., LL.M.; *Fannie Coplin Regents Chair, University Distinguished Teaching Professor.*
CALVIN H. JOHNSON, B.A., J.D.; *Andrews & Kurth Centennial Professor.*
SUSAN R. KLEIN, B.A., J.D.; *Alice McKean Young Regents Chair in Law.*
SANFORD V. LEVINSON, A.B., Ph.D., J.D.; *W. St. John Garwood and W. St. John Garwood, Jr. Centennial Chair in Law, Professor of Government.*
STEFANIE A. LINDQUIST, B.A., J.D., Ph.D.; *The Thomas W. Gregory Professorship in Law.*
BASIL S. MARKESINIS, LL.B., LL.D., DCL, Ph.D.; *Jamail Regents Chair.*
INGA MARKOVITS, LL.M.; *"The Friends of Joe Jamail" Regents Chair.*
RICHARD S. MARKOVITS, B.A., LL.B., Ph.D.; *John B. Connally Chair.*
THOMAS O. MCGARITY, B.A., J.D.; *Joe R. and Teresa Lozano Long Endowed Chair in Administrative Law.*
LINDA S. MULLENIX, B.A., M. Phil., J.D., Ph.D.; *Morris & Rita Atlas Chair in Advocacy.*

ROBERT J. PERONI, B.S.C., J.D., LL.M.; *James A. Elkins Centennial Chair in Law.*
H. W. PERRY, JR., B.A., M.A., Ph.D.; *Associate Professor of Law and Government.*
LUCAS A. POWE, JR., B.A., J.D.; *Anne Green Regents Chair in Law, Professor of Government.*
WILLIAM C. POWERS, JR., B.A., J.D.; *President, The University of Texas at Austin, Hines H. Baker and Thelma Kelley Baker Chair, University Distinguished Teaching Professor.*
DAVID M. RABBAN, B.A., J.D.; *Dahr Jamail, Randall Hage Jamail, and Robert Lee Jamail Regents Chair, University Distinguished Teaching Professor.*
ALAN S. RAU, B.A., LL.B.; *Burg Family Professorship.*
R. A. REESE, B.A., J.D.; *Arnold, White & Durkee Centennial Professor.*
DAVID W. ROBERTSON, B.A., LL.B., LL.M., J.S.D.; *W. Page Keeton Chair in Tort Law, University Distinguished Teaching Professor.*
JOHN A. ROBERTSON, A.B., J.D.; *Vinson & Elkins Chair.*
DANIEL B. RODRIGUEZ, B.A., J.D.; *Minerva House Drysdale Regents Chair in Law.*
WILLIAM M. SAGE, A.B., M.D., J.D.; *James R. Dougherty Chair for Faculty Excellence.*
JOHN J. SAMPSON, B.B.A., LL.B.; *William Benjamin Wynne Professor.*
THOMAS K. SEUNG, Ph.D.; *Jesse H. Jones Regents Professor in Liberal Arts.*
MICHAEL M. SHARLOT, B.A., LL.B.; *Wright C. Morrow Professor.*
CHARLES M. SILVER, B.A., M.A., J.D.; *Roy W. and Eugenia C. MacDonald Endowed Chair in Civil Procedure, Co-Director of Center on Lawyers, Civil Justice, and the Media.*
ERNEST E. SMITH, B.A., LL.B.; *Rex G. Baker Centennial Chair in Natural Resources Law.*
JANE STAPLETON, B.S., Ph.D., LL.B., D. Phil.; *Ernest E. Smith Professor.*
JORDAN M. STEIKER, B.A., J.D.; *Judge Robert M. Parker Endowed Chair in Law.*
MICHAEL F. STURLEY, B.A., J.D.; *Stanley D. and Sandra J. Rosenberg Centennial Professor in Property Law.*
SHIRLEY E. THOMPSON, A.B., A.M., Ph.D.; *Professor.*
GERALD TORRES, A.B., J.D., LL.M.; *Bryant Smith Chair in Law.*
GREGORY J. VINCENT, B.A., J.D., Ed.D.; *Professor.*
WENDY E. WAGNER, B.A., M.E.S., J.D.; *Joe A. Worsham Centennial Professor.*
LOUISE WEINBERG, A.B., J.D., LL.M.; *William B. Bates Chair for the Administration of Justice.*
OLIN G. WELLBORN, A.B., J.D.; *William C. Liedtke, Sr. Professor.*
JAY L. WESTBROOK, B.A., J.D.; *Benno C. Schmidt Chair of Business Law.*
ABRAHAM L. WICKELGREN, A.B., Ph.D., J.D.; *Bernard J. Ward Professor in Law.*
ZIPPORAH B. WISEMAN, B.A., M.A., LL.B.; *Thos. H. Law Centennial Professor in Law.*
PATRICK WOOLLEY, A.B., J.D.; *Beck, Redden & Secrest Professor.*

ASSISTANT PROFESSORS

OREN BRACHA, LL.B., S.J.D.
DANIEL M. BRINKS, A.B., J.D., Ph.D.
JENS C. DAMMANN, J.D., LL.M., Dr. Jur., J.S.D.
JUSTIN DRIVER, B.A., M.A., M.A., J.D.
ZACHARY S. ELKINS, B.A., M.A., Ph.D.
MIRA GANOR, B.A., M.B.A., LL.B., LL.M., J.S.D.
JOHN M. GOLDEN, A.B., J.D., Ph.D.
EMILY E. KADENS, B.A., M.A., Dipl., M.A., Ph.D., J.D.
JENNIFER E. LAURIN, B.A., J.D.
ANGELA K. LITTWIN, B.A., J.D.
MARY ROSE, A.B., M.A., Ph.D.
SEAN H. WILLIAMS, B.A., J.D.
HANNAH J. WISEMAN, A.B., J.D.

SENIOR LECTURERS, WRITING LECTURERS, AND CLINICAL PROFESSORS

WILLIAM P. ALLISON, B.A., J.D.; *Clinical Professor, Director of Criminal Defense Clinic.*
MARJORIE I. BACHMAN, B.S., J.D.; *Clinical Instructor.*
PHILIP C. BOBBITT, A.B., J.D., Ph.D.; *Distinguished Senior Lecturer.*
KAMELA S. BRIDGES, B.A., B.J., J.D.; *Lecturer.*
CYNTHIA L. BRYANT, B.A., J.D.; *Clinical Professor, Director of Mediation Clinic.*
SARAH M. BUEL, B.A., J.D.; *Clinical Professor.*
MARY R. CROUTER, A.B., J.D.; *Assistant Director of William Wayne Justice Center for Public Interest Law.*
TIFFANY J. DOWLING, B.A., J.D.; *Clinical Instructor.*
LORI K. DUKE, B.A., J.D.; *Clinical Professor.*
ARIEL E. DULITZKY, J.D., LL.M.; *Clinical Professor.*
DENISE L. GILMAN, B.A., J.D.; *Clinical Professor.*
BARBARA HINES, B.A., J.D.; *Director of Immigration Clinic.*
KRISTINE A. HUSKEY, B.A., J.D.; *Director of National Security & Human Rights Clinic.*
JEANA A. LUNGWITZ, B.A., J.D.; *Clinical Professor, Director of Domestic Violence Clinic.*
ROBIN B. MEYER, B.A., M.A., J.D.; *Lecturer.*
TRACY W. MCCORMACK, B.A., J.D.; *Director of Trial Advocacy Program.*
RANJANA NATARAJAN, B.A., J.D.; *Clinical Professor.*
ROBERT C. OWEN, A.B., M.A., J.D.; *Clinical Professor, Co-Director of Capital Punishment Clinic.*
SEAN J. PETRIE, B.A., J.D.; *Lecturer.*
WAYNE SCHIESS, B.A., J.D.; *Senior Lecturer, Director of Legal Writing.*
PAMELA J. SIGMAN, B.A., J.D.; *Lecturer, Director of Juvenile Justice Clinic.*
LESLIE L. STRAUCH, B.A., J.D.; *Clinical Professor.*
DAVID S. SOKOLOW, B.A., M.A., J.D., M.B.A.; *Distinguished Senior Lecturer, Director of Student Life.*
JAN SUMMER, B.A., M.A., J.D.; *Executive Director of Center for Public Policy Dispute Resolution.*
MELINDA E. TAYLOR, B.A., J.D.; *Senior Lecturer, Executive Director of Center for Global Energy, International Arbitration, and Environmental Law.*
HEATHER K. WAY, B.A., B.J., J.D.; *Lecturer, Director of Community Development Clinic.*
ELIZABETH M. YOUNGDALE, B.A., M.L.I.S., J.D.; *Lecturer.*

ADJUNCT PROFESSORS AND OTHER LECTURERS

WILLIAM R. ALLENSWORTH, B.A., J.D.
JAMAL K. ALSAFFAR, B.A., J.D.
MARILYN ARMOUR, B.A., M.S.W., Ph.D.
WILLIAM G. BARBER, B.S.Ch.E., J.D.
WILLIAM G. BARBER, III, B.A., LL.M.
NICOLAS G. BARZOUKAS, B.S., M.B.A., J.D.
SHARON C. BAXTER, B.S., J.D.
KATHERINE E. BEAUMONT, B.A., J.D.
KARA BELEW, B.A., B.B.A., J.D.
WILLIAM H. BEARDALL, JR., B.A., J.D.
JERRY A. BELL, B.A., J.D.
ALLISON H. BENESCH, B.A., M.S.W., J.D.
CRAIG R. BENNETT, B.S., J.D.
JAMES B. BENNETT, B.B.A., J.D.
MELISSA J. BERNSTEIN, B.A., M.L.S., J.D.
MURFF F. BLEDSOE, B.A., J.D.
WILLIAM P. BOWERS, B.B.A., J.D., LL.M.
ANTHONY W. BROWN, B.A., J.D.
JAMES E. BROWN, LL.B.
TOMMY L. BROYLES, B.A., J.D.
PAUL J. BURKA, B.A., LL.B.
W. A. BURTON, JR., B.A., M.A., LL.B.
AGNES E. CASAS, B.A., J.D.
RUBEN V. CASTANEDA, B.A., J.D.
EDWARD A. CAVAZOS, B.A., J.D.
CHARLES G. CHILDRESS, B.A., J.D.
DAN J. CHRISTENSEN, B.B.A., J.D.
JOSEPH A. CIALONE, II, B.A., J.D.
JEFF CIVINS, A.B., M.S., J.D.
JUDGE LEIF M. CLARK, B.A., M.Div., J.D.
DANA L. COBB, B.A., J.D.
GARY COBB, B.A., J.D.
MARK W. COCHRAN, B.J., LL.M., J.D.
ELIZABETH COHEN, B.A., M.S.W., J.D.
JAMES W. COLLINS, B.S., J.D.
TED CRUZ, A.B., J.D.
PATRICIA J. CUMMINGS, B.A., J.D.
WILLIAM H. CUNNINGHAM, B.A., M.B.A., Ph.D.
HECTOR DE LEON, B.S., J.D.
DICK DEGUERIN, B.A., LL.B.
MICHELE Y. DEITCH, B.A., M.S., J.D.
ADAM R. DELL, B.A., J.D.
CASEY D. DUNCAN, B.A., M.L.I.S., J.D.
PHILIP DURST, B.A., M.A., J.D., Ph.D.
ELANA S. EINHORN, B.A., J.D.
JAY D. ELLWANGER, B.A., J.D.
LOWELL P. FELDMAN, B.A., J.D.
KENNETH FLAMM, Ph.D.
JOHN C. FLEMING, B.A., J.D.
MARIA FRANKLIN, B.A., M.A., Ph.D.
DAVID C. FREDERICK, B.A., Ph.D., J.D.
GREGORY D. FREED, B.A., J.D.
ELIZABETH FRUMKIN, B.A., M.A., J.D.
FRED J. FUCHS, B.A., J.D.
MICHAEL GAGARIN, Ph.D.
GERALD J. GALOW, B.A., J.D.
JAMES B. GAMBRELL, B.S.M.E., M.A., LL.B.
FRANCIS J. GAVIN, B.A., M.S.T., Ph.D.
CHARLES E. GHOLZ, B.S., Ph.D.
MICHAEL J. GOLDEN, A.B., J.D.
JULIE E. GRANTHAM, B.A., J.D.
SHERRI R. GREENBERG, B.A., M.Sc.
ROBERT L. GROVE, B.S., J.D.
DAVID HALPERN, B.A., J.D.
JETT L. HANNA, B.B.A., J.D.
KELLY L. HARAGAN, B.A., J.D.
CLINT A. HARBOUR, B.A., J.D., LL.M.
AMY J. SCHUMACHER, B.A., J.D.
AARON R. SCHWARTZ, B.A., LL.B.
MARCUS F. SCHWARTZ, B.B.A., J.D.
SUZANNE SCHWARTZ, B.J., J.D.
ROBERT L. HARGETT, B.B.A., J.D.
BARBARA J. HARLOW, B.A., M.A., Ph.D.
JAMES C. HARRINGTON, B.A., M.A., J.D.
CHRISTOPHER S. HARRISON, Ph.D., J.D.
AMBER L. HATFIELD, B.S.E.E., J.D.
JOHN R. HAYS, JR., B.A., J.D.
PAUL M. HEBERT, A.B., J.D.
STEVEN L. HIGHLANDER, B.A., Ph.D., J.D.
SUSAN J. HIGHTOWER, B.A., M.A., J.D.
WILLIAM M. HINES, III, B.B.A., J.D.
JAMES C. HO, B.A., J.D.
KENNETH E. HOUP, JR., J.D.
RANDY R. HOWRY, B.J., J.D.
MONTY G. HUMBLE, B.A., J.D.
BOBBY R. INMAN, B.A.
PATRICK O. KEEL, B.A., J.D.
CHARI L. KELLY, B.A., J.D.
ROBERT N. KEPPLE, B.A., J.D.
MARK L. KINCAID, B.B.A., J.D.
KEVIN S. KUDLAC, B.S., J.D.
KURT H. KUHN, B.A., J.D.
AMI L. LARSON, B.A., J.D.
KEVIN R. LASHUS, B.A., J.D.
JODI R. LAZAR, B.A., J.D.
MAURIE A. LEVIN, B.A., J.D.
ANDREW F. MACRAE, B.J., J.D.
VIJAY MAHAJAN, M.S.Ch.E., Ph.D.
JIM MARCUS, B.A., J.D.
PETER D. MARKETOS, B.A., J.D.
FRANCES L. MARTINEZ, B.A., J.D.
RAY MARTINEZ, III, B.A., J.D.
ALOYSIUS P. MARTINICH, B.A., M.A., Ph.D.
LISA M. MCCLAIN, B.A., J.D., LL.M.
STEPHANIE G. MCFARLAND, B.A., J.D.
BARRY F. MCNEIL, B.A., J.D.
ANGELA T. MELINARAAB, B.F.A., J.D.
MARGARET M. MENICUCCI, B.A., J.D.
JO A. MERICA, B.A., J.D.
RANELLE M. MERONEY, B.A., J.D.
DARYL L. MOORE, B.A., M.L.A., J.D.
STEVEN A. MOORE, B.A., Ph.D.
EDWIN G. MORRIS, B.S., J.D.
MARCO M. MUNOZ, LL.B., M.C.J.
MANUEL H. NEWBURGER, B.A., J.D.
STEVEN P. NICHOLS, B.S.M.E., M.S.M.E., J.D., Ph.D.
JANE A. O'CONNELL, B.A., J.D.
PATRICK L. O'DANIEL, B.B.A., J.D.
GUILLERMO A. PADILLA, J.D., Ph.D.
M. A. PAYAN, B.A., J.D.
MARK L. PERLMUTTER, B.S., J.D.
LOUIS T. PIRKEY, B.S.Ch.E., J.D.
JUDGE ROBERT L. PITMAN, B.S., J.D.
ELIZA T. PLATTS-MILLS, B.A., J.D.
LAURA L. PRATHER, B.B.A., J.D.
JONATHAN PRATTER, B.A., M.L.S., J.D.
VELVA L. PRICE, B.A., J.D.
BRIAN C. RIDER, B.A., J.D.
GRETCHEN RITTER, B.S., Ph.D.
ROBERT M. ROACH, JR., B.A., J.D.
BRIAN J. ROARK, B.A., J.D.
BETTY E. RODRIGUEZ, B.S.W., J.D.
JAMES D. ROWE, B.A., J.D.
MATTHEW C. RYAN, B.A., J.D.
KAREN R. SAGE, B.A., J.D.
MARK A. SANTOS, B.A., J.D.
CHRISTOPHE H. SAPSTEAD, B.A., J.D.
SUSAN SCHULTZ, B.S., J.D.
WILLIAM F. STUTTS, B.A., J.D.
MATTHEW J. SULLIVAN, B.S., J.D.
GRETCHEN S. SWEEN, B.A., M.A., Ph.D., J.D.
BRADLEY P. TEMPLE, B.A., J.D.

RICHARD J. SEGURA, JR., B.A., J.D.
JUDGE ERIC M. SHEPPERD, B.A., J.D.
RONALD J. SIEVERT, B.A., J.D.
AMBROSIO A. SILVA, B.S., J.D.
LOUISE SINGLE, B.S., M.T.A., Ph.D.
BEA A. SMITH, B.A., M.A., J.D.
CORY W. SMITH, B.A., J.D.
DWAYNE W. SMITH, B.A., J.D.
TARA A. SMITH, Ph.D.
LYDIA N. SOLIZ, B.B.A., J.D.
JUSTICE ROSE B. SPECTOR, B.A., J.D.
LEWIS J. SPELLMAN, B.B.A., M.B.A., M.A., Ph.D.
DAVID B. SPENCE, B.A., J.D., M.A., Ph.D.
WILLIAM J. SPENCER, B.A., M.S., Ph.D.
LAURA L. STEIN, B.A., M.A., Ph.D.
JAMES B. STEINBERG, B.A., J.D.
PAUL J. STEKLER, Ph.D.
CHARLES H. STILL, B.B.A., J.D.
SHERINE E. THOMAS, B.A., J.D.
TERRY O. TOTTENHAM, B.S., LL.M., J.D.
JEFFREY K. TULIS, B.A., M.A., Ph.D.
ROBERT W. TURNER, B.A., LL.B.
TIMOTHY J. TYLER, B.A., J.D.
VALERIE L. TYLER, B.J., J.D.
SUSAN S. VANCE, B.B.A., J.D.
LANA K. VARNEY, B.J., J.D.
CLARK C. WATTS, B.A., M.D., M.A., M.S., J.D.
JANE M. WEBRE, B.A., J.D.
RANDALL B. WILHITE, B.B.A., J.D.
DAVID G. WILLE, B.S.E.E., M.S.E.E., J.D.
MARK B. WILSON, B.A., M.A., J.D.
JUDGE PAUL L. WOMACK, B.S., J.D.
NOLAN L. WRIGHT, B.A., M.A., M.L.I.S., J.D.
LARRY F. YORK, B.B.A., LL.B.
DANIEL J. YOUNG, B.A., J.D.

VISITING PROFESSORS

ANTONIO H. BENJAMIN, LL.B., LL.M.
PETER F. CANE, B.A., LL.B., D.C.L.
DAVID ENOCH, LL.B., Ph.D.
IAN P. FARRELL, B.A., LL.B., M.A., LL.M.
VICTOR FERRERES, J.D., LL.M., J.S.D.
FRANCESCO FRANCIONI, J.D., LL.M.

* * *

Texas Law Review

Volume 88 Number 6 May 2010

* * *

Texas Law Review

Volume 88, Number 6, May 2010

ARTICLES

The Incentives Matrix: The Comparative Effectiveness of Rewards, Liabilities, Duties, and Protections for Reporting Illegality
Yuval Feldman & Orly Lobel 1151

Keep Charity Charitable
Brian Galle 1213

The Taking/Taxing Taxonomy
Amnon Lehavi 1235

BOOK REVIEW

Some Moving Parts of Jurisprudence
Kevin Toh 1283

reviewing
BETWEEN AUTHORITY AND INTERPRETATION: ON THE THEORY OF LAW AND PRACTICAL REASON

NOTES

Post-9/11 Anti-terrorism Policy Regarding Noncitizens and the Constitutional Idea of Equal Protection Under the Laws
Daniel H. Cohen 1323

The Conman and the Sheriff: SEC Jurisdiction and the Role of Offshore Financial Centers in Modern Securities Fraud
Nick S. Dhesi 1345

A Comparative "Hard Look" at *Chevron*: What the United Kingdom and Australia Reveal About American Administrative Law
Robert C. Dolehide 1381

* * *

Texas Law Review

Volume 88, Number 6, May 2010

Articles

The Incentives Matrix: The Comparative Effectiveness of Rewards, Liabilities, Duties, and Protections for Reporting Illegality[†]

Yuval Feldman[*] & Orly Lobel[**]

Social enforcement is becoming a key feature of regulatory policy. Increasingly, statutes rely on individuals to report misconduct, yet the incentives they provide to encourage such enforcement vary significantly. Despite the clear policy benefits that flow from understanding the factors that facilitate social enforcement, i.e., the act of individual reporting of illegal behavior, the field remains largely understudied. Using a series of experimental surveys of a representative panel of over 2,000 employees, this Article compares the effect of different regulatory mechanisms—monetary rewards, protective rights, positive obligations, and liabilities—on individual motivation and behavior. By exploring the interplay between internal and external enforcement motivation, these experiments provide novel insights into the comparative advantages of legal mechanisms that incentivize compliance and social enforcement. At the policy-making level, the study offers important practical findings about the costs and benefits of different regulatory systems, including findings about inadvertent counterproductive effects of certain legal incentives. In particular, the findings

---

† The experiments for this Article were supported by a generous grant from the ABA Section on Litigation Research Fund. We thank Larry Alexander, Tsilly Dagan, Lauren Edelman, Nicholas Epply, Shachar Lifshitz, Bryant Garth, Faina Milman, Jacob Nussim, Christine Parker, Gideon Parchomovsky, Oren Perez, Mark Schuman, Susan Sturm, Fred Zacharias, and the participants of faculty workshops at USD, Vanderbilt, Chicago-Kent, the Behavioral Law and Economics Conference, Haifa University, and the Berkeley Conference on Building Theory through Empirical Legal Studies, 2009, for their valuable comments. The paper has been accepted for presentation at the 2010 American Law and Economics Association (ALEA) meeting. We thank Tammy Shterental for statistical consulting and Michelle Butler, Augustine Trezeguet, and Amir Shani for research assistance.

* Senior Lecturer, Faculty of Law, Bar-Ilan University. Ph.D. 2004, UC Berkeley; B.A., LL.B. 1998, Bar-Ilan University.

** Associate Professor, University of San Diego School of Law. S.J.D. 2006; LL.M. 2000, Harvard University; LL.B. 1998, Tel-Aviv University.

indicate that in some cases offering monetary rewards to whistle-blowers will lead to less, rather than more, reporting of illegality. At the more theoretical level, the findings contribute to several strands of inquiry, including motivational crowding-out effects, framing biases, the existence of a "holier-than-thou effect," and gender differences among social enforcers. Together, these findings portray a psychological schema that offers invaluable guidance for policy and regulatory design.

I. Introduction 1153
II. Protect—Command—Fine—Pay: The Legal Structures of Social Reporting 1157
A. Antiretaliation Protections 1161
B. Affirmative Reporting Duties and Liabilities for Failure to Report 1163
C. Monetary Incentives 1168
D. Understanding the Incentive Spectrum 1172
III. Linking Behavioral Economics with New Governance 1173
A. The Hypotheses of the Empirical Study 1176
B. Behavioral Theories on the Interplay between Intrinsic and Extrinsic Motivation 1178
C. The Framing of Legal Dollars 1181
IV. The Expressive Function of Law 1183
V. Misperception of Social Support for Blowing the Whistle 1185
VI. The Experiment 1187
A. Design and Sample 1187
1. Experimental Design 1187
2. The Sample 1189
B. Findings 1189
1. Regulatory Mechanisms and the Holier-Than-Thou Effect. 1189
2. The Interplay Between Internal and External Motivation 1192
3. Gender and Whistle-Blowing. 1196
4. Real Versus Perceived Motivation to Blow the Whistle 1198
C. Framing the Social Meaning of Whistle-Blowing 1200
VII. Discussion and Policy Implications 1201
A. The Inadvertent Effect of Rewards 1202
B. Motivation, Perception, and Whistle-Blowing 1202
C. One Policy Fits All? 1203
D. Incommensurability 1204
E. The Choice of Mechanism and Whistle-Blowing Status 1205
VIII. Conclusion 1206
Appendix I: Statistical Tables 1208
Appendix II: Excerpts from the Survey Instrument 1210

A. The Survey 1210
1. *Condition 1 (Duty + Low Reward)* 1210
2. *Condition 2 (Duty + High Reward)* 1211
3. *Condition 3 (Low Reward)* 1211
4. *Condition 4 (High Reward)* 1211
5. *Condition 5 (Duty + Protection)* 1211
6. *Condition 6 (Protection)* 1211
7. *Condition 7 (Duty)* 1211
8. *Condition 8 (Duty + Fine)* 1211

I. Introduction

Questions about social enforcement and the role of individual reporting in preventing corporate and governmental misconduct are at the forefront of current debates and reforms. Most recently, the 2009 stimulus bill introduced by President Obama for the recovery of our troubled economy includes elaborate antiretaliation rights for whistle-blowers who report financial misconduct.[1] Dozens of existing federal statutes and hundreds of state statutes include similar whistle-blower protections or incentives in a vast range of fields including tax regulation, environmental law, employment discrimination, health and safety, and trading standards.[2] Indeed, all regulatory systems have built-in mechanisms designed to promote legal compliance. However, the variation among these regulatory mechanisms and incentives is immense.[3] Some statutes are designed to protect employees against retaliation when they resist or report illegal activities.[4] Other statutes state an obligation of the individual to report and, at times, impose penalties for failure to report.[5] Yet another class of incentive-based systems encourages reporting by sharing part of the funds recovered from a report of corporate fraud.[6] In addition to the vast differences among the statutes themselves, there is significant debate about the application of the various laws. Most notably, recent case law interpreting whistle-blower protections

1. American Recovery and Reinvestment Act of 2009, Pub. L. No. 111-5, sec. 1553, 123 Stat. 115, 297–302. The whistle-blower protection clauses are often referred to as the "McCaskill Amendment." *E.g.*, Orly Lobel, *Citizenship, Organizational Citizenship, and the Laws of Overlapping Obligations*, 97 CAL. L. REV. 433, 455 n.135 (2009); Daphne Eviatar, *McCaskill Proposes Protection for Government Contractor Whistleblowers*, WASH. INDEP., Feb. 4, 2009, *available at* http://washingtonindependent.com/29016/mccaskill-proposes-protection-for-govt-contractor-whistleblowers.

2. *See infra* note 30 and accompanying text.

3. *See* Lobel, *supra* note 1, at 445, 447 (arguing that whistle-blower statutes vary widely regarding scope, content, and protection).

4. *See infra* subpart II(A).

5. *See infra* subpart II(B).

6. *See infra* subpart II(C).

has brought the field to a state of flux.[7] For example, in a recent decision decried as "the worst Supreme Court ruling on whistleblowing,"[8] the U.S. Supreme Court, in a 5–4 decision, refused to extend constitutional protections to employees who report illegal conduct when such reporting is "pursuant to their official duties."[9] Other recent cases similarly reveal a deep ambivalence and uncertainty about the role of individuals in resisting illegality in organizations.[10]

The legislative and adjudicative variations in the field of social enforcement—the act of individual reporting of illegality—indicate great uncertainty and undertheorizing about the comparative advantages and effectiveness of various reporting channels, protections, and incentives that affect the decision of individuals to report illegal conduct. In this Article, we offer new insights on the psychology of social enforcement. The Article presents original empirical research examining how incentives and variance in regulatory mechanisms affect individual motivation and behavior. By exploring the interplay between internal and external enforcement motivations, these experiments provide novel insights into the comparative advantages of legal mechanisms that incentivize compliance and social enforcement. Our findings reveal important differences in the effectiveness of existing mechanisms designed to incentivize reporting. Most strikingly, our findings suggest that legal incentives to report are frequently ill-designed and can in fact be inadvertently counterproductive.

Using a series of experimental surveys of a representative panel of more than 2,000 employees, the empirical study examines four prototypical legal mechanisms designed to promote individual reporting: (1) Antiretaliation Protection; (2) Duty to Report; (3) Liability Fines; and (4) Monetary Incentives. The experiments measure the value individuals attach to different types of regulatory mechanisms in deciding whether to react to illegality within their work environment. Through interactive regressions, the study further investigates the relationships between such incentives and the moral and practical considerations that underlie individual decisions to respond to a certain type of mechanism. The empirical study offers methodological advantages over existing studies. Departing from most empirical studies in the fields of behavioral economics and social psychology, the study offers a unique focus on the effect of regulatory approaches on individual behavior. Existing empirical research has largely neglected the role of the surrounding legal regimes as affecting individual behavior. Building on prior studies by

---

7. *See* Lobel, *supra* note 1, at 445–55 (discussing the disparate judicial interpretation of whistleblower statutes among the several states and the federal government).

8. Joyce Howard Price, *Justices Ease Whistleblower Protections*, WASH. TIMES, May 31, 2006, at A1 (internal quotations omitted).

9. Garcetti v. Ceballos, 547 U.S. 410, 421 (2006).

10. *See* Lobel, *supra* note 1, at 434 (asserting that there is judicial ambivalence about the role of individuals in exposing illegality in group settings); *infra* notes 56–58, 96–100, 135–38 and accompanying text.

the collaborators,[11] the study fills this research gap and argues that motivation to report is multidimensional and correlated with the operative legal incentive. The study thus develops a unique lens combining the study of behavioral economics with insights from new-governance theory, the school of thought that focuses on the significance of regulatory design. This lens enables the identification of problems in contemporary regulatory mechanisms. Legislators, as well as adjudicators, must consider tailoring the incentives embedded in the law to the misconduct and the individual that it targets as an enforcer. A major obstacle to such a comprehensive approach has been the difficulty of translating such factors into policy, including dollar metrics and effective incentive mechanisms. By and large, decisions about how to design whistle-blowing incentives and protections have not been grounded in a coherent set of valuations. The experiments that form the basis of this Article allow for the development of principled and effective legal design.

From a practical perspective, the Article presents new evidence about key factors that determine whether or not people will actively report and resist illegal conduct. Our findings suggest that a systematic approach to regulation must include an understanding of the fit between the adopted law, the misconduct it addresses, and the individuals it aims to incentivize. On a broader level, the findings of the study contribute to the scientific body of knowledge in a range of key social-science debates, including crowding-out effects, the implications of framing biases to the expressive function of the law, the "holier-than-thou effect," and gender differences in motivation and action. For policy development, the results suggest that in laws that are likely to trigger strong internal ethical motivation, offering monetary rewards may be unnecessary or, worse yet, counterproductive. In such circumstances, where legal violation is generally perceived as morally offensive, creating a duty to report may be sufficient. Where potential informants lack a moral imperative to report, our findings further indicate that offering low rewards is the worst mechanism that regulators can offer, as it neither motivates high levels of reporting nor is perceived by most individuals as constituting good citizenship behavior. In fact, offering low rewards triggers less reporting than merely offering protection or establishing a duty. Thus, the findings suggest that many existing laws may have inadvertent counterproductive effects by offering monetary incentives rather than triggering internal motivations of potential reporting individuals. More generally, this suggests that framing the reporting behavior as a commodity may actually crowd out, or suppress, internal moral motivation.

The findings further indicate important interactions between different types of legal incentives and the demographics of individuals for which they

11. *See* Yuval Feldman & Orly Lobel, *Decentralized Enforcement in Organizations: An Experimental Approach*, 2 REG. & GOVERNANCE 165, 171–81 (2008) (demonstrating the impact of institutional processes on individual decisions about whether to blow the whistle on illegality).

are designed, including gender, levels of income, job status, and professional roles. In particular, the results clearly show that while men care significantly more than women about the size of the monetary reward, women care more about protection against retaliation, as well as the impositions of a legal duty. Given the complexity observed in these interactions, we argue in this Article that that there is no one-size-fits-all solution for policy design. Rather, policy makers must consider the characteristics of the target population of social enforcers and incentivize them accordingly.

At a theoretical level, the findings reveal pervasive gaps in how individuals perceive the motivations driving their own behavior and the actions of the general population. In particular, the findings provide significant evidence of a holier-than-thou effect, in which individuals believe that their own morality is stronger than that of their peers and strangers. Overwhelmingly, the respondents in the study perceived their own reporting behavior as being more motivated by intrinsic ethical concerns than the actions of others. Unlike their own morally driven actions, individuals predicted that others would behave according to self-interest and would be motivated primarily by external rewards provided through law. However, this effect was reduced among women participants, who were found to be more ethically motivated and more confident about the ethical motivations of other female participants. Furthermore, the findings point to a general gap between people's perception of the motivation for their behavior and the actual effect of different incentives on their behavior. Respondents tend to overestimate their internal moral motivation and underestimate the extent to which monetary rewards dictate their behavior.

The Article proceeds as follows. Part II of the Article introduces the range of mechanisms that are employed by our legal system to encourage whistle-blowing, including antiretaliation protections, duties to report, liability fines, and monetary incentives. It further introduces the contemporary realities of social enforcement and the importance of incentives in helping whistle-blowers overcome their fears to undertake such action. Part III explains the framework of the study, which links behavioral economics with new-governance theory. The study is based on the understanding that decisions to report illegality will be affected by the individual characteristics of potential whistle-blowers and the legal and organizational environment in which they operate.[12] Building on prior studies of the collaborators, we argue that these aspects must be linked in order for experimental findings to offer policy insights. Part III thus provides an overview of the various behavioral theories and empirics on how incentives may

12. *See id.* at 175–81 (evaluating the effects of cultural variations on the probability of employees reporting illegal practices, including the following: individualism versus group solidarity; degree of job security; length of legal tradition of enforcement and types of legal protection; management responses to claims; and individual characteristics, such as gender and level of moral outrage).

affect individual action. These include theories about crowding-out effects in the interaction between internal and external motivation, framing biases when costs are presented as fines or rewards, and behavioral attitudes toward duties and protections. Part IV continues the theoretical framework of behavioral analysis by adding the dimension of expressive theories of law, hypothesizing that law can induce reporting by making certain behaviors salient and valued. Part V then offers a caveat about the ability of people to predict their own reporting behavior and discusses our hypothesis of a holier-than-thou effect in the context of legal compliance systems. Part VI of the Article presents the experimental design and the findings of the study. Part VII discusses the implications of these findings for both theory and practice. We conclude with a few suggestions for future research.

A better understanding of individual motivation and behavior can improve the currently chaotic choices between protection-based, duty-based, liability, and monetary incentives for reporting illegality. If policy makers knew which legal mechanisms trigger reporting action, a more tailored approach could be designed to provide employees with the needed motivation. Despite the developments in the legal protection of whistle-blowers, there is inadequate knowledge of the factors that contribute to effective private efforts to assist regulatory compliance and the ways in which behavioral economics can predict the comparative success of such efforts. The lack of empirical knowledge contributes to the many inconsistencies in legal protections and enforcement strategies in policy and adjudication. More generally, current debates about the desirability and effectiveness of private-enforcement approaches and their ability to replace traditional command-and-control regulation[13] would be better informed by more empirical knowledge about private individual behavior. At the broadest level, more knowledge about the behavior of individuals in reaction to illegality offers an important scholarly contribution to the interdisciplinary study of motivation, cooperation, norms, and institutional design.

II. Protect—Command—Fine—Pay: The Legal Structures of Social Reporting

The decision of whether to blow the whistle is a complex one and inevitably involves certain risks. In any organizational setting, employees

13. *Compare* CHRISTINE PARKER, THE OPEN CORPORATION: EFFECTIVE SELF-REGULATION AND DEMOCRACY 29–30 (2002) (discussing the promise of and strategies for corporate self-regulation as opposed to traditional command-and-control legislation), *and* Cynthia Estlund, *Rebuilding the Law of the Workplace in an Era of Self-regulation*, 105 COLUM. L. REV. 319, 377–83 (2005) (voicing new strategies for employee empowerment in the new age of self-regulation), *with* Anthony Ogus, *Rethinking Self-regulation*, 15 OXFORD J. LEGAL STUD. 97, 98–99 (1995) (summarizing traditional criticisms of self-regulation), *and* Darren Sinclair, *Self-Regulation Versus Command and Control? Beyond False Dichotomies*, 19 LAW & POL'Y 529, 532–33 (1997) (advocating for a regulatory approach that blends self-regulation and command-and-control regulation).

and managers must decide whether to take on some of these risks in order to report wrongdoing or, instead, ignore or participate in the illegal behavior. Empirical studies indicate that most employees will choose to "suffer in silence" in the face of wrongdoing for fear of retaliation in the form of termination and harassment.[14] Indeed, some of the biggest legal scandals were known to insiders long before they became public. For example, today, evidence exists that the causal link between asbestos and lung disease was clearly known to the manufacturing companies as early as 1924.[15] Yet, it was not until decades later that product-liability lawsuits were successfully launched against the industry, following years of active suppression of the damaging information within the companies.[16] In addition to direct employment retaliation, reporting often entails psychological and societal costs, including fear, guilt, and mistreatment by peers and community.[17] The fears of whistle-blowers have been substantiated by recent data showing that external whistle-blowers often experience retaliation by their supervisors and are shunned by their social circles.[18] One commentator has described whistle-blowing as "professional suicide."[19]

---

14. On "suffering in silence," see Brian Barry, *Review Article: 'Exit, Voice, and Loyalty,'* 4 BRIT. J. POL. SCI. 79, 97 (1974), for the assertion that if a consumer is too loyal to blow the whistle, and the status quo is more desirable than exiting the situation, then the employee will choose to "stay and be silent." For an empirical study indicating low levels of reporting, see Terance D. Miethe & Joyce Rothschild, *Whistleblowing and the Control of Organizational Misconduct*, 64 SOC. INQUIRY 322, 330–33 (1994), showing that across five studies of employee reactions to misconduct, on average, 58% remained silent and only 21% reported misconduct outside the company.

15. *See* P W J Bartrip, *History of Asbestos Related Disease*, 80 POSTGRAD MED. J. 72, 72 (2004) ("The first medical article on the hazards of asbestos dust appeared in the *British Medical Journal* in 1924."); Morris Greenberg, *Knowledge of the Health Hazards of Asbestos Prior to the Merewether and Price Report of 1930*, 7 SOC. HIST. OF MED. 493, 501 (1994) (stating that a wealth of evidence was available in the 1920s indicating that asbestos was associated with severe respiratory disease).

16. *See, e.g.*, Alan F. Westin, *Introduction* to WHISTLE BLOWING! LOYALTY AND DISSENT IN THE CORPORATION 1, 11–12 (Alan F. Westin et al. eds., 1981) (describing the pervasive suppression of medical reports and other industry findings linking asbestos to lung disease).

17. *See* FRANK ANECHIARICO & JAMES B. JACOBS, THE PURSUIT OF ABSOLUTE INTEGRITY: HOW CORRUPTION CONTROL MAKES GOVERNMENT INEFFECTIVE 64–69 (1996) (cataloguing past efforts to strengthen the federal and New York state protections for whistle-blowers, which were motivated in part by concerns for the harassment and collateral costs often incurred independent of initial retaliation); MARCIA P. MICELI & JANET P. NEAR, BLOWING THE WHISTLE: THE ORGANIZATIONAL & LEGAL IMPLICATIONS FOR COMPANIES AND EMPLOYEES 79–89 (1992) (describing hostile reactions to whistle-blowers by others within the organization and management).

18. Natalie Dandekar, *Contrasting Consequences: Bringing Charges of Sexual Harassment Compared with Other Cases of Whistleblowing*, 9 J. BUS. ETHICS 151, 152 (1990); Westin, *supra* note 16, at 2–3. *But see* Janet P. Near et al., *Explaining the Whistle-Blowing Process: Suggestions from Power Theory and Justice Theory*, 4 ORG. SCI. 393, 398 (1993) ("Contrary to the popular perception, most whistle-blowers do not suffer retaliation, at least among federal employees . . . .").

19. James Gobert & Maurice Punch, *Whistleblowers, the Public Interest, and the Public Interest Disclosure Act 1998*, 63 MOD. L. REV. 25, 35 (2000).

In the past, popular culture has generally portrayed whistle-blowers as "lowlife[s] who betray[] a sacred trust largely for personal gain."[20] In recent years, however, the act of whistle-blowing has been reshaped in the media as a heroic act that can bring deeply corrupt practices to a halt. In 2002, the whistle-blowers of the WorldCom and Enron financial debacles along with the government whistle-blower from the FBI were featured on the cover of *Time* magazine as the "Persons of the Year."[21] Moreover, whistle-blowing has become a focal point in legislative reform as a key means to preventing corporate illegality.[22] In a myriad of contexts, ranging from the prevention of financial misconduct to the prevention of discrimination and pollution, legislators and policy makers attempt to encourage individuals to step forward and report illegal conduct.[23]

Because of its inherent risks, whistle-blowing must be incentivized through regulatory policies that will encourage individuals to break the code of silence in corrupt organizations. However, identifying and understanding the various predictors of social enforcement in organizations is highly complex, as predictors are comprised of individual, organizational, and state-level factors.[24] This complexity has led policy makers to use a variety of enforcement and compliance strategies.[25] Corporations themselves have also implemented internal strategies for encouraging reporting, often with the promise that an employee will not be retaliated against for using these channels.[26] Some corporations even offer reward systems to their employees for reporting illegalities such as discrimination or harassment.[27] Taken together, this spectrum of regulatory strategies underscores the inherent

20. TERANCE D. MIETHE, WHISTLEBLOWING AT WORK: TOUGH CHOICES IN EXPOSING FRAUD, WASTE, AND ABUSE ON THE JOB 12 (1999).

21. *Persons of the Year: The Whistleblowers*, TIME, Dec. 30, 2002, at cover.

22. *See* Richard E. Moberly, *Sarbanes–Oxley's Structural Model to Encourage Corporate Whistleblowers*, 2006 BYU L. REV. 1107, 1108 (characterizing recent whistle-blower legislation as a response to corporate scandals meant to "encourage employees to become more effective corporate monitors").

23. *See* William Dorsey, *An Overview of Whistleblower Protection Claims at the United States Department of Labor*, 26 J. NAT'L ASS'N ADMIN. L. JUDICIARY 43, 48–50 (2006) (giving examples of statutes protecting individuals that report unlawful behavior such as unsafe environmental practices, discrimination in the trucking industry, and corporate fraud).

24. *See* Feldman & Lobel, *supra* note 11, at 171–74 (detailing the litany of factors found to affect social-enforcement choices made by individuals in whistle-blowing contexts including individual, organizational, and state-level factors).

25. *See* Lobel, *supra* note 1, at 467 (indicating that administrative agencies have built upon principles from established business strategies and regulatory approaches to expand their collaborative compliance programs, increasingly relying on internal self-regulation to complement traditional adversarial enforcement).

26. *See id.* at 496–97 (remarking on a "growing understanding that, in order to motivate employees to respond to unlawful behavior, employers must create procedures that allow third-party review and impartial judgments" and providing examples of such procedures).

27. *See id.* at 443–44 (describing "bounty programs" that offer monetary incentives for employees who externally report illegal behavior).

complexity of this area of law and provides capacity for fine-tuning these mechanisms to the particular problems they seek to address.

Increasingly, state and federal agencies such as the Securities and Exchange Commission (SEC), the Occupational Safety and Health Administration (OSHA), and the Environmental Protection Agency (EPA) rely on private enforcement.[28] This has meant that whistle-blowers have become a vital part of any "democratic, free enterprise system."[29] On the legislative side, there are dozens of federal statutes containing whistle-blower provisions, and many states have a statutory scheme in place that outlines whistle-blower protections.[30] Some of these statutes are general whistle-blowing laws,[31] while others are designed to protect employees who blow the whistle on specific types of alleged misconduct such as environmental pollution,[32] safety and health violations,[33] and financial fraud.[34] While this network of statutory provisions covers a significant range of social reporting, it is also "riddled with loopholes" that may leave individuals who report illegality unexpectedly vulnerable in certain cases.[35] Legislators and courts constantly struggle with defining which reporting activities should be protected either constitutionally or by statute.[36] In sum, although the significance of social enforcement and regulatory incentives is striking, there is little knowledge on the comparative advantages of the myriad regulatory tools available for providing such incentives. Despite this vast complexity, however, the current landscape of incentive programs nonetheless reveals several prototypes that may provide some structure to the regulatory toolbox. The most widely used strategies are providing employees with antiretaliation protections, creating a duty to report, imposing liability for failure to report, and incentivizing reporting with money. Some statutes include several of these legal categories, whereas others offer only one of these alternatives.

---

28. Feldman & Lobel, *supra* note 11, at 167–68.

29. Winters v. Houston Chronicle Publ'g Co., 795 S.W.2d 723, 730 (Tex. 1990).

30. STEVEN M. KOHN, CONCEPTS AND PROCEDURES IN WHISTLEBLOWER LAW 1 (2001).

31. *Id.* at 30, 99–104.

32. *Id.* at 141–42.

33. *Id.* at 94–95, 228–29.

34. *Id.* at 81–82.

35. *Id.* at 79.

36. *See* Lobel, *supra* note 1, at 447–50 (describing the addition of protections for internal whistle-blowing activities to classic external reporting protections, the widely varied scope of protections afforded to employees by legislatures, and the lack of consensus by jurists on the boundaries of protected employee actions). For examples of how other countries are similarly in the process of rethinking their whistle-blower laws and reforming their compliance and enforcement strategies, see Matthias Schmidt, *Whistle Blowing Regulation and Accounting Standards Enforcement in Germany and Europe—An Economic Perspective*, 25 REV. L. & ECON. 143, 153–61 (2005), comparing the status quo of U.S., U.K., and German whistle-blowing statutory schemes and discussing possible expansions of both external and internal whistle-blowing protections; and Elletta Sangrey Callahan et al., *Australian, U.K., and U.S. Approaches to Disclosure in the Public Interest*, 44 VA. J. INT'L L. 879, 889–95 (2004), contrasting the internal and external whistle-blowing protections afforded by U.S., U.K., and Australian statutory schemes.

A. Antiretaliation Protections

Like many other statutes in the fields of environmental, consumer, and financial regulation, the 2009 stimulus bill offers traditional antiretaliation protection: nonfederal employers may not retaliate against an individual who reasonably believes that there has been a legal violation in her organization and takes action to report the violation.[37] As such, the reporting individual is protected by law against any adverse action by her superiors, be it firing, demotion, or acts of harassment. Many of these reporting protections were developed in response to corporate scandals, such as Enron and WorldCom, where employers' retribution threats persuaded employees to "swallow the whistle."[38] Consequently, these statutes are designed to provide individuals with broad antiretaliation measures.[39] In contrast to the whistle-blower protections for federal employees, the American Investment and Recovery Act explicitly extends protection for disclosures made during the course of an employee's duties.[40] Another important example is found in the 2002 Sarbanes–Oxley Act (SOX), which protects corporate whistle-blowers when they report financial misconduct to the SEC or internally within their organization.[41] Hailed by scholars as "the gold standard" of whistle-blower protection,[42] SOX provides civil remedies for individuals who experience retaliation in reaction to such reporting and makes it a felony to act against such an individual.[43]

Despite their prominence, existing legal protections for reporting misconduct are largely unsettled and heatedly debated.[44] Antiretaliation protections have developed as a patchwork of state and federal statutory and common law exceptions to the employment-at-will regime, a century-old default rule that has allowed employers to terminate their employees "for

---

37. American Recovery and Reinvestment Act of 2009, Pub. L. No. 111-5, sec. 1553, 123 Stat. sec. 115, 297–302.

38. *See* Feldman & Lobel, *supra* note 11, at 168 ("[I]n response to recent corporate scandals . . . such as Enron and WorldCom, both federal and state legislatures have been strengthening whistle-blower protections."); *see also* Charles Derber, *Managing Professionals: Ideological Proletarianization and Mental Labor*, *in* PROFESSIONALS AS WORKERS: MENTAL LABOR IN ADVANCED CAPITALISM 167, 177 (Charles Derber ed., 1982) (labeling employees who decide not to report organizational misconduct as "swallow[ing] the whistle").

39. *See* American Recovery and Reinvestment Act sec. 1553(a) (prohibiting reprisals for a broad range of good-faith reporting conduct); *id.* sec. 1553(b) (establishing additional procedural protections during the investigation of reprisal complaints); *id.* sec. 1553(c)(3)–(5) (defining remedies and providing further administrative and judicial relief and appeals).

40. *Id.* sec. 1553(a).

41. Sarbanes–Oxley Act of 2002 § 806, 18 U.S.C. § 1514A (2006).

42. *E.g.*, Cynthia Eastland, *Rebuilding the Law of the Workplace in an Era of Self-regulation*, 105 COLUM. L. REV. 319, 376 (2005).

43. 18 U.S.C. §§ 1513(e), 1514A.

44. *See, e.g.*, Margit Cohn, *Fuzzy Legality in Regulation: The Legislative Mandate Revisited*, 23 LAW & POL'Y 469, 472 (2001) (claiming that no legislative mandate is necessary for regulators in a centralized government and inferring that future regulation by such regulators may disregard schemes considered settled by previous legislative action).

good cause, for no cause, or even for morally wrong cause."[45] As early as the 1930s, and even more significantly since the 1960s and 1970s, legislatures have carved away at this default by enacting laws that grant employees rights against discharge.[46] These federal statutes include antiretaliation provisions designed to enable employees to claim their rights and report illegal conduct without fear of retribution.[47] These statutes encompass a broad range of regulatory fields, including financial,[48] environmental,[49] consumer,[50] health,[51] and safety regulation.[52]

As a parallel development to legislative protections for whistle-blowers, courts have also developed the tort of wrongful termination, which allows plaintiffs to overcome the hurdle of at-will employment by claiming they were discharged for engaging in social enforcement in the face of corporate misconduct.[53] Thus, even in a context where there exists no statute that provides antiretaliation protection, courts have frequently held that individuals cannot be terminated by their employer for reporting legal violations.[54] In fact, retaliation is the fastest growing type of employment-law claim.[55] However, courts significantly disagree over the scope of such protections.[56]

---

45. Payne v. W. & Atl. R.R. Co., 81 Tenn. 507, 519–20 (1884), *overruled by* Hutton v. Watters, 179 S.W. 134, 137 (1915).

46. *See, e.g.*, 29 U.S.C. § 158 (2006) (protecting workers from "unfair labor practices," such as discharge for union membership or activities); 42 U.S.C. § 2000e-2(a)(1) (2006) (prohibiting employment discharge based on race, color, religion, sex, or national origin).

47. *See, e.g.*, 5 U.S.C. § 2302(b)(8)(A) (2006) (protecting federal employees from discharge or retaliation for good-faith whistle-blowing); 29 U.S.C. § 2002 (prohibiting employers from retaliating against employees who refuse to take a lie-detector test).

48. *See supra* notes 38–42 and accompanying text.

49. *See, e.g.*, 42 U.S.C. § 5851(a) (prohibiting retaliation against employees who report violations of the Atomic Energy Act of 1954).

50. *See, e.g.*, 15 U.S.C.A. § 2087(a)(1) (West 2009) (prohibiting employers from discharging or discriminating against an employee who reports a violation of the Consumer Product Safety Act).

51. *See, e.g.*, 20 U.S.C. § 3608 (2006) (prohibiting retaliation against whistler-blowers who report asbestos violations).

52. *See, e.g.*, 29 U.S.C. § 660(c) (protecting workers from retaliation for reporting potential violations under OSHA).

53. *See, e.g.*, Perks v. Firestone Tire & Rubber Co., 611 F.2d 1363, 1365 (3d Cir. 1979) (holding that firing an employee for exercising the statutory right not to submit to a polygraph test gives rise to a cause of action for tortious discharge); Reuther v. Fowler & Williams, Inc., 386 A.2d 119, 121–22 (Pa. Super. Ct. 1978) (holding that an employee fired for performing jury duty had a cause of action for wrongful termination).

54. *See, e.g.*, Ostrofe v. H.S. Crocker Co., 670 F.2d 1378, 1383–84 (9th Cir. 1982) (upholding claim of wrongful discharge for objecting to employer's violation of the Clayton Act, despite absence of an antiretaliation clause in order to promote "interests of antitrust enforcement").

55. For example, the EEOC reports that retaliation charges for discrimination complaints have nearly doubled in the past decade. *See* U.S. EQUAL EMPLOYMENT OPPORTUNITY COMM'N, CHARGE STATISTICS FY 1997 THROUGH FY 2009 (2009), *available at* http://www.eeoc.gov/eeoc/statistics/enforcement/charges.cfm (noting 19,694 retaliation charges in fiscal-year 1999 and 33,613 retaliation charges in fiscal-year 2009). In fiscal-year 2009, 36% of all discrimination claims contained retaliation charges. *Id.*

56. *See* Lobel, *supra* note 1, at 433 (discussing controversial whistle-blower cases).

At the constitutional level, the Supreme Court remains sharply divided on what kind of reporting by public employees is constitutionally protected, recently holding in a split 5–4 decision that "when public employees make statements pursuant to their official duties, the employees are not speaking as citizens for First Amendment purposes, and the Constitution does not insulate their communications from employer discipline."[57] At the tort level, courts vary in the extent to which they are willing to extend antiretaliation protections to different channels of reporting, different types of reported misconducts, and different categories of workers.[58]

B. Affirmative Reporting Duties and Liabilities for Failure to Report

Affirmative duties to report illegality have been imposed in several contexts including child abuse, elder abuse, domestic violence, environmental offenses, and financial crimes.[59] For the most part, duties to report are limited to either senior corporate officers or to members of certain professions such as lawyers,[60] accountants,[61] doctors,[62] and teachers.[63] At times, even when a duty is not explicitly assigned by law or corporate policy, courts may infer that individuals have an obligation to report misconduct or mismanagement.[64] These duties and liabilities are an exception to the general proposition that the law does not punish omission.[65] Therefore, duties to actively report are imposed in situations where the victim of misconduct is particularly vulnerable or the harm will be widespread.[66] Many of these duties impose criminal liabilities, and in many states, those required to report may be held civilly liable for their failure to report.[67] The statutes vary in the level of evidence at which the duty is imposed and the channels of reporting required.[68]

---

57. Garcetti v. Ceballos, 547 U.S. 410, 421 (2006).

58. *Id.* at 444–55.

59. *See* Sandra Guerra Thompson, *The White Collar Police Force: "Duty to Report" Statutes in Criminal Law Theory*, 11 WM. & MARY BILL RTS. J. 3, 3 (2002) (discussing the duty-to-report statutes in each of these areas and locating them in criminal law theory).

60. *See* MODEL RULES OF PROF'L CONDUCT R. 8.3(a) (2006) (laying out lawyers' duty to report professional misconduct).

61. *See* 15 U.S.C. § 78j-1(b)(3)(B) (2006) (requiring accounting firms to report a client corporation's illegal activities to the SEC if the board of directors fails to do so).

62. *See, e.g.*, ARIZ. REV. STAT. ANN. § 13-3620 (2009) (imposing upon doctors, teachers, and others a duty to report suspected child abuse).

63. *Id.*

64. Thompson, *supra* note 59, at 36–37.

65. *See, e.g.*, Williamson M. Evers, *The Law of Omissions and Neglect of Children*, 2 J. LIBERTARIAN STUD. 1, 1 (1978) (distinguishing between a legally binding obligation and one that is solely moral).

66. Thompson, *supra* note 59, at 37.

67. *Id.* at 18–19.

68. *Id.* at 16–17.

In recent legislation following the early twenty-first century financial debacles, a range of affirmative duties was put into place. For example, the affirmative reporting duties imposed under SOX extend to both attorneys and executives of public companies subject to SEC proceedings or investigations.[69] Under these compelled whistle-blower provisions, attorneys representing companies in an SEC proceeding must conduct an internal investigation into any evidence of securities fraud and report any relevant findings to the company's chief legal officer or its CEO.[70] The appropriate executive officer must then review the evidence and, upon finding any violations, must assume the duty of correcting the illegality within the company.[71] Where the officer fails to conduct an adequate investigation, the reporting attorney is then further required to report the evidence to the company's audit committee or board of directors.[72] Attorneys and executives who fail to comply in good faith with these SEC rules may be subject to civil liability for securities fraud.[73] However, by limiting the positive reporting obligations to the highest ranks within the corporation, SOX has narrowed its application to those actors who are most able to bear the potential retaliatory costs as well as prevent the violations at the earliest stage possible.[74]

Companies that handle environmentally hazardous substances are also subject to strict reporting duties under the relevant provision of the Comprehensive Environment Response, Compensation, and Liability Act (CERCLA).[75] Enacted in 1980, this act provides the most significant federal environmental reporting requirements for actual or threatened releases of hazardous substances. Section 9603(a) imposes an affirmative reporting duty on "[a]ny person in charge of a vessel or an offshore or an onshore facility"

---

69. Standards of Professional Conduct for Attorneys Appearing and Practicing Before the Commission in the Representation of an Issuer, 17 C.F.R. § 205.3 (2003); *see also* Orly Lobel, *Lawyering Loyalties*, 77 FORDHAM L. REV. 1245, 1265 (2009) ("[T]he Sarbanes–Oxley Act changed the attitude toward the obligations of attorneys representing publicly traded corporations. The SEC rules now permit lawyers to disclose a client's 'material violation' to the Commission, and failure to do so may carry significant sanctions."). In general, the ABA rules do not impose an affirmative duty to report; most of the reporting duties for in-house counsel are too qualified or discretionary to be considered mandatory. *See* MODEL RULES OF PROF'L CONDUCT R. 1.13(b) (2009) (requiring in-house counsel to report certain organizational misconduct but qualifying that such reporting is not necessary if the in-house counsel "reasonably believes that it is not necessary in the best interest of the organization to do so"). The only situation where an attorney has an unqualified duty to report involves peer misconduct, but even these provisions are vaguely limited to violations that raise "a substantial question" about another attorney's "honesty, trustworthiness or fitness as a lawyer." *Id.* R. 8.3.

70. 17 C.F.R. § 205.3.

71. *Id.*

72. *Id.*

73. *Id.* § 205.6.

74. *See* Elizabeth C. Tippett, *The Promise of Compelled Whistleblowing: What the Corporate Governance Provisions of the Sarbanes–Oxley Act Mean for Employment Law*, 11 EMP. RTS. & EMP. POL'Y J. 1, 50 (2007) (applying the Calabresian theory of torts, which holds that tort liability should be allocated to the cheapest cost avoider).

75. 42 U.S.C. § 9603 (2006).

that has knowledge of a hazardous release in excess of the permitted quantities.[76] The supervisor must immediately notify the National Response Center, which then assumes the remainder of the reporting duties.[77] These reports must ordinarily not exceed fifteen minutes after the person has knowledge of the hazardous release.[78] Failure to immediately report a release is a felony punishable by up to three years in prison for the first offense and up to five years for any subsequent convictions.[79] The EPA has categorized the severity of the violation along a three-tier scale, depending on the total delay in reporting the violation.[80] Courts have broadly extended the scope of the statute to apply equally to corporations and individual officers.[81] As the Eighth Circuit explained, "construction of CERCLA to impose liability upon only the corporation and not the individual corporate officers and employees who are responsible for making corporate decisions about the handling and disposal of hazardous substances would open an enormous, and clearly unintended, loophole in the statutory scheme."[82]

Banks and financial institutions maintain similar reporting obligations for any suspicious activity relating to money laundering and insider fraud. Under the terms of the Annunzio-Wylie Act (1992),[83] a reporting duty is imposed through Suspicious Activity Reports (SARs) that must be filed with the Financial Crimes Enforcement Network of the Department of the Treasury for any "known or suspected violation of Federal law or a

76. *Id.* § 9603(a).

77. *Id.*

78. EPA, ENFORCEMENT RESPONSE POLICY 12 (1999), *available at* http://www.epa.gov/compliance/resources/policies/civil/epcra/epcra304.pdf.

79. 42 U.S.C. § 9603(b).

80. EPA, *supra* note 78, at 11–13. The lowest penalties are issued for reports coming in within an hour but after the fifteen-minute notification period, intermediate penalties are issued for reports coming between one and two hours, and the highest penalties are reserved for reports exceeding two hours after the individual had knowledge of the release. *Id.* at 12–13. To avoid deterring reports involving potential self-incrimination, CERCLA includes an immunity provision that prevents the reported material from being used in any criminal case, except a prosecution for perjury or for giving a false statement. 42 U.S.C. § 9603(b). This immunity does not extend to civil liability, as self-reporters may still be liable for civil fines or damages to third parties. *Id.*

81. *See, e.g.*, United States v. Ne. Pharm. & Chem. Co., Inc., 810 F.2d 726, 744 (8th Cir. 1986) (holding that Illinois corporate law did not shield two corporate officers from their reporting liabilities under CERCLA where they exercised direct control over a hazardous spill). Some courts have limited this liability to cases where the corporate officer directly participated in the hazardous release. *See, e.g.*, United States v. Conservation Chem. Co., 619 F. Supp. 162, 190 (W.D. Mo. 1985) (finding that corporate officials who actively participate in the management of a disposal facility can be held liable under CERCLA); United States v. Mottolo, 629 F. Supp. 56, 59–60 (D.N.H. 1984) (holding that under CERCLA, individual officers may be held liable for tortious activity of their corporations if the officers participated in the activity). Other courts have imposed liability where the officer was in a position to prevent the hazardous conduct. *See, e.g.*, Michigan v. ARCO Indus. Corp., 723 F. Supp. 1214, 1219 (W.D. Mich. 1989) (requiring proof that an officer could have prevented or abated hazardous waste discharge in order to impose individual liability under CERCLA).

82. *Ne. Pharm.*, 810 F.2d at 743.

83. 31 U.S.C. § 5318 (2006).

suspicious transaction related to a money laundering activity or a violation of the Bank Secrecy Act."[84] For crimes requiring immediate attention, banks incur the additional duty of immediately notifying an appropriate law enforcement authority and the OCC in addition to filing a timely SAR.[85] Failure to file an SAR may subject individuals at all levels within the institution to supervisory action.[86]

Ironically, in some contexts, a duty to report serves as a reason for not applying antiretaliation protection. While an employee would normally be able to bring a wrongful-termination claim against her employer when the termination is based on her protected speech, the existence of a duty to report may put an employee's speech outside of the protection of the First Amendment and thus outside of antiretaliation protections. Two recent cases illuminate this paradox of an affirmative obligation offsetting protections. In *Ruotolo v. City of New York*,[87] the plaintiff was an officer with the NYPD for twenty years that was serving as Command Safety Officer for the 50th Precinct in the Bronx. Pursuant to his duties in this position, Ruotolo drafted a report on possible air and water contamination caused by spills from gasoline storage tanks within the precinct.[88] Shortly after Ruotolo submitted the report to his commanding officer, the precinct's environmental hazards appeared in the local media.[89] Ruotolo then experienced a series of retaliatory and demeaning acts by his superiors.[90] Despite the temporal proximity of the precinct's actions, the court refused to treat Ruotolo's report as protected activity because the environmental report was made pursuant to Ruotolo's official duties as an NYPD officer.[91] Similarly, in *Casey v. West Las Vegas Independent School District*,[92] a former school superintendent's claim for illegal termination for reporting financial oversights to federal authorities was dismissed because the court found that the report was made pursuant to the employee's duties.[93] The Tenth Circuit Court of Appeals reasoned that since the employee was aware that she "would be held legally responsible for having knowledge of something that was wrong and not reporting," any

84. 12 C.F.R. § 21.11 (1996). These reporting duties are triggered for: (1) insider abuses involving any amount; (2) violations over $5,000 where a suspect can be identified; and (3) crimes over $25,000 in which the bank believes it was an actual or potential victim or that it was used to facilitate a criminal transaction, even if there is no substantial basis for identifying a suspect or group of suspects. *Id.* § 21.11(c)(1)–(3).

85. *Id.* § 21.11(d).

86. *Id.* § 21.11(i).

87. 03 Civ. 5045, 2006 U.S. Dist. LEXIS 49903 (S.D.N.Y. July 19, 2006).

88. *Id.* at *4.

89. *Id.* at *4–5.

90. *See id.* at *5 (detailing Ruotolo's claims, including that he was denied time off, demoted, moved to a less desirable precinct, disciplined for trivial infractions, and received his first-ever negative performance evaluation).

91. *Id.* at *9–13.

92. 473 F.3d 1323 (10th Cir. 2007).

93. *Id.* at 1328–31.

statements issued in furtherance of these obligations were entirely within the scope of her duties, therefore excluding the employee's claim for First Amendment protections.[94] Thus, at least in the context of public-sector whistle-blowing, courts have recently interpreted the various regulatory mechanisms designed to encourage reporting as mutually exclusive: even where informants correctly expose misconduct pursuant to their official duties, the presence of a legal obligation may push such reports outside the realm of protected activity.

Even in statutes designed to encourage whistle-blower activity, several exceptions have emerged that carry the potential to chill potential reporting. For example, in the public-employment setting, various courts interpreting the federal Whistleblower Protection Act (WPA)[95] have developed similar "job duty" exceptions that exclude protection when a disclosure is made in the ordinary course of an employee's duties.[96] In other words, under the "job duty" defense, employees who report illegality within the scope of their work duties are not protected from retaliation.[97] For example, in *Huffman v. Office of Personnel Management*,[98] the Federal Circuit held that a law enforcement officer whose duties included the investigation and reporting of crime to his immediate supervisor was "a quintessential example" for exclusion from WPA antiretaliation protection.[99] Similarly, in *Langer v. Department of the Treasury*,[100] the Federal Circuit held that an IRS employee did not engage in protected activity under the WPA when he informed a Department of Justice official that a grand-jury investigation disproportionately targeted minorities because the employee's official duties included reviewing the actions of the IRS's criminal division.[101]

---

94. *Id.* at 1330–31.

95. Pub. L. No. 101-12, 103 Stat. 16 (1989) (codified in scattered sections of 5 U.S.C.).

96. *See, e.g.*, Gilder-Lucas v. Elmore County Bd. of Educ., 186 F. App'x 885, 887 (11th Cir. 2006) (finding that a public-school coach who responded to a questionnaire pursuant to her official duties was not making protected speech); Maturi v. McLaughlin Research Corp., 413 F.3d 166, 172–73 (1st Cir. 2005) (requiring a terminated employee to demonstrate that his actions went beyond his regular duties in order to alert his employers that he was engaging in protected conduct); Sasse v. U.S. Dep't of Labor, 409 F.3d 773, 780 (6th Cir. 2005) ("Sasse's investigation and prosecution of environmental crimes were not protected activities because he had a duty, as an Assistant United States Attorney, to perform them."); *cf.* Kodrea v. City of Kokomo, 458 F. Supp. 2d 857, 867–68 (S.D. Ind. 2006) (finding a genuine issue of material fact as to whether a terminated employee acted within the scope of his duties or as a "concerned citizen" and therefore whether his speech was protected).

97. *See, e.g.*, Erickson v. City of Orr, No. A05-481, 2005 WL 2277395, at *5 (Minn. Ct. App. Sept. 20, 2005) (inquiring into a terminated employee's job duty defense).

98. 263 F.3d 1341 (Fed. Cir. 2001).

99. *Id.* at 1352.

100. 265 F.3d 1259 (Fed. Cir. 2001).

101. *Id.* at 1267.

C. *Monetary Incentives*

In addition to protections and affirmative duties, monetary incentives exist in some instances to encourage reports of organizational illegality. Rewards are not as prevalent as antiretaliation protections, and, while they exist in several central federal programs in the United States, they are controversial, understudied, and have not been widely adopted in other countries.[102] The primary example of such programs is the qui tam process under the False Claims Act (FCA).[103] Qui tam claims encourage reporting of fraudulent government-contractor claims by providing informants with a percentage of the recovery.[104] Employees who file a qui tam suit on behalf of the government are compensated by up to 30% of the recovery in a successful suit.[105]

This bounty model has provided the basis for similar recovery programs such as that used by the IRS, which offers financial rewards to those who report tax evasion.[106] The IRS program, which already receives thousands of annual applications and has led to the recovery of billions in federal taxes, was further expanded under the Tax Relief and Health Care Act of 2006[107] to significantly increase the financial rewards paid to informants in high value cases and create a separate Whistleblower Office within the IRS.[108] These amendments responded to many of the uncertainties that have surrounded the IRS informant process.[109] Since the enactment of the bounty program in 1867,[110] the IRS has been highly conservative in providing rewards to informants—rewarding only about 8% of informants[111] and returning only 3% to 6% of the total recoveries.[112] In contrast to the previous program, where the Secretary maintained full discretion over the amount issued for all successful recoveries,[113] the 2006 amendment provided an alternative mandatory reward program for actions where "the tax, penalties, interest, additions

---

102. *See* Lobel, *supra* note 1, at 443–44, 489–91 (describing how federal agencies reward whistle-blowers and noting that Europe has been more resistant to adopting whistle-blower protections than the United States).

103. 31 U.S.C. § 3730 (2006).

104. *Id.* § 3730(d).

105. *Id.* § 3730(d)(2).

106. 26 U.S.C. § 7623(b) (2006).

107. Pub. L. No. 109-432, § 406, 120 Stat. 2922, 2958 (codified at 26 U.S.C. § 7623).

108. IRS, History of the Whistleblower/Informant Program, http://www.irs.gov/compliance/article/0,,id=181294,00.html (last updated Apr. 21, 2008) [hereinafter IRS].

109. *See id.* (explaining that before the 2006 Amendment, awards were discretionary, policies defined award percentages and caps, and awards were not paid if certain circumstances existed).

110. Marsha J. Ferziger & Daniel G. Currell, *Snitching for Dollars: The Economics and Public Policy of Federal Civil Bounty Programs*, 1999 U. ILL. L. REV. 1141, 1167; IRS, *supra* note 108.

111. Tom Herman, *Whistleblower Law Scores Early Success: Higher Rewards Attract Informants Submitting Tips*, WALL ST. J., May 16, 2007, at D3.

112. Ferziger & Currell, *supra* note 110, at 1167.

113. IRS, *supra* note 108; *see also* Ferziger & Currell, *supra* note 110, at 1203 (indicating that the 1867 bill authorizing payments to tax evasion informants allowed payment of such sums "as may in [the commissioner of internal revenue's] judgment be deemed necessary").

to tax, and additional amounts in dispute [including taxes, penalties and interest for those disputed amounts] exceeds $2,000,000."[114] For claims involving individual taxpayers, the new subsection only applies if the individual's gross income also "exceeds $200,000 for any taxable year subject to such action."[115] Claims that fall below this threshold are still subject to discretionary review and maximum reward caps, which generally do not exceed 15%.[116] For claims that meet these high-value criteria, the amended rewards program provides that successful informants shall receive a mandatory reward of between 15% and 30% of the collected proceeds.[117] Unlike the rewards issued for discretionary cases, these high-value rewards are further expanded by eliminating the $10 million reward cap, including any funds collected from settlements, and explicitly adding penalties and interest to the collected funds from which an informant may be rewarded.[118] The provision further ensured whistle-blower rewards by permitting informants to appeal the amount or denial of a reward in the U.S. Tax Court.[119]

Although the 2006 amendments did provide further incentives and certainty for potential whistle-blowers, Congress nonetheless limited their application to a small set of high-value cases.[120] Perhaps unsurprisingly, considering the little knowledge that currently exists on the comparative effectiveness of regulatory incentives, the legislative record does not provide any explanation for why these incentives were limited to this $2,000,000 mark.[121] In 2007 the Senate amended legislation introduced by the House to reduce the threshold amount from $2,000,000 to $20,000, but the amendment was not enacted.[122] Stephen Whitlock, Director of the IRS Whistleblower Office, has explained the 2006 amendments as a means of tapping into

> situations where there may be substantial tax noncompliance but the Service may have difficulty identifying it without assistance of a knowledgeable insider. When the statute was amended, it increased the percentage for rewards and removed the policy caps in place under the old statute. The basic idea here is to create a substantial financial

114. 26 U.S.C. § 7623(b)(5)(B) (2006); *see also* IRS, *supra* note 108 (noting that the Tax Relief and Health Care Act of 2006 made "fundamental changes" to the whistle-blower reward program).

115. 26 U.S.C. § 7623(b)(5)(A).

116. *See* IRS, *supra* note 108 (implying that the limitations of the program in place prior to the 2006 amendments, including a maximum award cap of 15%, still apply for actions that do not fall within the threshold requirements specified by 26 U.S.C. § 7623(b)(5)); *see also* 26 U.S.C. § 7623(b)(5) (specifying the requirements for the mandatory reward provisions to apply).

117. 26 U.S.C. § 7623(b)(1).

118. *Id.*

119. STAFF OF J. COMM. ON TAXATION, 109TH CONG., GENERAL EXPLANATION OF TAX LEGISLATION ENACTED IN THE 109TH CONGRESS 745–46 (Comm. Print 1997).

120. 26 U.S.C. § 7623(b)(5).

121. *See* 152 CONG. REC. H9032–33 (2006) (proposing without explanation language that would limit the statute's application to $2,000,000).

122. H.R. 1591, 110th Cong. § 543(a) (2007) (engrossed amendment as agreed to by the Senate, Mar. 29, 2007).

> incentive for people to come forward with information that will help the Service make cases that we might not be able to make without them.[123]

The Amendment appears to have had some success in drawing its high-value targets. As Whitlock explains:

> Some of the things we've received over the past few months are consistent with the statutory purpose, and people who were in a position to know what was going on inside a corporation have come forward and told us about it. In some cases, they're talking about tens and hundreds of millions of dollars. That's not what the program was getting very often in the preamendment days. Many preamendment cases were much smaller issues.[124]

With regard to the proposed amendment to reduce the threshold to $20,000, Whitlock has also noted that low requirements could create a "significant problem" for the current policy of concentrating on large-dollar cases,[125] and another commentator points out reduced thresholds could allow more "weak claims and vindictive cases among neighbors."[126]

Another such system is the SEC's bounty program,[127] though the primary utility of this program has come from its ability to reveal the unsuccessful and unappealing features of would-be reward systems. Enacted under the Insider Trading and Securities Fraud Act of 1988,[128] this program was designed to draw upon the IRS model to increase successful prosecutions against inside traders.[129] However, by all measures, the program has failed to make any significant contribution toward this end. In fact, in the decade following its enactment, it is estimated that only a single bounty was paid out to an informant.[130] Considering the relative success of its FCA and IRS counterparts, there appear to be several explanations for the complete failure of the SEC program. With regard to the total bounty compensation, the SEC's 10% cap falls considerably behind the 15–30% available under the FCA and the updated IRS plan, thereby eliminating a large class of externally

---

123. Interview by Jeremiah Coder with Stephen Whitlock, Dir., IRS Whistleblower Office (July 3, 2007) (transcript available at 116 TAX NOTES 98, 98 (2007)).

124. *Id.*; *see also* Erika A. Kelton, *To Catch a Tax Cheat*, N.Y. TIMES, Aug. 7, 2008, *available at* http://www.nytimes.com/2008/08/07/opinion/07iht-edkelton.4.15087010.html (lauding the IRS whistle-blower program for making recovery in large cases more likely and citing successful examples in international tax whistle-blower programs, including Germany and Britain).

125. Dennis J. Ventry, Jr., *Whistleblowers and Qui Tam for Tax*, 61 TAX LAW. 357, 385 n.153 (2008).

126. *Id.*

127. 15 U.S.C. § 78u-1(e) (2006).

128. Pub. L. No. 100-704, sec. 3(a)(1), 102 Stat. 4677, 4677–80 (codified as amended at 15 U.S.C. § 78u-1).

129. Ferziger & Currell, *supra* note 110, at 1144.

130. *Id.*

motivated informants.[131] This reward gap is further underscored by the fact that the SEC limits its rewards to the penalties imposed under the Act, whereas the FCA also permits its qui tam litigants to recover their shares of any settlements.[132] Furthermore, the SEC rewards are entirely discretionary and not subject to judicial review.[133] In other words, even in the event of a successful prosecution, SEC informants are not guaranteed any proportion of the recovery.[134]

Although the success of these bounty programs has led states to enact mini-FCAs,[135] courts have struggled to define the effective boundaries of this reward mechanism. Recently, the Supreme Court opined in a 6–2 vote that a qui tam plaintiff cannot rely upon information previously disclosed to the public.[136] In this case, the Court overturned a qui tam claim of a former employee who won a $4.1 million judgment for reporting radioactive contamination at a nuclear-weapons plant.[137] In a dissenting opinion, Justice Stevens argued that "the Court has misinterpreted these provisions to require that an 'original source' in a qui tam action have knowledge of the actual facts underlying the allegations on which he may ultimately prevail."[138] Once again, these recent debates signify the large uncertainty as to the function and effect of existing social-enforcement mechanisms. Such uncertainty seems to be especially troubling given the huge resources being allocated—both by the government and the industry—to the different incentivizing

131. *Compare* 15 U.S.C. § 78u-1(e) ("[T]here shall be paid from amounts imposed as a penalty under this section and recovered by the Commission or the Attorney General, such sums, not to exceed 10 percent of such amounts . . . to the person or persons who provide information leading to the imposition of such penalty."), *with* 26 U.S.C. § 7623(b)(1) (2006) ("If the Secretary proceeds with any administrative or judicial action . . . based on information brought to the Secretary's attention by an individual, such individual shall . . . receive as an award at least 15 percent but not more than 30 percent of the collected proceeds . . . .").

132. *Compare* 15 U.S.C. § 78u-1(e) ("[T]here shall be paid from amounts imposed as a penalty under this section and recovered by the Commission or the Attorney General, such sums, not to exceed 10 percent of such amounts, as the Commission deems appropriate, to the person or persons who provide information leading to the imposition of such penalty."), *with* 31 U.S.C. § 3730(d) (2006) ("If the Government proceeds with an action brought by a person under subsection (b), such person shall . . . receive at least 15 percent but not more than 25 percent of the proceeds of the action or settlement of the claim . . . .").

133. *See* 15 U.S.C. § 78u-1(e) (noting that the Commission's determination as to an appropriate reward "shall be final and not subject to judicial review").

134. Still, given that the IRS program also includes a discretionary provision for recoveries under $2 million, it seems this feature in itself cannot account for the failure of the program. Rather, even if the 5% difference between the SEC and IRS reward caps could be dismissed as negligible, it appears that the SEC's stingy history compared to the IRS's pro-informant policies has done much more to decrease informant confidence than its reserved right to confiscate potential rewards. In 1997 the IRS raised its reward ceiling from $100,000 to $2 million, and again increased this cap to $10 million in 2006.

135. *E.g.*, Florida False Claims Act, FLA. STAT. ANN. §§ 68.081–.09 (West 2009); New Jersey False Claims Act, N.J. STAT. ANN. §§ 2A:32C-1 to -15, -17 (West 2009).

136. Rockwell Int'l Corp. v. United States, 549 U.S. 457, 470–71 (2007).

137. *Id.* at 466, 478–88.

138. *Id.* at 479.

mechanisms without clear enough understanding of their efficacy and social implications.

D. Understanding the Incentive Spectrum

Interestingly, the FCA is an example of a law that combines several mechanisms, using both bounty rewards and retaliation protections as incentives to encourage reporting.[139] In fact, many of our existing laws combine more than one incentive mechanism to promote reporting. Yet, while regulators and courts use this broad spectrum of regulatory mechanisms to ensure compliance in most areas of law, there is little knowledge about the comparative advantage of each mechanism or the effectiveness of combining various mechanisms. As Professor Roberta Romano has commented, "[C]ongressional initiatives rarely are constructed from whole cloth; rather, successful law reform in the national arena typically involves the recombination of old elements that have been advanced in policy circles for a number of years prior to adoption."[140] Commenting on the adoption of SOX, Romano concludes that the legislation was enacted under conditions of limited legislative debate with provisions that were "poorly conceived, because there was no basis to believe they would be efficacious."[141] Bringing together the various developments in incentivizing social enforcement, the study of whistle-blowing provides an ideal context for studying the interplay between individual compliance behavior, the organizational setting in which it is detected, and the regulatory regime that defines the contours of legality.[142] Surprisingly, despite the widespread recognition of the importance of social enforcement and the potential application of different behavioral predictors to regulatory policy, questions about these fundamental interactions between individual, organizational, and state-level factors have received relatively little research attention.[143]

---

139. *See* 31 U.S.C. § 3730(d) (2006) (providing for rewards for qui tam plaintiffs); *id.* § 3730(h) (providing for reinstatement and damages when an employee who has brought an action under the FCA is discharged, demoted, or harassed because of the action).

140. Roberta Romano, *The Sarbanes–Oxley Act and the Making of Quack Corporate Governance*, 114 YALE L.J. 1521, 1524 (2005).

141. *Id.* at 1594.

142. *See generally* MICELI & NEAR, *supra* note 17, at 179–230 (exploring the interaction between various factors influencing whistle-blowing, including the individual characteristics of whistle-blowers, situational variables, long-term consequences, and legal approaches).

143. *See* Yuval Feldman & Oren Perez, *How Law Changes the Environmental Mind: An Experimental Study of the Effect of Legal Norms on Moral Perceptions and Civic Enforcement*, 36 J.L. & SOC'Y 501, 502 (2009) (noting that the "expressive influence of the law in the context of a multi-faceted regulatory environment" had not yet been explored); Marcia P. Miceli & Janet P. Near, *Standing Up or Standing By: What Predicts Blowing the Whistle on Organizational Wrongdoing*, 24 RES. PERSONNEL & HUM. RESOURCES MGMT. 95, 97 (2005) (noting that despite enduring interest in whistle-blowing, there is relatively little research on the topic in leading management journals).

III. Linking Behavioral Economics with New Governance

The study explores motivations that underlie the willingness to engage in whistle-blowing. While there have been important developments in the research of human motivation in the field of behavioral research, legal inquiry has lagged behind in adopting contemporary insights. The underlying framework of our empirical study is a linkage between two areas: behavioral research and new-governance theory. From the behavioral perspective, most studies on social enforcement have been conducted by experimental economists using laboratory simulations and games.[144] These studies have yielded important findings, yet have been largely detached from the real-world context.[145] Indeed, while the study of motivation has been the subject of academic inquiry for over a century, with literally thousands of studies and experiments on the subject, very few of these studies have examined the effects of legal mechanisms.[146] Moreover, although the contribution of behavioral studies has been highly significant, lab-based models have "largely lack[ed] social or organizational context."[147] This deficiency has in turn limited their potential to inform concrete law and policy. In particular, the importance of more textured studies is highly pronounced in the context of reporting behavior where a range of factors—individual, organizational, and legal environments—may impact decision making. Although existing behavioral studies have successfully pointed to key factors affecting reaction

144. *See, e.g.*, Feldman & Perez, *supra* note 143, at 505–06 (describing an experiment in which the authors exposed individuals to corporate-pollution scenarios and examined the extent to which the individuals were willing to engage in social enforcement).

145. *See, e.g.*, Gary E. Bolton & Alex Ockenfels, *ERC: A Theory of Equity, Reciprocity, and Competition*, 90 AM. ECON. REV. 166, 189 (2000) (describing the study as limited "in the sense that it explains *stationary patterns* for relatively *simple games*, played over a *short time span* in a *constant frame*"); Ernst Fehr & Urs Fischbacher, *Third-Party Punishment and Social Norms*, 25 EVOLUTION & HUM. BEHAV. 63, 85 (2004) (contrasting the experiment with "real life"); Ernst Fehr & Simon Gächter, *Altruistic Punishment in Humans*, 415 NATURE 137, 140 (2002) (describing the experiment as being conducted on computer screens using experimental software); Ernst Fehr & Simon Gächter, *Cooperation and Punishment in Public Goods Experiments*, 90 AM. ECON. REV. 980, 982 (2000) (explaining that the experiment was conducted in a computerized laboratory with no interaction between the subjects); David Hirshleifer & Eric Rasmusen, *Cooperation in a Repeated Prisoners' Dilemma with Ostracism*, 12 J. ECON. BEHAV. & ORG. 87, 90 (1989) (summarizing the experiment as a game presenting a situation with limited responses and outlining several assumptions in the examination of results); Yoram Kroll & Liema Davidovitz, *Inequality Aversion Versus Risk Aversion*, 70 ECONOMICA 19, 26–27 (2003) (disclaiming that the study's sample was not "representative of the general population" and cautioning that in order to draw certain conclusions, the study should be replicated with adults); Michael E. Price et al., *Punitive Sentiment as an Anti-Free Rider Psychological Device*, 23 EVOLUTION & HUM. BEHAV. 203, 226 (2002) (admitting that the experiment's hypothesis should be tested among people in living conditions more representative of the environments provided in the study).

146. *See, e.g.*, Thomas S. Bateman & J. Michael Crant, Revisiting Intrinsic and Extrinsic Motivation (unpublished manuscript, on file at http://74.125.47.132/search?q=cache:4KzmCTxDn64J:gates.comm.virginia.edu/Mirror/faculty_research/Research/Papers/IMOBHDP24.pdf+revisiting+intrinsic+and+extrinsic+motivation&cd=1&hl=en&ct=clnk&gl=us) (counting thousands of studies on extrinsic motivation).

147. Feldman & Lobel, *supra* note 11, at 167.

by those witnessing wrongdoing—primarily self-interest and moral outrage[148]—the literature to date has failed to differentiate between various regulatory mechanisms. For example, it is highly difficult to draw conclusions from the existing literature about mechanism selection: Are statutory antiretaliation protections more compatible with supporting individual reporting than government bounty programs? Will different levels of monetary rewards increase or diminish the likelihood of reporting? Is a combination of various incentive mechanisms more effective than selecting a single mechanism to encourage whistle-blowing? While these questions are crucial for evaluating the optimal design of mechanisms that would encourage social enforcement, they remain largely unanswered by the behavioral literature.

Departing from the behavioral approach, the field of new governance in sociolegal studies rejects the idea that employee behavior within organizations is "reducible to individuals and their characteristics."[149] It further rejects the idea that institutions are simply the object of regulation.[150] Instead, the lens of new-governance theory is a systemic mapping of the range of possibilities in the interaction between regulation and regulated parties.[151] New-governance scholars call for an understanding of the regulatory process as consisting of a range of possible tools and mechanisms, each with its comparative advantages and costs.[152] The new-governance lens thus helps to frame the inquiry about the possible incentives that can be applied by law to encourage certain behaviors. Increasingly, sociolegal research is introducing evidence that procedures and incentives can induce illegal conduct by limiting oversight or, alternately, promote ethical behavior by emphasizing social responsibility.[153] Behavior is shaped not simply by

148. *See* Price et al., *supra* note 145, at 203 (concluding that providers of a public good harbor punitive sentiments against free riders to mitigate the fitness advantage such parties achieve through their free-riding behavior).

149. Robert Tillman & Henry Pontell, *Organizations and Fraud in the Savings and Loan Industry*, 73 SOC. FORCES 1439, 1459 (1995).

150. *See id.* at 1458–59 (suggesting that organizations and their practices may be the mechanisms through which illegal schemes are conducted, rather than the mere context in which those schemes exist).

151. LESTER M. SALAMON, THE TOOLS OF GOVERNMENT: A GUIDE TO THE NEW GOVERNANCE 1–14 (2002). For an overview of new governance scholarship as a new school of thought see Orly Lobel, *Formulating a New Paradigm: Newness and the Ripeness of the Moment*, 2005 WIS. L. REV. 479; Orly Lobel, *The Paradox of "Extra-Legal" Activism: Critical Legal Consciousness and Transformative Politics*, 120 HARV. L. REV. 937 (2007); Orly Lobel, *The Renew Deal: The Fall of Regulation and the Rise of Governance in Contemporary Legal Thought*, 89 MINN. L. REV. 342 (2004); Orly Lobel, *Setting the Agenda for New Governance Research*, 89 MINN. L. REV. 498 (2004).

152. *See* SALAMON, *supra* note 151, at 1–2 (discussing new governance in terms of the "*tools* of public action" and explaining that those tools vary in their "operating procedures, skill requirements, and delivery mechanism[s]").

153. *See, e.g.*, JOHN BRAITHWAITE, CRIME, SHAME AND REINTEGRATION 9 (1989) ("Shame is conceived as a tool to allure and inveigle the citizen to attend to the moral claims of the criminal law, to coax and caress compliance, to reason and remonstrate with him . . . .").

isolated decisions of individuals but also by the institutional and legal environment.[154] There is a widespread consensus among organizational theorists that "structures, processes, and tasks are opportunity structures for misconduct because they provide (*a*) normative support for misconduct, (*b*) the means for carrying out violations, and (*c*) concealment that minimizes detection and sanctioning."[155] For example, studies show organizations that constantly pressure their employees to meet unreasonable expectations can lead employees to resort to illegal means to achieve these goals.[156] Similarly, recent studies point to counterproductive effects of regulation, as when the regulatory action is perceived as illegitimate and creates resistance among private actors.[157]

Building on these insights, the current study thus fills a research gap by bringing together the individual aspects of behavior and motivation with the institutional perspective of regulatory variance. The integration of behavioral and new-governance approaches understands individuals, their work environment, and the legal regime in which they operate as symbiotic. In sum, the integrative methodological approach enables better investigation of the effect of law and policy on individual motivation and behavior. The study is based on the understanding that at the enforcement stage, there are certain conditions and incentives that encourage and support action by individuals, while others are more likely to invoke inaction and silence in the face of illegality.

In a previous study that aimed to provide an integrative approach to the study of social enforcement, we measured enforcement behaviors by manipulating common misconducts in five different domains: financial, environmental, sexual harassment, safety, and employee theft.[158] We found that individual characteristics, such as gender, income, and cultural

154. *See* IRVING L. JANIS, GROUPTHINK: PSYCHOLOGICAL STUDIES OF POLICY DECISIONS AND FIASCOES 243 (1982) (arguing that a "sizable percentage" of the decisions made by any single policy-making group will "prove to be at least partly attributable to groupthink tendencies"); Diane Vaughan, *Toward Understanding Unlawful Organizational Behavior*, 80 MICH. L. REV. 1377, 1377 (1982) ("[A]ll social control efforts encounter natural constraints because of the ways in which the social structure continuously and systematically generates unlawful organizational behavior."); On Amir & Orly Lobel, *Stumble, Predict, Nudge: How Behavioral Economics Informs Law and Policy*, 108 COLUM. L. REV. 2098, 2109 (2008) (book review) (citing behavioral economics in claiming that measures adopted by regulators, even when noncoercive, may nonetheless steer individuals to make the "right" choices for themselves).

155. Diane Vaughan, *The Dark Side of Organizations: Mistake, Misconduct, and Disaster*, 25 ANN. REV. SOC. 271, 289 (1999).

156. MARSHALL B. CLINARD, CORPORATE ETHICS AND CRIME 140–44 (1983).

157. *See, e.g.*, Orly Lobel, *The Four Pillars of Work Law*, 104 MICH. L. REV. 1539, 1552 (2006) (describing the benefits of collective-labor laws, which allow for flexibility and dynamic learning, as opposed to traditional command-and-control regulations); Orly Lobel, *Interlocking Regulatory and Industrial Relations: The Governance of Workplace Safety*, 57 ADMIN. L. REV. 1071, 1089–91 (2005) (describing how the OSHA's dense, unreasonable rules enforced by spot inspections have deterred employers from complying).

158. Feldman & Lobel, *supra* note 11, at 170.

differences, have strong effects on the willingness of employees to engage in social enforcement.[159] These findings demonstrate the variation in the role of intrinsic and extrinsic motivation with regard to each of those misconducts.[160] Moreover, the type of motivation activated by particular misconducts affected the choice of enforcing behaviors by employees, ranging between direct confrontation, corporate procedures, and government reporting.[161] While this prior collaboration was an important step in understanding social enforcement in action, perhaps the most important question—both theoretically and practically—is still missing: How can policy makers ensure that the incentives offered fit the motivational range of enforcing employees? In other words, while the motivation of the individual employees in institutions was the focus of the previous collaborative research, it did not answer pressing questions on how law and policy can translate motivational and organizational insights into concrete regulatory mechanisms. By integrating the two disciplinary approaches of behavioral and institutional studies, our new study brings into focus the question of social enforcement in relation to the spectrum of concrete regulatory mechanisms available to policy makers.

The incentive matrix at the basis of the experiments builds on the four categories that currently exist in our legal system: protections, duties, fines, and monetary incentives. We hypothesize that each mechanism will have varying levels of effectiveness across different categories of individuals and settings. In the following section, we shall set the ground for the empirical part of the Article by focusing on the theoretical assumptions that underlie our predictions for the behavioral consequences of each one of the mechanisms described thus far.

A. The Hypotheses of the Empirical Study

In developing the experimental studies, we hypothesized that the effectiveness of regulatory incentives for the initiation of enforcement litigation is related to both the objective price the individual would pay and to the individual's subjective willingness to pay it. In some cases the motivation is primarily instrumental and utilitarian, and, hence, the level of incentives needs to be high. However, in other contexts, where the motivation to report is intrinsic and moralistic, offering remuneration may undermine the likelihood that misconduct will be reported by producing a crowding-out effect in which the presence of external rewards dilutes the moral dimension of the act.[162] In such cases, it is reasonable to assume that

159. *Id.* at 180–81.

160. *See id.* ("[F]or different types of misconduct, behavioral incentives and institutional environments varied in their impact on the decision to socially enforce . . . .").

161. *See id.* (describing how reporting of misconduct is predictable by factors such as "perceived norms, moral outrage, and legitimacy" and that various forms of misconduct activate different motivations, which can influence the choice of enforcement behavior).

162. *See* Amir & Lobel, *supra* note 154, at 2135 (observing that rewards may motivate compliance but that their use beyond a certain level may "backfire").

antiretaliation protections may prove more effective than monetary incentives as these regulatory mechanisms, at best, place the individual in a situation where she is not worse off than she was prior to the report. This dilemma demonstrates the importance of identifying *ex ante* what motivates the prospective whistle-blower in a given situation. Nonetheless, the goal of policy is complicated by the fact that increased enforcement can be counterproductive.[163] For example, overprotection may encourage bad-faith reporting and exaggerated, or even false, accusations. It can also diminish the positive ties and organizational citizenship behavior (OCB) of institutional players.

An additional question that our study attempts to answer is whether all employees are affected by various mechanisms in similar ways. For example, some scholars have theorized that women are more likely to report misconduct because of their status as organizational outsiders.[164] It has been suggested that women's shorter and more intermittent work histories may lead them to feel less organizational loyalty.[165] At the same time, the existing literature reveals the complexity of accounting for gender differences in social enforcement. Some studies suggest that women are more likely to blow the whistle because of different moral and ethical viewpoints than men.[166] Others postulate that the majority of whistle-blowers are white males in part due to their higher esteem and higher positions within corporations.[167] Further, one study found that women whistle-blowers suffer more retaliation than men.[168] Our own previous experiments reveal that women are more likely to report wrongdoing to law-enforcement authorities.[169] Yet, we could not conclusively explain the motivational source of such difference. A factor that is completely missing in prior studies is whether men and women are expected to react similarly to different legal mechanisms. For example, it might be expected that the greater retaliation against women would lead to

163. *Id.* at 2131.

164. Cindy Schiapani & Terry Dworkin, *Women and the New Corporate Governance: Pathways for Obtaining Positions of Corporate Leadership*, 65 MD. L. REV. 504, 530–31 (2006).

165. *See* Patricia A. Gwartney-Gibbs & Denise H. Lach, *Gender and Workplace Dispute Resolution: A Conceptual and Theoretical Model*, 28 LAW & SOC'Y REV. 265, 273 (1994) (suggesting that women's shorter and more intermittent work histories due to family responsibilities may prompt conflict with managers and supervisors).

166. *See* Mary Brabeck, *Ethical Characteristics of Whistle Blowers*, 18 J. RES. PERSONALITY 41, 50 (1984) (describing a study in which females scored higher in moral reasoning); Terance D. Miethe & Joyce Rothschild, *Whistleblowing and the Control of Organization Misconduct*, 64 SOC. INQUIRY 322, 334 (1994) (suggesting that women may feel more compelled to speak out against wrongdoing).

167. *See* Philip Jos et al., *In Praise of Difficult People: A Portrait of the Committed Whistle-Blower*, 49 PUB. ADMIN. REV. 552, 556 (1989) (justifying an overwhelmingly white and male sample size in a survey of whistle-blowers by noting that demographic group's disproportionate distribution among higher ranking corporate positions).

168. Michael T. Rehg et al., *Antecedents and Outcomes of Retaliation Against Whistleblowers: Gender Differences and Power Relationships*, 19 ORG. SCI. 221, 233 (2008).

169. Feldman & Lobel, *supra* note 11, at 179.

greater reliance on protection for women. In contrast, men's greater reliance on financial incentives in the decision making may suggest that using rewards might be more beneficial for them.

The behavioral effects of the different legal mechanisms used to encourage reporting are highly relevant for optimal regulatory design. An important insight from our previous study was that when the misconduct is harmful to the public but useful to the organization as a whole or to its senior managers, reporting to a state agency is viewed by employees as the most effective and justifiable course.[170] Studying the interaction between the manipulated factors—i.e., different legal incentives to encourage reporting—and the behavioral characteristics of the employee helps answer important policy questions.

B. Behavioral Theories on the Interplay between Intrinsic and Extrinsic Motivation

Based on a review of the literature, we hypothesize that each of the four mechanisms will affect motivation differently. Motivation is often divided into "intrinsic" or "extrinsic."[171] Extrinsic motivation is linked to actions that are driven by external commands or rewards such as payments.[172] Conversely, intrinsic motivation is when the behavior is chosen from within the individual out of a sense of moral or civic duty.[173] There is an ongoing, heated debate in the behavioral-economics literature about the relationship between the intrinsic and extrinsic aspects of motivation.[174] While some view extrinsic rewards as undermining intrinsic motivation, others find that the two mechanisms can be reinforcing.[175] Most generally, the crowding-out

---

170. *Id.* at 180–81.

171. *See, e.g.*, Edward L. Deci & Richard M. Ryan, *The "What" and the "Why" of Goal Pursuits: Human Needs and the Self-Determination of Behavior*, 11 PSYCHOL. INQUIRY 227, 233–36 (2000) (distinguishing intrinsic motivation, which concerns "active engagement with tasks that people find interesting," from extrinsic motivation, in which people's behavior is controlled by "specific external contingencies"); Judith M. Harackiewicz & Carol Sansone, *Rewarding Competence: The Importance of Goals in the Study of Intrinsic Motivation*, *in* INTRINSIC AND EXTRINSIC MOTIVATION: THE SEARCH FOR OPTIMAL MOTIVATION AND PERFORMANCE 79, 79–80 (Carol Sansone & Judith M. Harakiewicz eds., 2000) ("When humans freely engage in an activity for its own sake, their behavior is considered intrinsically motivated [yet] interventions, communications, and incentives represent extrinsic intrusions that can affect subsequent intrinsic motivation . . . .").

172. Carol Sansone & Judith M. Harackiewicz, *Looking Beyond Rewards: The Problem and Promise of Intrinsic Motivation*, *in* INTRINSIC AND EXTRINSIC MOTIVATION, *supra* note 171, at 1, 1–2.

173. Edward L. Deci et al., *A Meta-analytic Review of Experiments Examining the Effects of Extrinsic Rewards on Intrinsic Motivation*, 125 PSYCHOL. BULL. 627, 627 (1999); T. Kasser & R.M. Ryan, *Further Examining the American Dream: Differential Correlates of Intrinsic and Extrinsic Goals*, 22 PERSONALITY & SOC. PSYCHOL. BULL. 280, 280–87 (1996).

174. Harackiewicz & Sansone, *supra* note 172, at 1–2.

175. For one of the early studies on the subject, see Edward L. Deci, *Effects of Externally Mediated Rewards on Intrinsic Motivation*, 18 J. PERSONALITY & SOC. PSYCHOL. 105 (1971). Since then, the debate has continued. *See, e.g.*, Thomas S. Bateman & J. Michael Crant, *Extrinsic*

literature suggests that when people attribute their behavior to external rewards, they discount any moral incentives for their behavior, thereby lowering the perceived effect of intrinsic motivation.[176] As applied to the regulatory incentives, crowding-out theory predicts that external incentives that utilize monetary rewards or punishments may undermine intrinsic motivations.[177] For instance, paying people in return for their blood might lead donors to view the event as a transaction rather than a charitable act, thereby eroding altruistic blood donations.[178] In a series of lab-based experiments, Deci found that tangible rewards undermine intrinsic motivation for a range of activities.[179] Deci, Koestner, and Ryan have argued in their research that "tangible rewards tend to have a substantially negative effect on intrinsic motivation."[180] They thus warn that attempts to externally control people's behavior may yield considerable long-term counterproductive results.[181] Such findings have led some organizational theorists to strongly advise against the use of extrinsic rewards when trying to achieve behavior.[182]

Rewards and Intrinsic Motivation: Evidence from Working Adults 17 (McIntire Sch. of Commerce, Working Paper, 2005), *available at* http://www.commerce.virginia.edu/faculty_research/faculty directory/documents/IntrinsicJPSP.pdf (discussing the findings of their survey showing that extrinsic rewards have some positive relationships with intrinsic motivation).

176. *See* BRUNO S. FREY, NOT JUST FOR THE MONEY: AN ECONOMIC THEORY OF PERSONAL MOTIVATION 14 (1997) ("A group of social psychologists have identified that under particular conditions monetary (external) rewards undermine intrinsic motivation.").

177. *See, e.g.*, Ernst Fehr & Armin Falk, *Psychological Foundations of Incentives*, 46 EUR. ECON. REV. 687, 687–88 (2002) (postulating that a narrow view based exclusively on economic incentive without consideration for other psychological factors is likely to result in an incomplete and inaccurate understanding of human motivation); Ernst Fehr & Bettina Rockenbach, *Detrimental Effects of Sanctions on Human Altruism*, 422 NATURE 137, 137 (2003) ("Here we show experimentally that the prevailing self-interest approach has serious shortcomings because it overlooks negative effects of sanctions on human altruism."); Ernst Fehr & Simon Gächter, *Do Incentive Contracts Undermine Voluntary Cooperation?* 1 (Univ. of Zurich, Inst. for Empirical Research in Econ., Working Paper No. 34, 2002), *available at* http://papers.ssrn.com/sol3/papers.cfm?abstract_id=313028 ("Our results show that incentive contracts may undermine voluntary cooperation."). For a review, see generally FREY, *supra* note 176; George A. Akerlof, *Labor Contracts as Partial Gift Exchange*, 97 Q.J. OF ECON. 543 (1982); and Bruno S. Frey & Reto Jegen, *Motivation Crowding Theory: A Survey of Empirical Evidence* (Ctr. for Econ. Studies, Working Paper No. 245, 2000), *available at* http://ssrn.com/abstract=203330.

178. *See* RICHARD M. TITMUS, THE GIFT OF RELATIONSHIP: FROM HUMAN BLOOD TO SOCIAL POLICY 157, 205–06 (1971) (arguing that monetary payments to givers of blood could diminish the amount of blood given voluntarily).

179. *See* Deci et al., *supra* note 173, at 658–59 (providing a meta-analysis of 128 studies concluding that "tangible rewards tend to have a substantially negative effect on intrinsic motivation" within certain limitations).

180. *Id.*

181. *See id.* ("[S]trategies that focus primarily on the use of extrinsic rewards do, indeed, run a serious risk of diminishing rather than promoting intrinsic motivation . . . .").

182. *See* Alfi Kahn, *By All Available Means: Cameron and Pierce's Defense of Extrinsic Motivators*, 66 REV. EDUC. RES. 1, 3 (1996) ("[E]ven a casual survey of the literature reveals research . . . that attests to the detrimental effects of rewards.").

A refinement of the crowding-out literature has demonstrated that the framing of incentives may affect whether such incentives interact with the moral aspects of reporting. Indeed, studies increasingly show that framing has an effect on how people perceive legal regulation. For example, although traditional economic analysis would consider fines and pricing as equal if they entail the same amount of payment by an individual, researchers have shown that in reality the way payments are framed matters significantly.[183] For example, Fehr and Gächter found that when monetary incentives were framed as a price reduction, they had a greater effect than when they were framed as a bonus.[184] Similarly, Frey and Stutzer have argued that tradable emission rights and emission taxes could create a different crowding-out effect, bringing about different behavior.[185] This refinement of the crowding-out literature guides our hypothesis that the legal framing of payments could impact whether such rewards crowd out alternative reasons for action, thereby affecting people's behavior. Of course, the very idea that monetary incentives will crowd out internal motivations and decrease action runs counter to classic economic predictions. Normally, the introduction of money is a major push for action.[186] This idea has found some empirical support in studies showing that in some instances rewards can in fact increase perceived self-determination.[187]

Thus, a key question is how do the two effects of "buying" behavior operate together: will a classic price effect emerge in which higher economic incentives lead to increased action, or will a crowding-out effect take hold, resulting in less action after the introduction of a financial reward? In other words, monetary and ethical motivations can be complements that mutually reinforce or, alternately, exclusive elements that inversely affect one another. The debate about whether extrinsic rewards undermine intrinsic motivation has largely been neglected in legal theory and practice. Mostly, the literature

183. *See* Fehr & Gächter, *supra* note 177, at 1 (providing an example of the effect that framing can have on incentives).

184. *Id.*

185. *See* Bruno S. Frey & Alois Stutzer, *Environmental Moral and Motivation* 14–16 (Univ. of Zurich, Inst. for Empirical Research in Econ., Working Paper No. 288, 2006), *available at* http://papers.ssrn.com/sol3/papers.cfm?abstract_id=900370 (arguing that though government intervention through emission taxes reduces individuals' feelings of personal moral responsibility for protecting the environment, tradable emission rights create a stronger crowding-out effect as individuals are led to believe that they have paid for the right to pollute rather than seeing pollution as a harm for which they are monetarily punished).

186. *See* STEVEN E. LANDSBURG, THE ARMCHAIR ECONOMIST: ECONOMICS & EVERYDAY LIFE 3 (1993) ("Most of economics can be summarized in four words: 'People respond to incentives.' The rest is commentary.").

187. *See, e.g.*, Robert Eisenberger & Judy Cameron, *Detrimental Effects of Reward: Reality or Myth?*, 51 AM. PSYCHOL. 1153, 1154 (1996) ("We argue here that claimed negative effects of reward on task interest and creativity have attained the status of myth, taken for granted despite considerable evidence that the conditions producing these effects are limited and easily remedied.").

described above focuses on work behaviors.[188] In the context of incentive mechanisms, it seems reasonable to expect that antiretaliation measures will have a more limited negative effect on the role of internal motivation than monetary incentives.[189] Because the former type of measure is designed to put the employee in the same position she was prior to her action, it is less likely to be interpreted as a type of external motivation. Similarly, with regard to bounty incentives, the study aims to provide valuable insight into the circumstances under which such monetary rewards will lead to increased reporting. Much of the literature on crowding out recognizes that the effect of incentives is not linear and that intermediate levels of incentives are most likely to curb value-driven behavior. In the context of incentives, there is a documented difference between small, intermediate, and high payoffs, such that intermediate payoffs trigger crowding-out effects most often.[190]

C. *The Framing of Legal Dollars*

The potential effect of introducing monetary rewards depends not only on its interaction with internal motivation but also on the conditions that are set to trigger its effect. In that regard, there is a growing body of studies both in social psychology and in behavioral economics indicating that people respond more strongly to incentives than penalties.[191] Bateman and Crant explain that "[r]eceiving rewards that we have earned means that we are no longer at the mercy of a capricious or over controlling environment, and we have gained control over our outcomes."[192] Gneezy and Rustichini used an experimental setting to explore whether fines may actually be interpreted as placing price tags on certain misconduct.[193] In their study, they imposed a monetary fine on parents who were late picking up their child from a day-care center.[194] After the introduction of the fine, they observed a steady *increase* in the number of parents coming in late.[195] Again, this result runs contrary to traditional deterrence models that predict that increasing the cost

188. *See generally, e.g.*, Judith L. Komaki et al., *The Role of Performance Antecedents and Consequences in Work Motivation*, 67 J. APPLIED PSYCHOL. 334 (1982).

189. FREY, *supra* note 176, at 31.

190. *See* Frey & Stutzer, *supra* note 185, at 14–16 (using intermediate-sized environmental taxes as an example to illustrate that such situations incur the negative effect of eliminating the intrinsic-motivation factor while not being sufficiently strong to achieve particular behavior by means of extrinsic motivation).

191. *See* Raymond De Young, *Changing Behavior and Making It Stick: The Conceptualization and Management of Conservation Behavior*, 25 ENV'T & BEHAV. 485, 489 (1993) ("In general, environmental psychology argues against the use of punishment as a conservation behavior change technique.").

192. Bateman & Crant, *supra* note 175, at 8.

193. Uri Gneezy & Aldo Rustichini, *A Fine Is a Price*, 29 J. LEG. STUD. 1, 3–9 (2000).

194. *Id.* at 4–5.

195. *Id.* at 5–8.

of an activity will necessarily decrease the rate at which it is performed.[196] Gneezy and Rustichini offer two explanations for their surprising results. First, the introduction of the fine may have changed parents' perception of the social dynamic between themselves and the day-care center.[197] That is, whereas the act of arriving late was previously wrong in itself, the introduction of a fine may have allowed parents to rationalize the fine as a price for arriving late. According to this logic, as long as they paid the price for such behavior, parents felt comfortable being late.[198] Second, the fine may have revealed information to parents regarding the expected sanction for tardiness.[199] Thus, parents who were previously punctual out of fear of incurring a costly sanction may have exercised less caution after learning the actual cost of the behavior as revealed by the fine.[200] As applied to the context of regulatory compliance then, it appears that the use of fines may similarly interfere with the moral dimension of compliance activity: where certain types of misconduct were once inherently wrong, the introduction of a fine may inadvertently specify the financial tipping point at which the costs of reporting misconduct outweigh the moral and social benefits.

Correctly assessing the optimal level and type of rewards also has strong policy implications as it will allow policy makers to more effectively apply their limited resources toward ensuring that the proper incentives are in place to encourage the external reporting of organizational illegality.[201] Therefore, in addition to measuring the immediate crowding-out effects and willingness to report, the study also measures the participants' subjective evaluation of the severity of the misconduct, which we interpret to reveal their ethical stance on the illegal behavior they witnessed. The behavioral impact of the different mechanisms is expected to interact with the social aspects of whistle-blowing, which may also serve as facilitators of or impediments to social enforcement. To account for the social impact for the law, we now turn to the expressive-law literature.

196. *See, e.g.*, *id.* at 2–3 (reviewing the general literature predicting that negative consequences should produce deterrence); *see also* Christopher R. Leslie, *Trust, Distrust, and Antitrust*, 82 TEXAS L. REV. 515, 651 (2004) ("The traditional deterrence model assumes that criminal law deters crimes by prohibiting specific acts and by assigning a particular penalty to each specific act, and that this makes firms less likely to commit those acts lest they get caught and suffer the punishment.").

197. Gneezy & Rustichini, *supra* note 193, at 13–14.

198. *Id.* at 13–15.

199. *Id.* at 10.

200. *Id.* at 10–13.

201. In fact, the 2006 update to the IRS informant program may be a nod to the crowding-out literature. Disputes involving less than $2 million are capped at 15% and not subject to appeal. IRS, WHISTLEBLOWER OFFICE, FY 2008 ANNUAL REPORT TO CONGRESS ON THE USE OF SECTION 7623, 2 (2009). For penalties over $2 million, the recovery limit increases to 30%, the rewards are guaranteed, and the informant can appeal decisions to a tax court, suggesting these cases may be more externally driven. *Id.* at 3.

IV. The Expressive Function of Law

Developing the lens of intrinsic versus extrinsic motivation in the context of legal compliance connects the behavioral literature with an important insight of legal scholars in recent years, known as *expressive law*. Expressive-law theorists view law as having more than a simple, direct effect on human behavior.[202] Beyond simple calculative effects, individuals respond to the expressive signals embodied within our legal system.[203] For traditional economists, the starting point of predicting behavior is the belief that individuals respond to rewards and sanctions.[204] According to this traditional view, individuals will report noncompliance if the benefits from legal rewards, or the costs of legal liability, exceed the costs of reporting.[205] However, social scientists increasingly recognize that the motivation for compliance, as well as reporting noncompliance, frequently defies a simplis-

---

202. *See* Cass R. Sunstein, *On the Expressive Function of Law*, 144 U. PA. L. REV. 2021, 2024 (1996) ("[T]he expressive function of law [is] the function of law in 'making statements' as opposed to controlling behavior directly.").

203. *See* Elizabeth S. Anderson & Richard H. Pildes, *Expressive Theories of Law: A General Restatement*, 148 U. PA. L. REV. 1503, 1504 (2000) ("[E]xpressive theories tell actors—whether individuals, associations, or the State—to act in ways that express appropriate attitudes towards various substantive values."); Dan M. Kahan, *What Do Alternative Sanctions Mean?*, 63 U. CHI. L. REV. 591, 592 (1996) ("The political unacceptability of alternative [criminal] sanctions . . . reflects their inadequacy along the expressive dimension of punishment. The public rejects [them] . . . because they fail to express condemnation as dramatically and unequivocally as imprisonment."); Richard H. Pildes & Elizabeth S. Anderson, *Slinging Arrows at Democracy: Social Choice Theory, Value Pluralism, and Democratic Politics*, 90 COLUM. L. REV. 2121, 2142 (1990) ("Rationality must be understood to be a matter of interpretation and evaluation, not merely of aggregation and calculation. And social institutions, including democratic ones, must play an active role in structuring individuals' preferences to enable those preferences to rationally express the diverse values that individuals actually experience and affirm."); Sunstein, *supra* note 202, at 2022–25, 2031–32, 2048, 2051 (explaining that laws, having moral weight, make statements that impact social norms, which in turn signal appropriate behavior, give legal consequences social meaning, and control the actions of some individuals). For a critical view, see Matthew D. Adler, *Expressive Theories of Law: A Skeptical Overview*, 148 U. PA. L. REV. 1363, 1375 (2000) ("I furnish reason to be skeptical about each and every [expressive theory]. Each theory involves some variant of the claim that the linguistic meaning of governmental action possesses foundational moral relevance; but in no case does the claim turn out to be persuasive or even particularly plausible."). For empirical studies of the expressive function of the law, see generally Feldman & Perez, *supra* note 143 and Richard H. McAdams & Janice Nadler, *Testing the Focal Point Theory of Legal Compliance: Expressive Influence in an Experimental Hawk/Dove Game*, 2 J. EMPIRICAL LEGAL STUD. 87 (2005). For empirical evaluations of the different models for the expressive function of the law in the context of trade-secret laws, see generally Yuval Feldman, *The Expressive Function of Trade Secrets Laws, Legality, Cost, Intrinsic Motivation and Consensus*, 6 J. EMPIRICAL LEGAL STUD. 177 (2009).

204. *See* LANDSBURG, *supra* note 186, at 3–7 (describing how individuals respond to incentives, which take the form of positive and negative outcomes).

205. Most basically, this model is elaborated in the context of compliance and crime. *See* Gary S. Becker, *Crime and Punishment: An Economic Approach*, 76 J. POL. ECON. 169, 176 (1968) ("[A] person commits an offense [and fails to comply with the law] if the expected utility to him exceeds the utility he could get by using his time and other resources on other activities.").

tic cost–benefit analysis.[206] Instead, people appear to evaluate legal compliance under a more nuanced cost–benefit scale that includes elements that are foreign to pure economic analyses: duty and legitimacy.[207]

Along those lines, theoretical and empirical expressive-law studies indicate that when the law presents either prohibitions or obligations, the very act of expressing this norm within the legal regime provides a reason for people to act. Even in the absence of sanctions or protections, the mere existence of the law helps to shape and define our world views.[208] Expressivists argue that moral and legal evaluation and conduct depend on normative expressions embedded in the law.[209] For example, the law can have a positive effect on norms because citizens view law as information that helps them make decisions about whether to engage in particular behaviors.[210] In this sense, the law solves a pluralistic ignorance problem by signaling the underlying attitudes of a community or society.[211]

---

206. *See* Lawrence M. Friedman, *Coming of Age: Law and Society Enters an Exclusive Club*, 1 ANN. REV. L. SOC. SCI. 1, 14 (2005) ("There is no question that human beings do react to the carrot and the stick. However, people are not blindly mechanical cost-benefit machines. . . . The moral sense—conscience and the sense of legitimacy, the feeling that we ought to obey the law . . . —is also a significant cluster of motives that surely affect how people behave.").

207. *See, e.g.*, James L. Gibson, *The Legitimacy of the U.S. Supreme Court in a Polarized Polity*, 4 J. EMPIRICAL LEGAL STUD. 507, 513 (2007) ("Some scholars *equate* legitimacy with compliance; others treat legitimacy as one of the many causes of compliance. I take the latter tack . . . ."); Kristina Murphy, *Regulating More Effectively: The Relationship Between Procedural Justice, Legitimacy, and Tax Non-compliance*, 32 J.L. & SOC'Y 562, 567 (2005) ("Taken together, the procedural justice research appears to indicate that fair procedures play a significant role in people's perceptions of legitimacy, and that these perceptions of legitimacy can in fact go on to influence subsequent cooperation and compliance with authority decisions and rules."); Tom R. Tyler & Jeffrey Fagan, *Legitimacy and Cooperation: Why Do People Help the Police Fight Crime in Their Communities?*, 6 OHIO ST. J. CRIM. LAW 231, 235 (2008) ("What is legitimacy? Legitimacy is a feeling of obligation to obey the law and to defer to the decisions made by legal authorities. Legitimacy, therefore, reflects an important social value, distinct from self-interest, to which social authorities can appeal to gain public deference and cooperation."); Michael Wenzel, *The Impact of Outcome Orientation and Justice Concerns on Tax Compliance: The Role of Taxpayers' Identity*, 87 J. APPLIED PSYCHOL. 629, 640 (2002) ("Taxpayers [in Australia] were more compliant with tax laws when they identified with Australians and thought they were treated fairly and respectfully by the tax authorities. . . . When authorities convey . . . respect and acknowledgement, taxpayers are more compliant irrespective of whether decision outcomes are favorable to them."). For an empirical demonstration of the limits of traditional economic models in the context of legal compliance, see Yuval Feldman & Doron Teichman, *Are All Legal Probabilities Created Equal?*, 84 N.Y.U. L. REV. 980, 985–86 (2009), which compares attitudes toward compliance in situations of uncertainty.

208. For an empirical demonstration of this point, see generally McAdams & Nadler, *supra* note 203.

209. *See* Anderson & Pildes, *supra* note 203, at 1504 ("[E]xpressive theories tell actors . . . to act in ways that express appropriate attitudes towards various substantive values . . . . [M]uch of our existing practices of moral and legal evaluation are best understood through expressivist perspectives . . . [and] expressivism is thus an internal account of existing normative practices.").

210. *See* Richard H. McAdams, *An Attitudinal Theory of Expressive Law*, 79 OR. L. REV. 339, 358–72 (2000) [hereinafter McAdams, *Attitudinal Theory*] (describing expressive law as an information-signaling system); Richard H. McAdams, *The Expressive Power of Adjudication*, 2005 U. ILL. L. REV. 1043, 1068–74 [hereinafter McAdams, *Expressive Power*] (describing adjudication as an information-signaling system); McAdams & Nadler, *supra* note 203, at 117 ("That players are

The expressive function of law underscores the importance that individuals generally place on the opinion of others. Existing research suggests that such adherence to the social norm is important to compliance.[212] The importance of group identity and the need of the individual to feel part of the group is well-established in the field of psychology.[213] The need to belong is widely recognized as a motivation that countervails immediate self-interest.[214] McAdams and Nadler contend that law can induce compliance by making a particular behavior salient.[215] Continuing this line of studies, we hypothesize that law can induce reporting by making the action of reporting seem significant and valued. In other words, imposing a duty to report is expected to send an important message of the social desirability of whistle-blowing and alter participants' behavior above and beyond the monetary consequences of conforming to the law.

V. Misperception of Social Support for Blowing the Whistle

The potential for social norms to shape enforcement behavior, discussed above,[216] gives them a particular significance in the context of whistle-blowing. Social attitudes towards blowing the whistle and predictions on how others would behave when faced with similar dilemmas are key to understanding individual decision making. However, the perception of what others will do is complex and often biased. In our study, we hypothesized that people would predict and evaluate their own actions more favorably than the actions of their peers and the general public. In a recent article, social

---

likely to choose a strategy [based on information clearly supplied] at random suggests that expression influences behavior even when it lacks legitimacy or intentionality. Legal expression always 'points to' an outcome. Law influences behavior in many ways, but we infer from this result that, at the most basic level, law *also* influences behavior ostensively.").

211. *See* McAdams, *Attitudinal Theory*, *supra* note 210, at 340 ("[L]aw changes behavior by signaling the underlying attitudes of a community or society . . . [and] individuals have imperfect information about what others approve.").

212. *See id.* ("[T]here is a motivational assumption that an individual's behavior depends, in part, on what actions she believes others will approve or disapprove."); *id.* at 342 ("[A social choice theory] framework can explain social regularities, such as norms, that were at one time thought to be outside the range of economic theory . . . . [I]ndividuals have a preference for *esteem*; they care what others, even strangers, think of them as an end in itself."); *id.* at 349–58 (describing at length the "pluralistic ignorance" problem).

213. *See, e.g.*, Roy F. Baumeister & Mark R. Leary, *The Need to Belong: Desire for Interpersonal Attachments as a Fundamental Human Motivation*, 117 PSYCHOL. BULL. 497, 497 (1995) (listing numerous psychological inquiries into the need to belong and concluding that "[t]he existence of a need to belong is thus a familiar point of theory and speculation").

214. *See* Marilynn B. Brewer & Roderick M. Kramer, *The Psychology of Intergroup Attitudes and Behavior*, 36 ANN. REV. PSYCHOL. 219, 227 (1985) ("*[I]ntergroup* comparisons must be considered above and beyond *interpersonal* factors in predicting social behavior in group contexts.").

215. *See* McAdams & Nadler, *supra* note 203, at 108–18 (presenting results of an empirical study supporting the claim that third-party legal expression influences behavior by creating a focal point around which individuals coordinate).

216. *See supra* notes 208–14 and accompanying text.

psychologists Nick Epley and David Dunning found that most people view themselves as morally superior to the average person.[217] They found that people show a holier-than-thou effect in which they perceive themselves as being fairer, more altruistic, and more self-sacrificing.[218] Epley and Dunning hypothesized that two causes could be at the bottom of this bias: either people perceive themselves as better and more moral than they really are, or they view others as morally inferior.[219] In their experiment, participants were asked to predict what they and others would do in certain circumstances.[220] In most cases, participants predicted they would do the right thing a lot more often than their peers.[221] Later, they compared findings of what participants would do in real life circumstances.[222] They found that individuals overestimated their own moral character but were quite accurate when it came to predicting how others would behave.[223] However, Epley and Dunning's study examined moral behavior, such as giving to charity and civic voting, and thus did not involve regulatory contexts.[224] Our experiment is the first to inquire into such comparisons in the context of reporting illegal misconduct.

---

217. Nicholas Epley & David Dunning, *Feeling "Holier Than Thou": Are Self-Serving Assessments Produced by Errors in Self- or Social Prediction?*, 79 J. PERSONALITY & SOC. PSYCHOL. 861 (2000) [hereinafter Epley & Dunning, *Holier Than Thou*]. For example, they asked their participants to predict whether they and others will buy flowers in an upcoming charity drive to benefit the American Cancer Society. *Id.* at 862. Among the 250 students who participated in the experiment, over 80% said they would buy a flower, but they predicted that only half their peers would be as generous. *Id.* After the actual charity drive took place, the study found that only 43% of the participants actually bought flowers at the event, indicating that students had a pretty good idea of how the group would behave, but not a good sense of their own individual behavior. *Id.* at 863. As for the students' estimation of voting behaviors, students similarly overestimated their civic responsibility, with 84% of them predicting their own participation, but expecting an overall participation of 67%. Nicholas Epley & David Dunning, The Mixed Blessings of Self-Knowledge in Behavioral Prediction (2004) (unpublished manuscript, on file with the authors). Again, the actual rates among the group turned out to be very close to the predictions about peer participation, as 68% of the students actually voted. *Id.*

218. Epley & Dunning, *Holier Than Thou*, *supra* note 217, at 862.

219. *Id.*

220. *Id.*

221. *See id.* ("Whereas people may be somewhat reluctant to say they are smarter than their peers, they have no difficulty noting that they are more generous, fair, ethical, or moral.").

222. *Id.*

223. *Id.* at 868.

224. There are formal analyses of the long-term effect of misperception of social norms on ethical behavior of people. *See* Robert D. Cooter, Michal Feldman & Yuval Feldman, *The Misperception of Norms: The Psychology of Bias and the Economics of Equilibrium*, 4 REV. L. & ECON. 889, 889–90 (2008) (studying the effects of applying the equilibrium concept to psychological studies of cognitive biases); *see also* Epley & Dunning, *Holier Than Thou*, *supra* note 217, at 862 (analyzing why individuals make flattering self-assessments); Chip Heath, *On the Social Psychology of Agency Relationships: Lay Theories of Motivation Overemphasize Extrinsic Incentives*, 78 ORGANIZATIONAL BEHAV. & HUM. DECISION PROCESSES 25, 54–55 (1999) (chronicling that subjects based job-satisfaction decisions on intrinsic values such as skill building or doing something worthwhile over extrinsic factors including job security and benefits but predicting incorrectly that their peers would be motivated by extrinsic factors); Justin Kruger, *Lake Wobegon Be Gone! The "Below-Average Effect" and the Egocentric Nature of Comparative Ability Judgments*, 77 J. PERSONALITY & SOC. PSYCHOL. 221, 223–29 (1999) ("[M]ost of us appear

In sum, social psychology and behavioral economics have produced an important array of insights that could inform regulators in designing compliance systems. Yet, by and large, these insights have not been incorporated into legal inquiry and regulatory practices. Our study is designed to build on existing insights and provide insights that are specifically geared to answer some of the most important questions for regulatory compliance: When and under what conditions will individuals chose to report illegality? Can the legal policy maker take into account the behavioral ramifications of each one of the competing mechanisms? Is there a legal mechanism that minimizes the inadvertent behavioral effect that incentivizing whistle-blowers might cause? Can we identify ex ante the optimal match between the legal mechanism, the individual, the organization, and the misconduct?

In an attempt to make the answer to these questions applicable to current legal policy, the focus of the current study will be on the four main mechanisms described above: antiretaliation protections, duties to report, liability fines, and monetary incentives. As reviewed in the previous Part, for each one of these mechanisms, there is a vast amount of behavioral research that suggests potential pros and cons of the mechanism.[225] By comparing their efficacy[226] in one setting, combined with data on relevant organizational and demographic characteristics, we hope to offer a comprehensive policy analysis.

VI. The Experiment

A. Design and Sample

1. Experimental Design.—The experimental survey consists of eight different questionnaires randomly assigned to eight subgroups. While each

---

to believe that we are more athletic, intelligent, organized, ethical, logical, interesting, fair-minded, and healthy—not to mention more attractive—than the average person."); Justin Kruger & David Dunning, *Unskilled and Unaware of It: How Difficulties in Recognizing One's Own Incompetence Lead to Inflated Self-Assessments*, 77 J. PERSONALITY & SOC. PSYCHOL. 1121, 1121 (1999) (surveying studies showing that people overestimate their abilities); Michael McCall & Katherine Nattrass, *Carding for the Purchase of Alcohol: I'm Tougher Than Other Clerks Are!*, 31 J. APPLIED SOC. PSYCHOL. 2184, 2190–91 (2001) (examining bias related to predictions of whether identification would be requested for purchasers of alcohol); Dale T. Miller & Rebecca K. Ratner, *The Disparity Between the Actual and Assumed Power of Self-Interest*, 74 J. PERSONALITY & SOC. PSYCHOL. 53, 53 (1998) (discussing the results of five studies related to the overestimation of the influence of self-interest on attitudes and behaviors); Dale T. Miller & Rebecca K. Ratner, *The Power of the Myth of Self-Interest*, *in* CURRENT SOCIETAL CONCERNS ABOUT JUSTICE 25, 29–30 (Leo Montada & Melvin J. Lerner eds., 1996) (explaining that subjects demonstrated higher volunteerism when reward included payment to charity than when reward included payment to the volunteer, but predicted that their peers would choose the latter option).

225. *See supra* Part IV.

226. As will be explained in the following Part, the four mechanisms were tested in a few combinations, making it into eight experimental conditions.

questionnaire depicts a different legal mechanism, all groups were given the same factual scenario, as follows:

> Imagine you are an employee of Roadblock LTD, one of the largest construction companies in the country. Roadblock has recently secured a fixed-price government contract to build a major highway in your city. One day, while staying late in the office, you run across a document that reveals that the company has been substituting lower grade and inferior quality parts from those specified in the contract. The document also reveals that the company has been omitting required testing and quality procedures. You estimate that as a result the government is overpaying your employer approximately $10,000,000.

Each group of survey takers was asked to predict their own actions as well as the actions of others when encountering the described situation. The legal mechanisms were derived from the four leading incentive categories reviewed above: protection, duty, fine, and reward.[227] Given the variation in the level of rewards, we have manipulated the size of the reward across the different bounty programs. In addition, in the context of duty, its combination with the other mechanisms was tested, resulting in the following eight categories of incentives:[228]

1) High Reward ($1,000,000)
2) Low Reward ($1,000)
3) Duty + High Reward
4) Duty + Low Reward
5) Antiretaliation Protection (1 year)
6) Duty + Antiretaliation Protection
7) Duty + Fine ($10,000)
8) Duty

Following the factual scenario and the legal mechanism provided in each category, we measured the following variables:[229]

1) Intention of self and others to report
2) Evaluation of effect of the legal mechanism on the decision to report
3) Perceived morality, harm, and severity of the misconduct
4) Expected social and career ramifications
5) Organizational features and individual status

227. *See supra* subparts II(A) (protection), II(B) (duty and fine), and II(C) (reward).

228. The vignettes used for manipulating the eight conditions are accumulated in Appendix II.

229. In addition, we have measured an elaborated demographic profile.

2. The Sample.—The data collection was web based, using the survey firm Zoomerang.[230] Zoomerang holds a panel of three million volunteers who earn credits toward online rewards in return for filling a survey.[231] The sample of the survey was drawn from a diverse, representative panel of working adults in the United States; the sample included participants of different genders and races, as well as various education levels, professions, and workplace structures. The 2,081 participants of our survey consisted of a representative panel based on census demographics in the United States.[232]

B. Findings

The findings of the study are presented in the three following sections. We begin by describing the effect of manipulating the eight categories of mechanisms on the individual's self-stated intentions to take action. We further compared individual behavior to one's predictions about the behavior of others under the same circumstances. In the second section, we present findings on the gap between the *actual* (experimental) effect on decisions to report, as measured in the first section, and the *perceived* (self-reported) effect the selected legal mechanism had on respondents' decision to report. In the third section, we focus on the social aspects of reporting. These include perception of social norms, social support, career effects, and the social meaning of reporting under the competing mechanisms.

1. Regulatory Mechanisms and the Holier-Than-Thou Effect.—The first set of findings focuses on the experimental effect of the manipulated legal mechanism on participants' intention to report the misconduct.[233] The type of legal mechanism described in the scenario was entered as an independent

---

230. *See generally* Zoomerang, http://www.zoomerang.com.

231. Interview by Lynda Resnick with Dave Deasy, Gen. Manager of Zoomerang (June 23, 2009) (edited transcript available at http://blog.lyndaresnick.com/2009/06/market-research-affordable-user-friendly-real-time/).

232. Participants were 2,081 individuals, 47% women and 53% men, whose mean age was 45 years (SD=17 years) with 78% reporting completion of higher education (college or above) and only 3% reporting less than a high-school education. The majority of participants identified their racial or ethnic background as European-American (88%) with very small percentage of individuals who identified themselves as African-American (5%), Hispanic (3%), Asian (3%), or Native American (1%). Seventy percent were employed in professional industries (e.g., food service, manufacturing, construction, and agriculture), 52% in the public sector, 54% in companies with 100 or more employees, and 76% reporting the work place as not unionized. The majority of participants (61%) reported that there was no internal compliance department at their workplace. Data was collected in 2008 and 2009.

233. To test whether the perceived likelihood of action against the company differed among the legal mechanism subgroups, a one-way MANCOVA test was conducted. Three variables were included as covariates in the analyses, and they were controlled to reduce their value as competing explanations for the participants' outcomes: age, education (academic/nonacademic), and the existence of internal compliance mechanisms at the workplace. The Bonefferoni post-hoc test was used to determine the significance of the differences among the adjusted mean scores. To prevent interference with the flow of the Article, the statistical tables were moved to Appendix I.

variable and the actions taken toward the company were entered as dependent variables.[234] The findings indicate significant differences among the legal mechanisms.[235]

Respondents reported the highest self-reporting rates when faced with a combination of a *Duty* and *High Reward*, whereas respondents in the *Low Reward* condition were least likely to report.[236] In line with the familiarity bias, peers were also predicted to report at the highest levels in the *Duty* + *High Reward* condition.[237] By contrast, participants predicted that most people would take action in the highest levels when there were high rewards, thus revealing a perception that a stranger's decision to report is more likely to be externally driven.[238] However, this crowding-out effect virtually disappears when participants are asked to predict what most other people will do under the various mechanisms. Respondents in the *High Reward* and *Duty* + *High Reward* conditions provided the highest scores for likelihood that others would report the company's misconduct.[239]

The findings reveal that individuals predict higher reporting action by themselves than by their peers and others. Using a two-way MANCOVA test,[240] we find that respondents predicted that they themselves would be more likely to report[241] than their peers and that their workplace peers[242] would be more likely to report than most other people.[243] In other words, individuals believe that they themselves will behave more ethically in the face of misconduct than others and that people with whom they are familiar will behave more ethically than the general population. This pattern was similar across the various subgroups. That is, for the most part, the legal

234. The actions taken toward the company include: (1) The participants' intention to report the misconduct [hereinafter *self-action*]; (2) their perception of the likelihood that most people will engage in reporting the misconduct [hereinafter *most people*]; and (3) the perceived likelihood that most employees in their own company would report the misconduct [hereinafter *most peers*].

235. Multivariate $F(21,5445) = 4.73, p < .001, \eta^2 = .02$.

236. In addition, participants in the *Duty* + *Fine* group were more likely than *Low Reward* to take self-action against the company.

237. *See infra* fig.1. *Duty* + *High Reward* graded higher than *Duty* in likelihood that "most employees" would report the employer.

238. *High Reward* saw the highest report rates, followed by *Duty* + *High Reward*—both of these categories graded higher than all the other subgroups in likelihood that "most people" would report the employer.

239. $F(7,1898) = 9.96, p < .001, \eta^2 = .03$.

240. The type of legal mechanism was entered as an independent variable, the initiator of the action taken against the company (*Self, Most People, Most Peers*) was entered as a within-subject factor, and background factors were held constant. Such analysis enables us to compare the three responses of each one of the participants when controlling for the other factors that differentiated the participants.

241. Mean and Standard Deviation: ($M = 8.41$, $SD = 2.16$).

242. Mean and Standard Deviation: ($M = 6.34$, $SD = 2.73$).

243. Mean and Standard Deviation: ($M = 5.52$, $SD = 2.58$), Multivariate $F(2,1897) = 49.75$, $p < .001, \eta^2 = .03$.

mechanism did not have an effect on the gaps between the three types of initiators (*Self*, *Peers*, *General Population*).

Interestingly, the only exception was the difference in perceived reporting behavior between the *Low Reward* and *High Reward* subgroups, which was particularly pronounced for the respondents' estimation of *Most People*.[244] In this case, respondents believed that increasing the level of financial rewards is likely to have the greatest effect on the reporting behavior of the general population,[245] suggesting a belief that the average person's actions are more externally motivated than the one's own actions.

These gaps support the research on a holier-than-thou effect, which suggests that people perceive their own actions as more ethically guided than those of others.[246] The fact that respondents did not display this bias in the same intensity towards their peers suggests a reduced effect among familiar circles, supporting the psychological phenomenon of familiarity leading to increased perceptions of likeness.[247] Importantly, according to previous holier-than-thou studies,[248] and supported by other behavioral findings of our study discussed below concerning the gap between action and perception, participants' estimation of the reward-driven behavior of others may be more accurate than the estimation of their own expected behavior.

244. This gap was expected according to the holier-than-thou phenomenon discussed above. *See supra* notes 218–24 and accompanying text.

245. Notably, the general population consists of the people whom they know the least in comparison to their coworkers.

246. Epley & Dunning, *Holier Than Thou*, *supra* note 217, at 861.

247. Constantine Sedikides & John J. Skwronski, *The Law of Cognitive Structure Activation*, 2 PSYCHOL. INQUIRY 169, 179 (1991) ("Moreover, we would claim that familiarity and proximity contribute to the formation of relationships through their indirect effects on similarity. Familiarity and proximity provide the individual an opportunity either to discover commonalities . . . or to create commonalities through gradual changes in cognitive content.").

248. Epley & Dunning, *Holier Than Thou*, *supra* note 217, at 861.

Figure 1: The Relative Ranking of Reporting by Self, Peers, and General Population[249]

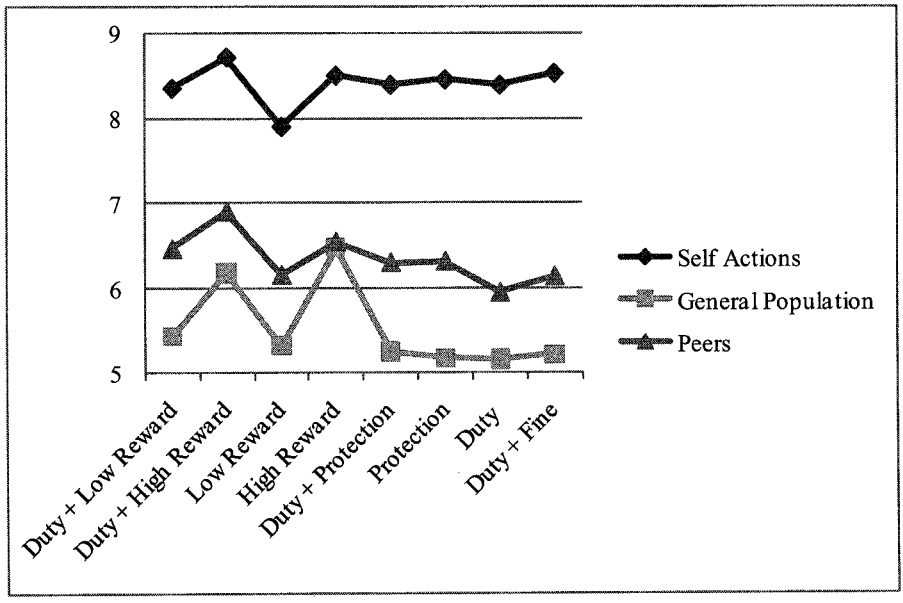


Figure 1 clearly demonstrates the holier-than-thou effect. For self-reporting, an affirmative duty to report remained an influential factor for increasing reporting rates even when it was not paired with fines or protection. However, with regard to others, especially nonpeers, respondents felt that high rewards would serve as a much stronger motivator than a duty to report.

2. The Interplay Between Internal and External Motivation.—The findings reveal significant differences between the subgroups in the evaluation of misconduct.[250] As is expected intuitively, the more outraged respondents feel about the illegal behavior, the more likely they are to report and to predict reporting by others. Conversely, the less severe the misconduct, the less likely respondents are willing to take action. In order to test these predictions, participants were divided into two subgroups based on their evaluations of the misconduct. The first group included the respondents who evaluated the severity and immorality of the misconduct as high. The second group included those respondents who gave lower evaluations of the

249. Adjusted mean scores for the initiator of action against the company across the eight subgroups. The score ranged from 1–10, with higher values indicating higher likelihood of action taken against the company ($n = 1{,}909$).

250. In evaluation of the misconduct, we refer to a factor that is based on questions 19–26. Within it, we include factors such as moral outrage, percept of risk to the public government from the misconduct, legitimacy, and acceptability. Multivariate $F(3,1808) = 54.14$, $p < .001$, $\eta^2 = .08$.

severity of the misconduct.[251] The first group—the high-severity group—can be understood as including those individuals who are internally motivated to report (*High Internal*). Conversely, the second group—the low-severity group—can be understood as the group with lower levels of internal motivation (*Low Internal*).[252]

As expected, participants with *High Internal* motivation were more likely than participants with *Low Internal* motivation to predict self-reporting[253] and to predict higher reporting rates for both peers[254] and the general population.[255]

To examine the interplay between internal motivation and external legal mechanisms, we further tested the interaction between evaluation of the misconduct and intention to report. Among our two evaluation groups, a significant interaction was found between the type of legal mechanism and the perceived severity of the misconduct.[256] Among the *Low Internal* subgroup, participants in the *High Reward*, *Duty + High Reward*, and *Duty + Fine* were more likely than *Low Reward* to report an intention to act. However, in the *High Internal* subgroup, no difference emerged between the subgroups; the pattern of differences among means was similar across the eight legal mechanisms.[257] This signifies that, while the choice of legal mechanisms was important for those who viewed the misconduct to be relatively insignificant, there was largely no difference for those who viewed the misconduct to be severe. The findings suggest that when individuals recognize an ethical stake in an issue, policy-design variances are diminished.

---

251. Whether a respondent rated the misconduct as high or low was based upon their rating compared to the median value of the evaluation of the misconduct. A two-way MANCOVA test was then conducted. The type of legal mechanism and the evaluation of the misconduct subgroups were entered as independent variables. The actions taken against the company were entered as three dependent variables. Background factors were held constant.

252. The association between internal motivation to report (as opposed to reporting due to external rewards) and perception of the misconduct's severity is clear given that the items measuring perceived severity included moral and legitimacy concerns.

253. Univariate tests: $F(1,1890) = 159.10$, $p < .001$, $\eta^2 = .08$; *Low Evaluation of the Misconduct*: $M = 7.83$, $SD = 2.35$; *High Evaluation of the Misconduct*: $M = 9.03$, $SD = 1.72$.

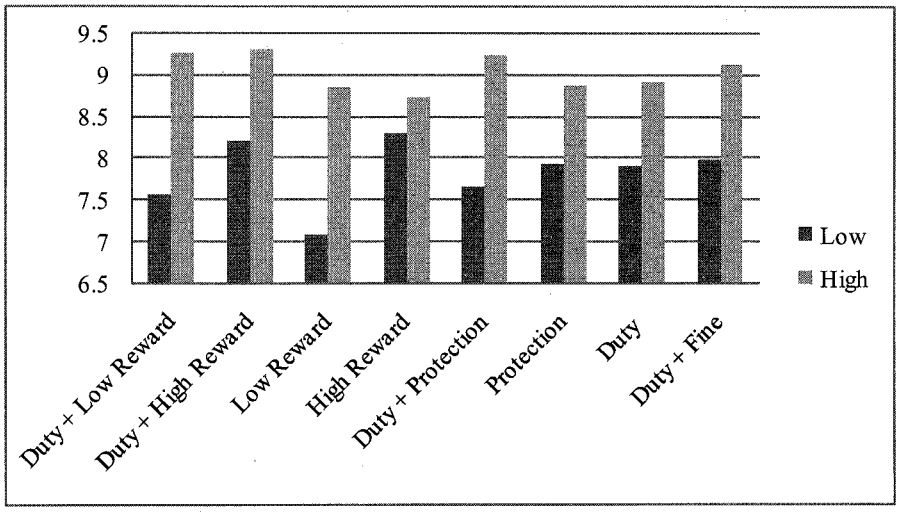
254. $F(1,1890) = 3.86$, $p < .05$, $\eta^2 = .00$; *Low Evaluation of the Misconduct*: $M = 5.41$, $SD = 2.45$; *High Evaluation of the Misconduct*: $M = 5.64$, $SD = 2.71$.

255. $F(1,1890) = 3.86$, $p < .05$, $\eta^2 = .00$; *Low Evaluation of the Misconduct*: $M = 5.41$, $SD = 2.45$; *High Evaluation of the Misconduct*: $M = 5.64$, $SD = 2.71$.

256. $F(7,1890) = 2.89$, $p < .01$, $\eta^2 = .01$.

257. *See infra* app. I, tbl.2 (showing no statistical difference between the different groups among the various legal mechanisms).

Figure 2: Report Against the Company by Participants with High and Low Perception of Severity of the Act



To illustrate the interplay between internal and external motivation, Figure 2 presents the responses of participants across condition and perceived severity of the misconduct. The graph demonstrates the sharp contrast between incentive mechanisms once the moral impulse to report is removed as a potential source of motivation. When internal motivation is missing—i.e., the severity of the misconduct is perceived as low—external incentives matter much more when deciding whether or not to report misconduct. Respondents least likely to report were those offered a low reward while they had a low perception of misconduct severity. Reporting in those circumstances was even lower than situations where no incentive was present. Moreover, the perceived severity of the misconduct produced the greatest disparity in the reporting levels for respondents in the *Low Reward* category while it had the smallest effect on the *High Reward* category. These findings suggest that the introduction of an external reward interferes with the moral dimension of reporting.

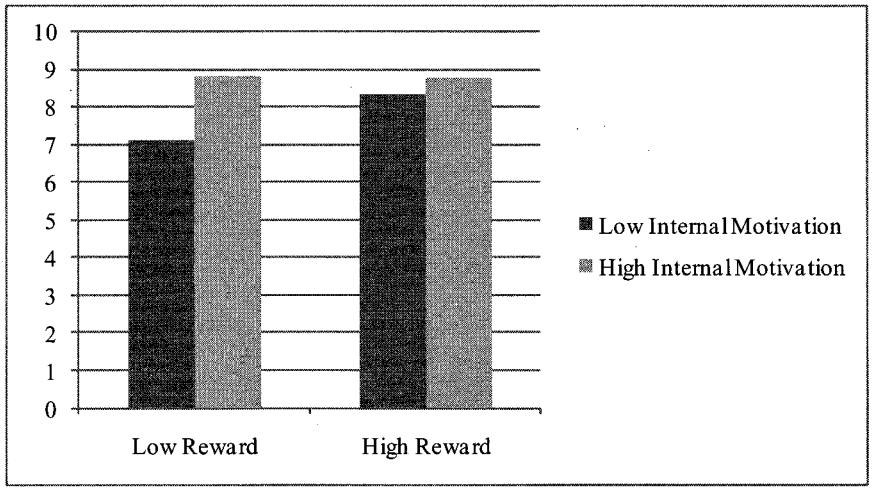
Another way to look at the importance of using high rewards when there is a fear that some of the workers—and hence potential enforcers—might have low internal motivation, is through the following analysis. Conducting a two-way ANCOVA revealed a significant interaction effect between the legal mechanisms (*Low Reward* and *High Reward*) and the internal motivation subgroups indicating that the monetary reward reduced the self-reporting differences between the internal motivation subgroups.[258] Within the *Low Reward* condition, participants with a *High Internal* motivation were more

258. $F(1,474) = 10.60, p < .001, \eta^2 = .02$.

likely to report than participants with a *Low Internal* motivation.[259] However, within the *High Reward* condition, there was no significant difference between the internal motivation subgroups ($p > .05$). As demonstrated in Figure 3, paying a high reward basically decreases the differences between high and low internally motivated individuals.[260] While we do not have a measure of quality of reporting in this Article—a measure that might point out the inferiority of money-driven reporting[261]—it is clear that high rewards are able to overcome internal-motivation impediments to social reporting.

Importantly, the decrease in reporting for employees who were not internally motivated was less pronounced for the other mechanisms: fines and protection. This suggests a key normative insight: both fines and protection are less likely to be perceived as external motivation and therefore carry less of a risk of crowding out internal motivation. Finally, and equally importantly, the findings indicate that creating a duty to report, which emphasizes the moral dimension of the reporting action, enhances the effect of severity on one's intention to engage in whistle-blowing. Namely, the combination of duty and a high level of internal motivation results in the highest levels of reporting behavior.

Figure 3: Effect of High and Low Extrinsic Rewards on Participants with High and Low Perception of Severity of the Act



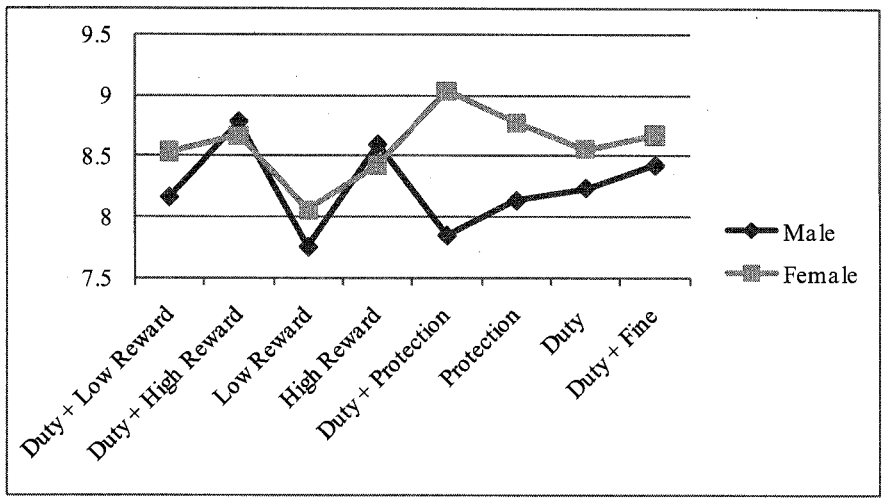
259. $F(1,474) = 35.00, p < .001, \eta^2 = .07$.

260. A three-way ANCOVA revealed that there was no significant interaction effect between the gender, the legal mechanisms (*Low Reward* and *High Reward*), and the internal motivation subgroups indicating that the same pattern of results was found among women and men.

261. This is so both in terms of motivation—e.g., honesty of the report—and in terms of quality of the report.

3. Gender and Whistle-Blowing.—Gender plays a role in predicting whistle-blowing behavior. Our findings indicate significant differences between men and women.[262] Women were far more likely than men to blow the whistle.[263] At the same time, among women, the legal mechanism did not have a significant effect. In particular, women were significantly more likely than men to take action against the company in the subgroups *Duty + Protection* and *Protection.* Moreover, men were far more affected by the mechanism manipulations: while the level of monetary reward did not seem to matter to women, the higher reward significantly increased the likelihood of reporting among men.[264]

Figure 4: Gender and the Effect of the Alternative Incentive Mechanisms



As can be clearly seen from Figure 4, in conditions of *Duty* and *Duty + Protection*, women were far more likely than men to engage in whistle-blowing. These findings corroborate previous findings of a gender effect in that regard. However, the graph illustrates an important perspective regarding the interplay between choice of legal mechanisms and gender. Whereas men care significantly more than women about the size of the monetary

262. To examine whether the differences among the legal mechanisms varied according to gender, we conducted a two-way MANCOVA. The type of legal mechanism and gender were entered as independent variables, and the actions taken against the company were entered as three dependent variables: self action, actions by most people, and actions by most peers. Background factors were held constant. Multivariate $F(3,1875) = 6.03$, $p < .001$, $\eta^2 = .01$.

263. $F(1,1877) = 12.61$, $p < .001$, $\eta^2 = .01$. *Men*: $M = 8.24$, $SD = 2.28$; *Women*: $M = 8.59$, $SD = 2.02$.

264. $F(7,1877) = 2.40$, $p < .05$, $\eta^2 = .01$. Similarly, men in *Duty + High Reward* were also more likely than participants in the *Duty + Protection* to take self-action against the company. Furthermore, only among men, *Duty + High Reward* and *High Reward* were more likely than *Low Reward* to take self-action against the company.

reward, women are more incentivized by antiretaliation protections and legal duties.[265]

The experiment also revealed gender differences in the reported preferences of respondents concerning the reporting process itself. Women respondents cared more about maintaining anonymity both in relation to the employer[266] and to the public.[267] This finding could be explained by a greater job insecurity experienced by women, evident also in their greater reliance on protection mechanisms. There was also a gender effect for perceived social support.[268] Women respondents graded higher than men the importance of reactions from both family[269] and friends[270] to their decision to report.

The gender differences with regard to women provide the following crowding-out effect that emerges from using monetary rewards. Using an ANCOVA, we see that among women within the *Low Internal* subgroup, participants in the *Duty* group were more likely than *Low Reward* to report an intention of most people to act.[271] However, among men within the *Low Internal* subgroup, participants in the *Duty* group were less likely than *Low Reward* to report an intention of most people to act.[272]

This reduction is a clear demonstration of a crowding-out effect where low rewards harm, rather than benefit, the willingness of individuals to engage in whistle-blowing when a duty is present. To demonstrate the effect graphically, Figure 5 presents the responses of men and women focusing on two of the conditions within the lower perceived severity of misconduct.

---

265. This finding accounts for differences in economic status of the respondents.

266. $F(1,1869) = 4.54$, $p < .05$, $\eta^2 = .002$; *Male*: $M = 7.44$, $SD = 2.91$; *Female*: $M = 7.72$, $SD = 2.78$.

267. $F(1,1869) = 4.08$, $p < .05$, $\eta^2 = .002$; *Male*: $M = 7.62$, $SD = 2.83$; *Female*: $M = 7.89$, $SD = 2.76$.

268. Multivariate $F(2,1874) = 8.99$, $p < .001$, $\eta^2 = .01$.

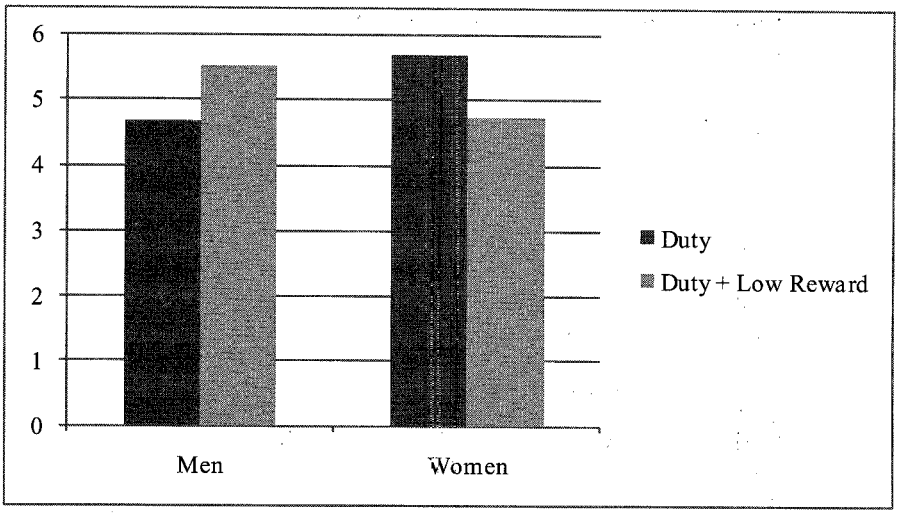
269. $F(1,1875) = 17.71$, $p < .001$, $\eta^2 = .009$; *Male*: $M = 7.87$, $SD = 2.16$; *Female*: $M = 8.28$, $SD = 2.06$.

270. $F(1,1875) = 4.61$, $p < .05$, $\eta^2 = .002$; *Male*: $M = 5.41$, $SD = 2.41$; *Female*: $M = 5.65$, $SD = 2.29$.

271. $F(1,105) = 4.16$, $p < .05$, $\eta^2 = .04$.

272. $F(1,125) = 3.41$, $p < .05$, $\eta^2 = .03$.

Figure 5: Actions Taken Against the Company by Most People as a Function of the Legal Mechanism and Gender Within the Low Internal Subgroup



4. Real Versus Perceived Motivation to Blow the Whistle.—The experiments had two stages. Stage 1 tested predictions about action under the various conditions. In Stage 2, we asked respondents to assess their own response in Stage 1. Thus far, we have focused on the results of Stage 1, the *experimental* effects of the different legal mechanisms. In this section, we analyze the results of Stage 2, the *perceived* effects and compare perception to the experimental findings. In this second stage, we investigate how people evaluate their own motivations and those of others when asked to reflect on the decision to report misconduct. We ask what individuals believe motivated reporting and whether those motivations are primarily internal or external.

To examine the overall ranking of the mechanisms based on their perceived influence on intention to enforce, a series of *t*-tests was conducted.[273] The combined analysis showed that *Duty* and *Fine* were graded as major factors in a participant's decision to report, followed by *Protection*, *High Reward*, and *Low Reward*.[274] Not surprisingly, the size of the reward was perceived to be influential. Respondents in *High Reward* conditions graded higher the influential effect of the reward on their decision to report than those offered low rewards.[275] Similarly, respondents believed

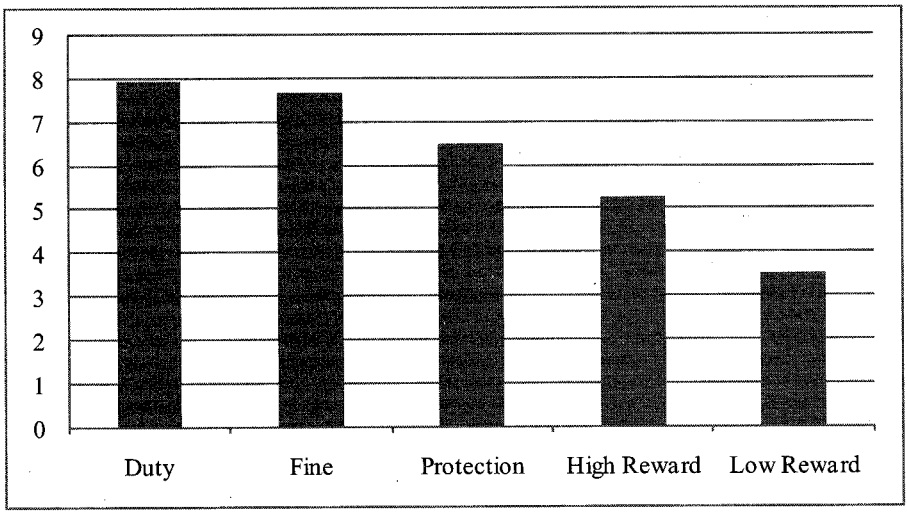
273. Using a simulator, which is available at http://www.dimensionresearch.com/resources/calculators/ttest.html.

274. The differences among the means were all significant (all p's < .001) with the exception of *Duty* and *Fine* ($p > .05$).

275. $F(3,930) = 31.32, p < .001, \eta^2 = .09$.

their peers were more affected by the reward when the reward was high.[276] Most interestingly, respondents thought that the size of the rewards was far more influential when it came to others and less influential when it came to themselves.[277]

Figure 6: Perceived Influence of the Different Legal Mechanisms



While the *High Reward* mechanism was more influential than *Duty* and *Fine* in the experimental setting, people perceived the imposition of a duty to be the most dominant mechanism in their decision to report, suggesting a misperception or a social-desirability bias in their own motivation to report. Similarly, in the experimental subgroups, where respondents were offered both rewards and were told they had a legal duty to report, respondents attributed their decision to report to the duty. Given that this disparity could not be seen in the actual effect of duty as measured in the experimental part,[278] we attribute the gap to a behavioral tendency to undermine the role of money in one's own behavior. Fines, and especially protection, were not perceived to be more or less influential according to the two subgroups, suggesting that they were less likely to be seen as representing external

276. $F(3,930) = 17.93, p < .001, \eta^2 = .05$. We asked respondents how influential the size of the reward was to their peers' decision to report relative to their own: more important than to themselves, same as themselves, or less than themselves. These responses were then categorized to the 1–10 scale.

277. Indeed, the difference in percentage of participants who graded higher the influence of the reward on their friend's decision to report in comparison to their own decision was found to be significantly higher in the *Duty + High Reward* and *High Reward* subgroups (29.6% and 29.5%, respectively) than in the *Duty + Low Reward* and *Low Reward* subgroups (7.5% and 1.5%, respectively). $\chi^2 (6, N = 1033) = 51.37, p < .001$.

278. In other words, the difference between *Duty + High Reward* and *High Reward* subgroups was marginal.

motivation. Overall, the dominance of the imposition of *Duty*, particularly in its perceived effect between the *Low* and *High* reward subgroups, suggests a tendency to overestimate the perceived effect of duty and to underestimate the perceived effect of reward.

C. *Framing the Social Meaning of Whistle-Blowing*

Given the importance of social norms and social approval in the decision to engage in reporting illegality, this subpart describes how participants predicted social perceptions of their decision to report. These perceptions are important, not only for understanding respondents' immediate decision to report but also in predicting long-term effectiveness of a given regulatory mechanism. If certain mechanisms are associated with negative status, their long-term efficacy will be jeopardized. As discussed in Part IV, law carries expressive meanings, and the mechanisms it chooses to employ are likely to shape the social understandings of whistle-blowing.

In our experiments, we measure social norms in various ways, including evaluating respect, attribution of virtue, law abidingness, and patriotism. The differences in the appreciation of actions taken against the company by others across the legal mechanisms are captured in Figure 7. Most strikingly, respondents viewed peers who reported without the promise of a reward as more worthy of their respect. The four *Reward* subgroups received the lowest evaluation levels, revealing a certain stigma[279] against those who blow the whistle in response to these mechanisms. In this sense, employees who blow the whistle in low monetary-rewards regimes are the worse off: they receive less glory for a limited amount of money that, as shown in the first section, was unlikely to be influential in their decision to report.

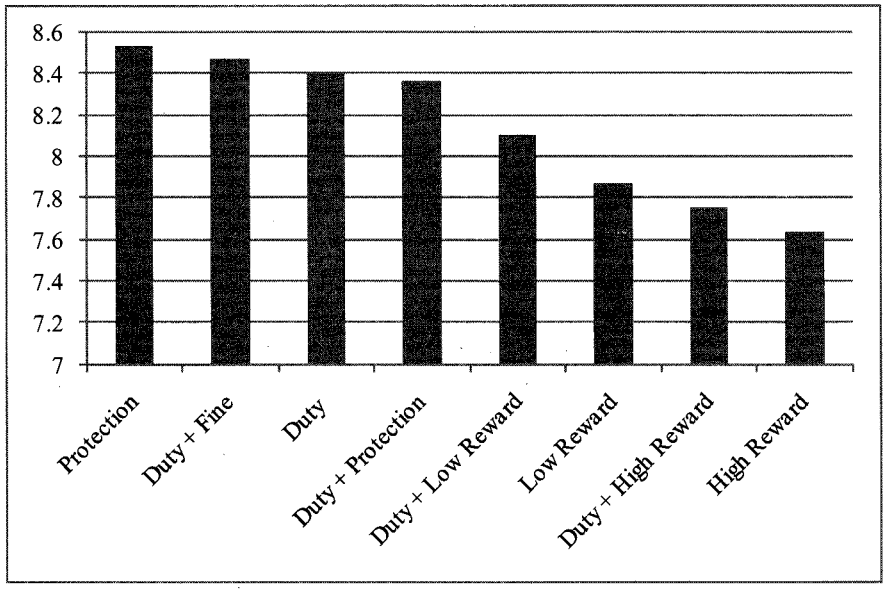
A further finding is that the existence of a duty caused participants to increase their appreciation for people who blew the whistle, while the introduction of a reward had the opposite effect of reducing respect for the informant. This result can be explained in two ways. Under one explanation, as hypothesized, the law has an expressive effect and a legal duty signals that certain behaviors are important.[280] Alternatively, we can understand this effect as a general appreciation that people have for others who are motivated by respect for the law rather than by money. This hypothesis draws some support from the finding that a duty was in fact more meaningful when a low, rather than high, reward was attached. This may

279. In fairness, we must be very careful not to overstate this argument. Given that even for the "less popular" conditions, values were above 7 in a 1–10 scale, it is hard to define such scores as being negative. For the most part, this Article focuses on trends and comparison among subgroups, and in that regard, some incentive mechanisms are more likely to create more or less positive evaluation than others. Presumably, if this study had been conducted in a different setting—i.e., involving financially related misconduct—we might have seen bigger differences given the presented gap.

280. *See supra* text accompanying notes 207–215.

suggest that high rewards overwhelm the attribution of others' motivations such that the duty is perceived as a minor consideration.

Figure 7: Evaluation of Others



An additional measure of social meaning for whistle-blowing is related to the perceived career impact on whistle-blowers. Respondents viewed the receipt of low rewards as especially likely to create the most severe career problem for individuals. The tendency for low rewards to only minimally influence the likelihood of reporting and, at the same time, to have highly negative career effects once again underscores the social price and inefficiency of this mechanism.[281]

VII. Discussion and Policy Implications

Through the experimental manipulation of different regulatory mechanisms attached to a common misconduct scenario, we sought to answer key questions about the optimal design of reporting mechanisms: Under which circumstances do individuals find themselves and others most likely to report misconduct? What are the ethical or professional motivations

281. Other than career effect, the other central aspect of social norms is the reaction of friends at work and family. We found that support by family was graded significantly higher. That is, in general, people expected a much stronger support for the enforcement behaviors from their family than from their friends at work. However, in contrast to career effect, there was no significant interaction effect between the legal mechanism and the type of support. This means that in contrast to career effects, the reactions of family and friends were not perceived by people to be sensitive to the mechanisms that were used in that context. Family ($M = 8.06$, $SD = 2.13$) than support by friends at work ($M = 5.53$, $SD = 2.36$) $F(1,1896) = 103.41$, $p < .001$, $\eta^2 = .05$.

driving these differences in reporting behavior? How are different types of reporting incentives perceived, and what kind of environments can be created to induce reporting?

The findings of the study provide both practical and theoretical insights. The study clearly demonstrates that the choice of mechanism significantly influences decisions to report illegal conduct in the workplace. Moreover, our findings provide insight into the subtle interactions between legal mechanisms and factors such as gender, the perceived severity of the misconduct, and the social consequences of whistle-blowing. These findings can help policy makers refine their regulatory toolbox and design tailored and efficient incentive mechanisms.

A. The Inadvertent Effect of Rewards

The experimental study clearly indicates that levels of monetary rewards are consequential, frequently affecting levels of reporting. Only when actions are perceived as morally offensive do reward levels have minimal effects on actions. When the ethical significance attached to the reporting act is absent, the level of monetary compensation offered through the regulatory system is decisive. Most surprisingly, our experiment shows that where misconduct is expected to evoke a lower level of moral outrage, the introduction of small bounties may actually decrease the rate at which it is reported. In the experiment, reporting under such circumstances was even lower than in situations where no incentive was present.

In some respects, these findings continue the studies regarding a crowding-out effect. The experiment demonstrates[282] that framing reporting as a commodity with a price tag attached may actually suppress internally motivated action. At the same time, the findings demonstrate that the crowding-out effect largely disappears with the introduction of sufficiently high monetary rewards. From a practical perspective, legislatures can use these insights in developing and adjusting existing bounty programs to appropriate levels of monetary rewards, so as not to produce counterproductive effects by the introduction of money as a legal mechanism to induce compliance.

B. Motivation, Perception, and Whistle-Blowing

From a more theoretical perspective, our study provides significant contributions to the literature on motivation and behavior. The experiments show that individuals perceive their own actions as more motivated by intrinsic ethical concerns than the actions of others. In the absence of a legal duty to report, the introduction of a higher financial reward is perceived to have the greatest impact on the reporting behavior of others rather than on self-

282. Albeit, the experiment demonstrates this effect in a limited way—the crowding-out effect was significant only with regard to women and with regard to the perception of others.

behavior. Respondents felt that such higher bounties would only moderately increase their own reporting rates, as well as the reporting of their peers, yet they predicted that bounties would lead to significantly higher whistle-blowing activity for nonpeers. These findings are generally in line with the psychological holier-than-thou effect—the general belief of individuals that they themselves are more ethically driven than others.[283] The respondents in our experiments expected others to exhibit more self-interested attitudes in their decision to report wrongdoing. Overall, people overestimated the effect of duty on their own behavior and underestimated the effect of reward on their behavior. In fact, high rewards were highly influential at the experimental stage, but those high rewards were ranked as the least influential factors when respondents were consciously estimating what factors influenced their own decisions to report.

We believe that this perceptual gap should receive more attention in regulatory discourse in general and particularly in the whistle-blowing context. Often, the decision to blow the whistle depends on whether one believes others will come forward as well. For example, if a financial reward is offered, employees who think that others would have preceded them may behave strategically and try to be the first to collect the reward. These questions, exploring such strategic behavior and the predictions of others' behavior, are highly relevant to recent legal debates, as exemplified by the split decision in *Rockwell International Corp. v. United States*.[284] Furthermore, given the overestimation of the instrumental self-interested motivation of others, the introduction of monetary incentives by regulatory agencies may exacerbate perceptions of greed to a greater extent than it encourages desirable individual decision making.

C. One Policy Fits All?

The study points to major variation in the importance of regulatory design in relation to different illegalities. The greater the perception of severity of the misconduct, the less important the choice among the regulatory mechanisms. In our analysis, we used severity of the misconduct as a proxy for internal motivation.[285] In the group of participants that viewed the illegality as highly offensive, and who hence had high levels of internal motivation for reporting, the type of mechanism available to informants was largely irrelevant. Respondents expected the reporting levels of themselves and others to remain consistently high across all categories of legal

283. *See* Emily Balcetis & David Dunning, "Holier Than Thou" Self-impressions and Social Attribution—And How Experience Diminishes the Fundamental Attribution Error 12 (2006) (unpublished manuscript, on file at http://oak.cats.ohiou.edu/~balcetis/attributions.pdf) ("[P]eople appear to hold the strongest holier-than-thou self-impressions when it comes to the social virtues of altruism and morality.").

284. 549 U.S. 457 (2007).

285. These included items such as moral outrage, legitimacy, and perceived risk from the misconduct.

mechanisms. However, when illegalities witnessed by potential enforcers were perceived as less severe, the use of high rewards and fines produced considerably higher levels of reporting than the use of low rewards. These findings suggest the importance of legal-mechanism selection in instances where individuals do not have an ethical stake in compliance. In such cases, triggering external motivation through regulatory policy takes on a far greater role in promoting reporting activity. Therefore, regulatory agencies may consider providing high monetary rewards when the goal is to incentivize reports in contexts that evoke less moral outrage, such as tax evasion.

In addition to the contextual variation, our study revealed dramatic gender differences. On balance, women were more likely than men to report misconduct. Women were also more likely to be motivated by internal ethical considerations and to view other people's behavior as similarly ethically motivated. For men, the decision to report was largely externally driven, such that the introduction of a higher reward significantly increased the rate at which men decided to report, both with and without a duty. Between the gender groups, the greatest difference existed where whistle-blowers were provided protection from retaliation, for which women were significantly more likely to report than men. On balance then, rewards were more effective to incentivize reporting by men, while imposing duties, and to some extent granting antiretaliation protections, were more important to women.

More generally, our findings on the interaction between internal motivation and the choice of legal mechanism suggest the following important policy insight. In areas where the misconduct is expected to trigger high internal motivation, there is less need to invest in incentive mechanisms. This is contrary to the basic intuition of the legal policy maker to give higher rewards as the misconduct is more severe (given its likely correlation with greater harm to society). Our findings suggest that when the misconduct is viewed to be severe, the type of mechanism used is less important. In areas where the misconduct is likely to be viewed, at least by some of the people, as severe, there is less need to use rewards that carry both monetary costs for the state and some social cost for the whistle-blower herself.

D. Incommensurability

The findings also bear on the general problem of "incommensurability," the difficulty people have in using monetary or other quantitative metrics to assess morality, beliefs, and behavior.[286] In the context of whistle-blowing, this problem takes on a crucial role as individuals must necessarily make

286. *See, e.g.*, ELIZABETH ANDERSON, VALUE IN ETHICS AND ECONOMICS 66–73 (1993) (explaining the human propensity to conceptualize goods hierarchically, with higher goods, such as justice and friendship, being immeasurable by comparison to lower goods, such as money and convenience); JOSEPH RAZ, THE MORALITY OF FREEDOM 321 (1986) ("[B]oth values and valuables are to a large degree incommensurable . . . .").

normative and financial judgments in deciding whether to report misconduct. Many thinkers are skeptical about the ability of people to translate such qualitative judgments into a price.[287] Our findings reveal that this problem may be further compounded in the context of incentive-driven reporting, where people appear to think differently about their own motivations and the behavior of others. Where people tend to portray themselves as internally motivated, duty-bound citizens, the reporting decisions of others are viewed as opportunistic actions that are primarily driven by external rewards. That these Good Samaritan self-assessments do not comport with actual behavior, which tends to resemble the reward-driven calculus that people attribute to others, may suggest that it is easier for people to predict commensurability for others but not for themselves. Related to this issue of incommensurability is the problem of translating moral judgments into action. In an experimental study about penalties, Cass Sunstein, Nobel Laureate Daniel Kahenman, and their colleagues described what they call "the translation problem" as a significant source of incoherence in various legislative penalties.[288] They explain the translation problem as "the distinctive problem involved in translating a moral judgment of some kind into the terms made relevant by the legal system, such as monetary penalties, civil fines, or criminal punishment."[289] Their study indicates that such translation is frequently problematic because it is grounded neither in principle nor on a set of shared intuitions.[290] While the translation problem is pervasive, our findings suggest that when a large enough gap exists between high and low rewards, people's perceptions and behaviors will differ significantly, affecting judgments and internal motivations.

E. The Choice of Mechanism and Whistle-Blowing Status

Additionally, the results of our experiments reveal significant tendencies in the evaluation of reporting behaviors. Duty-based reporting provides the highest status to a whistle-blower, while rewards carry the highest levels of social stigma. Even when the rewards are low, are combined with a duty, or both, decisions to report in response to these mechanisms uniformly receive the lowest levels of social respect and appreciation. On the other hand, fines and duty were relatively neutral in this regard, drawing consistently high levels of social admiration for the reporter. These findings contribute an additional dimension to the process of designing effective regulatory mechanisms: even where certain mechanisms may be intrinsically

287. *Cf.* ANDERSON, *supra* note 286, at 68–69 (postulating that common-sense morality prohibits the comparison of value between higher goods and lower goods and using the example of the practical impossibility of valuing the cost of another's humiliation in terms of the fun derived by bystanders at the person's humiliation).

288. Cass R. Sunstein et al., *Predictably Incoherent Judgments*, 54 STAN. L. REV. 1153, 1155 (2002).

289. *Id.*

290. *Id.*

or financially appealing, the social praise and stigmas attached to certain reporting behaviors may add unexpected weight to the whistle-blowing scales. Such a view, which will take into account the social aspects of the incentive mechanisms used, might shift the cost–benefit balance of the mechanism-choice dilemma. For example, our study suggests that the existence of a duty is likely to increase the social status of whistle-blowers to a greater degree than it affected the levels of reporting. Yet, if there are long-term effects on the social status of whistle-blowers in a given community, imposing a duty to disclose is important even beyond its immediate ability to promote more reporting. Similar effects were found with regard to fines.[291] Our study therefore suggests the desirability of imposing duties and fines, as they tend to enhance the social status of whistle-blowers and reduce the stigma associated with rewards and legal protections.

However, the association between the choice of legal mechanisms and the social status of the reporter varies significantly in the various aspects of our study. Participants viewed their own behavior as less affected by external mechanisms and viewed their own perceptions of others as less affected by choice of mechanism. Furthermore, participants did not attach high importance to anonymity with regard to their own actions. These gaps between self and others in the interpretation of reporting behavior underscore the existence of a holier-than-thou effect. This effect further suggests that people's self-perception of their own higher moral standards may lead them to believe that others view them in the same light. Alternately, this effect may be explained as a type of fundamental attribution error in which people construe their own reporting behavior as an expression of their character, irrespective of the legal mechanisms involved. At the same time, we do find that individuals fear negative implications for their career even as they perceived their actions and the reaction to their reporting by their peers and family to be positive.

VIII. Conclusion

Social enforcement has become a central feature in the design of effective regulatory systems. In many areas of law, government relies on private individuals to trigger detection of misconduct through reporting. This study demonstrates that empirical research, comparing the behavioral and social ramifications of the regulatory mechanisms used to incentivize social enforcement, carries important practical insights for policy. By exploring the interplay between internal and external enforcement motivation, these experiments provide novel understandings about the comparative strengths and

291. Fines were seen as reducing the perception of greed associated with blowing the whistle for a reward. Common to both mechanisms is their removing the voluntary nature of whistle-blowing. Apparently, at least in the context of reporting fraudulent behavior, participants thought more of people who blew the whistle because they were obliged to do so, rather than where they did so voluntarily.

advantages of different legal mechanisms. An important implication of the study is that no one-size-fits-all policy design exists, but rather, policy makers must evaluate the full scope of psychological and situational factors in order to design the most efficient incentive structures. In doing so, policy makers must consider several factors. First, policy makers must assess the nature and severity of the anticipated conduct. Where levels of moral outrage are expected to be low, financial rewards will likely be a decisive factor, and the inquiry may shift to discovering the true price tag of the reporting behavior. For inherently offensive misconduct, policy design must take a more nuanced approach that integrates the moral dimension of the situation. In such cases, where the informant is expected to have a greater ethical stake in the outcome, regulation must fully appeal to the informant's sense of duty. This may mean that financial incentives are not only unnecessary but are counterproductive and offset internal motivations to report. Identifying such crowding-out effects in regulatory design is particularly beneficial, as it can save public dollars while simultaneously pointing to better mechanisms to induce reporting.

The study, moreover, demonstrates that informed policy makers must factor in the possibility that informants may underestimate the role of financial incentives in their own decision to report. Whereas people perceive others as reporting mainly for money, they tend to perceive their own social-enforcement actions as more ethically driven. Moreover, in choosing among the various mechanisms available, the study demonstrates that stigma levels attached to reporting vary along the selected mechanism. Finally, policy makers must also consider the target individuals for which the regulatory mechanisms are provided. In particular, gender differences are highly pronounced in the realm of whistle-blowing and further research should consider the reasons for which women care much more than men about antiretaliation protections.

At a broader level, the study contributes to the empirical literature about individual and group behavior, including debates on motivational crowding out, trust, misperception of norms and attribution, and the ability of individuals to rationally balance the costs and benefits of their own decisions and to assess the behavior and interests of others. Despite the growing interest in the legal academy and practice in new-governance approaches to law and policy, the study of individual motivation and behavior as directly connected to legal design is in its developmental stages. The study reported in this Article offers an important scholarly contribution and a step forward in this nascent interdisciplinary field.

Appendix I: Statistical Tables

Table 1: Adjusted Mean Scores and Standard Deviations for the Actions Taken Against the Company as a Function of the Legal Mechanism

		Imitators of Whistle-blowing					
		Most Peers		Most People		Self Actions	
		Mean	St. Dev.	Mean	St. Dev.	Mean	St. Dev.
Legal Mechanism	Duty + Low Reward	6.46	2.75	5.43	2.64	8.36	2.27
	Duty + High Reward	6.90	2.64	6.18	2.52	8.71	1.94
	Low Reward	6.15	2.81	5.31	2.55	7.91	2.45
	High Reward	6.54	2.66	6.48	2.59	8.50	2.14
	Duty + Protection	6.30	2.55	5.23	2.47	8.39	2.20
	Protection	6.31	2.81	5.16	2.48	8.46	1.95
	Duty	5.94	2.87	5.15	2.62	8.39	2.11
	Duty + Fine	6.14	2.67	5.21	2.44	8.53	2.17

Note: The score ranged from 1–10. Higher values indicate higher likelihood of action taken against the company (n = 1909).

Table 2: Adjusted Mean Scores and Standard Deviations for Action as a Function of the Evaluation of the Misconduct Across Legal Mechanisms

		Actions Taken Against the Company					
		Most Peers		Most People		Self Actions	
Legal Mechanism	Evaluation of the Misconduct	Mean	St. Dev.	Mean	St. Dev.	Mean	St. Dev.
Duty + Low Reward	Low	6.24	2.55	5.16	2.37	7.55	2.45
	High	6.71	2.94	5.72	2.89	9.26	1.63
Duty + High Reward	Low	6.67	2.54	5.92	2.44	8.21	2.21
	High	7.16	2.73	6.48	2.60	9.30	1.33
Low Reward	Low	5.71	2.51	5.14	2.36	7.08	2.59
	High	6.65	3.03	5.50	2.74	8.85	1.85
High Reward	Low	6.42	2.47	6.30	2.44	8.30	2.17
	High	6.70	2.89	6.70	2.76	8.73	2.08
Duty + Protection	Low	6.09	2.60	5.12	2.43	7.66	2.40
	High	6.53	2.48	5.36	2.53	9.23	1.51
Protection	Low	6.09	2.52	5.09	2.34	7.92	2.23
	High	6.49	3.01	5.21	2.59	8.86	1.58
Duty	Low	5.99	2.75	5.18	2.59	7.90	2.15
	High	5.89	3.01	5.11	2.66	8.92	1.95
Duty + Fine	Low	6.16	2.56	5.34	2.40	7.98	2.40
	High	6.10	2.82	5.05	2.49	9.12	1.66

Note: The scores ranged from 1–10. Higher values indicate higher likelihood of action taken against the company (n = 1909).

Table 3: Adjusted Mean Scores and Standard Deviations for the Actions Taken Against the Company as a Function of Gender Across the Legal Mechanisms

		Actions Taken Against the Company					
		Most Peers		Most People		Self Actions	
Legal Mechanism	Gender	Mean	St. Dev.	Mean	St. Dev.	Mean	St. Dev.
Duty + Low Reward	Male	6.48	2.62	5.73	2.66	8.16	2.46
	Female	6.41	2.90	5.07	2.56	8.54	2.05
Duty + High Reward	Male	6.73	2.69	6.13	2.56	8.78	1.92
	Female	7.12	2.59	6.26	2.51	8.67	1.97
Low Reward	Male	6.11	2.75	5.42	2.46	7.75	2.49
	Female	6.19	2.91	5.19	2.69	8.05	2.41
High Reward	Male	6.94	2.58	6.79	2.65	8.59	2.11
	Female	6.07	2.69	6.12	2.48	8.42	2.18
Duty + Protection	Male	6.22	2.56	5.21	2.36	7.85	2.42
	Female	6.42	2.56	5.27	2.62	9.04	1.75
Protection	Male	6.27	2.92	5.22	2.57	8.13	2.05
	Female	6.36	2.68	5.09	2.36	8.77	1.84
Duty	Male	5.79	2.97	4.75	2.75	8.24	2.40
	Female	6.13	2.81	5.57	2.46	8.56	1.79
Duty + Fine	Male	6.21	2.64	5.43	2.32	8.42	2.21
	Female	6.04	2.72	4.92	2.58	8.67	2.13

Note: The score ranged from 1–10. Higher values indicate higher likelihood of action taken against the company (n = 1896).

Table 4: Adjusted Mean Scores and Standard Deviations for Respect of Others Decision to Report the Misconduct as a Function of the Legal Mechanism

		Evaluation of Others' Decision to Report	
		Mean	St. Dev.
Legal Mechanism	Duty + Low Reward	8.10	1.66
	Duty + High Reward	7.75	1.67
	Low Reward	7.87	1.63
	High Reward	7.64	1.78
	Duty + Protection	8.36	1.66
	Protection	8.53	1.55
	Duty	8.40	1.69
	Duty + Fine	8.47	1.62

Note: The scores ranged from 1–10. Higher values indicate higher evaluation of others' decision to report the misconduct (n = 1922).

Appendix II: Excerpts from the Survey Instrument[292]

A. The Survey

The next survey deals with the readiness of an employee to report on an employer. We will now present you with a hypothetical situation. Please read it carefully and answer the questions that follow.

> Imagine you are an employee of Roadblock LTD, one of the largest construction companies in the country. Roadblock has recently secured a fixed-price government contract to build a major highway in your city.
>
> One day, while staying late in the office, you run across a document that reveals that the company has been substituting lower-grade and inferior-quality parts from those specified in the contract. The document also reveals that the company has been omitting required testing and quality procedures. You estimate that as a result the government is overpaying your employer approximately $10,000,000.

1. Condition 1 (Duty + Low Reward).—According to the Government Fraud Act, as an employee of a government contractor, you have a legal duty to inform the government of any fraudulent behavior of your employer.

292. The full text of the survey questions are on file with the authors and will be sent upon request.

Furthermore a person who files an action against federal contractors claiming fraud against the government, stands to receive a portion of the recovered fraudulent billing. In this case, you are estimated to receive a reward of $1,000.

2. *Condition 2 (Duty + High Reward).*—According to the Government Fraud Act, as an employee of a government contractor, you have a legal duty to report to the government any fraudulent behavior of your employer. Furthermore, a person who files an action against federal contractors claiming fraud against the government stands to receive a portion of the recovered fraudulent billing. In this case, you are estimated to receive a reward of $1,000,000.

3. *Condition 3 (Low Reward).*—According to the Government Fraud Act, a person who files an action against federal contractors claiming fraud against the government stands to receive a portion of the recovered fraudulent billing. In this case, you are estimated to receive a reward of $1,000.

4. *Condition 4 (High Reward).*—According to the Government Fraud Act, a person who files an action against federal contractors claiming fraud against the government stands to receive a portion of the recovered fraudulent billing. In this case, you are estimated to receive a reward of $1,000,000.

5. *Condition 5 (Duty + Protection).*—According to the Government Fraud Act, as an employee of a government contractor, you have a legal duty to report to the government any fraudulent behavior of your employer. Furthermore, according to the Government Fraud Act, a person who files an action against federal contractors claiming fraud against the government cannot be fired because of the reporting.

6. *Condition 6 (Protection).*—According to the Government Fraud Act, a person who files an action against federal contractors claiming fraud against the government cannot be fired because of the reporting.

7. *Condition 7 (Duty).*—According to the Government Fraud Act, as an employee of a government contractor, you have a legal duty to report to the government any fraudulent behavior of your employer.

8. *Condition 8 (Duty + Fine).*—According to the Government Fraud Act, as an employee of a government contractor, you have a legal duty to report to the government any fraudulent behavior of your employer. Furthermore, failure to report will result in a fine of $10,000.

* * *

Keep Charity Charitable

Brian Galle*

This Article responds to recent claims, most prominently by Anup Malani, Eric Posner, and Todd Henderson, that much of the work of the charitable sector should be farmed out to for-profit firms. For-profit firms are said to be more efficient because they can offer high-powered incentives to cut costs. I argue, however, that because of the high costs of monitoring and the presence of externalities, low-powered incentives are preferable for firms that produce public goods, as most charities do. Further, allowing some for-profit firms to receive charitable subsidies would raise the cost of producing those goods in government or other firms because it would diminish the "warm glow" workers enjoy from being recognized as self-sacrificing.

I. Introduction

Everyone likes charity. In the United States, a measure of our warm regard is the substantial tax support we offer to those who make cash or in-kind contributions towards charitable endeavors. Ordinarily such largesse is tied to an expectation that those who receive the funds will operate in a charitable way—that they will commit through their organization charter to forgo distribution of profits to themselves or their officers. Leading theorists of the field have maintained that charity could not long maintain its popularity, or even function, without such promises.

Recently, though, some commentators have begun to argue in favor of what might be called "for-profit" charity.[1] For example, Anup Malani and Eric Posner argue that philanthropic services could be carried on equally well by for-profit firms.[2] Pointing to the charitable efforts of Google and other money-making enterprises, they claim that the deduction is just another form of government contract for the delivery of public goods, much like payments to developers of alternative energy or private security for the Department of

* Visiting Associate Professor, George Washington University Law School; Assistant Professor, Florida State University College of Law. I am grateful for helpful comments from and conversations with Ellen Aprill, Rob Atkinson, Bill Bratton, Joe Dodge, Daniel Halperin, Don Langevoort, Sarah Lawsky, Ben Leff, Brendan Maher, Dan Markel, Gregg Polsky, Eric Posner, Katie Pratt, David Schizer, Ted Seto, Kirk Stark, and attendees of the American Association of Law Schools Section on Scholarly Paper Award Winners, as well as of presentations at Georgetown, George Washington, Harvard, Loyola-L.A., and Prawfsfest-FSU.

1. *E.g.*, Todd Henderson & Anup Malani, *Corporate Philanthropy and the Market for Altruism*, 109 COLUM. L. REV. 571, 576 (2009); Anup Malani & Eric A. Posner, *The Case for For-Profit Charities*, 93 VA. L. REV. 2017, 2019–23 (2007).

2. Malani & Posner, *supra* note 1, at 2019–23.

State.[3] They then offer a series of challenges to the conventional wisdom that these contracts can only be carried out by firms subject to the nondistribution constraint.[4] A few other commentators have joined with tentative endorsements,[5] albeit sometimes with suggestions for careful internal governance.[6]

Although these efforts are appealingly counterintuitive, this is one instance in which intuition is not only right but also determinative. As I will argue, the fact that society *perceives* an organization as charitable is a critical element of the entity's success. By opening philanthropy to potential profiteering, Malani, Posner, and their allies would dilute the power of these perceptions for every firm, including those that remain wholly charitable. And by inviting firms to reward their employees with top payouts for top "performance," they risk seriously compromising the quality of core charitable services—those that cannot be produced in a traditional profit-seeking market precisely because they cannot easily be measured. These dangers far outweigh any potential efficiency to be gained from encouraging charity to cut costs.

---

3. *Id.* at 2019–21. For other recent (and largely positive) discussions of the noncharitable-charity model, see DAN PALLOTTA, UNCHARITABLE: HOW RESTRAINTS ON NONPROFITS UNDERMINE THEIR POTENTIAL 17, 116–25 (2008); Dana Brakman Reiser, *For-Profit Philanthropy*, 77 FORDHAM L. REV. 2437, 2438 (2009); James J. Fishman, *Wrong Way Corrigan and Recent Developments in the Nonprofit Landscape: A Need for New Legal Approaches*, 76 FORDHAM L. REV. 567, 598 (2007); Victor Fleischer, *Urban Entrepreneurship and the Promise of For-Profit Philanthropy*, 30 W. NEW ENG. L. REV. 93, 93 (2007); and Shruti Rana, *From Making Money Without Doing Evil to Doing Good Without Handouts: The Google.org Experiment in Philanthropy*, 3 J. BUS. & TECH. L. 87, 89 (2008). The debate is long-standing in the hospital field. *See, e.g.*, M. Gregg Bloche, *Corporate Takeover of Teaching Hospitals*, 65 S. CAL. L. REV. 1035, 1092 (1992) (arguing that for-profit hospitals can exist side-by-side with nonprofits). Also related is the rise of collaborations between nonprofits and for-profit firms; for a survey, see MARTHA MINOW, PARTNERS, NOT RIVALS: PRIVATIZATION AND THE PUBLIC GOOD 7–22 (2002).

Additionally, while this Article was in press, Jim Hines, Jill Horwitz, and Austin Nichols provided the author with a copy of their own take on for-profit charity. James R. Hines, Jr. et al., *The Attack on Nonprofit Status: A Charitable Assessment*, 108 MICH. L. REV. 1179 (2010). Their approach is similarly critical of Malani and Posner, but largely for different reasons than those I outline here. *See id.* at 1207–18 (arguing that nonprofits have some incentives to be cost-effective and possess tools to encourage employees to do so; that nonprofits do not necessarily overproduce; and that allowing for-profit charities could allow tax arbitrage).

4. Malani & Posner, *supra* note 1, at 2029–56. The nondistribution constraint is a scholarly term for the prohibition on the distribution of a firm's net revenue to any person or entity. *See infra* notes 17–22 and accompanying text. All 501(c)(3)-eligible firms must abide by the nondistribution constraint. 26 U.S.C. § 501(c)(3) (2006). This means they cannot issue stock or pay their top-tier managers with a share of the firm's revenues.

5. *See, e.g.*, Fishman, *supra* note 3, at 598 (celebrating the emergence of for-profit charities as an "encouraging development"); Brakman Reiser, *supra* note 3, at 2438 (arguing that for-profit charity challenges the assumptions of current charity law).

6. *See, e.g.*, Rana, *supra* note 3, at 89 (suggesting that Google.org "must develop innovative accountability structures to match its ambitious goals"); Dana Brakman Reiser, *Charity Law's Essentials* 34–40 (Brooklyn Law Sch. Legal Studies Paper No. 167, 2009), *available at* http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1479572 (advocating group governance to reduce risks in for-profit charities).

Thus, I argue here that federal law should continue to insist that only true nonprofit organizations should be eligible to receive deductible charitable contributions. This claim also implies that firms eligible to receive the deduction cannot pay their key employees with a share of the organization's profits. I do not consider several related but distinct questions, such as whether firms that engage in charity should be exempt from the corporate income tax or whether, at a minimum, a firm should get a corporate tax deduction for the money it spends on philanthropic activity it carries out itself.

The Article proceeds in five Parts. Part II sets out in more detail the theoretical basis for government support for the charitable sector and the traditional explanation—usually associated with the work of Henry Hansmann—for why such support must be tied to a promise not to distribute profits. As Part III explains, Malani and Posner argue that this promise can be replaced with a contract with a private auditing firm. Part III claims that this proposal would increase the social costs of monitoring and inefficiently shift those costs to the charitable sector. Part IV takes issue with claims by for-profit charity proponents that the nonprofit sector is a less efficient producer of charitable services, emphasizing the role of warm glow and low-powered incentives. Part V considers briefly another alternative to traditional charity floated by Malani and Posner. Part VI concludes.

II. Theories of Subsidized Charity

Modern commentators view the deduction for charitable contributions as a federal subsidy to the recipient firms and argue that the subsidy is justified as a tool for encouraging the production of goods that would otherwise be underproduced by the private market.[7] Mostly these consist of public goods—goods whose use can be shared by many consumers and for which it

7. *See, e.g.*, John D. Columbo, *The Marketing of Philanthropy and the Charitable Contributions Deduction: Integrating Theories for the Deduction and Tax Exemption*, 36 WAKE FOREST L. REV. 657, 698 (2001) (contending that one should view the charitable-contribution deduction as a subsidy remedying a private market failure); Mark P. Gergen, *The Case for a Charitable Contributions Deduction*, 74 VA. L. REV. 1393, 1394 (1988) (accepting that one can justify the charitable-contributions deduction as a subsidy); Henry Hansmann, *The Rationale for Exempting Nonprofit Organizations from Corporate Income Taxation*, 91 YALE L.J. 54, 68, 71 (1981) (explaining that one can justify a charitable-contribution deduction because it stimulates production of an otherwise scarce good); David E. Pozen, *Remapping the Charitable Deduction*, 39 CONN. L. REV. 531, 547 (2006) (stating that modern scholars have justified charitable-contribution deductions as a matter of public policy as a public subsidy); Daniel Shaviro, *Assessing the "Contract Failure" Explanation for Nonprofit Organizations and Their Tax-Exempt Status*, 41 N.Y.L. SCH. L. REV. 1001, 1007 (1997) (reviewing the basic argument that the charitable-contributions deduction encourages the production of a public good); *see also* OFFICE OF MGMT. & BUDGET, EXECUTIVE OFFICE OF THE PRESIDENT, ANALYTICAL PERSPECTIVES: BUDGET OF THE UNITED STATES GOVERNMENT FISCAL YEAR 2008, at 257–59 (2007), *available at* http://www.gpoaccess.gov/usbudget/fy08/pdf/spec.pdf (arguing that the federal government should provide tax incentives for charitable giving).

would be relatively difficult for the producer to exclude users.[8] Because of these features, there is a private-market failure in the production of public goods.[9] Individuals have a strong incentive to free ride on others' consumption, and thus it is difficult for any producer to sell the good at a profit.[10] Even if some goods can be sold, the market will probably produce less than the socially optimal amount of the good.[11]

One classic example of a public good is the fireworks display.[12] Once I pay for the show, all of my neighbors can camp out and enjoy the evening without paying; therefore, no one of us has much reason to be the one who pays. Even if I am willing to pay, I may purchase fewer fireworks than would satisfy everyone because I am indifferent to others' consumption. However, if there is a subsidy for purchasing fireworks, I will buy more of them due to both the income and substitution effects of the subsidy. Ideally, the subsidy amount would be set so that I will purchase exactly the amount of fireworks that would maximize every viewer's preferences.[13]

This public-goods explanation for the charitable-contribution deduction accords well with federal law. Current law generally conditions eligibility for deductible contributions on private-market failures.[14] The IRS can audit

---

8. On the general theory of public goods, see RICHARD A. MUSGRAVE & PEGGY B. MUSGRAVE, PUBLIC FINANCE IN THEORY AND PRACTICE 42–48 (5th ed. 1989).

9. *Id.* at 44–45; *see also* Gergen, *supra* note 7, at 1397–98 ("Left to its own devices without a tax subsidy, [private charity] may not be able to overcome its freerider problems to provide the appropriate amount of the good."); Hansmann, *supra* note 7, at 72 (arguing that without incentives charitable contributions to nonprofits are below socially optimal levels).

10. MUSGRAVE & MUSGRAVE, *supra* note 8, at 44.

11. Hansmann, *supra* note 7, at 72.

12. *See, e.g.*, JONATHAN GRUBER, PUBLIC FINANCE AND PUBLIC POLICY 184–85 (2d ed. 2006) (explaining that a fireworks display is a public good because it can be shared by many consumers and it is difficult to exclude consumers from the display).

13. Although I refer throughout this Article to "public goods," the reader should understand that the discussion for the most part also applies more generally to any private good with a significant positive externality attached. *See id.* at 171 ("It is helpful to think about a public good as one with a large positive externality."); MUSGRAVE & MUSGRAVE, *supra* note 8, at 49–50 (noting that the two are largely equivalent, albeit that a smaller subsidy is needed to produce private goods with positive externalities).

Similarly, in some cases, some of what I refer to as public goods might more precisely be described as "club goods" because in theory they could be "fenced" and made accessible only to "members." Again, some club goods may also produce externalities for nonmembers, such as an exclusive park with historic social significance. Here, again, there will be underprovision of the nonprivate aspects of consumption; the club will not necessarily use the park in a way that is consistent with historic preservation. For a survey of other arguments for government production of potential club goods, see Amnon Lehavi, *Property Rights and Local Public Goods: Toward a Better Future for Urban Communities*, 36 URB. LAW. 1, 16–24 (2004).

14. *See* BRUCE R. HOPKINS, THE LAW OF TAX-EXEMPT ORGANIZATIONS § 4.10, at 103–08 (9th ed. 2007) (explaining that the IRS may deny eligibility on the basis that an "organization's operation is similar to a commercial enterprise operated for profit" and providing examples).

firms and revoke exemptions or impose excise taxes on firms and managers if they produce only private goods or distribute profits.[15]

As Henry Hansmann argues, this justification for the deduction also implies that beneficiaries of the deduction must likely be limited to nonprofit organizations.[16] A nonprofit, in Hansmann's formulation, is one that is subject to the "nondistribution constraint": it can make a profit, but it must use these profits for internal development rather than distributing them to investors or managers.[17] The logic is that donors cannot easily judge the quality of public goods, especially where those goods are delivered to someone other than the donor.[18] As a result, it would be easy for managers of the firm to divert donations to their own personal benefit.[19] Knowing this, donors will not give to the firm.[20] Hansmann calls this phenomenon "contract failure."[21] In response, the firm voluntarily takes on the nondistribution constraint so that donors know that their giving will not be wasted.[22]

Several factors contribute to the difficulty of monitoring the quality of public goods and hence to contract failure. First, many public goods are inherently ineffable or controversially defined.[23] For instance, donating to a soup kitchen might contribute to a more just society. While we can easily

---

15. *See* 26 U.S.C. §§ 4941–4942, 4958 (2006) (authorizing the IRS to revoke exemptions); IRS, PUB. NO. 4221-PC, COMPLIANCE GUIDE FOR 501(C)(3) PUBLIC CHARITIES 2–3 (2009), *available at* http://www.irs.gov/pub/irs-pdf/p4221pc.pdf (outlining IRS enforcement procedures for charitable entities).

16. *See* Henry B. Hansmann, *Reforming Nonprofit Corporation Law*, 129 U. PA. L. REV. 497, 623 (1981) (arguing that government subsidy to charitable organizations justifies federal oversight of nonprofit status); *see also* Rob Atkinson, *Altruism in Nonprofit Organizations*, 31 B.C. L. REV. 501, 617–18 (1990) (explaining that under Hansmann's analysis "only nonprofit firms address the kinds of contract failure and experience the kinds of capital constraints that warrant" a subsidy).

17. Hansmann, *supra* note 16, at 501; Henry B. Hansmann, *The Role of Nonprofit Enterprise*, 89 YALE L.J. 835, 838 (1980).

18. Hansmann, *supra* note 16, at 506; *see also* Atkinson, *supra* note 16, at 572 (arguing that managers of a for-profit firm would need to provide an "altruistic investor" with an enforceable commitment "that they will use the imputed return on her donated capital to subsidize consumption").

19. Hansmann, *supra* note 7, at 68–70; *cf.* Kenneth J. Arrow, *Uncertainty and the Welfare Economics of Medical Care*, 53 AM. ECON. REV. 941, 964–65 (1963) (suggesting that the profit-making aspect of hospitals is relatively unimportant because doctors' commitment to professionalism helps to assure patients unable to measure quality that the doctors will prioritize care over profits).

20. Hansmann, *supra* note 7, at 69–70.

21. *Id.* at 69; Hansmann, *supra* note 16, at 506–07.

22. Hansmann, *supra* note 16, at 507; Malani & Posner, *supra* note 1, at 2033–34; *see also* Susan Rose-Ackerman, *Altruism, Nonprofits, and Economic Theory*, 34 J. ECON. LIT. 701, 716 (1996) ("[D]onors may be willing to donate only to nonprofit institutions. Given the difficulty of monitoring charitable work, donors may fear that for-profit firms will convert gifts into profits for the owners.").

23. *See* John D. Donahue, *Market-Based Governance and the Architecture of Accountability*, *in* MARKET-BASED GOVERNANCE 1, 5–6 (John D. Donahue & Joseph S. Nye, Jr. eds., 2002) (describing the difficulties of arriving at universal definitions of success in areas of public interest, such as schools and taxes).

measure the calories in the soup or the number of persons served, it remains difficult to say to what degree those outcomes have helped effect justice.

More pragmatically, as Hansmann observes, many public goods involve a separation between purchaser and beneficiary.[24] The donor to a soup kitchen is a purchaser of the public good of social justice, but it is the kitchen's clients who are direct recipients of the services. Accordingly, it can be hard for the purchaser to monitor the quality of the services provided. I would add that the monitoring problem is especially acute in the common case of public goods with large gaps of time or distance between donor and direct beneficiary. To take one example, education directly benefits students, but the public-good component of education is that it produces a more informed electorate, a better trained workforce, and perhaps a more just society.[25] All of these gains arise many years after a donor contributes and often in places far removed from the school itself. Thus, donors get very little meaningful short-term feedback on the quality of their donation. Even if feedback were available, each individual donor has strong incentives to free ride on the monitoring efforts of others, since monitoring by any one provides benefits to all.

In short, government provides subsidies for charity as a way to encourage production of public goods, and these subsidies are reserved for nonprofit organizations because otherwise donors could not know whether their funds—and the government's dollars—were being diverted to private profit.

III. Monitoring

Malani and Posner begin their defense of for-profit charity by disputing Hansmann's link between contract failure and the need for a nondistribution constraint.[26] They acknowledge the possibility of contract failure but suggest that instead of taking the nonprofit form, each firm could simply "promise donors, by contract, that it will not distribute profits."[27] The firm can guarantee compliance by hiring "an auditor such as PricewaterhouseCoopers to police the contract."[28] In the event of default, the donor "could sue the entrepreneur for breach of contract."[29] The duo acknowledges that this arrangement creates costs for donors, who must monitor and sue the firm for

24. Hansmann, *supra* note 17, at 846–47.

25. GRUBER, *supra* note 12, at 287–89.

26. Malani & Posner, *supra* note 1, at 2035.

27. *Id.* at 2035–36.

28. *Id.* at 2036; *see also* Michael Krashinsky, *Transaction Costs and a Theory of the Nonprofit Organization*, *in* THE ECONOMICS OF NONPROFIT INSTITUTIONS: STUDIES IN STRUCTURE AND POLICY 114, 117 (Susan Rose-Ackerman ed., 1986) (suggesting that warranties and middlemen are alternatives to the nondistribution constraint); Geoffrey A. Manne, *Agency Costs and the Oversight of Charitable Organizations*, 1999 WIS. L. REV. 227, 229 (offering "creation of private, for-profit monitoring companies" as a solution to failures in nonprofit monitoring).

29. Malani & Posner, *supra* note 1, at 2036.

noncompliance.[30] But, they say, there are also costs of government-enforced compliance, and it is not clear which is cheaper from a societal perspective.[31]

A. Distribution of Monitoring Costs

This analysis overlooks the distribution of the costs of compliance. The expense of government enforcement is borne by all taxpayers, while the burden of private auditing and bringing contract claims would fall on the firm and its stakeholders.[32] Thus, public enforcement is an additional subsidy for the charitable form.[33] As a practical matter, this subsidy may be crucial for small charities, where the costs of private guarantees might represent a large portion of the annual budget. Also, if each additional firm generates its own contracting and monitoring costs, private monitoring would tend to encourage charities to consolidate rather than specialize, distorting what would otherwise be the optimal choice of policy focus. And donors would change their behavior to give more money to a few firms rather than spreading their largesse more widely, thereby threatening the health of small charities.

On a more theoretical level, the costs of monitoring charity should be borne by the public because that mirrors the distribution of costs when public goods are provided by government. Shifting these costs onto the firms would place them at a disadvantage relative to the government sector. The problem with this arrangement is that, as I argue elsewhere, legal rules should leave taxpayers indifferent between purchasing public goods from government or from charity so that the two can compete on the basis of quality alone.[34]

B. The Mixed-Firm Problem

In addition to the distribution problem, mixing charitable and noncharitable enterprises in the same firm increases monitoring costs.[35]

---

30. *Id.* at 2037–38.

31. *Id.* at 2038.

32. *Cf.* Atkinson, *supra* note 16, at 519 (observing that enforcing the nondistribution constraint is costly for monitors). At the state level, attorneys general are the main enforcers of nonprofit firm charters. *See* James J. Fishman, *Improving Charitable Accountability*, 62 MD. L. REV. 218, 259–65 (2003) (tracing the past and present role of attorneys general in enforcing accountability in the nonprofit sector).

33. To the extent that the burden of higher monitoring costs is borne by the recipients of charity, such as through lower total expenditures, this shift in distribution might be the equivalent of a regressive tax. That is, if beneficiaries are poorer on average than the average taxpayer, the shift in incidence redirects wealth from poor to rich. In most economic analysis, that shift would reduce overall social welfare because of the diminishing marginal utility of wealth: money is more valuable to those who have less of it. MUSGRAVE & MUSGRAVE, *supra* note 7, at 83–85.

34. Brian Galle, The Role of Charity in a Federal System 77 (Dec. 1, 2009) (unpublished manuscript), *available at* http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1473107. Making the consumption of public goods more costly through charity would also make lobbying for government provision relatively more attractive, which is a poor outcome if we think that charity is justified as a cure for some of the problems of public choice theory. *Id.* at 27–33.

35. Malani & Posner, *supra* note 1, at 2052–53.

Malani and Posner acknowledge that expanding eligibility to for-profits may increase the burden on the IRS to screen applicants, but they claim that this problem can be solved by charging application fees or imposing penalties on "abusive" applicants.[36] These user charges, like private monitoring, are simply another mechanism for shifting costs onto the charitable provider and are subject to the same distributional problems I have just mentioned.

More problematically, the overall social cost of monitoring is higher in for-profit charity, regardless of distribution, because mixed-purpose organizations present extra difficulties not present in pure charitable entities. As David Schizer and Victor Fleischer have pointed out, making for-profit firms eligible for contribution deductions would require the IRS to identify worthy charitable functions with considerably more specificity than it does today.[37] Consider scientific journals. A journal looks a lot like a for-profit magazine, but under current law it can establish that it is a charity by showing a general editorial policy of selecting articles on the basis of scientific merit rather than commercial appeal.[38] But under the Malani and Posner proposal, a for-profit magazine like *Popular Mechanics* could accept deductible donations and use them to fund the publication of boring but scientifically sound articles in its otherwise profitable magazine.[39] We then would presumably have to weigh, article by article, whether each piece was selected on the basis of scientific merit. The costs to both government and charity of these endless determinations would increase many times over.

Malani and Posner might argue in response that this example is a *reductio ad absurdum* and that what they have in mind instead is something like a for-profit company that operates a charitable branch. Then the activities of the subdivision can be monitored at a general level, like any other charity. But that situation, if anything, is even worse from a monitoring perspective. Related businesses can easily pass value to each other in ways that are extremely difficult to observe.[40] Organizations can share staff, space,

---

36. *Id.* at 2052–54.

37. *See* David M. Schizer, *Subsidizing Charitable Contributions: Incentives, Information and the Private Pursuit of Public Goals*, 62 TAX L. REV. 221, 254 (2009) (advocating subsidizing "for-profit" charities only if the definition of eligible activities is tightened considerably); Victor Fleischer, *"For Profit Charity": Not Quite Ready for Prime Time*, 93 VA. L. REV. IN BRIEF 231, 231–32 (2008), http://www.virginialawreview.org/inbrief/2008/01/21/fleischer.pdf (noting that if the nondistribution constraint were removed, the broad language of § 501(c)(3) would open the floodgates for companies to claim to be charities).

38. *See* HOPKINS, *supra* note 14, § 8.5, at 292 (describing rules under which entities can qualify for 501(c)(3) status through their "advancement of education or science").

39. It is useful to distinguish the deduction the firm takes for its own costs from the deduction outsiders could receive for purchasing charitable services from the firm. The firm itself can always take a deduction for its ordinary and necessary expenses. *See* Fleischer, *supra* note 37, at 232–33 (pointing out that a corporation may use the losses from an unprofitable subsidiary—such as a charity—to offset the taxable income of a profitable subsidiary). So the key issue here is really whether the firm's "customers" can take a deduction.

40. *See* Brakman Reiser, *supra* note 3, at 2466 (acknowledging this concern about for-profit charities); Hines et al., *supra* note 3, at 1212 (making this point with regard to donor efforts to

mailing lists, phone trees, and good will.[41] While staff time can be logged, logs alone cannot account for the economies of scale and scope that come with working for the two organizations in tandem. A regulatory system capable of detecting these transfers would be complex and likely involve much effort spent resolving disputes over close cases.[42]

Transfers from a charitable branch to a corporate parent merit this close scrutiny because they raise the threat of significant distortions. If deductible dollars can be leveraged by a for-profit parent entity, it will gain a cost advantage over its competitors.[43] Therefore, the hybrid form would encourage the creation of plausible but largely fake charities or philanthropic operations whose benefits to the public would be small relative to the gains to the firm. The possibility of cross-subsidies would also distort investors' choice of efficient firm organization, encouraging entrepreneurs to bring charity in-house rather than simply supporting a variety of outside philanthropy. Or, again, society could spend large sums policing the boundaries. Either way, the for-profit charity is costlier than the pure charity.

Henderson and Malani turn this argument on its head, claiming economies of scope make charity more efficient in a for-profit firm, with the charity benefiting from expenditures on the for-profit side.[44] But this possibility highlights the difficulty of monitoring mixed-purpose firms. Any firm would be able to offer the Henderson and Malani reasoning as cover for its use of deductible dollars to build up its for-profit infrastructure, leaving regulators with few tools to determine when deductions have been put to bad purposes. And whether waste or efficiency predominates will depend in large part on the efficacy of enforcement efforts—if enforcement is ineffective, then competition will oblige virtually every firm to have a captive charity supplying it with cross-subsidies.

IV. Cost and Quality

The effects of mixing the profit motive with charity also undercut Malani and Posner's claim that nonprofits are less efficient. They suggest

---

monitor firms); Schizer, *supra* note 37, at 255 n.82 ("[T]ransfer pricing is already quite malleable, and without separate legal entities it becomes even harder to monitor.").

41. *See* Atkinson, *supra* note 16, at 570 (describing the cost-lowering benefits of horizontally and vertically integrating for-profit and nonprofit organizations).

42. For a sense of the difficulties here, consider the law of joint ventures between nonprofit and for-profit hospitals, which covers more than one-hundred pages in a leading treatise. MICHAEL I. SANDERS, JOINT VENTURES INVOLVING TAX-EXEMPT ORGANIZATIONS 491–609 (3d ed. 2007).

43. *Cf.* Hansmann, *supra* note 17, at 883 (noting that the deduction "gives a financial advantage to [firms] that qualify for it"). This is already a serious problem with many current collaborations between nonprofits and for-profits, as with hospitals and affiliated for-profit care centers owned by physicians who practice at the hospital. The doctors can refer patients from the hospital to their own businesses, in effect leveraging the good will and contacts of the hospital to their own gain. Council on Ethical & Judicial Affairs, Am. Med. Ass'n, *Conflicts of Interest: Physician Ownership of Medical Facilities*, 267 J. AM. MED. ASS'N 2366, 2366 (1992).

44. Henderson & Malani, *supra* note 1, at 590–93.

that for-profit firms will be more skilled at keeping costs low,[45] that nonaltruistic managers may work in nonprofits and enjoy leisure rather than high profits,[46] and that even altruistic managers face little competition and so have little incentive to perform effectively.[47] Relatedly, Henderson and Malani argue that for-profits offer lower agency costs than charities.[48] These claims all overlook the possible effect of the nonprofit form itself on managers' behavior and neglect the existence of rivalries between nonprofits and government.

A. Warm Glow Effects

One distinctive feature of charity is the good feeling, or "warm glow," that some people get from participating in it.[49] Warm glow can derive from moral satisfaction, social approbation, or simply the status signal of being able to spend generously.[50] Empirical studies suggest that warm glow is a significant motivation for many charitable gifts.[51]

---

45. Malani & Posner, *supra* note 1, at 2048–50; *see also* Myron J. Roomkin & Burton A. Weisbrod, *Managerial Compensation and Incentives in For-Profit and Nonprofit Hospitals*, 15 J.L. ECON. & ORG. 750, 752–53 (1999) (noting that nonprofits may pay higher base salaries than for-profits to compensate for the lack of profit-based compensation).

46. Malani & Posner, *supra* note 1, at 2034–35; *see also* Marco A. Castaneda et al., *Competition, Contractibility, and the Market for Donors to Nonprofits*, 24 J.L. ECON. & ORG. 215, 222 (2007) (discussing the ways that promotional expenditures by nonprofits may provide utility to donors and asserting that increased promotional expenditures lead to increased donations); Richard A. Posner, *What Do Judges and Justices Maximize? (The Same Thing Everybody Else Does)*, 3 SUP. CT. ECON. REV. 1, 8–9 (1993) (positing that people who choose to work in the nonprofit sector may willingly trade income for increased job security or other sources of nonmonetary utility).

47. Malani & Posner, *supra* note 1, at 2054–56.

48. Henderson & Malani, *supra* note 1, at 598–600. Henderson and Malani also offer other similar arguments that warrant less discussion. For example, they suggest that for-profit firms acquire unique "leverage" over charities' managers by collating many different donations. *Id.* at 599. In the charitable literature, this is a function commonly attributed to foundations. *See* Atkinson, *supra* note 16, at 583–84 (arguing that sizable private foundations may have adequate leverage to achieve economical deals with for-profit suppliers that smaller donors could not). They also point to the greater societal scrutiny of for-profit firms generally. Henderson & Malani, *supra* note 1, at 599–600. But they offer no reason to suspect that this scrutiny is aimed at the philanthropic activities of firms—as opposed to, say, the other, vastly larger, activities of the firm. *Cf. id.* at 622 (noting that the average firm contributes 1.5% of profits to charity).

49. *See* B. Douglas Bernheim & Antonio Rangel, *Behavioral Public Economics: Welfare and Policy Analysis with Nonstandard Decision-Makers*, *in* BEHAVIORAL ECONOMICS AND ITS APPLICATIONS 7, 62–65 (Peter Diamond & Hannu Vartiainen eds., 2007) (evaluating the difficulty of quantifying the effects of warm glow); Rose-Ackerman, *supra* note 22, at 712–13 (observing that donors may donate in order to feel a warm glow that is separate from and in excess of their desire that others benefit from their contribution).

50. *See* James Andreoni, *Impure Altruism and Donations to Public Goods: A Theory of Warm-Glow Giving*, 100 ECON. J. 464, 473–74 (1990) (predicting that, all other things being equal, donors prefer a transfer of income that will result in the most warm glow); Kenneth J. Arrow, *Gifts and Exchanges*, 1 PHIL. & PUB. AFF. 343, 348–49 (1972) (describing blood donors as being partially motivated by the moral satisfaction gleaned from having individually contributed to the welfare of society); Bernheim & Rangel, *supra* note 49, at 62–65 (providing a partial list of "warm-glow mechanisms," including responses to perceived public perception of the donor's behavior and

Warm glow changes the incentives of a charity's employees. Just as giving to charity produces a warm glow, so too may working for one. We should expect the possibility of this noncash compensation to lower the actual cost of wages for nonprofits.[52] In effect, the employees are making donations of the difference between their salary and the market salary for someone of their talents and realizing the psychic rewards from the gift. Note that the value here is not simply the warm glow from producing the good, as Malani and Posner assume,[53] but also the warm glow from producing it *at a charity*, where peers will know that the employee is making a sacrifice.[54] Even if working in the charity division of a for-profit created some kind of comparable glow, any cost savings would quickly be competed away. Because the lower wages offer a competitive price advantage, new firms would enter to capture the available surplus.[55] Each additional firm would increase the demand for warm glow workers until the point at which enough firms had entered to bid up the price to the market salary.[56]

The warm glow point goes at least some way towards answering all three of the Malani and Posner criticisms. Obviously, firms with lower labor

---

positive and negative reciprocity); Peter Diamond, *Optimal Tax Treatment of Private Contributions for Public Goods with and Without Warm Glow Preferences*, 90 J. PUB. ECON. 897, 917 (2006) (arguing that warm glow is partially a result of the decrease in social pressure to donate); Amihai Glazer & Kai A. Konrad, *A Signaling Explanation for Charity*, 86 AM. ECON. REV. 1019, 1019–21 (1996) (presenting evidence for the proposition that altruism is often motivated by a desire to demonstrate wealth); Rose-Ackerman, *supra* note 22, at 712–13 (pointing out that the satisfaction people derive from philanthropy stems from both seeing positive changes in the community and from the warm glow associated with giving).

51. *See, e.g.*, Bruce R. Kingma & Robert McClelland, *Public Radio Stations Are Really, Really Not Public Goods: Charitable Contributions and Impure Altruism*, 66 ANNALS PUB. & COOP. ECON. 65, 66–67 (1995) (citing studies showing that impure altruism, including warm glow, explains some charitable giving); John Peloza & Piers Steel, *The Price Elasticities of Charitable Contributions: A Meta-analysis*, 24 J. PUB. POL'Y & MKTG. 260, 265 & tbl.1 (2005) (compiling studies and concluding that tax deductions are treasury efficient, suggesting that noneconomic motivations are involved in charitable giving).

52. *See* Bruce R. Kingma, *Public Good Theories of the Non-profit Sector: Weisbrod Revisited*, 82 VOLUNTAS 135, 142 (1997) (arguing that employee control over the organization's mission and services explains empirical evidence of lower wages in the nonprofit sector); Roomkin & Weisbrod, *supra* note 45, at 753–54 (noting studies finding higher wages in for-profits but suggesting explanations other than a warm glow offset in complex industries); Shaviro, *supra* note 7, at 1003 (suggesting that certain fields are more likely to benefit from warm glow in the potential labor pool). Professor Malani has acknowledged this possibility in his other work. *See* Henderson & Malani, *supra* note 1, at 583–84, 619 (acknowledging that workers would be likely to accept lower wages in exchange for jobs that provide warm glow).

53. *See* Malani & Posner, *supra* note 1, at 2047–48 ("[A]ltruists obtain additional value from producing public goods . . . .").

54. *See* Glazer & Konrad, *supra* note 50, at 1020–21 (explaining that donors donate to charity at least partly for signaling purposes rather than simply to obtain satisfaction unrelated to status).

55. *See* WILLIAM J. BAUMOL ET AL., CONTESTABLE MARKETS AND THE THEORY OF INDUSTRY STRUCTURE 192 (1988) (asserting that monopoly rents permit new entrants).

56. Nonprofit entrepreneurs would not enter a field simply to leverage submarket wages into profits because by definition they cannot extract profit. Thus, it is more likely that nonprofit entrepreneurs launch their endeavors to meet unmet social needs or to find employment, and accordingly they would not be attracted by low wages.

costs will tend to have lower total costs, although some charities may still initially lack the skill to hold costs down. Malani and Posner state that there is no competitive pressure to force nonprofit managers to learn to cut costs, but they assume that the only source of competitive pressure is the threat of job loss.[57] Yet managers who are motivated by the desire to fulfill their mission will have reason to minimize costs that are unrelated to the mission.[58] In a firm that pays below market salary, managers will self-select to those who in fact receive warm glow from their work.[59] Likewise, to the extent that they are self-selected for commitment, nonprofit managers are less likely to shirk on quality.[60]

Mixing charitable enterprise with the for-profit form would undermine the benefits of warm glow for everyone.[61] A portion of warm glow likely

---

57. Malani & Posner, *supra* note 1, at 2055–56. Empirical evidence also suggests that competition for donor funds at least partially disciplines nonprofit firms. *See* Castaneda et al., *supra* note 46, at 245 (concluding that competition, along with contracting, reduces reported administrative expenses).

58. *See* Krashinsky, *supra* note 28, at 117; Rose-Ackerman, *supra* note 22, at 719 (arguing that ideological commitment plays an important role in overcoming inefficient incentives within professions); George G. Triantis, *Organizations as Internal Capital Markets: The Legal Boundaries of Firms, Collateral, and Trusts in Commercial and Charitable Enterprises*, 117 HARV. L. REV. 1102, 1147 (2004) (asserting that informational asymmetry between donors and managers creates an obstacle in fundraising situations). On the other hand, prestige motivations introduce another possible agency cost, as Triantis argues. *Id.* at 1116–17.

59. *See* Kingma, *supra* note 52, at 142 (opining that employees and entrepreneurs in nonprofits prefer serving the public good to increased wages and social output); Rose-Ackerman, *supra* note 22, at 719 (noting that ideologue founders seek employees that share their ideals and vision). *But see* Linda Sugin, *Resisting the Corporatization of Nonprofit Governance: Transforming Obedience into Fidelity*, 76 FORDHAM L. REV. 893, 907 (2007) (noting that low pay may also simply attract those who are less skilled). I assume, in contrast to Professor Sugin's view, that firms can screen for employee quality. Whether there are likely to be sufficient numbers of skilled employees is a subject I return to shortly.

It could be argued that some individuals will also choose nonprofit employment as a way to gain leisure time in exchange for a lower salary. For that to be true, the employee would have to expect that the firm's principals are unable to monitor her effectively. That is plausible. *See* Hansmann, *supra* note 16, at 507, 568 (observing that patrons of nonprofits have little authority to oversee the management and work of the nonprofit). On the other hand, coworkers can monitor their fellow employees' efforts more easily. *See* Hansmann, *supra* note 17, at 876 (remarking that normative constraints operate better in certain nonprofits than in others). If coworkers are motivated by their mission or resent slacking in others, this monitoring may have some bite. *Cf.* Mark Barenberg, *Democracy and Domination in the Law of Workplace Cooperation: From Bureaucratic to Flexible Production*, 94 COLUM. L. REV. 753, 881–903 (1994) (describing self-monitoring workplaces). As a result, it would seem that those who value leisure above salary could obtain the same tradeoff between salary and leisure without the risk of sanction simply by taking a part-time job.

60. *See* Hansmann, *supra* note 17, at 876 (noting that nonprofits are likely to choose as managers those whose ideals match the firm's).

61. The argument here is similar to, but distinct from, Bruno Frey's point that extrinsic motivations may sometimes "crowd out" people's internal motivations to do good. Bruno Frey, *A Constitution for Knaves Crowds Out Civic Virtues*, 107 ECON. J. 1043, 1044–47 (1997). Like Frey, I argue that an inaptly designed subsidy may actually reduce incentives to engage in the subsidized behavior, but the mechanism I suggest is not the same as the psychological factors he surveys, *id.* at 1045–46.

derives from donors' ability to signal their own largesse.[62] To gain this benefit, others must be able to perceive the gift. Thus, if casual outside observers cannot be sure whether a given worker is toiling for a nonprofit or for a money-making enterprise, the social rewards of working for any charity are lower. Malani and Posner's proposal would create this confusion by making it unclear whether any given firm producing charitable services was paying its employees a share of profits.[63]

As a result, the mixed form lowers nonmonetary rewards not only at the for-profit but also at all enterprises carrying out a charitable mission.[64] Charities then will be obliged to pay higher wages and to screen and monitor more aggressively for self-dealing. Thus, even if for-profits are more efficient at their own work, encouraging for-profits to do charity could on net make the charitable sector as a whole less efficient.

Malani and Posner recognize to some extent the benefit of relying on employees who are committed to their mission, but they argue that the labor market may not supply enough of those individuals.[65] Their solution, though, actually cools the glow, reducing the number of available workers motivated by warm glow. The status quo, in which government providers of public goods constrain inefficient charities, is preferable. As the supply of warm-glow-driven workers dries up and costs in the nonprofit sector rise, government will become a more attractive alternative for donors and taxpayers choosing their service provider. This competition will help to drive highly inefficient nonprofits out of business. In addition, there is little danger that outside observers will confuse working for a charity with working for city or state government, so that dividing the provision of public goods solely between those two sectors maintains the purity of the signal sent by nonprofit employment and the power of warm glow motivation.

B. The Advantage of Low-Powered Incentives

Next, the fact that the nondistribution constraint largely bars any form of high-powered incentives for nonprofit employees to perform is a feature, not a bug, of the sector.[66] Typically, owners of for-profit firms align their

---

62. *See* Glazer & Konrad, *supra* note 50, at 1020–21 (stating that, according to empirical data, donations serve as a means to signal income to those who do not view the individual's consumption of luxury goods).

63. *See* Malani & Posner, *supra* note 1, at 2065 (suggesting that the IRS should permit nonprofit managers to receive incentive pay correlated to profits, revenues, or the operating costs of the organizations).

64. *Cf.* Hansmann, *supra* note 16, at 524 (suggesting that changed norms in some parts of the nonprofit sector could "undermine the collective morality of that sector as a whole").

65. Malani & Posner, *supra* note 1, at 2047–49.

66. Hines, Horwitz, and Nichols argue that nonprofits can probably lawfully offer their employees performance incentives. Hines et al., *supra* note 3, at 1193–97. They acknowledge, though, that there is no clear guidance on the question, *id.*, and that anecdotal evidence suggests most firms are unwilling to push to find the edge of the legal limits, *id.* at 1196–97. Also, there is little even in their careful unpacking of the applicable law to suggest that firms can pay managers

employee incentives with their own by offering pay structures that have a very powerful influence on the behavior of their employees.[67] Giving the employees a share of profits or a general equity stake in the firm is the quintessential example of these high-powered incentives: employees now care more about ownership's interests because they are also owners.[68] In contrast, nonprofits and government are limited to offering "low-powered" incentives—such as performance bonuses, promotions, and job security—that have a somewhat weaker influence on behavior.[69] Critics of government have long argued that this weak incentive structure is what allows the private sector to outperform government.[70]

1. Measurement Problems.—High-powered incentives, like high-powered explosives, are dangerous if left in the wrong place.[71] As has been widely recognized, in the recent financial crisis firms offered their managers powerful incentives to take on massive amounts of risk.[72] And the managers followed their self-interest, to our collective sorrow.[73] Powerful incentives,

---

for cutting costs, which I will argue here are the most problematic incentives (and the incentives of greatest interest to Malani and Posner). In any event, my analysis suggests that any strong expansion towards permitting powerful incentives in nonprofit compensation would be a mistake.

67. *See* Lucian Arye Bebchuk et al., *Managerial Power and Rent Extraction in the Design of Executive Compensation*, 69 U. CHI. L. REV. 751, 761–63 (2002); Michael C. Jensen & William H. Meckling, *Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure*, 3 J. FIN. ECON. 305, 328 (1976) (both discussing use of employee compensation structures as tools for controlling employee incentives).

68. *See* Bebchuk et al., *supra* note 67, at 775 (describing stock options as a crucial means of encouraging executives to maximize firm performance and shareholder value); David Weisbach, *Tax Expenditures, Principal-Agent Problems, and Redundancy*, 84 WASH. U. L. REV. 1823, 1846–47 (2006) (discussing how assigning risk from the principal to the agent can alter agents' incentives).

69. *See* Weisbach, *supra* note 68, at 1846–48 (asserting that without large cash incentives, incentives given to public servants are considered low-powered and thus less effective).

70. *See, e.g.*, Armen A. Alchian & Reuben A. Kessel, *Competition, Monopoly, and the Pursuit of Money*, *in* ASPECTS OF LABOR ECONOMICS 157, 166 (Nat'l Bureau of Econ. Research ed., 1962) (noting that a public utility does not possess the ability to distribute wealth as dividends to its owners and thus can only offer weak incentives).

71. *See* EDWARD E. LAWLER III, STRATEGIC PAY 58 (1990) (stating that when used incorrectly, incentive systems enable many employees to spend their energies outsmarting the system rather than increasing the value of the products); Triantis, *supra* note 58, at 1114 (explaining that shifting risk to managers may increase risk-averse decision making, resulting in inefficient allocation of capital).

72. *See, e.g.*, John C. Coffee, Jr. & Hillary A. Sale, *Redesigning the SEC: Does the Treasury Have a Better Idea?*, 95 VA. L. REV. 707, 729–30 (2009) (explaining that because of the transition in compensation from cash to stock options, senior management undertook riskier strategies); Jeff N. Gordon, *"Say on Pay": Cautionary Notes on the U.K. Experience and the Case for Shareholder Opt-In*, 46 HARV. J. ON LEGIS. 323, 363–64 (2009) (remarking that the compensation structures of senior executives that were based on high-powered incentives enabled excessive risk taking).

73. This is not to say that managerial incentives were nearly the whole story. *See, e.g.*, INT'L MONETARY FUND, GLOBAL FINANCIAL STABILITY REPORT: FINANCIAL STRESS AND DELEVERAGING 14 (2008), *available at* http://www.imf.org/external/pubs/ft/gfsr/2008/02/pdf/text.pdf (pointing to the U.S. housing market); Coffee & Sale, *supra* note 72, at 731–49 (exploring

in other words, are risky. If incentive targets are misaligned—for example, where management cannot define in advance what behavior will maximize outputs—then the firm can expend large resources paying out bonuses for actions that can actually hurt the firm, or, at best, waste both the firm resources sunk into incentives and the manager time spent pursuing them.[74]

As difficult as designing an incentive structure can be in a for-profit firm, nonprofits, under my account of the nonprofit sector, present special problems because of the difficulty of measuring nonprofits' outputs. Many commentators, from Hansmann on, have noted that public goods are hard to value and that this creates a dilemma for monitoring efforts.[75] It follows that a high-powered incentive structure would be unlikely to produce manager behavior that matches the firm's goals. Jacob and Levitt, for instance, have documented how rewards for teachers based on test results lead teachers to "teach to the test," or to just give their students the answers.[76] Acemoglu and his coauthors argue more generally that when "bad" motivations such as teaching to the test dominate the positive rewards of high-powered incentives, the firm would be better off providing only low-powered incentives.[77] Putting these points together, firms that produce public goods should not use high-powered incentives because the risks of mismeasurement, waste, and bad incentives are prohibitively high.[78]

---

regulatory failure, the structure of the investment-banking industry, moral hazard by mortgage originators, and behavior by bank executives as contributing possibilities).

74. *See* George P. Baker, *Incentive Contracts and Performance Measurement*, 100 J. POL. ECON. 598, 599–600, 606 (1992) (asserting that if the agent's actions do not directly conform to the principal's objectives, there will inevitably be inefficiencies).

75. *E.g.*, Hansmann, *supra* note 17, at 898 n.160; Manne, *supra* note 28, at 239; Posner, *supra* note 46, at 10; Triantis, *supra* note 58, at 1147; *see also* William W. Bratton & Joseph A. McCahery, *The New Economics of Jurisdictional Competition: Devolutionary Federalism in a Second-Best World*, 86 GEO. L.J. 201, 236 (1997) (arguing that the evaluation problem prevents meaningful interjurisdictional competition based on quality of outputs).

76. Brian A. Jacob & Steven D. Levitt, *Rotten Apples: An Investigation of the Prevalence and Predictors of Teacher Cheating*, 118 Q.J. ECON. 843, 844 (2003). Similarly, an earlier study of educational incentives found that measuring student outputs led to "cream skimming": more aggressive recruitment of students who would score high on the evaluation tool without any additional schooling, greatly reducing the usefulness of the money spent on training. Michael Cragg, *Performance Incentives in the Public Sector: Evidence from the Job Training Partnership Act*, 13 J.L. ECON. & ORG. 147, 161–62 (1997).

77. Daron Acemoglu et al., *Incentives in Markets, Firms, and Governments*, 24 J.L. ECON. & ORG. 273, 274 (2008); *see also* Oliver E. Williamson, *Public and Private Bureaucracies: A Transaction Cost Economics Perspective*, 15 J.L. ECON. & ORG. 306, 325 (1999) ("[A]dded incentive intensity undermines probity."). David Weisbach offers a related point, building on Holmstrom and Milgrom and others. Weisbach points out that if some but not all of an agent's outputs can be measured and the agent is offered high-powered incentives, the agent will tend to overproduce the measurable outputs. Weisbach, *supra* note 68, at 1848–49 (citing AVINASH K. DIXIT, THE MAKING OF ECONOMIC POLICY 96 (1996); Bengt Holmstrom & Paul Milgrom, *Multitask Principal–Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design*, 7 J.L. ECON. & ORG. 24 (1991)).

78. *See* Hines et al., *supra* note 3, at 1197–98, 1205–06 (arguing that difficulty of measuring nonprofit outputs can make it hard to design incentives for managers and might allow them to cut costs at expense of quality); Jill R. Horwitz, *Why We Need the Independent Sector: The Behavior,*

The privatization literature reaches similar conclusions, although not all the lessons of that debate translate to the nonprofit context.[79] Critics of privatization argue that for-profit firms will cut costs—which are measurable—at the expense of quality—which often is not.[80] Other commentators suggest that there may be some services for which the risks of corner cutting are small and the cost savings large, so that privatization is a safer bet.[81] Perhaps, then, the IRS should authorize for-profit charity in those fields where low quality would result in little social harm. Theater, opera, and art all come to mind as potential examples.

The problem with this proposal is that it runs contrary to the basic premise of subsidies for charity. In order to determine which public goods would gain a net benefit from for-profit production, the government would first have to weigh the relative harms of low quality for each charity. That, in turn, would require explicit government judgments about the value of a charity's output. Yet that is precisely what the law of charity, as currently constructed, is designed to prevent.[82] Charitable law eliminates case-by-case evaluations of a charity's worth for fear that controversial, unpopular, or

---

Law, and Ethics of Not-for-Profit Hospitals, 50 UCLA L. REV. 1345, 1410 (2003) (arguing that low-powered incentives are key to ensuring that nonprofits deliver high-quality services); *cf.* Jonathan R. Macey, *Transaction Costs and the Normative Elements of the Public Choice Model: An Application to Constitutional Theory*, 74 VA. L. REV. 417, 491–93 (1988) (claiming that public officials cannot be controlled adequately by standard incentive-based pay because there exist no useful measures on which to base the incentives). This point is also consistent with the familiar claim in the incentives literature that the usefulness of incentives as a monitoring tool declines with the ease of measuring performance accurately. Baker, *supra* note 74, at 609–11.

79. "Privatization" refers to the shift of government activities to the private—usually the private for-profit—sector. MINOW, *supra* note 3, at 1–3.

80. *See, e.g.*, *id.* at 64–65 (asserting that schooling systems governed by a privatized, incentive-based model do not take into account the fact that education is not about academic achievement alone but also involves the transmission of civic ideals to students so that they can become "productive workers and responsible citizens," and thus education "has crucial features that depart from privately consumed goods and services"); John D. Donahue, *The Transformation of Government Work: Causes, Consequences, and Distortions*, *in* GOVERNMENT BY CONTRACT: OUTSOURCING AND AMERICAN DEMOCRACY 41, 43–45 (Jody Freeman & Martha L. Minow eds., 2009) (claiming that shifting production of services to for-profit firms leads to cost-cutting as "the prime directive"); Oliver D. Hart et al., *The Proper Scope of Government: Theory and an Application to Prisons*, 112 Q.J. ECON. 1127, 1136–41 (1997) (arguing that private ownership leads to a strong incentive to strive for cost reduction and a weak incentive to engage in quality improvement).

81. *E.g.*, Karen N. Eggleston & Richard J. Zeckhauser, *Government Contracting for Health Care*, *in* MARKET-BASED GOVERNANCE, *supra* note 23, at 29, 42–43, 55–56; Hart et al., *supra* note 80, at 1141–43, 1154–55.

82. *See* JOEL L. FLEISHMAN, THE FOUNDATION: A GREAT AMERICAN SECRET 22–24 (2007) (warning that tax exemptions for donations must be available to a broad array of organizations, or else unpopular entities would not receive support); Boris Bittker & George Rahdert, *The Exemption of Nonprofit Organizations from Federal Income Taxation*, 85 YALE L.J. 299, 342 (1976) (explaining that tax-exempt organizations "enjoy the privilege of spending 'government money'" to further their independent objectives while being protected from legislative pressure regarding this spending).

novel endeavors would be judged harshly by the powers that be.[83] Thus, absent new forms of governmental decision making that might mitigate the problem of discretion, transferring some charitable functions to for-profit firms would threaten to curtail the pluralism and experimentalism of the philanthropic sector.[84]

2. Externalities.—As the financial-crisis example suggests, another problem with high-powered incentives is externalities.[85] Many charitable services are controversial or redistributive.[86] In both of these cases, the firm's production affects outsiders in ways that may reduce the outsiders' welfare. For example, redistributive services can reduce overall social welfare by crowding out other productive activity; recipients may consume the service rather than engage in some other transaction, such as working, that would increase welfare for the counterparty or others.[87] Since any of these effects are largely externalities for the firm and its stakeholders, allowing the firm to use high-powered incentives would increase the risk that the firm would produce too much redistribution relative to the socially optimal point.

Even if the firm does not intend to incentivize its employees to produce externalities, it may do so by accident. That seems to be the financial-crisis story: firms were each individually indifferent to the risk that their high-powered incentives encouraged their executives to take on too much systemic risk because the dangers of that systemic risk were largely externalities for the firm's stakeholders.[88] Again, to the extent that producing public

---

83. *See* Atkinson, *supra* note 16, at 636–37 (warning that one of the risks of basing the tax exemption of charities on altruism theory is that "only certain favored purposes [will] be allowed to thrive"); Dean Pappas, Note, *The Independent Sector and the Tax Law: Defining Charity in an Ideal Democracy*, 64 S. CAL. L. REV. 461, 476 (1991) (acknowledging that the IRS's use of a single, broad criterion of "charitable"—that an organization "serve the public and not significantly contravene public policy"—may produce a conformist view of "charitable"); *see also* NORMAN I. SILBER, A CORPORATE FORM OF FREEDOM: THE EMERGENCE OF THE NONPROFIT SECTOR 5–6, 31–66 (2001) (tracing the history of the judicial evaluation of charity in an effort to show that judicial decisions reflect the judiciary's policy preferences).

84. It is worth emphasizing that I believe such governance innovations are possible. *See* Galle, *supra* note 34, at 80 (suggesting that organizations "should be obliged to explain why their organization deserves a subsidy, taking into account opportunity costs, moral hazard, and other possible harms to others" and that "[t]reasury officials should be empowered to consider these arguments and reject the application of organizations that lack merit, albeit while subject to careful oversight by layers of administrative and judicial review").

85. *See* Steven L. Schwarcz, *Systemic Risk*, 97 GEO. L.J. 193, 198 (2008) ("[N]o individual market participant has sufficient incentive . . . to limit its risk taking in order to reduce the systemic danger to other participants and third parties.").

86. *See* Pappas, *supra* note 83, at 476 (noting the potential for charitable organizations to use the benefits of tax exemption "to contravene public policy and create injustice").

87. *See* MUSGRAVE & MUSGRAVE, *supra* note 8, at 190–91 (summarizing empirical findings that social insurance crowds out work).

88. Gordon, *supra* note 72, at 364–66; *see also* Schwarcz, *supra* note 85, at 206 (noting that firms are motivated to protect themselves, not the system as a whole).

goods is even more uncertain than producing private goods, this danger of firms ignoring misaligned incentives would be higher in the nonprofit sector.

C. Is For-Profit Charity Just Different?

Finally, even if Henderson and Malani are correct that mixed firms are on average more efficient,[89] it is not obvious that this efficiency is good for philanthropy. If charity is more efficient in mixed firms, we should expect that form to crowd out pure charities.[90] That shift in the location of charity may also give rise to substantive changes. Corporations, for example, are vulnerable to objections to their work in ways that simpler nonprofits are not. Nonprofits do not have unrelated commercial product lines that can be subjected to boycott or capital stock that can be divested.[91] Thus, we might predict that corporate philanthropy will be more tepid, less willing to offend or push the boundaries of social norms.[92] Of course, donors who prefer their charity "edgy" can still donate to traditional charities, but those with limited resources may well redirect their money to less controversial but more efficient choices. In the absence of government-enforced rules for transparency and auditing, consumers might also struggle to choose the charity that matches their preferences.[93] Thus, on net, efficient corporate philanthropy might reduce the entrepreneurial character of charity as a whole.

***

As a result, maintaining the current charitable sector is likely more efficient overall than Malani and Posner's proposal. Mixing the for-profit and charitable sectors increases the costs and reduces the efficacy of charities, while introducing high-powered incentives to the production of public goods exposes all of society to the risks of those incentives. There are also a variety of other negative effects, such as diminishing the diversity of charities' size and focus, that could result from the introduction of a for-profit element into the charitable realm.

---

89. Henderson & Malani, *supra* note 1, at 598–600.

90. *Cf.* Bloche, *supra* note 3, at 1096–97; Eggleston & Zeckhauser, *supra* note 81, at 45 (both noting that, under competition from for-profit hospitals, nonprofits begin to behave like for-profits).

91. *See* RICHARD POSNER, ECONOMIC ANALYSIS OF LAW 455 (7th ed. 2007) (claiming that "corporations avoid controversial charities" because of potential adverse reactions from shareholders).

92. *Cf.* Sharon Dolovich, *How Privatization Thinks: The Case of Prisons*, *in* GOVERNMENT BY CONTRACT: OUTSOURCING AND AMERICAN DEMOCRACY, *supra* note 80, at 128, 134–35 (claiming that private firms will not account for values that are incommensurable with cost savings); Rose-Ackerman, *supra* note 22, at 721 (suggesting that peaceful coexistence between for-profit investor interests and existence of the nondistribution constraint would be "fragile" because of "short-term opportunism"); Williamson, *supra* note 77, at 331–32 (arguing that private firms cannot duplicate bureaucratic performance because of tensions between profit motive and public mission).

93. *See* Archon Fung, *Making Social Markets: Dispersed Governance and Corporate Accountability*, *in* MARKET-BASED GOVERNANCE, *supra* note 23, at 145, 155 (arguing that consumers have difficulty selecting products based on social preferences because there are few resources that facilitate the play of social values in economic markets).

V. Donor Mistakes

Lastly, Malani and Posner posit that tying the deduction to the nondistribution constraint may depend on a claim that donors to nonprofits lack the capacity to choose effectively between charities.[94] Irrespective of the costs to rational donors of monitoring the uses of their money, irrational or poorly informed donors may be incapable of making correct choices no matter how low the costs.[95] Alternatively, some supposed charities might be skilled at misleading donors.[96] If so, then misguided donors may allocate the government's subsidy wastefully.[97] But, Malani and Posner say, these problems are adequately dealt with by existing state and federal laws prohibiting consumer fraud.[98] And, if not, then the government should simply remove donors from the allocation process, such as through a system of government quasi-grants in which donors select a charitable activity but not a particular organization (a proposal I will refer to as the "activity-only plan").[99] While Malani and Posner's fraud argument can be pushed aside fairly easily, their challenge to the role of donors in allocating money for public-goods production is a fundamental problem for supporters of the deduction.

First, on the fraud point, there are many ways for a charity to mislead donors short of outright fraud that would therefore be beyond the reach of current antifraud laws. Credit-card companies, to take one instance, are subject to antifraud statutes like every other industry, but it is now a widely accepted finding that consumers have a very poor understanding of the terms of their credit contracts.[100]

A better argument Malani and Posner might have made—but did not—would be that additional disclosure requirements would be superior to the nondistribution constraint. That, in fact, is a proposal offered by many critics

---

94. Malani & Posner, *supra* note 1, at 2050. Among those who advance this argument are John Donahue and Martha Minow. *See* JOHN DONAHUE, THE PRIVATIZATION DECISION: PUBLIC ENDS, PRIVATE MEANS 33–34 (1991) (noting the tendencies of different individuals to be informed about different topics and the overall impact this has on dividing public business between the public and private sectors); MINOW, *supra* note 3, at 34 ("It is too often empirically false to assume the existence of sufficiently informed consumers.").

95. *See* Malani & Posner, *supra* note 1, at 2050–51 (arguing that some donors are unable to correctly choose between charities and, therefore, governments must be aware of this when directing or matching subsidies and donations).

96. *Id.* at 2050.

97. *Id.* at 2051.

98. *Id.*

99. *Id.* at 2051–52.

100. *See, e.g.*, Angela Littwin, *Beyond Usury: A Study of Credit-Card Use and Preference Among Low-Income Consumers*, 86 TEXAS L. REV. 451, 499 (2006) (suggesting that consumers are unlikely to understand common credit-card contract terms); Ronald J. Mann, *"Contracting" for Credit*, 104 MICH. L. REV. 899, 911 (2006) (positing that consumers are unlikely to look at credit-card contracts and consequently do not rationally consider credit-card contract terms).

of the credit-card industry.[101] Whether improved disclosure will be sufficient in the case of truly irrational consumers is unknown.[102]

More importantly, in the context of charitable donations, regulation is necessary not only to protect consumers but also other firms. Again, the nonprofit sector competes with government for the business of providing public goods.[103] Self-dealing by managers of some nonprofits can create reputational externalities for the whole sector.[104] And disclosure and enforcement actions may actually be self-defeating if they spread the perception of self-dealing or undermine the norm of self-sacrifice among other nonprofit managers.[105] The nondistribution constraint, therefore, may be more effective than disclosure, as it enables government to undertake enforcement quietly when that is the optimal solution.

Turning next to the activity-only plan, the traditional rationales for the deduction can offer only logistical quibbles to the Malani and Posner alternative. Consistent with these rationales, the activity-only plan would foster a diverse array of public goods, with taxpayers unsatisfied by the existing level of public goods able to obtain a deduction to acquire more.[106] Defenders of traditional rationales for the deduction might argue that the activity-only plan would lead each charity to reduce its own fundraising efforts in order to free ride on the efforts of others. This reduction may be a good thing if we think fundraising is wasteful, but either way, charities could check this effect by forming trade associations.[107] Another result would be that government, apparently, would have to judge the quality or at least

---

101. *See, e.g.*, Oren Bar-Gill, *Seduction by Plastic*, 98 NW. U. L. REV. 1373, 1378 (2004) (arguing that additional disclosure requirements should be implemented to protect consumers from excessive credit-card interest rates).

102. *Cf.* Oren Bar-Gill et al., Product Use Information and the Limits of Voluntary Disclosure 3–4 (2009) (unpublished manuscript, on file with Texas Law Review) (arguing that some forms of mandatory disclosure may not be useful to consumers because consumers themselves have "imperfect information about how they will use a product").

103. Henderson & Malani, *supra* note 1, at 575.

104. *See* Deborah A. DeMott, *Self-Dealing Transactions in Nonprofit Corporations*, 59 BROOK. L. REV. 131, 134, 146–47 (1993) (highlighting the different standards imposed on nonprofit self-dealing and explaining why such externalities likely exist); Fishman, *supra* note 3, at 576 (noting that self-dealing and other civil and criminal wrongdoings by charitable fiduciaries have led to regulatory, legislative, public, and media scrutiny of the nonprofit sector).

105. *Cf.* Dan M. Kahan, *Trust, Collective Action, and Law*, 81 B.U. L. Rev. 333, 334–35 (2001) (arguing that conspicuous rewards and punishments can create the perception that the regulated individuals are not inclined to comply voluntarily). This point is controversial. *See, e.g.*, Leandra Lederman, *The Interplay Between Norms and Enforcement in Tax Compliance*, 64 OHIO ST. L.J. 1453, 1484–99 (2003) (arguing that the evidence is inconsistent with Kahan's hypothesis).

106. Malani & Posner, *supra* note 1, at 2051.

107. *See* MANCUR OLSON, THE LOGIC OF COLLECTIVE ACTION: PUBLIC GOODS AND THE THEORY OF GROUPS 145 (1971) (noting that trade associations are one tool for overcoming the political free-rider problem). Trade associations are the groups that produced, for example, the "Got Milk?" and "Pork: The Other White Meat" campaigns. Jon D. Hanson & Douglas A. Kysar, *Taking Behavioralism Seriously: A Response to Market Manipulation*, 6 ROGER WILLIAMS U. L. REV. 259, 363 (2000).

expenditures efficiency of charities, leading to some increased government oversight with the accompanying danger of bias.[108]

In contrast, the Malani and Posner idea is flatly inconsistent with goals of the deduction as I have elsewhere outlined them.[109] The core of my new rationale is the competition between charities and government entities on quality.[110] It is hard to see how the activity-only plan could allow society to direct funds only to charities that outperform a specific governmental entity and to shift money away from charities that underperform government or other charities. Competition may also depend for its efficacy on warm glow feelings that accompany personal connections between the donor's actions and the resulting public good; the activity-only plan inserts a government bureaucrat between the donor and the result, which by most accounts of the warm glow phenomenon reduces its potency.

VI. Conclusion

Overall, extending the charitable-contribution deduction to include contributions to for-profit firms creates risks that are not worth the putative benefits. For-profit charity threatens to shift costs to charities, weaken the warm glow of giving, distort managerial incentives, and diminish or confuse donor choice.

This is not to say that for-profit firms have no role in the production of public goods. Firms can always contribute resources to other charities; I take no view here on whether using the firm's resources in that way would be consistent with a manager's fiduciary duty to shareholders. And government can always contract with for-profit firms to carry out select governmental functions. Oversight, accountability, and public perception all distinguish contracting from § 170 eligibility.[111] No one will likely confuse Blackwater with the United Way, either in their personnel, their fundamental goals, or in the ways in which they are responsive to their stakeholders. And it is just these factors, I have argued, that make for-profit charity problematic. Whether contracting public goods out to for-profit firms is ever attractive is a larger debate I leave for a different day.

---

108. *See* Atkinson, *supra* note 16, at 636–37 (cautioning that government oversight may restrict charity only to "favored purposes").

109. *See* Galle, *supra* note 34, at 52–78 (outlining the six goals).

110. *Id.* at 77.

111. *See* HOPKINS, *supra* note 14, §§ 20.1–27.17 (outlining limits on nonprofit activities and tools for government oversight of them).

* * *

The Taking/Taxing Taxonomy

Amnon Lehavi*

> I do not propose either to purchase or to confiscate private property in land. The first would be unjust; the second, needless. Let the individuals who now hold it still retain, if they want to, possession of what they are pleased to call *their* land. Let them continue to call it *their* land. Let them buy and sell, and bequeath and devise it. We may safely leave them the shell, if we take the kernel. *It is not necessary to confiscate land; it is only necessary to confiscate rent.*
>
> —Henry George, 1879[1]

Takings jurisprudence is engaged in a constant paradox. It is conventionally portrayed as chaotic and muddy, and yet attempts by the judiciary to create some sense of order in it by delineating the field into distinctive categories that each have a different set of rules are often criticized as analytically incoherent or normatively indefensible.

This Article offers an innovative approach to the taxonomic enterprise in takings law by examining what is probably its starkest and most entrenched division: that between taking and taxing. American courts have been nearly unanimous in refusing to scrutinize the power to tax, viewing this form of government action as falling outside the scope of the Takings Clause. Critics have argued that the presence of government coercion, loss of private value, and potential imbalances in burden sharing mandate that the two instances be conceptually synchronized and subject to similar doctrinal tests.

The main thesis of this Article is that this dichotomy, and other types of legal line drawing in property, should be assessed not on the basis of a point-blank analysis of allegedly comparable specific instances, but rather on a broader view of the foundational principles of American property law and of the way in which takings taxonomies mesh with the broader social and jurisprudential understanding of what "property" is.

Identifying American property law as conforming to two fundamental principles—formalism of rights and strong market propensity—but at the same time devoid of a constitutional undertaking to protect privately held value against potential losses as a self-standing strand in the property bundle, this

---

* Senior Lecturer and Director of Real Estate Studies, Radzyner School of Law, Interdisciplinary Center (IDC) Herzliya. J.S.D., LL.M., Yale Law School. For helpful comments, I thank Hanoch Dagan, Eric Kades, Gerald Korngold, Daphna Lewinsohn-Zamir, Jacob Nussim, Gideon Parchomovsky, Joan Youngman, and participants at the *Building Market Institutions: Property Rights, Business Formalization, and Economic Development* conference. Financial support for this project by the Lincoln Institute of Land Policy is gratefully acknowledged.

1. HENRY GEORGE, PROGRESS AND POVERTY 405 (Robert Schalkenbach Found. 1975) (1879).

Article explains why prevailing forms of taxation do seem to be disparate from other forms of governmental interventions with private property. Focusing on property taxation, this Article shows why taxation is considered a lesser-evil type of government coercion, how the taking/taxing dichotomy better addresses the public–private interplay in property law, and why taxation is often viewed as actually empowering property rights and private control of assets.

I. Introduction 1236
II. The Core of American Property Law 1242
A. The Formality of Property 1243
B. . . . and Its Market Propensity 1249
III. Property-System Analysis of the Taking/Taxing Taxonomy 1255
A. The Judicially Created Divide and Its Critique 1255
B. Reevaluating the Taking/Taxing Line Drawing 1257
1. *The Divide and the Constitutional Protection of Rights, Not Value* 1257
2. *Incorporating the Public/Private Interface in Property* 1262
3. *Taxation as an Empowerment of Property Rights* 1266
IV. The Virtues and Vices of Taxonomy in Property 1276
V. Conclusion 1281

I. Introduction

Takings jurisprudence faces enormous, nearly Sisyphean challenges in trying to design legal doctrines to fit the complexities of government acts that affect private property. This body of law is often criticized as being ad hoc, vague, and unpredictable.[2] Yet whenever courts do try to come up with some allegedly bright-line categorical rules within this field by viewing some instances of public intervention with property as takings per se, others as subject to multifactor, case-specific tests, and yet others as generally falling outside the scope of this strand of constitutional protection, such taxonomies are then criticized as being too rough, conceptually inconsistent, or normatively indefensible.[3] Although such dilemmas about conceptual and doctrinal line drawing are quite familiar to other legal fields, the law of takings seems

2. *See, e.g.*, William P. Barr et al., *The Gild that Is Killing the Lily: How Confusion over Regulatory Takings Doctrine Is Undermining the Core Protections of the Takings Clause*, 73 GEO. WASH. L. REV. 429, 431–33 (2005) (discussing the patchwork nature of takings jurisprudence concerning public utilities); Carol M. Rose, Mahon *Reconstructed: Why the Takings Issue Is Still a Muddle*, 57 S. CAL. L. REV. 561, 594–97 (1984) (suggesting that the source of muddled takings jurisprudence is in the tension between Lockean natural rights and Aristotelian civic republican conceptions of property, each of which underpins areas of the law); Susan Rose-Ackerman, *Against Ad-Hocery: A Comment on Michelman*, 88 COLUM. L. REV. 1697, 1697–1700 (1988) (rebutting Frank Michelman's argument that the Supreme Court's most recent term signaled a move towards a coherent, unified takings jurisprudence).

3. *See* discussion *infra* Parts II(A), III.

to be particularly vulnerable to perpetual discontent over the way in which its landscape is being shaped.[4]

This Article takes up what traditionally purports to be the clearest division within this alleged entanglement: the distinction between taking and taxing. American courts have been practically unanimous in viewing taxation as a chief and essential state power, and have generally refused to strictly scrutinize tax legislation and regulation.[5] As the epigraph demonstrates, the taking/taxing divide has also been hailed as normatively worthy by numerous thinkers—though for diverse reasons—throughout American history.

The dichotomy between the sweeping deference to taxation and the extensive judicial preoccupation with other forms of government-based adverse effects on private property has been, however, increasingly criticized. Various theorists have pointed to the strong conceptual similarity between the compulsory levy and collection of a tax and the nonconsensual transfer of ownership or other key rights in a privately owned asset for a public purpose. Very simply argued, in both types of cases, government forces an owner to hand over privately held value.[6] Some scholars have taken this argument further by calling to formally synchronize the normative and jurisprudential framework for these currently distinctive legal spheres, albeit with differing views about the appropriate direction that this reunification should take.[7]

In this Article, I argue that despite the intuitive appeal in collapsing categorical distinctions between different forms of governmental interventions with private property and in searching for a universal formula that would allegedly rub out arbitrary boundaries, such an approach misses the more fundamental role that these typologies play in the property-law system in general. American property law, so I will argue, is conventionally driven by a formal and market-oriented approach that assigns certain roles to government as provider and regulator of a property-rights system and others to private-property owners, relevant market players, and other stakeholders. Such institutional components are inherently intertwined with the jurisprudential structure of American property law and may accordingly explain

---

4. *See* Thomas W. Merrill, *The Landscape of Constitutional Property*, 86 VA. L. REV. 885, 889 (2000) (reviewing four Supreme Court cases decided in the space of two years, all involving the definition of "property," and criticizing the fact that none of these cases "makes any reference to any of the others, or makes any effort to integrate its innovations . . . into the preexisting fabric of the law").

5. *See* Stephen W. Mazza & Tracy A. Kaye, *Restricting the Legislative Power to Tax in the United States*, 54 AM. J. COMP. L. (SUPPLEMENT) 641, 644–46 (2006) (surveying the sweeping judicial deference to tax legislation and administrative regulation).

6. *See, e.g.*, William B. Stoebuck, *A General Theory of Eminent Domain*, 47 WASH. L. REV. 553, 571 (1972) (arguing that taxation "is not merely similar to eminent domain; it is the same, as far as the power itself goes").

7. *See* discussion *infra* subpart III(A).

legal concepts that otherwise seem to make little sense under a point-blank analysis of specific property doctrines.

In essence, American law comprehends the governmentally provided system of property as charged mainly with the duty to provide and enforce readily identifiable sets of entitlements and obligations in regard to resources—ones that endow property owners with the security of holding on to features that stress visible bundles such as formal title, possession, use, and control over decision making. But at the same time, property law is largely devoid of an independent, firm undertaking to preserve a definite economic value for assets. In other words, whereas the American property system, as construed by the Supreme Court, considers the power to exclude to be "one of the most treasured strands in an owner's bundle of property rights,"[8] and similarly views rights of possession, control, and disposition as "valuable rights that inhere in the property,"[9] no such clear commitment exists for any particular benchmark of value. Counterintuitive as it may sound, value in itself is not one of the strands of constitutionally guarded property.

I argue that this is the case not because American society is indifferent to asset values—quite the contrary. It is so because in a free-market-oriented yet organized society, a property-rights system created and enforced by the state simply cannot commit itself simultaneously to (a) strong, constitutionally based protection of certain property bundles such as exclusionary possession, use, control over decision making, or free alienability as inherently grounded in formal title; and (b) some objective, entrenched stream of economic benefits deriving from property ownership. This basic insight has enormous implications for the way in which property law is structured, including the various demarcations drawn out in takings jurisprudence.

This construction of American property law is far from inevitable. In many national and subnational economic systems, the rules pertaining to the control and use of resources are more oriented toward ensuring a certain value for stakeholders, but this comes at the price of stronger ongoing intervention with formal property strands. This is the case not only with traditional communities in the developing world or with centrally controlled national economies,[10] but it can also be traced in other market economies as well, as in alternative subsociety structures within the United States.[11]

Nowadays, following the financial crisis, American society is undergoing a dramatic process of aggressive governmental intervention with what were considered to be the basic tenets of markets and private rights.

---

8. Loretto v. Teleprompter Manhattan CATV Corp., 458 U.S. 419, 435 (1982).

9. Phillips v. Wash. Legal Found., 524 U.S. 156, 170 (1998).

10. *See* Amnon Lehavi, *The Global Law of the Land* (pt. 1), 81 U. COLO. L. REV. (forthcoming 2010) (manuscript at 3–7, on file at http://ssrn.com/abstract=1357731) (surveying the history of property law in Europe and Asia).

11. *See infra* notes 83–86 and accompanying text.

The American Recovery and Reinvestment Act of 2009,[12] viewed as a "striking return of big government" with its $787 billion worth of government spending, expansion of social programs, and federal tax cuts,[13] and its predecessor the Emergency Economic Stabilization Act of 2008,[14] which mandated the massive purchase of mortgage-related securities and capital investments in banks,[15] along with the outright takeover of key financial institutions and companies to prevent asset meltdown,[16] may all have profound long-term effects on the fundamentals of American property law.

Importantly, this Article does not aim at suggesting which property paradigms should be considered normatively superior. What this work does is recognize the fact that American law made some very meaningful choices—not at all universally "inherent"—in the way it has conventionally constructed its property system, and argue that the resulting jurisprudence and its specific doctrines and line drawing that emerged over the years should be understood in view of these foundational principles. Accordingly, this Article does not offer a normative defense of the current taxonomies in takings law. It instead takes on the innovative analytical enterprise of illuminating the broader perspectives against which current legal rules have been shaped, thus tying together what are so often considered to be loose ends within the takings jurisprudence.

How does the characterization of the grand structure of American property law help to better explain the logic behind the intricate web of takings law and, specifically, the broad gap in the constitutional approach to taxation vis-à-vis other forms of state coercion against property owners?

First, a formal, market-oriented system that consecrates certain sticks as indispensable seems to view as a lesser evil those forms of governmental extraction of private value that minimize the explicit derogation of such prominent property incidents. Whenever a governmental act coercively acquires entitlements—such as ownership, leasehold, or easement—either explicitly, by registering such rights in the government's or a third party's name, or implicitly, by using rights and prerogatives that are regularly considered to represent the core of such rights—by entering land to set up

12. Pub. L. No. 111-5, 2009 U.S.C.C.A.N. (123 Stat.) 115.

13. David M. Herszenhorn, *A Smaller, Faster Stimulus Plan, but Still with a Lot of Money*, N.Y. TIMES, Feb. 14, 2009, at A14.

14. Pub. L. No. 110-343, 122 Stat. 3766 (to be codified at various sections of 5, 12, 31 U.S.C.).

15. *See* David M. Herszenhorn, *Bush Signs Bill: House Votes 263 to 171—Markets Down on Jobs Data*, N.Y. TIMES, Oct. 4, 2008, at A1 (describing how the legislation enabled the U.S. Treasury to buy troubled securities to ease the credit crisis); Mark Landler, *Stock Markets Rally Worldwide—Biggest Intervention Since '30s*, N.Y. TIMES, Oct. 14, 2008, at A1 (announcing the Treasury Department's plan to invest in banks and guarantee new debt issued by banks for three years).

16. *See, e.g.*, Stephen Labaton & Edmund L. Andrews, *Mortgage Giants Taken Over by U.S.: A Costly Bailout.*, N.Y. TIMES, Sept. 8, 2008, at A1 (discussing the bailout of "the nation's two largest mortgage finance companies," Fannie Mae and Freddie Mac).

public facilities and thus undermining the right to exclude,[17] by making certain interventionist decisions about the use of the resource,[18] or by prohibiting or limiting certain forms of asset transfers[19]—the owner's remaining rights are viewed as crippled. This is so even if the pure economic consequences of such governmental acts are not harsher than those inflicted by a newly imposed tax on the property. The property system has been better accustomed to view taxation as a background institution, which although financially significant, creates less uncertainty in figuring out who the owner is and what she owns.[20]

Second, legal concepts and doctrines controlling governmental interventions with private property are obviously not hermetically detached from the private law of property, especially in a market-oriented system. Although the interface between the private and public realms in property is highly intricate and avoids clear demarcation,[21] and although I definitely do not argue that the law of governmental intervention with private property should necessarily aspire for harmony with the law governing property relations among private parties,[22] it would be safe to say that the law of takings does have some bearing on the way people broadly understand property entitlements and obligations.

Thus, for example, the public and legal outrage over the *Kelo v. City of New London*[23] decision, as vividly expressed in Justice O'Connor's assertion in her dissent that "[n]othing is to prevent the State from replacing any Motel 6 with a Ritz-Carlton, any home with a shopping mall, or any farm with a factory,"[24] expressed a deep concern that the overbroad construction of "public use" to facilitate a condemn-and-transfer practice for economic development was not only a matter of governmental abuse, but one that also undermined the fundamental understanding of what it means to be a property owner, including vis-à-vis other persons.[25] Indeed, in a number of cases, the

17. *See* Loretto v. Teleprompter Manhattan CATV Corp., 458 U.S. 419, 434–36 (1982) (arguing that a cable TV company's installation of wires in an apartment building constituted a taking from the building's owners).

18. *See* Phillips v. Wash. Legal Found., 524 U.S. 156, 170 (1998) (asserting that the right of disposition over even minimal interest income generated by funds is essential to the concept of private property and that government interference with that disposition may constitute a taking).

19. *See* Hodel v. Irving, 481 U.S. 704, 715–17 (1987) (holding that the right to will one's property to one's heirs is essential to the meaning of ownership and that the total abrogation of this right constitutes a taking).

20. *See infra* section III(B)(1).

21. Amnon Lehavi, *The Property Puzzle*, 96 GEO. L.J. 1987, 2000–12 (2008).

22. *See id.* at 2017–18 (arguing that governmental intervention through eminent domain is often justifiable).

23. 545 U.S. 469 (2005).

24. *Id.* at 503 (O'Connor, J., dissenting).

25. *See id.* at 505 (arguing that property rights are fundamentally insecure once a legislative body may take one party's private property for the benefit of another private party); Thomas W. Merrill & Henry E. Smith, *The Morality of Property*, 48 WM. & MARY L. REV. 1849, 1879–84

Supreme Court has made cross-references between the public law and private law of property, for example by referring to its private law jurisprudence in defining the "treasured" right to exclude in takings cases such as *Loretto v. Teleprompter Manhattan CATV Corp.*[26] and *College Savings Bank v. Florida Prepaid Postsecondary Education Expense Board.*[27] Viewed through this prism, one might understand why taxation is generally considered to be less controversial than the governmental taking of property strands that are intuitively analogized to the core concepts of private law. In this sense, it is more convenient for courts to view taxation as a qualitatively distinct type of governmental intervention with private property.

Third, taxation may often be viewed as actually entrenching and validating formal ownership, thus strengthening the security of title and formal rights. As this Article shows, this is especially the case with the property tax, which has received only scant attention in the taking/taxing academic debate, but is nevertheless considered a major source of revenue for government, as well as a chief determinant of local governance and property-owner-based collective control. In fact, one underlying characteristic that seems to broadly differentiate legal systems with state-dominated formal private-property rights in land from those that have a less comprehensive formal regime is the extent to which the imposition and collection of property taxation is fiscally significant and administratively feasible, since such a tax inevitably depends on a centrally coordinated recordation (or at the least a governmental validation) of lands and title holdings.

An important caveat is in order at the outset. Even if one accepts the categorization of taxation as a distinguishable type of governmental action in the American setting, this does not necessarily mean that judicial review of such acts must always be lenient. Specifically, the ability of courts to divert their attention in such matters to other constitutional channels, most prominently to procedural and substantive due process, may be considered a potential blessing rather than a matter of confusion or undue fragmentation. Since the "property" component of due process is quite consistently considered to be more detached from the private law of property than is the case with the "private property" of the Takings Clause,[28] due process jurispru-

(2007) (portraying a *Kelo*-type condemn-and-transfer use of the eminent domain power as contradicting popular conceptions about the overall morality of property rights).

26. 458 U.S. 419, 436 n.12 (1982) (referring to PruneYard Shopping Ctr. v. Robins, 447 U.S. 74 (1980), which discussed the private property rights of a shopping mall that banned the handing out of antiwar pamphlets).

27. 527 U.S. 666, 673 (1999) (referring to K Mart Corp. v. Cartier, Inc., 485 U.S. 176, 185–86 (1988), which was a trademark law dispute).

28. *See* Merrill, *supra* note 4, at 969–70 (identifying takings jurisprudence as especially befitting discrete assets that are the object of ownership and that include in turn a right to exclude others, hence doing "a good job of identifying those interests that we may loosely call common-law property rights").

dence may be better suited for doctrinal severance in reviewing such types of governmental actions.

This Article is structured as follows: Part II identifies the core ingredients of a property system committed to the formality of private property rights and to free-market trade, and explains why such a legal regime cannot purport to protect both firm private control and guaranteed value for such assets. Part III presents the doctrinal differentiation between taxation and other forms of governmental intervention with property. It briefly discusses prominent critiques of current doctrine and then explains why this taxonomy does seem to make better sense when viewed through the larger framework of the American property system. Part IV reflects briefly on the potential pros and cons of creating categories in property law by reevaluating other types of legal line drawing in takings jurisprudence and the more general nature of legal taxonomy in property law. This Article concludes that because legal taxonomy is necessarily embedded in broader normative and institutional considerations, any major shifts in the fundamental paradigms of American property law that ensue in the aftermath of the financial crisis are bound to reconfigure the line drawing of property doctrines.

II. The Core of American Property Law

Reducing American property law to a clear-cut paradigm is obviously highly challenging. First, local and state property laws may substantially diverge among different jurisdictions within the United States,[29] and federal law in itself is highly complicated and often obscure, with federal constitutional property law being a special source of intricacy.[30] Second, the law of property is also highly contingent on the type of resource that is the object of property rights, both in defining the scope of rights and in providing remedies to protect them, such that the laws of land, chattels, intellectual property, or securities may significantly differ from one another.[31] Third, on a normative level, it is highly doubtful whether American property law adheres to

29. *See, e.g.*, Abraham Bell & Gideon Parchomovsky, *Of Property and Federalism*, 115 YALE L.J. 72, 74–75 (2005) (discussing competition over property regimes among different jurisdictions); Christopher Serkin, *Local Property Law: Adjusting the Scale of Property Protection*, 107 COLUM. L. REV. 883, 889 (2007) (explaining how local zoning restrictions change over time due to electoral changes or shifts in demographics).

30. *See, e.g.*, Merrill, *supra* note 4, at 889–90 (discussing four Supreme Court decisions in 1998 and 1999 that were "likely to produce bewilderment among lower courts and practicing lawyers" in terms of how to demonstrate that an interest is property deserving protection under the Due Process Clause or the Takings Clause).

31. The question of the degree to which there is—or should be—similarity in defining property rights for different resources is highly contentious and will not be discussed here. For the tension between tangible and intangible property, compare Peter S. Menell, *Intellectual Property and the Property Rights Movement*, REGULATION, Fall 2007, at 36, 37, with Richard A. Epstein, *A Response to Peter Menell: The Property Rights Movement and Intellectual Property*, REGULATION, Winter 2008, at 58, 61–63.

any predominant goal, as various values such as societal welfare, liberty, personhood, equity, and social responsibility battle it out not only in scholarly discourse[32] but also in the actual design of property doctrines.[33] This latter aspect is particularly pertinent in dramatic times such as the current ones, in which government engages in a major restructuring of property markets and institutions.[34]

That said, in analyzing the current landscape of American property law, it seems that at least two traits can be discerned as typifying its grand structure, even if they are applied somewhat differently in various contexts.

A. The Formality of Property . . .

Property law in the United States is by and large formal, meaning that state institutions set out to create and enforce sets of private and public entitlements and obligations pertaining to resources,[35] to make them publicly known and transparent to the extent necessary and feasible,[36] and to generally subject other types of property arrangements to the overall supervision and control of the centrally coordinated property system.[37]

Formality of property is obviously no novelty. It is a fundamental feature of the social contract underlying modern organized society and government, by which the protection and stability of private property is both cause and effect in the entrustment of rule making and enforcement at the hands of the sovereign—even if the various prominent theories in Western thought substantially diverge on the proper scope of government power in shaping the procedural and substantive ingredients of such a property system.[38]

---

32. *See* Lehavi, *supra* note 21, at 1997–98 (contrasting arguments regarding appropriation of resources based on economic-efficiency considerations with arguments based on ideas of social justice and equality). For an important recent statement in favor of social responsibility in property, see Gregory S. Alexander, *The Social-Obligation Norm in American Property Law*, 94 CORNELL L. REV. 745, 769–70 (2009).

33. *See, e.g.*, THOMAS W. MERRILL & HENRY E. SMITH, PROPERTY PRINCIPLES AND POLICIES 243–91 (2007) (discussing the influences of personhood considerations on designing the objects and contents of property).

34. *See, e.g.*, Labaton & Andrews, *supra* note 16 (reporting the federal government's seizure of Fannie Mae and Freddie Mac in an effort to stabilize the financial crisis). For a listing and description of all the interventions by the government during the financial crisis, see FED. RESERVE BANK OF N.Y., FINANCIAL TURMOIL TIMELINE 4–5 (2010), http://www.ny.frb.org/research/global_economy/Crisis_Timeline.pdf.

35. *See* YORAM BARZEL, A THEORY OF THE STATE: ECONOMIC RIGHTS, LEGAL RIGHTS, AND THE SCOPE OF THE STATE 23 (2002) (arguing that the state uses third-party power to enforce contractual agreements and that in doing so it delineates legal rights).

36. *See infra* notes 41–43 and accompanying text (discussing the standardization of property entitlements and obligations and the publication of these rights and duties).

37. For the intricate relationships between formal and informal subsociety property regimes, see Amnon Lehavi, *How Property Can Create, Maintain, or Destroy Property*, 10 THEORETICAL INQUIRIES L. 43, 67–74 (2009).

38. *See, e.g.*, JEREMY WALDRON, THE RIGHT TO PRIVATE PROPERTY 5–16 (1989) (offering a critical assessment of prominent property theories, from Marxist to libertarian ones).

In the case of land, the most significant step in the evolution of property in the Anglo-American tradition was the gradual shift of the land tenure system from the original, early medieval interpersonal web of direct services and duties owed through the chain of feudal hierarchy to the impersonal, permanent, and inheritable system of land entitlements.[39] This process reflected and further entrenched the centralization of political power, the shifting focus from the family to the individual as the subject of law, and the constant expansion of market—rather than status—society.[40] Accordingly, two prominent principles that emerged from the formalization of property were, first, standardization of the types of officially recognized forms of property entitlements and obligations (i.e., the *numerus clausus* principle explicit in the civil law system[41] but also highly indicative of the Anglo-American one[42]) and, second, the creation of mechanisms for publicizing such rights and duties, mainly by public recordation or registration of entitlements in land.[43]

The resulting structure of the American legal regime is thus that property rights—ownership, leasehold, servitude, mortgage, etc.—comprise a set of rights and duties that are "endorsed" by the state—to use Felix Cohen's famous depiction of property[44]—and are accordingly enforced and remedied when these rights are being breached.

The specific content of the property bundle is a source of fierce debate, mainly between essentialists—those who believe that certain sticks inhere in property rights with the right to exclude being most often associated with the inevitable core of ownership—and those who take the bundle concept to have a normative meaning such that the array of rights and duties can and should be contextually crafted by state institutions.[45] But regardless of this debate, property is typified by the fact that the various entitlements and obligations,

---

39. *See* JESSE DUKEMINIER & JAMES E. KRIER, PROPERTY 199–204, 209–11 (5th ed. 2002) (describing the early system of feudal tenures, services, and incidents and the rise of the fee simple).

40. Lehavi, *supra* note 10, at 4.

41. UGO MATTEI, BASIC PRINCIPLES OF PROPERTY LAW: A COMPARATIVE LEGAL AND ECONOMIC INTRODUCTION 39 (2000).

42. *See* Thomas W. Merrill & Henry E. Smith, *Optimal Standardization in the Law of Property: The* Numerus Clausus *Principle*, 110 YALE L.J. 1, 12–24 (2001) (describing the relative strength of *numerus clausus* in property rights, from the five fixed types of estates in land and concurrent interests to the wider variety of more customizable intellectual property interests).

43. *See* Benito Arruñada, *Property Enforcement as Organized Consent*, 19 J.L. ECON. & ORG. 401, 425–28 (2003) (justifying governmental monopoly in land recording and registration as facilitating private contracts and protecting third parties).

44. *See* Felix S. Cohen, *Dialogue on Private Property*, 9 RUTGERS L. REV. 357, 374 (1954) (summarizing Cohen's analysis of property in terms of a simple label, which is signed by a private citizen and endorsed by the state, that says "[k]eep off X unless you have my permission, which I may grant or withhold").

45. For an overview of this debate, see Lehavi, *supra* note 21, at 2000–07, providing the history and distinct viewpoints of essentialists and disintegratives, and then examining the weaknesses of each school of thought.

the strands in the bundle that are recognized by the state institutions, are then formally enshrined, respected, and enforced.[46]

One needs, however, not to confuse the formal trait of property with any of the terms "absolute," "clear-cut," or "complete." First, the absolutistic conception of ownership as a "sole and despotic dominion"[47] is long considered obsolete, normatively unworthy, and practically unfeasible.[48] Second, property entitlements and obligations are often not clear-cut either in terms of the nature of the right or in the scope of its remedial enforcement. As Henry Smith shows, whereas some property doctrines follow an "exclusion" strategy, others—such as nuisance conflicts—often adopt a "governance" approach that breaks up property rights into more specific-use entitlements and also tend to contextualize the remedy awarded.[49] This does not undermine, however, the formality of property in the senses elaborated above. Third, and related, the formality of property law does not necessarily mean that property rights are complete, such that law is able to conceive of every possible conflict in advance, explicitly allocate every potential attribute of the resource,[50] or predict every relationship that will develop among persons with respect to the resource.[51] Accordingly, in recognizing that potential loopholes are probably inevitable, property law often resorts to legal standards that leave such conflicts to *ex post* judicial rulings that would fill such incomplete norms with content.[52] Yet this too does not undercut the overall formal structure of property.

---

46. Importantly, the specific composition of the bundle is a source of argument even among those who subscribe to an essentialist viewpoint. *See, e.g.*, Larissa Katz, *Exclusion and Exclusivity in Property Law*, 58 U. TORONTO L.J. 275, 277–78 (2008) (arguing that the core feature of ownership is not physical exclusion but rather the owner's exclusive right to "set the agenda" for the resource).

47. 2 WILLIAM BLACKSTONE, COMMENTARIES *2.

48. *See, e.g.*, Carol M. Rose, *Canons of Property Talk, or, Blackstone's Anxiety*, 108 YALE L.J. 601, 603–06 (1998) (describing the limitations and qualifications of Blackstone's notion of property as exclusive dominion and noting that Blackstone himself was aware of many of these problems with his theory).

49. *See* Henry E. Smith, *Exclusion and Property Rules in the Law of Nuisance*, 90 VA. L. REV. 965, 972–73 (2004) (discussing modes of delineating property rights and comparing "exclusion" regimes, in which the law grants owners a "gatekeeper" right to exclude others from a resource, to "governance" regimes, which focus on proper use of a resource).

50. *See* YORAM BARZEL, ECONOMIC ANALYSIS OF PROPERTY RIGHTS 90–91 (2d ed. 1997) (stating that describing and protecting property consumes resources, and perfect delineation of property rights is prohibitively costly, so property rights are never perfectly delineated).

51. *See* Amnon Lehavi, Legal Standards in Property 19–32 (Feb. 2, 2010) (unpublished manuscript, on file with Texas Law Review) (laying out a nonexhaustive taxonomy of incompleteness in property to describe why legal designs cannot "predict, allocate, and decide in advance all possible states-of-the-world regarding the bundle of property rights").

52. A seminal work on the rules-versus-standards debate in law is Louis Kaplow, *Rules Versus Standards: An Economic Analysis*, 42 DUKE L.J. 557 (1992). For a slightly different depiction of this type of difference in the context of property, see Carol M. Rose, *Crystals and Mud in Property Law*, 40 STAN. L. REV. 577, 592–95 (1988).

That said, American property law does place enormous weight on defining types of property interests, crafting the bundle of entitlements and obligations for each one of them, and viewing property rights as worthy of legal validation and protection *as such*[53]—so that the jurisprudential inquiry starts with the identification of rights and duties and whether these were violated, and only then moves to evaluate the effects of the infringement for designing the appropriate remedy. Hence, for example, in the constitutional setting, it is not the loss of value in itself that triggers constitutional scrutiny and intervention, but rather the identification of constitutionally protected rights and a resolution that such rights have been infringed by state action.[54]

This trait of American property law is vividly demonstrated in a couple of seminal Supreme Court cases, which quite dramatically separated the component of the taking of a constitutionally protected property right from the different question of loss of value.

In *Loretto*, the Court reviewed § 828 of the 1973 New York Executive Law enacted to facilitate tenant access to cable television.[55] Section 828 provided that a landlord may not "interfere with the installation of cable television facilities upon his property" and limited compensation to an amount later set by the State Commission on Cable Television at $1.[56]

The Court accepted the New York Court of Appeals' determination that § 828 serves a legitimate purpose.[57] Yet, portraying the power to exclude as "one of the most treasured strands in an owner's bundle of property rights," the Court viewed permanent physical invasion to land as qualitatively different from other types of intervention, noting that a special injury occurs when such an invasion and occupation is made by "a stranger,"[58] and concluded that state-authorized permanent invasions constitute a taking per se.[59] On remand, the New York Court of Appeals upheld the State Commission's determination regarding the $1 compensation, relying, *inter alia*, on the relatively insignificant market-value damage to an owner's property by attachment of cable facilities.[60] And yet, *Loretto* remains deeply rooted in

53. *See, e.g.*, Chi., Burlington & Quincy R.R. Co. v. Chicago, 166 U.S. 226, 241 (1897) (finding that a violation of a property owner's rights in his property is, on principle, an unconstitutional violation of due process).

54. *Id.*

55. *See* Loretto v. Teleprompter Manhattan CATV Corp., 458 U.S. 419, 421 (1982) (summarizing the impact of N.Y. EXEC. LAW § 828 (McKinney Supp. 1981–1982)). Prior to 1973, Teleprompter obtained installation permits from property owners along the cable route in return for a standard rate of 5% of the gross revenues that it realized from the particular property. *Id.* at 422–24.

56. *Id.* at 423–24.

57. *Id.* at 425.

58. *Id.* at 435–36.

59. *See id.* at 441 (affirming the traditional rule that a permanent physical occupation is a taking).

60. *See* Loretto v. Teleprompter Manhattan CATV Corp., 446 N.E.2d 428, 434–35 (N.Y. 1983) (concluding that the due process requirement of just compensation had been met).

American takings jurisprudence as constituting a rule that any type of permanent government invasion to land, regardless of its actual effects, violates constitutionally protected property rights.[61]

Even more instructive in this respect are the Court's decisions in *Phillips v. Washington Legal Foundation*[62] and later in *Brown v. Legal Foundation of Washington*.[63] The two cases dealt with the Interest on Lawyers' Trust Accounts (IOLTA) programs adopted in different states. Under these programs, certain client funds held by an attorney in connection with his practice of law are deposited in a bank account, with the interest income generated by the funds being paid to foundations that finance legal services for low-income individuals.[64] The Court generally recognized the respondents' argument that each one of the separate client funds was too small to generate interest income in itself, such that there was no direct economic loss,[65] but at the same time held that:

> We have never held that a physical item is not "property" simply because it lacks a positive economic or market value. For example, in *Loretto* . . . we held that a property right was taken even when infringement of that right arguably *increased* the market value of the property at issue. Our conclusion in this regard was premised on our longstanding recognition that property is more than economic value; it also consists of "the group of rights which the so-called owner exercises in his dominion of the physical thing," such "as the right to possess, use and dispose of it." While the interest income at issue here may have no economically realizable value to its owner, possession, control, and disposition are nonetheless valuable rights that inhere in the property.[66]

In *Brown*, in a 5–4 decision, the Court once again held that the IOLTA programs constituted a taking, since the interest of the bank accounts' beneficial owners was "taken for a public use when it was ultimately turned over to the Foundation."[67] But the Court then went on to say that no "just compensation" was due for the taking because "compensation is measured by the owner's pecuniary loss—which is zero whenever the Washington law is obeyed" so that "there has been no violation of the Just Compensation Clause of the Fifth Amendment in this case."[68]

---

61. *See, e.g.*, Eric Kades, *Drawing the Line Between Taxes and Takings: The Continuous Burdens Principle, and Its Broader Application*, 97 NW. U. L. REV. 189, 199 (2003) (acknowledging that permanent state invasions are, per se, unconstitutional takings).

62. 524 U.S. 156 (1998).

63. 538 U.S. 216 (2003).

64. *Phillips*, 524 U.S. at 156.

65. *See id.* at 169 (observing that the one IOLTA provision that might result in a generation of net interest was not being enforced).

66. *Id.* at 169–70 (citations omitted).

67. *Brown*, 538 U.S. at 235.

68. *Id.* at 240.

One may be left to wonder—as the dissenting opinion in *Brown* did[69]—what point is there in recognizing an infringement of property rights as a "taking" but at the same time holding that no compensation is due. Puzzling and controversial as this ruling may be,[70] it does seem to reflect a persisting leitmotif in U.S. property law, by which formal property rights, and not value, are the subject of legal protection,[71] whereas lost private value serves as a benchmark—though not the only possible measure—in designing the remedy.[72]

As a matter of fact, the less clear the issue of formal rights, the "muddier" the applicable legal doctrine. Consider, for example, the body of law that deals with nonconfiscatory regulations that adversely affect private assets. It is governed by the extremely complicated and ad-hoc test developed in *Penn Central Transportation Co. v. New York City*,[73] according to which the court examines: (1) "[t]he economic impact of the regulation on the claimant"; (2) the extent of interference with "distinct investment-backed expectations"; and (3) "the character of the governmental action."[74]

Although at least one prong of the test seems to focus on value as an independent factor that should be considered to determine whether a taking has occurred, it seems that the enormous confusion that governs regulatory takings can be attributed to the fact that the Court has been unable to address a more fundamental, straightforward question: What kind of legal right, if any, does a person have to develop her privately owned land?[75] The difficulty in defining the nature and extent of such a strand and the ensuing ad hocery are understandable, yet they emphasize that when the Court moves away from the notion of rights, it truly struggles in shaping its takings jurisprudence.[76]

---

69. *See id.* at 252 (Scalia, J., dissenting) ("Perhaps we are witnessing today the emergence of a whole new concept in Compensation Clause jurisprudence: the Robin Hood Taking . . . . [T]o extend to the entire run of Compensation Clause cases the rationale supporting today's judgment . . . would be disastrous.").

70. Debate lingers as to the finding of a taking in this case. *See* MERRILL & SMITH, *supra* note 33, at 1333 (questioning whether the Court engaged in inappropriate "conceptual severance" of the right to the interest as distinctive of the overall right to the principal in a single account balance). Debate lingers even more so as to the question of compensation. *See, e.g.*, Christopher Serkin, *Valuing Interest: Net Harm and Fair Market Value in* Brown v. Legal Foundation of Washington, 37 IND. L. REV. 417, 418 (2004) (criticizing *Brown*'s compensation principle of "net harm" as inconsistent with just compensation precedents).

71. *See* MATTEI, *supra* note 41, at 172 ("Western legal tradition has developed almost entirely around the protection of property rights.").

72. *See id.* at 178–79 (discussing the use of an injunction as a remedy when monetary damages do not guarantee the property right).

73. 438 U.S. 104 (1978).

74. *Id.* at 124.

75. *See* Gideon Kanner, *Making Laws and Sausages: A Quarter-Century Retrospective on* Penn Central Transportation Co. v. City of New York, 13 WM. & MARY BILL RTS. J. 679, 701–02 (2005) (expressing concern that an owner's right to build on her own land has in some cases effectively ceased to exist).

76. *See infra* text accompanying notes 255–59.

B. . . . and Its Market Propensity

The relationship between property rights and markets might seem straightforward at first glance, but it is far more subtle and intricate, and avoids inherent causality in either direction.

The right of alienability, i.e., the right to transfer property rights in assets to others, is considered to be a standard ingredient of the institution of property. Beyond the clear economic benefit that it endows on the property owner by allowing her to realize the asset's long-term value at the timing of her choice, it is also perceived as enhancing her autonomy in controlling the identity of the successor to the rights, be it in case of a transfer for consideration (sale) or for none (gift, inheritance).[77] In some outstanding cases, the legal system prohibits alienability when it considers the general societal benefits of allowing assets to end up in the hands of those who value them most to be much more than offset by particular moral, societal, or economic considerations—prohibitions on most types of transfers of body parts being a prominent example.[78]

However, the options for legal ordering of alienability do not necessarily narrow down to either authorizing property owners to act in an unfettered market or prohibiting owners altogether from engaging in any sort of transfer. Alienability may be legally sanctioned but at the same time be denied certain features of the free market, e.g., by restricting the identity of potential buyers or sellers, limiting overall supply, substantially intervening in the terms of transference, or otherwise constructing the bundle of rights in certain resources so as to constrain the development of wholly decentralized, impersonal markets (consider instances such as tradable allowance schemes,[79] taxi medallions,[80] tenancy by the entirety,[81] or a partner's interest in the standard business partnership[82]).

Other legal structures for transfers of rights may include pricing mechanisms that deviate from free-market rules. This is the case, for example, with the rapidly growing sector of "shared equity housing,"[83] most

77. MERRILL & SMITH, *supra* note 33, at 531.

78. For a discussion of inalienability in the body-parts context, see Guido Calabresi, *An Introduction to Legal Thought: Four Approaches to Law and to the Allocation of Body Parts*, 55 STAN. L. REV. 2113, 2132–52 (2003).

79. *See* Carol M. Rose, *Expanding the Choices for the Global Commons: Comparing Newfangled Tradable Allowance Schemes to Old-Fashioned Common Property Regimes*, 10 DUKE ENVTL. L. & POL'Y F. 45, 51, 68 (1999) (analyzing the effectiveness of tradable environmental-allowance systems, which allow environmental rights to be traded to conserve resources).

80. *See* Katrina M. Wyman, Is Bentham Right? The Case of New York City Taxi Medallions 18–23 (Dec. 15, 2008) (unpublished manuscript, on file with Texas Law Review) (exploring the rights and regulations of transferring taxi medallions in New York City).

81. *See* MERRILL & SMITH, *supra* note 33, at 635–36 (describing tenancy by the entirety).

82. *See infra* text accompanying note 112.

83. *See* JOHN EMMEUS DAVIS, NAT'L HOUS. INST. SHARED EQUITY HOMEOWNERSHIP: THE CHANGING LANDSCAPE OF RESALE RESTRICTED, OWNER OCCUPIED HOUSING 2 (2006), *available at* http://www.nhi.org/pdf/SharedEquityHome.pdf (explaining that nonprofit organizations, private

dominantly Community Land Trusts (CLTs).[84] One of the underlying features of CLTs is that upon resale of an individual housing unit, the CLT repurchases the property itself or monitors its direct transfer from seller to buyer, restricting the resale price to a formula that aims at giving the departing homeowner a fair return on her investment while giving future income-eligible homebuyers fair and affordable access to this housing.[85] Hence, CLTs formally circumvent free-market pricing in transfers of housing units.[86]

The main point here is that a societal choice to resort to the full-blown features and effects of free markets regarding property rights is neither automatic nor indispensable. It reflects a conscious determination, according to which the implementation of property rights through largely unfettered markets will optimize the attainment of organized society's goals and values. This, so I argue, is the case with constitutional, statutory, and judicially created American property law. What this means is that for most types of resources, property law and policy rely on free markets not only as a measure to implement existing rights but also as the normative paradigm upon which property rights are designed *ab initio* by the legal system.

A prominent example of property law's market propensity is the evolution of the land system. In subpart II(A), I mentioned the major shift in English land law, by which the original feudal system was gradually replaced by an impersonal, inheritable, and marketable system of property rights, reflecting major sociopolitical changes.[87]

The intensification of such processes typified American law's early endeavors to break ranks with those elements of English land-law heritage that were still considered archaic in the American context.[88] Thus, for example, Thomas Jefferson's view that the fee tail and primogeniture were detestable means of perpetuating a hereditary aristocracy led him to persuade

---

lenders, and governmental agencies have all increased their involvement with resale-restricted owner-occupied housing in recent years).

84. Briefly, the CLT is a community-based nonprofit that acquires land for the purpose of retaining ownership in it forever for affordable housing. *Id.* at 18. The individual homeowner leases the land for a long period of time and is the owner of the building that is erected on the land. *Id.* "The lease agreement on the land divides the property bundle between the individual and the CLT both during the tenancy and upon its transfer by inheritance or resale." Amnon Lehavi, *Mixing Property*, 38 SETON HALL L. REV. 137, 201 (2008). Thus, for example, the homeowner must occupy the land as his primary residence, may not sublease the land without the CLT's consent, and is obligated to properly maintain the building. DAVIS, *supra* note 83, at 19. "If the homeowner fails to pay the mortgage, his interests may be taken over by the CLT." Lehavi, *supra*, at 201.

85. DAVIS, *supra* note 83, at 19.

86. *See* Lehavi, *supra* note 84, at 199–202 (describing the alternative process by which CLTs transfer housing units).

87. *See supra* text accompanying notes 39–40.

88. *See* LAWRENCE FRIEDMAN, A HISTORY OF AMERICAN LAW 413 (2d ed. 1985) ("Inherited doctrines did not last if they seemed to clash with the needs of the American economy."). For the immigration of English land law to the American colonies and later to the United States, see *id.* at 58–65, 230–45, 412–21.

the Virginia legislature to abolish these mechanisms around the time of the American Revolution, with most other state legislatures soon following suit.[89] More broadly, American law consistently worked to entrench concepts of standardization of estates and types of property rights, promote free alienability, coordinate registration of transactions in land, and facilitate a broad, impersonal free market for real estate.[90] Not surprisingly, the fee simple soon came to dominate the landscape of American real estate, as it seemed to epitomize the idea of clearly delineated, strong property rights that allow for easy recordation, facilitation of credit, and broad mandate for transfer of rights.[91]

Moreover, in view of the fact that public housing traditionally has not played a dominant role in shaping land development,[92] the real-estate economy that developed over the years was one of a decentralized market that is governed mainly by the forces of supply and demand,[93] and is regulated chiefly by local governments that in turn rely extensively on value-based property tax as a major source of public revenue.[94] Real property has thus been dominated, for better and for worse, by market forces and trends.

This market propensity is evident in the crafting of property institutions for other resources. Intellectual property law, for example, is an immensely broad field that does not follow a single blueprint either in its normative underpinnings or in the doctrinal rules applying to each one of its different branches.[95] But it seems safe to say that the market is not only the mechanism through which intellectual property rights are being implemented and

89. JESSE DUKEMINIER ET AL., PROPERTY 188 (6th ed. 2006).

90. *See, e.g.*, Uriel Reichman, *Toward a Unified Concept of Servitudes*, 55 S. CAL. L. REV. 1177, 1189–90 (1982) (explaining the development of new types of servitudes in America both by the reduced fear of impeding assignability in a country with vast resources of uncultivated land and by the existence of an efficient recording system in the United States from early on).

91. *See, e.g.*, Robert C. Ellickson, *Property in Land*, 102 YALE L.J. 1315, 1368–71 (1993) (discussing the benefits of fee ownership in land).

92. *See* Robert C. Ellickson, *The Mediocracy of Government Subsidies to Mixed-Income Housing Projects* 4–7 (Yale Law & Econ. Research Paper No. 360, 2008), *available at* http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1217870 (noting the varied public-housing programs that have been adopted since the 1930s).

93. Major exceptions to this have been the federally sponsored mortgage insurance and secondary mortgage market starting as of the mid-1930s. *See* ALEX F. SCHWARTZ, HOUSING POLICY IN THE UNITED STATES: AN INTRODUCTION 47–68 (2006) (outlining the history of housing finance in the United States).

94. Amnon Lehavi, *Intergovernmental Liability Rules*, 92 VA. L. REV. 929, 948–52 (2006).

95. *See, e.g.*, Yochai Benkler, *Through the Looking Glass: Alice and Constitutional Foundations of the Public Domain*, 66 J.L. & CONTEMP. PROBS. 173, 175 (2003) (arguing that copyright law is defined by a tension between doctrines rather than by a single rule); Daniel A. Crane, *Intellectual Liability*, 88 TEXAS L. REV. 253, 254 (2009) (arguing that recent developments in intellectual property law represent "a shift from a property regime to a liability regime"). There are of course other approaches to intellectual property law, emphasizing values such as democracy, openness, and pluralism, calling in turn to extend the scope of public domain in regard to such resources. *See* Benkler, *supra*, at 181–82 (discussing the various values of the public domain, a concept exemplified by the Internet and television).

given economic substance, but is also a dominant goal in its own right in the very creation of certain intellectual property rights. U.S. intellectual property law has thus been portrayed as fulfilling two distinct fundamental functions: first, promoting innovation in technological or expressive works (being a principal motive behind patent, copyright, and other laws), and second, "ensuring the integrity of the market place" (pertaining to trademark law and issues of unfair competition law).[96]

As for innovation, the choice to promote it through the allocation of exclusive property rights in the information output, rather than through other potential legal mechanisms for reward—such as a governmental grant for the innovative effort—is by no means self-evident and has been the subject of increasing debate in legal and economic literature.[97] This is especially so in view of the concern that the benefits of awarding exclusive property rights may be offset by problems such as consumer deadweight loss,[98] inefficient underutilization of information for further development by others,[99] and transaction costs that may prohibit efficient reallocation of the rights.[100]

Yet irrespective of the normative debate on whether the mechanism of exclusive property rights is better than others, by awarding innovators exclusive rights such as using, selling, displaying, or reproducing the protected information, the legal system consciously absolves itself of the need to measure and legally entrench the value of the input or of reasonable expected returns to it.[101] Rather, law awards the innovator with an exclusive right to capitalize on her innovation through the market for the period of protection, thus granting the forces of market demand the power to decide the economic fate of the information's realized value. While a small percentage of patent- or copyright-protected information turns out to enjoy large, long-enduring streams of incomes, most others turn out to be commercially insignificant or

---

96. *See* Peter S. Menell & Suzanne Scotchmer, *Intellectual Property Law*, *in* 2 HANDBOOK OF LAW AND ECONOMICS 1473, 1475 (A. Mitchell Polinsky & Steven Shavell eds., 2007) (exploring the "economic dimensions of the intellectual property field" and the "principle objective[s] of intellectual property law").

97. *See generally* Steven Shavell & Tanguy van Ypersele, *Rewards Versus Intellectual Property Rights*, 44 J.L. & ECON. 525 (2001) (explaining both sides of the scholarly debate on preferable mechanisms to facilitate innovation).

98. *See* Menell & Scotchmer, *supra* note 96, at 1476–77 (arguing that the protection of intellectual property is necessary because a competitive market is unable to support an efficient level of innovation yet acknowledging that this very protection may result in "dead weight loss to consumers").

99. *See* MICHAEL HELLER, THE GRIDLOCK ECONOMY: HOW TOO MUCH OWNERSHIP WRECKS MARKETS, STOPS INNOVATION, AND COSTS LIVES 1–22 (2008) (arguing that the creation of too many separate owners of a single resource leads to a failure of cooperation because each owner can block the others' use).

100. *See* Menell & Scotchmer, *supra* note 96, at 1477 (pointing out that intellectual property does not "guarantee that the research effort will be delegated efficiently to the most efficient firms, or even the right number of firms").

101. *See* Henry E. Smith, *Intellectual Property as Property: Delineating Entitlements in Information*, 116 YALE L.J. 1742, 1748 (2007) (arguing that exclusive property rights allow returns from inputs without the need to value the uses of the inputs).

outright failures.[102] Hence, notwithstanding any intrinsic autonomy-based benefits that a creator may enjoy when she is recognized as formal owner of her innovation, the actual economic value of such protected information is not in any way enshrined or guaranteed by the state, as opposed to the protection of the legal right in it.

The second major pillar of intellectual property law, that of "protecting the integrity of the marketplace," is obviously no less inclined towards the market. It is in fact the cause for the creation of the rights. The conventional economic rationale for protecting trademarks or restricting certain types of unfair competition is to ensure the "quality of information in the marketplace" so that consumers are not misled or confused about the source of goods and accumulated information on producers' goodwill and products' quality.[103] In this sense, granting legal rights and subsequent causes of action to producers is conceived largely as a vehicle to promote the functionality of the market.[104]

As a final example, the unique property structure of business organizations, and especially of the modern corporation, is the subject of much analysis.[105] My interest here is in the property rights of corporate shareholders, in view of the corporation's separate legal entity and its ownership of the corporation's assets.[106] This "asset partitioning"[107] between the corporation's separate pool of assets and the personal assets of the firm's owners calls into question what it is exactly that the shareholder "owns."

In their classic work, *The Modern Corporation and Private Property*, Adolf Berle and Gardiner Means define the shareholder interest as "passive property," endowing him with the beneficial interest of "an expectation that a portion of the profits remaining after taxes will be declared as dividends, and

---

102. *See* Samson Vermont, *The Economics of Patent Litigation*, *in* FROM IDEAS TO ASSETS: INVESTING WISELY IN INTELLECTUAL PROPERTY 327, 332 (Bruce Berman ed., 2002) (noting that 97% of U.S. patents generate no revenues).

103. *See* Laura R. Bradford, *Emotion, Dilution, and the Trademark Consumer*, 23 BERKELEY TECH. L.J. 1227, 1241 (2008) ("Trademarks are an efficient and simple means of communicating information. Sellers use advertising and trade symbols to inform likely buyers about desirable qualities and characteristics of their goods. Trademarks ensure that consumers associate these characteristics with the right product.").

104. The legal system therefore protects formally recognized rights but not an independent commitment to guarantee some benchmark of economic value for the producer of information. *See* WILLIAM M. LANDES & RICHARD A. POSNER, THE ECONOMIC STRUCTURE OF INTELLECTUAL PROPERTY LAW 354–56 (2003) (stating that the only remedy for lost trade secrets is through conventional common law; the law gives no remedy when a trade secret has been leaked or unmasked by an opponent).

105. The two foundational works on this topic are probably ADOLF A. BERLE & GARDINER C. MEANS, THE MODERN CORPORATION & PRIVATE PROPERTY (Transaction Publishers 1991) (1932), and Ronald H. Coase, *The Nature of the Firm*, 4 ECONOMICA 386 (1937).

106. *See generally* ROBERT CHARLES CLARK, CORPORATE LAW 15–19 (1986) (explaining the legal bases for partnerships' and corporations' legal personalities).

107. *See* Henry Hansmann & Reinier Kraakman, *The Essential Role of Organizational Law*, 110 YALE L.J. 387, 392–96 (2000) (defining "asset partitioning" as "the partitioning off of a separate set of assets in which creditors of the firm itself have a prior security interest").

that in the relatively unlikely event of liquidation each share will get its allocable part of the assets."[108] This is in addition to the right to vote (which they deemed to be diminishing to negligible importance) and the right to bring action against the corporation in cases of theft, fraud, or certain wrongdoing by the managers.[109]

According to Berle and Means, shareholders in the corporation have "exchanged control for liquidity."[110] Although the argument that power and control in the corporation has shifted away from the common shareholders is subject to much criticism and indeed seems to be overbroad in reality, the second part of their insight seems to adequately reflect what is perhaps the most striking feature of shareholding: the nearly unconstrained ability to sell the shares in the market "for ready cash."[111] As Robert Clark notes, the free transferability of shares and the existence of organized stock markets make the shareholder's bundle of rights easily sold and realized, in stark contrast to the property interest of a partner in a business partnership.[112]

Somewhat surprisingly, the doctrinal development of the shareholder's right to sell his shares has been rather sparse,[113] perhaps because it has been seen as self-evident in the absence of specific circumstances that justify the imposition of limits on this right. In practice, the right to sell shares in the market is considered to be much more significant and readily viable than the right to receive actual dividends, and thus seems to reflect the economic core of property rights in corporate shares.[114] This means that similar to other types of resources discussed above, property rights in corporate shares are very much market-oriented. In other words, while the legal system protects the right of shareholders to approach the market and not to be abused by other shareholders or the corporate managers,[115] the price of the share is generally determined by the market and no set or minimal value is enshrined by law as part of the property right. The decentralized market setting of the value of ownership obviously has its price tag, as we are witnessing in the

108. BERLE & MEANS, *supra* note 105, at xxxi.

109. *Id.*

110. *Id.* at xvi.

111. *Id.* at xxxi.

112. CLARK, *supra* note 106, at 14–15.

113. *See* J.P. Ludington, Annotation, *Validity of Restrictions on Alienation or Transfer of Corporate Stock*, 61 A.L.R.2d 1318, § 7 (1958) (noting that no American decisions have addressed the validity of an absolute prohibition on the transfer of stock contained in articles of incorporation).

114. *See* Julian Velasco, *The Fundamental Rights of the Shareholder*, 40 U.C. DAVIS L. REV. 407, 407 (2006) (conceptualizing the right to sell alongside the right to elect directors as the two "fundamental rights").

115. *See id.* at 421–24 (describing the rights of shareholders to sue shareholders and managers under state and federal law).

current financial downturn.[116] But this principle has been, at least until now, the key to the social and legal understanding of what stock ownership means.

III. Property-System Analysis of the Taking/Taxing Taxonomy

A. The Judicially Created Divide and Its Critique

The judicial treatment of taxing as a governmental power that is inherently different from other types of economic deprivations of private wealth is one of the most long-standing and entrenched concepts of American constitutional law.[117] The constitutional "power to lay and collect taxes"[118] had been depicted by the Court from early on as "essential to the very existence of government,"[119] and it has consequently been viewed as located well within the domain of the legislature.[120] Without going into a detailed chronology of the fate of different channels of constitutional challenges to tax legislation, this basic conception of taxation means that courts broadly defer to legislatures and guard only against rare instances in which the act demonstrates a gross abuse of the taxing power.[121]

Moreover, the Court has made clear that taxation for a public purpose does not even trigger the Takings Clause because taxation, "however great," is not considered "the taking of private property for public use, in the sense of the Constitution."[122] Thus, although the Court has stated that it will intervene in "rare and special instances" in which the tax is "so arbitrary as to constrain to the conclusion that it was not the exertion of taxation but a confiscation of property; that is, a taking of the same in violation of the Fifth Amendment,"[123] the judicial review of taxation has been conceptually and doctrinally divorced from takings jurisprudence, with the Court emphasizing time and again that these two realms are "essentially different."[124] Put somewhat differently, tax legislation is being scrutinized only when the governmental act is considered to be illegitimate on its own terms: the focus is on the appropriateness of the public action and less on the nature and extent of harm to the taxpayer. This is unlike takings law, under which an

116. *See, e.g.*, Steven Gjerstad & Vernon L. Smith, Editorial, *From Bubble to Depression?*, WALL ST. J., Apr. 6, 2009, at A15 (explaining that price trends and momentum can drive bubbles even when traders know the true value of an asset).

117. Mazza & Kaye, *supra* note 5, at 641.

118. U.S. CONST. art. I, § 8, cl. 1.

119. McCulloch v. Maryland, 17 U.S. (4 Wheat.) 316, 428 (1819).

120. Leo P. Martinez, *"To Lay and Collect Taxes": The Constitutional Case for Progressive Taxation*, 18 YALE L. & POL'Y REV. 111, 113–14 (1999).

121. Kades, *supra* note 61, at 204–06; *see also* Calvin R. Massey, *Takings and Progressive Rate Taxation*, 20 HARV. J.L. & PUB. POL'Y 85, 102–11 (1996) (explaining the conceptual distinction between taxation and takings); *cf.* Martinez, *supra* note 120, at 126–44 (explaining how this premise would permit congressional implementation of a progressive taxation scheme).

122. County of Mobile v. Kimball, 102 U.S. 691, 703 (1880).

123. Brushaber v. Union Pac. R.R. Co., 240 U.S. 1, 24–25 (1916).

124. *Kimball*, 102 U.S. at 703.

infringement of property rights may require constitutional compensation even if the governmental act is otherwise legitimate, reasonable, and furthers a public purpose.[125]

This traditional divide has been increasingly criticized. Commentators, looking at the substantive effects of tax legislation—the amounts extracted, the proportionate sacrifice of the taxpayer vis-à-vis others, and the reciprocity of governmental benefits in return for the tax—have argued that the cases in which a taxpayer is forced to surrender substantial value on a disproportionate or nonreciprocal basis are not inherently different from the taking of property.[126] In other words, whatever, if any, the institutional or power-conferring-source differences between tax legislation and other government measures are, overlaying substantive principles such as fairness, proportionality, or efficiency in burden sharing for the public benefit are those that should govern legal delineation.[127]

Accordingly, numerous writers, albeit with very different normative agendas, have called to formally unify the legal principles pertaining to takings and taxings. At one end of the spectrum, Richard Epstein has seen the conceptual similarity as vindicating the case for circumventing any type of governmentally imposed burden that does not conform to strict proportionality, most notably progressive income taxation.[128] Calvin Massey, driven by a similar normative agenda, calls to extend the various takings-doctrine tests to progressive taxation,[129] such that in appropriate cases the court could "conclude that the portion of income taken by progressive taxation is an uncompensated taking either because it is a permanent dispossession or because it deprives the taxpayer of all economically viable use of that severed strand of property."[130]

At the other end, progressive writers take a different route in calling for such a synthesis. Eric Kades identifies an overreaching constitutional principle of preventing the singling out of a few property owners for an unfair share of public burdens while allowing reasonably constructed progressivism, and thus calls to apply a "Continuous Burden Principle" that

---

125. *See* Daphna Lewinsohn-Zamir, *Compensation for Injuries to Land Caused by Planning Authorities: Towards a Comprehensive Theory*, 46 U. TORONTO L.J. 47, 55–56 (1996) (chronicling takings law jurisprudence).

126. *See* Eduardo Peñalver, *Regulatory Taxings*, 104 COLUM. L. REV. 2182, 2185–91 (2004) (reviewing the scholarly literature representing the different sides of the taking/taxing debate).

127. *See, e.g.*, Saul Levmore, *Just Compensation and Just Politics*, 22 CONN. L. REV. 285, 292 (1990) (suggesting that a system of government that protects property rights would minimize the use of broadly based taxes that do not directly entitle the taxpayer to corresponding benefits).

128. *See* RICHARD A. EPSTEIN, TAKINGS: PRIVATE PROPERTY AND THE POWER OF EMINENT DOMAIN 283–305 (1985) (analyzing special assessments, windfall-profits taxes, severance taxes on the extraction of minerals, income taxes, and estate and gift taxes to determine their validity based on the principle of proportional impact).

129. *See* Massey, *supra* note 121, at 111–23 (arguing that a constitutional rule of proportional taxation is required to make taxation consistent with the Takings Clause).

130. *Id.* at 124.

would monitor against two discontinuous "jumps" in marginal burdens imposed on owners or taxpayers.[131] Eduardo Peñalver identifies two principles that should guide all inquisitions as to the constitutionality of governmental burdens: (1) effect on nonfungible property interests and (2) singling out of owners for disparate treatment.[132] Yet Peñalver works in exactly the opposite way from Epstein or Massey. Viewing taxation doctrines as ones that enjoy "greater consensus," he suggests that it is rather takings law that should make systematic adjustments, such that many types of governmental interventions that are currently being considered as takings—physical or regulatory—should be regarded as legitimate forms of "regulatory taxation," thus narrowing the scope of the takings doctrine.[133]

Irrespective of these normative divergences, all of these writers are probably correct in their basic intuition that an isolated, "one on one" comparison of taking and taxing cases yields little support for the type of deep and uncompromising divide created by the courts over the years. Possible "reconciling theories"[134] trying to distinguish either the facts of specific cases, the economic consequences of certain taxing schemes versus confiscatory or regulatory acts, or the doctrinal sources of government authority as a ground for such a categorical disconnect are bound to encounter some substantial degree of incoherence or overlapping.

Yet the key to understanding why the courts have been so consistent in developing such an allegedly inconsistent jurisprudence lies in figuring out the broad-based pillars of American property law—most specifically, the two traits of rights formalism and market propensity. This Part now moves to explicate how these features better explain the development of the taking/taxing taxonomy. It might be worthwhile, however, to flag a preannounced conclusion that will be discussed further in Part IV: The taking/taxing divide is not inherent to the institution of property as such; it was created and is maintained in American law as a progeny of the general paradigms of its property system, such that a change of paradigms in American jurisprudence may in turn influence this seemingly persistent enclave of a judicially created categorical distinction.

B. Reevaluating the Taking/Taxing Line Drawing

1. The Divide and the Constitutional Protection of Rights, Not Value.— What exactly is it about property that the Constitution protects under the Fifth Amendment? My argument is that it does not protect the asset's value

131. Kades, *supra* note 61, at 224–47.

132. Peñalver, *supra* note 126, at 2215–18, 2223–28.

133. *Id.* at 2248–51.

134. *See id.* at 2192 (arguing that "Reconciling Theories" concentrating on only one element of takings law fail to account for "subsidiary fixed points," such as the need to compensate those whose real property is appropriated by the government).

in itself against government-inflicted losses, but rather that it shields those legally recognized rights contained within the statutorily or judicially crafted bundle of rights in regard to such assets, with the question of restoring lost value coming into play mostly during the second stage of remedying the infringement.[135] The thrust of the judicial enterprise of creating content in constitutional property thus lies in delineating the type of protected rights and entitlements and the kind of circumstances under which governmental invasion of such rights amounts to a constitutional violation that requires a remedy.

Moreover, the unequivocal embracement of fair market value as the measure of constitutional "Just Compensation"[136] further illustrates that the question of value is not only contingent on the identification of an otherwise protected right, but also that the quantification of compensated-for value is not determined a priori. The top-down constitutional protection of rights is not followed by an enshrinement of a certain socially determined stream of benefits or a capitulation to the subjective demands of the injured owner; value is set by aiming to mimic the market and trying to identify what would have been agreed upon between a "willing buyer" and a "willing seller."[137] Opting for fair market value frees the government from making difficult policy choices about what is the proper value that a person is entitled to enjoy as owner of a certain resource, a determination that may have enormous implications on the government/individual property relationships in many ways beyond the specific instance of a taking.[138]

In these two fundamental respects, the public law of property very much resembles the American private law of property. In setting up a system of formal, enforceable private property rights that applies among members of society, the law determines what interests or strands are enshrined and under what circumstances they would be enforced and remedied in case of a breach

135. This is not to say that the question of loss is irrelevant to the first stage of inquiry. As I showed in subpart II(A), the amount of loss is one of the prongs of the *Penn Central* test for regulatory takings—but as I argued, this is exactly why regulatory takings law is so muddy. *See supra* text accompanying notes 73–76.

136. U.S. CONST. amend. V, cl. 4.

137. *See* United States v. 564.54 Acres of Land, 441 U.S. 506, 511 (1979) ("[T]he owner is entitled to receive 'what a willing buyer would pay in cash to a willing seller' at the time of the taking." (quoting United States v. Miller, 317 U.S. 369, 374 (1943))).

138. *See* Katrina M. Wyman, *The Measure of Just Compensation*, 41 U.C. DAVIS L. REV. 239, 244 (2007) (criticizing the sweeping adherence to "subjective" measures and calling to incorporate objective parameters for compensation that will reflect broader societal perspectives about the goals of property). Other legal systems do tie up the question of just compensation for takings to fundamental societal concepts about property. For example, § 25(3) of the 1996 South African Constitution, dealing with compensation for expropriation, creates a multi-factor test, in which the market value is but one component, and aims at achieving "an equitable balance between the public interest and the interests of those affected." S. AFR. CONST. 1996 § 25(3).

by another.[139] But unlike the state endorsement of such rights and correlative duties, no such guarantee exists in regard to the actual value that the legally undisturbed owner would enjoy.[140] As Lee Anne Fennell notes, whereas a homeowner more or less controls the "onsite factors" of the resource (and I would add the control of onsite-related legal infringements such as nuisances), the owner can do no more than hope for the best regarding the substantial value influences of "offsite factors" such as "neighborhood changes and larger housing market trends."[141]

The private law of property thus enshrines a certain bundle of rights and access to the market. But it does not vouch for the actual stream of benefits that the owner derives, either as consumption or as investment within the market. What one therefore legally owns is the set of exclusive rights allocated to her, not the resource's value.[142] Accordingly, although different in many ways from the public law realm, the private law of property similarly protects rights, not value.

This, I think, is a point that is largely overlooked in many of the current critiques that seek to analyze and "deconstruct" the basic taking/taxing taxonomy. A comparative examination that isolates the amount of economic loss or the spread of economic deprivation among different owners/citizens as the all-embracing legal watermark cutting through different categories of government action misses the enormous importance that the American property system places on identifying the different types of property strands, the diverging ways in which these strands may be infringed upon, and the signals that a certain governmental action affecting property sends to the entire property system.

Disregarding these elements makes it very difficult to understand why the Court is taking pains to hold that the governmental acts in the above-discussed *Loretto* and *Washington Legal Foundation* cases constitute a taking, and why it hails the significance of constitutionally recognized strands such as possession, use, or alienability and accordingly refrains from awarding compensation when it is obvious at the outset that no actual economic or market value loss has occurred.[143] But this is the way in which American property law works: it attributes enormous significance to identifying the formal features and attributes of property rights and to pointing out the cases in which these recognized rights are infringed upon—although it is

---

139. *See* Wyman, *supra* note 138, at 246 (explaining the process of formulating a system of property law by analogizing the common justifications for takings law, justice and deterrence, to similar justifications for tort law).

140. *See* EPSTEIN, *supra* note 128, at 3 (discussing the uncertainty of the value of the enjoyment of property rights apart from state endorsement).

141. Lee Anne Fennell, *Homeownership 2.0*, 102 NW. U. L. REV. 1047, 1049 (2008).

142. *See* J.E. Penner, Value, Property, and Wealth 1–7 (2007) (unpublished manuscript, on file with Texas Law Review) ("To own something is to be in the position of one with a particular property right to it, the property right (or interest) called 'ownership.'").

143. *See supra* subpart II(A).

at the same time careful and pragmatic in selecting the actual modes of intervention, relying as it does broadly on the dominance of markets in value setting.

It is this broader perspective of American property law that helps to explain why taxation is generally considered the lesser evil among the different forms of governmental extraction of private value. The property system is better accustomed to viewing taxation as a background institution, which, although financially significant, creates less uncertainty in figuring out who the owner is and what is the bundle she owns as compared to other types of governmental interventions with property rights in resources.[144]

The starting point of this differentiation is the pragmatic understanding that government must act at times through coercion to finance public goods, solve other collective-action problems, or promote values and goals that cannot be advanced solely through the market.[145] The qualitative nature and extent of such government activity is of course a matter of a fundamental normative resolution, be it a "night-watchman state,"[146] a highly progressive interventionist welfare state, or anywhere in between, but the essentiality of some level of resource coercion in itself cannot be denied.

Given this upfront dictate, the way that typical taxing schemes work, including progressive income or business taxation, may be very irritating to those who pay them, but they do not tend to undermine the broader understanding of who the formal owner and the person who otherwise controls decision making, use, and other legally recognized strands for a given resource is.[147] Taxes also tend to impinge less on the basic freedom that a person has to act in the market (in most cases, taxation is a result of a person's otherwise-autonomous decision to sell or buy, although some specific types of taxes do seek to change actors' incentive structure).[148] Moreover, these are not generally interpreted as reshuffling common

144. *See* JOHN RAWLS, A THEORY OF JUSTICE 274–84 (1971) (arguing that taxation is a conditional institution that exists to further the more essential aspects of governance).

145. *See* MANCUR OLSON, THE LOGIC OF COLLECTIVE ACTION: PUBLIC GOODS AND THE THEORY OF GROUPS 101 (1965) ("Coercion is a means of assuring the full effectiveness of the communal spirit, which is not equally developed in all members of the community.").

146. *See* ROBERT NOZICK, ANARCHY, STATE, AND UTOPIA 26–27 (1974) (announcing the subtle distinction between the minimal state and the ultraminimal state).

147. *See* WALTER J. BLUM & HARRY KALVEN, JR., THE ANATOMY OF JUSTICE IN TAXATION 4–5 (1973) ("[T]he taking of specific property by the state is more intrusive than the creation of obligations to be satisfied in money. . . . [P]erhaps it is suspected that the taking of property will be not systematic or disciplined by principle.").

148. *See* Roy Bahl et al., *The Property Tax in Practice*, *in* MAKING THE PROPERTY TAX WORK: EXPERIENCES IN DEVELOPING AND TRANSITIONAL COUNTRIES 3, 5–8 (Roy Bahl et al. eds., 2008) (reconciling the standard view that property taxes should be designed merely to raise revenues with the conflicting theory that property taxes have the potential to "influence social policy and economic decisions").

understandings about other property institutions that are not directly affected by the tax.[149]

It would be safe to say that although progressive income taxation in the United States is not a matter of consensus, it is not understood by its proponents and opponents alike as giving government carte blanche to similarly intervene with all other types of privately held resources.[150] Even with its various taxing schemes intact, the United States has been perceived from both outside and within as the paradigm of a formal property rights, free-market society.[151] In this respect, taxes are more easily regarded as an isolated phenomenon that does not undercut, and at times—such as with property taxation discussed below—even entrenches ownership, enforcement and protection of rights, and free-market propensity. Obviously, a 100% income tax or anything close to it would be viewed differently by everyone,[152] but this is exactly the kind of "rare and special instances" that take such government deprivation way outside the scope of conventional taxation in American law.[153]

This state of affairs is very different for high profile cases regarding other types of government interventions with private property. I mentioned above the public outcry and legal backlash following the *Kelo* case,[154] but this is in no way an isolated phenomenon. Other key takings cases typically have much broader effects beyond the contours of the specific dispute, and it thus seems clear why the Court is paying such close attention to reviewing these cases and why it retains its ability to intervene in designing and re-designing through them the constitutional landscape of property. Thus, for example, the Court places enormous weight on portraying the nature and scope of the various "strands" such as possession (e.g., *Loretto* or the recent

149. In referring to such understandings or perceptions of taxes versus takings, one may be left to wonder who is the subject of such property views: Is it what Bruce Ackerman calls the "Scientific Policymaker" (i.e., the learned legal professional) or rather the "Ordinary Observer" (i.e., the layman)? BRUCE A. ACKERMAN, PRIVATE PROPERTY AND THE CONSTITUTION 26–31, 97–103 (1977). I shall not delve here into a discussion of Ackerman's theory of property as deriving from the tension between these two viewpoints, but would rather make the argument—that would have to be articulated elsewhere—that legal categories of property tend to be more sustainable when they do not consistently clash with laymen concepts.

150. *See* Marjorie E. Kornhauser, *The Rhetoric of the Anti-Progressive Income Tax Movement: A Typical Male Reaction*, 86 MICH. L. REV. 465, 480–81 (1987) (recognizing that "consensus on tax reform may be impossible because of disagreements on underlying basic values" but noting that tax policy must respond to certain normative values).

151. *See* Kades, *supra* note 61, at 196 n.32 (contrasting the formalism of some judicial opinions with classical thought).

152. *See id.* at 200 ("The classic view suggests that, at some point, a narrowly focused tax becomes a taking; however, the discrete-asset model does not apply the Takings Clause to such a general liability."); Massey, *supra* note 121, at 104–05 (considering it self-evident that such a tax would violate the Takings Clause).

153. *See* A. Magnano Co. v. Hamilton, 292 U.S. 40, 44 (1934) (collecting authority applying this exception).

154. *See supra* text accompanying notes 23–25.

Wilkie v. Robbins[155] case),[156] use and control (e.g., the *Washington Legal Foundation* cases),[157] or alienability (e.g., the *Hodel v. Irving*[158] ruling on the invalidity of prohibitions on descent of fragmented individual ownership within Indian tribal land).[159]

Whereas all of these cases greatly differ from one another, and thus often lack orderly intrafield rules and classifications in the delineation of the spectrum of takings, they do seem to share one distinctive feature. They all touch in some significant manner on the core understandings of the institution of property and the underlying features of the American property system, and thus go well beyond questions of the legitimacy of government authority or the sheer economic consequences of the forced contribution of privately held value. This, so I argue, substantially distinguishes them from tax disputes.

2. Incorporating the Public/Private Interface in Property.—The interrelationship between the public and private law of property is extremely complex, although it receives surprisingly scant attention within the broader public law/private law discourse.[160] Whereas a detailed analysis is outside the scope of this Article, a few observations are in order to explicate why this intricate issue may further support the taking/taxing differentiation.

I referred above to an underlying similarity between the public and private law of property regarding the two basic traits of rights formalism and market propensity.[161] This does not mean, of course, that these two branches of law are synchronic or anything close to it. There are good reasons to

---

155. 551 U.S. 537 (2007).

156. *See id.* (rejecting a claim by a Wyoming rancher, who argued that six years of trespass and harassment by a federal agency looking to attain a free easement constituted a taking); Loretto v. Teleprompter Manhattan CATV Corp., 458 U.S. 419, 435 (1982) (declaring that the "power to exclude," which is inherently bound up with the power to possess, "has traditionally been considered one of the most treasured strands in an owner's bundle of property rights").

157. *See* Brown v. Legal Found. of Wash., 538 U.S. 216, 240 (2003) (rejecting a challenge to the application of IOLTA rules to "Limited Practice Officers (LPOs), nonlawyers who are licensed to act as escrowees in real estate closings"); Phillips v. Wash. Legal Found., 524 U.S. 156, 172 (1998) ("While IOLTA interest income may have no economically realizable value to its owner, its possession, control, and disposition are nonetheless valuable rights.").

158. 481 U.S. 704 (1987).

159. The Court invalidated as a taking the provision of § 207 of the Indian Land Consolidation Act, Pub. L. No. 97-459, tit. II, 96 Stat. 2519 (1983). *Hodel*, 481 U.S. at 718. Section 207 prohibited the descent or devise of any individual interest in Native American tribal land that represented less than two percent of the total tract, so that upon the death of the fractional owner the shares would escheat to the tribe. *Id.* at 704.

160. *See, e.g.*, N.E. Simmonds, *Justice, Causation and Private Law*, *in* PUBLIC & PRIVATE: LEGAL, POLITICAL AND PHILOSOPHICAL PERSPECTIVES 149, 149 (Maurizio Passerin d'Entrèves & Ursula Vogel eds., 2000) ("[T]he distinction between public and private law will only rarely play a dispositive role in the decision of litigated disputes . . . . Legal practitioners and judges are concerned with the immediate business of representing clients or deciding particular cases. It is therefore left to the legal theories to speculate about the public/private distinction.").

161. *See supra* text accompanying notes 141–42.

award government with certain powers that should not be granted to private persons who interact with private property owners—including the power to exercise coercion at times—just as there are solid arguments to impose certain duties and restrictions on government that should not apply, at least not in the same magnitude, to individuals, including limits on discrimination against owners or nonowners,[162] due process requirements,[163] transparency in property dealings,[164] and so forth.

Yet even given these differences, it is clear enough that no hermetic separation exists between public law and private law in just about any specific doctrine.[165] First, as I have shown elsewhere, simply drawing the line between public and private in property is especially complicated as compared to other fields of law, in a way that mandates that decisions shaping the core aspects of property law must be made by state entities entrusted with the power and duty of collective decision making—chiefly legislative and administrative bodies supervised in turn by judicial review.[166]

Second, in dealing with specific doctrines in takings law, it is evident that the Court is aware of the potential spillover effects and interrelationships between the two realms of property law. In Part I, I mentioned the way in which the Court, in deciding the key takings cases of *Loretto* and *College Savings Bank*, referred to its private law jurisprudence in defining the "treasured" right to exclude as a mainstay of constitutional property.[167] Hence, although takings cases touch on the distinctive government power to coercively take or regulate property, the way in which the bundle of rights is

---

162. The borders between public and private action in this respect are not, however, clear-cut. In the famous *Shelley v. Kraemer* case, the Court invalidated race-based restrictive covenants in privately owned houses, reasoning that "in granting judicial enforcement of the restrictive agreements . . . the States have denied petitioners the equal protection of the laws and that, therefore, the action of the state courts cannot stand." 334 U.S. 1, 20 (1948). An even earlier example, the roots of which are located in common law, is the set of limits on exclusion from privately owned public accommodations. *See* Joseph William Singer, *No Right to Exclude: Public Accommodation and Private Property*, 90 NW. U. L. REV. 1283, 1348–51 (1996) (discussing the scope of these limitations).

163. *See* MERRILL & SMITH, *supra* note 33, at 1164–91 (describing how due process provides both procedural and substantive protection of private property rights from government interference).

164. Consider, for example, the prohibition on government to secretly purchase land for public purposes—a limit that does not apply to private actors. *See* Daniel B. Kelly, *The "Public Use" Requirement in Eminent Domain Law: A Rationale Based on Secret Purchases and Private Influence*, 92 CORNELL L. REV. 1, 31–33 (2006) ("[W]hile private parties can choose not to disclose the nature or location of their projects, government projects are subject to public accountability and thus publicly known in advance.").

165. *See* Simmonds, *supra* note 160, at 150 ("Unable to proceed from the common understanding of [the public/private] distinction, therefore, we can only proceed from theoretical articulations that are inherently contentious.").

166. *See* Lehavi, *supra* note 21, at 2012–25 ("[T]he integrity of property as a legal system designed to send broad signals for specific properties and resources is preserved when rival ethical and other considerations are battled out publicly through the prism of society's collective institutions.").

167. *See supra* text accompanying notes 26–27.

defined in such cases has bearing—conceptual and practical—on the way in which we basically understand property in seemingly parallel conflicts among private stakeholders.

Thus, for example, the power of eminent domain, i.e., the coercive taking of possession and title in privately owned land, is allegedly unique to government.[168] However, conflicts about the limits of the power, such as the *Kelo*-type contested application of the public-use requirement for economic development—meaning in effect the condemning and transferring of land to private entities—have obvious implications on the scope of property rights and the type of legal protection that a person is entitled to against potential incursions by others. Beyond the oft-made point that politically powerful private actors would be motivated to circumvent the market and turn to the government as merely a vehicle to facilitate a coerced private transaction,[169] the delineation of the power of eminent domain for such nonquintessential public uses may have implications for the ways in which persons understand the laws of trespass, building encroachments, servitudes, adverse possession, etc. If the Court in *Loretto* defines the constitutional property rights against uncompensated, permanent government invasion based in part on private law doctrines, might not jurisprudential connections be drawn from the expansion of the power of eminent domain for economic development to a potential erosion of the traditional, property-rule protection against private invasions (i.e., injunction) and to a switch toward liability-rule protection (i.e., compensation)? This is not to say that such a private law switch is necessarily wrong.[170] But the point here is that the potential interconnectivity is not something that courts can ignore, especially when they have drawn their own public/private parallels in past cases.

The same can be said about cross-effects in other takings doctrines. Take, for example, the famous "nuisance exception" to takings, as articulated in late nineteenth- and early twentieth-century cases such as *Hadacheck v. Sebastian*[171] and *Miller v. Schoene*,[172] according to which diminution in property value caused by nuisance-control measures never requires compensation.[173] In the 1992 *Lucas v. South Carolina Coastal Council*

168. However, this power extends to entities such as common carriers that are specifically authorized by statute to do so. *See* Amnon Lehavi & Amir N. Licht, *Eminent Domain, Inc.*, 107 COLUM. L. REV. 1704, 1710–11 (2007) (illustrating the situations in which governments delegate their powers of eminent domain to private corporations).

169. *See, e.g.*, Kelly, *supra* note 164, at 34–41 (describing how private parties with disproportionate influence may subvert the eminent domain process for their own advantage).

170. Numerous commentators have advocated such a switch. *See, e.g.*, IAN AYRES, OPTIONAL LAW: THE STRUCTURE OF LEGAL ENTITLEMENTS 38 (2005) (concluding by suggesting an alternative to the injunction approach).

171. 239 U.S. 394 (1915).

172. 276 U.S. 272 (1928).

173. *See generally* MERRILL & SMITH, *supra* note 33, at 1323–28 (describing how the state's police power can interfere with private property rights without providing just compensation under these precedents).

case,[174] the Court limited the rule only to those noxious uses that constitute common law nuisances—hence not only referring to private law but also directly relying on it for the purpose of reconstructing the doctrine.[175]

One can also think about the alleged similarity between the *Washington Legal Foundation* cases and the law of unjust enrichment that deals, *inter alia*, with scenarios in which one person benefits from another person's property and the owner demands to be restituted for such benefits even though she suffered no direct economic harm.[176] Again, although the formal legal fields are distinctive, the conceptual question of whether a nonowner may use property without receiving the owner's consent even if the latter does not suffer direct economic loss is one that crosses the boundaries between public and private law.

Hence, in such takings cases, although the courts are dealing with the use of distinctive governmental powers, the judiciary must and does take into consideration the broad-based effects that such doctrines have on the entire system of property law.[177]

Taxes are different. Taxation does seem to be a unique government power that has no apparent or intuitive parallel in the private law of property.[178] It may be annoying or socially contestable, but it does not have a substantial bearing on the entire property system in the way that various takings cases do. This characteristic allows courts to view taxation as a genuinely distinctive sphere of government activity that must not be inherently conjoined with the takings doctrine but rather can be classified and evaluated on its own terms.

Once again, the argument that I make here is not a normative one, by which courts should not be allowed to scrutinize tax legislation in appropriate cases, but rather a jurisprudential-analytic claim: the genuine distinctiveness of taxation from private property conflicts serves as yet another ground for separating this sphere of government action from the grasp of the complicated-as-it-is takings jurisprudence and the public/private entanglement that is an inherent part of it.

---

174. 505 U.S. 1003 (1992).

175. *Id.* at 1029–31. Accordingly, future challenges to regulatory measures, such as whether the *Lucas* rule applies to regulation of noxious uses that do not constitute a common law nuisance but nevertheless do not wipe out all economically viable use of the property, are bound to consider developments in private nuisance law. John A. Humbach, *Evolving Thresholds of Nuisance and the Takings Clause*, 18 COLUM. J. ENVTL. L. 1, 1–2 (1993).

176. *See generally* HANOCH DAGAN, THE LAW AND ETHICS OF RESTITUTION (2004) (surveying the current practice and theory of unjust enrichment in American law). For such a cross-field analysis in the context of land use, see Lehavi, *supra* note 94, at 986.

177. *See* Humbach, *supra* note 175, at 13–17 (discussing how the Supreme Court has instructed lower courts with regard to takings and their impact on the property law system through the lens of nuisance cases).

178. *See supra* notes 147–49 and accompanying text.

3. Taxation as an Empowerment of Property Rights.—Counterintuitive as it might seem at first glance, some forms of taxation, as opposed to other types of governmental intervention with private property, may rather be viewed as validating, entrenching, and solidifying property rights in assets. The argument here goes beyond the often-contested general claim that the financial contribution of taxpayers is more than offset by the level of overall public benefits provided directly or indirectly to the taxpayer, including through the preservation of a public system of law and order.[179] The extent to which this "tax benefit" theory is valid and can further justify progressive taxation remains deeply contested, though empirically heavily under-researched.[180]

My thesis is different and argues for a link between the payment of taxes and the empowerment of property rights in the specific *assets* that are the object of taxation. In so doing, I focus my attention on what is perhaps the tax that shows the closest link between tax and property rights: the property tax. Although it has been almost completely neglected in the taking/taxing debate—which tends to focus on progressive income taxation—the study of property tax can provide important insights for purposes of the taxonomy discussion.

a. Why the Property Tax Is So Contested.—The property tax in the United States is within the province of state and, chiefly, local governments.[181] Although the state authorizes the imposition of the tax and at times sets certain limits on local decision making (such as on the maximum tax rate or on the overall growth rate of annual tax levies),[182] the various local-level governments enjoy the overwhelming majority of property tax revenues[183] and rely heavily on this tax as the most important source of own-revenue, second only to state aid as a general-revenue resource.[184]

---

179. *See* Kades, *supra* note 61, at 221 (summarizing the arguments that the wealthy benefit disproportionately from goods and services provided by the state).

180. *See id.* at 221 (explaining the difficulty of valuing public goods and services).

181. ROBERT B. DENHARDT & JANET V. DENHARDT, PUBLIC ADMINISTRATION: AN ACTION ORIENTATION 242 (2008).

182. *See* MARK HAVEMAN & TERRI A. SEXTON, PROPERTY TAX ASSESSMENT LIMITS: LESSONS FROM THIRTY YEARS OF EXPERIENCE 9 (2008), *available at* http://www.lincolninst.edu/pubs/dl/1412_733_PFR%20Property%20Tax%20Limits.pdf (discussing the inadequacies of limits on assessment values).

183. *See* RONALD FISHER, STATE AND LOCAL PUBLIC FINANCE 319 (3d ed. 2007) (reporting that these localities—counties, municipalities, townships, school districts, and special districts—enjoy 96.5% of all property tax revenues).

184. *See id.* at 319–20; JOAN YOUNGMAN, LEGAL ISSUES IN PROPERTY VALUATION AND TAXATION: CASES AND MATERIALS 3–6 (2d ed. 2006) (discussing the advantages of using a property-tax system to create revenue for state and local government). The distribution of the property tax among the various types of local governments that have overlapping jurisdiction and taxing power is a highly complicated, state-specific issue that need not be addressed here. *See*

Although the property-tax scheme changes among the different states and localities, land and buildings typically form the major part of the property-tax base, alongside a few items of personal property.[185] The tax base for a specific piece of property is its assessed value ("taxable value"), derived from an estimate of the property's market value according to a set of formal procedures and formulas established by state law.[186] The tax rate is calculated as a certain fraction of the assessed value.[187] The tax rate typically differs among different sorts of assets, with commercial or industrial properties usually subjected to higher tax rates than are residential properties.[188] The property-tax levy for a specific asset is thus calculated by multiplying its assessed tax base by the applicable tax rate.

The property tax is often considered to be unpopular among taxpayers.[189] Probably most famously, in 1978, after several years of major annual increases in property-tax bills following the rapid rise in real-estate prices, voters in California approved Proposition 13, which broadly limited the growth of property taxes.[190] Proposition 13 stated, *inter alia*, that (1) the assessed value of all properties would be set back to their 1975–1976 values; (2) the property-tax rate would not exceed one percent of the assessed value; and (3) the assessed value of any property can increase at no more than two percent per year, unless there is change of ownership, in which case the property is reassessed at its new market value.[191] In the years that followed, many states similarly imposed limits on property tax.[192] Interestingly, in

---

Lehavi, *supra* note 94, at 948–49 (outlining the various state-specific approaches of allocating property-tax revenues among local governments).

185. YOUNGMAN, *supra* note 184, at 7. The movables include certain equipment, inventories, and vehicles. *Id.* at 6–9.

186. The assessed value is set by law or government practice at some specific percentage of market value, called the "assessment ratio rule." This means that the assessed value in itself may often be lower than the full market value of property. FISHER, *supra* note 183, at 321.

187. *Id.*

188. *See* YOUNGMAN, *supra* note 184, at 17 ("[T]he law may in either case call for a nonuniform system of taxation, with the effective tax rates different according to property category, such as residential, commercial, or industrial.").

189. Periodic surveys show that the local property tax is voted by Americans as one of the two most "unfair" taxes, alongside the federal income tax. *See, e.g.*, Richard L. Cole & John Kinkaid, *Public Opinion and American Federalism: Perspectives on Taxes, Spending and Trust*, PUBLIUS: J. FEDERALISM, Winter/Spring 2000, at 189, 194 ("Historically, the local property tax has been ranked the next worst tax. In ACIR's first survey in 1972, it was regarded as, by far, the worst tax. In 1999, 29.4 percent of the respondents rated the local property tax as the worst tax, in line with the predominant trend.").

190. HAVEMAN & SEXTON, *supra* note 182, at 5.

191. *Id.*; *see also* Arthur O'Sullivan, *Limits on Local Property Taxation: The United States Experience*, *in* PROPERTY TAXATION AND LOCAL GOVERNMENT FINANCE 177, 177–78 (Wallace E. Oates ed., 2001) (illustrating the effect of the "modern tax revolt" on the types of property-tax limits in use throughout the nation).

192. *See* HAVEMAN & SEXTON, *supra* note 182, at 15 (reporting that in nineteen states and the District of Columbia there are limits on increases in the assessment of specific properties and that sixteen of these jurisdictions also place limits on the overall growth of tax revenues or cap the tax rate).

2006–2007, after steep price increases for a decade and just prior to the current plunge in the real-estate market, lawmakers in twenty-seven states introduced tax-relief measures to address taxpayer discontent over increased tax burdens.[193]

Despite the alleged discontent with property taxation and the broad disparities between and within different jurisdictions, the Court has broadly deferred to the legislatures in the face of constitutional attacks no less than it has with respect to other taxes. In *Nordlinger v. Hahn*,[194] the Court rejected an equal protection claim made by taxpayers who purchased properties in California after Proposition 13 came into effect and were thus required to pay substantially higher property taxes than their neighbors on otherwise comparable properties.[195] Even though the Court itself admitted that "the differences in tax burdens are staggering,"[196] it emphasized that the standard of review "is especially deferential in the context of classifications made by complex tax laws."[197] It saw "no difficulty in ascertaining at least two rational or reasonable considerations of difference" in the California tax-burden scheme, the first being that "the State has a legitimate interest in local neighborhood preservation, continuity, and stability," and the second being the state's legitimate conclusion that "a new owner at the time of acquiring his property does not have the same reliance interest warranting protection against higher taxes as does an existing owner."[198]

Nordlinger has been depicted as representing the "zenith of the power to tax."[199] My intention here is not to delve into the question of whether *Nordlinger* was rightly or wrongly decided on the merits of its equal protection claim but rather to point to what critics of the taking/taxing taxonomy may view as probably the most artificial categorical distinction between the two realms.[200] Whether one views such a highly differential treatment of property owners as normatively justified or not, what conceptual difference is there between this case and other instances in which landowners claim to have been forced to sacrifice a grossly uneven share of their private assets to promote the public welfare? On the face of it, the property tax seems to be the tax that most resembles other types of governmental intervention with property and should thus be judged accordingly.

My argument is that although a point-blank comparison may indeed leave one to wonder whether such a distinction is genuine or merely a self-

193. *Id.* at 8.

194. 505 U.S. 1 (1992).

195. *Id.* at 10–18.

196. *Id.* at 6.

197. *Id.* at 11.

198. *Id.* at 12–13.

199. Martinez, *supra* note 120, at 140.

200. Peñalver considers *Nordlinger* as justified and well-grounded, and then calls to view regulatory-taking cases in a similar manner, based on the criteria of nonfungibility and singling out. Peñalver, *supra* note 126, at 2202–05, 2212–28.

perpetuating fallacy, a property-system analysis reveals implicit and explicit groundings for the different view of the exerting of property taxation vis-à-vis the taking of property rights in the asset that is subject to taxation. Slanted and disliked as the property tax may often seem, it is in fact part and parcel of the unique way in which the American property system is structured in terms of both individual entitlements and collective governance through property.

b. And Is Yet a Signifier of Formal Rights and Markets.—The existence of an individually assessed, market-based property-tax system that is levied on property owners is far from self-evident. It is in fact probably one of the most identifying features of the American-type, formal, market-oriented property system. Accordingly, property taxation is typified by major differences not only between developed countries vis-à-vis developing and transitional ones, but also among various Westernized countries.[201] My focus is not so much on the quantitative differences in tax revenues (the property tax in OECD countries is on average about 2% of the GDP as opposed to less than 0.7% in developing and transitional countries,[202] with Australia, Canada, and the United States collecting more property taxes than Germany or Britain[203]), but rather on the institutional and organizational features of the establishment, organization, and enforcement of property taxation. The characteristics of the property tax mechanism both reflect and further entrench the fundamentals of a certain property legal system.

Consider the legal and administrative prerequisites for a property-tax system that is shouldered by property owners and based on individual assessment of properties and the imposition of an ad valorem tax set at a certain fraction of the assessed market value.

First, one requires a comprehensive formal recording system of property rights in land. Briefly, the American system of land-title record is one of recordation of deeds.[204] Although the official recordation of a land transaction is not itself binding on third parties, it nevertheless makes public such a constructive notice and endows the recorder with legal priority over

201. *See* Richard M. Bird & Enid Slack, *Introduction and Overview to* INTERNATIONAL HANDBOOK OF LAND AND PROPERTY TAXATION 1, 4 (Richard M. Bird & Enid Slack eds., 2004) ("The diversity in the application of land and property taxes even among the 25 countries [examined] is striking.").

202. Roy Bahl & Jorge Martinez-Vazquez, *The Determinants of Revenue Performance*, *in* MAKING THE PROPERTY TAX WORK, *supra* note 148, at 35, 39.

203. *See id.* at 40 (stating that total property taxes in United States amount to roughly 3% of the GDP); Bird & Slack, *supra* note 201, at 8 (listing Australian property taxes as 2.49% of their GDP, Canadian property taxes as 4.07%, Germany as 1.05%, and Britain as 1.43%).

204. *See* JOSEPH H. BEALE, JR., THE ORIGIN OF THE SYSTEM OF RECORDING DEEDS IN AMERICA 1 (1907) (summarizing the distinguishing characteristics of the American recording system). For more information about the U.S. title-recording system, see MERRILL & SMITH, *supra* note 33, at 917–36.

subsequent parties seeking to register a conflicting transaction.[205] Together with the instrument of private title assurance and the possibility of initiating a quiet-title action to clear away fears of earlier-though-not-registered conflicting claims, the American system provides stable and secure formal property rights.[206]

Beyond the issue of clearly identifying ownership, the major administrative challenge for setting up a property-tax mechanism is to construct a comprehensive, reliable, and constantly updated system of public information on the attributes of each individual parcel: its exact size and geographical boundaries, its legally authorized land uses and other pertinent planning or zoning data, the nature and scope of buildings and other fixtures located on the land, and so forth.[207] Tracking and documenting these features within some sort of a cadastral system[208] is essential not only for resolving potential property disputes among neighbors, or for allowing government to carry out its land use and planning policies, but is also the basis for an efficient, fair, and enforceable property-tax system.[209] All of these traits have a direct bearing on value, and without them, any type of individual-based assessment of the property for tax purposes is virtually impracticable.

The challenges of developing and transitional countries in this context may shed light on what may be considered self-evident in the American setting. Irrespective of the question as to whether a sweeping Western formalization of property rights is in fact constructive for such countries, or can otherwise claim normative superiority over alternative forms of resource control, it seems clear enough that the underutilization of property taxation in many of these countries stems primarily from the lack of a comprehensive system of rights formalization and cadastral information that is paramount to the construction of an individual, parcel-based property-tax system.

Thus, whereas economist Hernando de Soto depicts countries such as the Philippines, Peru, or Haiti as dominated by extralegal land holdings,[210] he

205. MERRILL & SMITH, *supra* note 33, at 918.

206. *Id.* at 907.

207. *See* Robert M. Bird & Enid Slack, *Land and Property Taxation in 25 Countries: A Comparative Review*, *in* INTERNATIONAL HANDBOOK OF LAND AND PROPERTY TAXATION, *supra* note 201, at 19, 41–42 (asserting that for a property-tax system to function appropriately, information about property must be periodically updated and made consistent).

208. *See* Chrit Lemmen & Peter van Oosterom, *Cadastral Systems II*, 26 COMPUTERS, ENV'T & URB. SYS. 355, 355 (2002) (explaining that a cadastral system is the environment in which the process of land administration takes place, which includes land registration and the cadastral map).

209. *See* Gerhard Navratil & Andrew U. Frank, *Processes in a Cadastre*, 28 COMPUTERS ENV'T & URB. SYS. 471, 472–73 (2004) (arguing that the sufficient condition for a cadastre is the inclusion of data on the technical and legal traits of privately owned land needed to assess the tax liability for the owner of that land).

210. *See* HERNANDO DE SOTO, THE MYSTERY OF CAPITAL: WHY CAPITALISM TRIUMPHS IN THE WEST AND FAILS EVERYWHERE ELSE 32–33 (2000) (reporting that in the Philippines, 57% of city dwellers and 67% of the rural population live in housing that is dead capital; in Peru, 53% of city dwellers and 81% of the rural population live in extralegal dwellings; and in Haiti, 68% of city dwellers and 97% of the rural population live in housing with no clear title).

seeks to refute the conception that those who take cover in such sectors do so to avoid paying taxes. He argues that the costs imposed on those who act extralegally generally outweigh the amount of potential taxes paid, and thus points the finger at the "bad legal and administrative system" that denies persons in such countries broad-based access to legally protected property rights.[211]

Other writers, who are more ambivalent about an essential positive link between formal property rights and individual prosperity, nevertheless stress the significant interconnectivity between property taxation and a societal validation of claims to assets. Attempts to establish a viable system of property taxation, even in the face of incomplete formality of rights, absence of a comprehensive cadastre, or yet-to-develop property markets, regard property taxation as distinct from other sources of revenue in that it serves to promote individual or group stability and security in the control over assets.[212]

The comparison to emerging or struggling systems of property taxation may thus shed light on the way in which property taxation is considered in the United States to be inherent to a formal, legally entrenched property system. One can say that despite all political murmurs and alleged unpopularity, the prevailing notion is one of "I'm taxed, therefore I own."

A value-based (ad valorem) taxation system relies at its core on the existence of a vibrant, transparent, and privately dominated property market—one in which actual market, arm's-length transactions serve as the best indicator for the true market value of properties, even if the formal assessment procedure later adds certain interventionist or restricting factors to arrive at the "assessed value" for property-tax purposes.

As Joan Youngman notes, the thousands of taxing jurisdictions in the United States exhibit just about every possible approach to property taxation, with some states such as California, Michigan, Oregon, and Florida allegedly moving away in the last decade from accurate market-based assessments.[213] But at the same time, one can quite safely generalize that the American system, especially as compared to other legal systems, is still very much reliant on property markets as the foundational stone for the design of property taxation.

To start with, even in the aftermath of Proposition 13, most states do not impose statutory limits of market-based assessment, thus maintaining market

211. *See id.* at 153–59 ("[B]eing free from the costs and nuisance of the extralegal sector generally compensates for paying taxes.").

212. *See, e.g.*, Martim O. Smolka & Claudia M. De Cesare, *Property Taxation and Informality: Challenges for Latin America*, LAND LINES, July 2006, at 14, 18 ("Finally, there is an advantage for the property tax to cover informal property because its application requires specific knowledge of the area, which has immensurable value to the city management.").

213. Joan Youngman, *The Property Tax in Development and in Transition*, *in* MAKING THE PROPERTY TAX WORK, *supra* note 148, at 19, 23–24; *see also supra* text accompanying notes 189–93.

evaluation of assets for tax purposes as the rule rather than the exception.[214] Even within those jurisdictions that have adopted restrictive measures, it would be wrong to conclude that market values have no bearing on the way governments construct their public-finance policy or that property taxation is otherwise considered to be alienated from property market values. The formulas that are set by legislatures in these different states do take market values as the basis for assessing taxable value, even if these are somewhat "modified" during the assessment process (typically by capping the annual growth rate of the assessed value).[215]

No better proof can be provided for this ongoing interconnectivity than the recent turn of events following the sharp downturn of real-estate prices throughout the United States since 2007. Property owners from all states (including California, for that matter) have been appealing to their local governments to reassess their home values, and official reassessments that have been made since then have indeed resulted in often-dramatic reductions in assessed values, tax bills, and overall property-tax revenues.[216] Thus, the caps that have been placed during eras of steep rises in market values may have exhibited the genuine liquidity problem stemming from the rapid increases in the property-tax burden relative to residents' incomes for yet-unrealized paper gains in real-estate prices, as well as the all-too-familiar human tendency to try to exploit the political process to shift the burden onto others (as is the case with Proposition 13's "welcome stranger" provision, reassessing property values at the new market price in the case of transfer).[217] But even during times of mercurial market tides, the connection was never broken. Property taxation is thus still inherently intertwined, in the eyes of both governments and residents, with the workings of property markets.

It is in this sense that property taxes are viewed as fundamentally different from takings and other types of governmental interventions with private property. Even if implicitly, the design and administration of prop-

214. *See supra* text accompanying notes 190–93.

215. FISHER, *supra* note 183, at 334–39.

216. *See, e.g.*, Jill P. Capuzzo, *Homeowners Fight Back as Market Cools Off*, N.Y. TIMES, June 15, 2008, NJ, at 1 (chronicling New Jersey homeowners' requests for property-value reassessments in the face of dropping real-estate prices); Michael Mansur, *Kansas Property Assessments Begin to Reflect Real-Estate Market Downturn*, KAN. CITY STAR, June 19, 2008, at B1 (reporting that assessments following the real-estate market's sharp decline better reflected those battered property prices); Jennifer Steinhauer, *Taxes Reassessed in Housing Slump as Prices Decline*, N.Y. TIMES, Dec. 23, 2007, at A1 ("Homeowners across the nation are looking to county governments to reassess the values of their homes in the face of flattening and falling prices that have befallen scores of markets."); Neil Gonzales, *Money for Schools Hit by Property Value Drop*, INSIDE BAY AREA, Mar. 27, 2008, LEXIS, News Library, INSBAY File (explaining how the reassessment of software mogul Larry Ellison's house in California cost local school districts and agencies significant tax revenues); Richard Halstead, *Thousands in Marin Asking for Property Tax Cuts*, MARIN INDEP. J., June 13, 2008, LEXIS, News Library, MARIN File (describing how many Marin County homeowners were seeking lower property taxes based on reassessment of greatly depreciated property values).

217. *See supra* text accompanying note 191.

erty taxation validates the concept of a market-driven property system and can thus be seen as an integral part of it. Contested and highly visible as it is, the perseverance of the property-tax system as an ad valorem, market-based mechanism and the broad judicial deference to these attributes cannot be viewed as a mere coincidence.

In contrast, takings and especially cases of eminent domain generally represent a deviation from the principles of property markets. The coercive transfer of property rights from an individual owner to the government, even if specifically essential to promote general welfare, is conceived as a challenge or as a constitutionally mandated exception to the normal state of affairs by which properties are voluntarily transferred through markets.[218] Moreover, although constitutional due compensation is based on calculating fair market value at the time of the taking, as I noted above,[219] it cannot practically guarantee that the condemned will be subjectively indifferent to the taking as would be the case in a genuine market transaction.[220] In addition, it inherently denies property owners future-value appreciation stemming from the government action. This tension has obviously been the source of tremendous public conflict, as the vivid examples of Susette Kelo and her counterparts in other high-profile takings cases demonstrate,[221] along with the numerous calls for legal reform.[222] It thus seems clear that takings instances do and will remain a breakaway from the ordinary workings of the market, in a manner that further helps to explain the fundamental difference between taking and taxing when viewed through the prism of the American property system in its entirety.

There is yet another prominent aspect in which the structure of property taxation in the United States works to strengthen property rights and other private interests as an inherent part of the very same mechanism that imposes a financial burden on property owners.

As mentioned, the American property tax is almost exclusively the province of local governments.[223] Although the basic economic features of real-estate taxes do seem to be natural to local governance (due to immobility of land, the more intimate acquaintance of the locality with real-estate values, and so forth), it is not self-evident that property taxes would be set,

---

218. *See* Loretto v. Teleprompter Manhattan CATV Corp., 458 U.S. 419, 434–35 (1982) (holding that a permanent physical occupation by the government constituted a taking, regardless of the benefit to the public).

219. *See supra* text accompanying notes 136–38.

220. *See* United States v. 564.54 Acres of Land, 441 U.S. 506, 511 (1979) ("Although the market-value standard is a useful and generally sufficient tool for ascertaining the compensation required to make the owner whole, the Court has acknowledged that such an award does not necessarily compensate for all values an owner may derive from his property.").

221. *See* Lehavi & Licht, *supra* note 168, at 1708–09, 1714–17 (collecting and describing several such cases).

222. *See* Wyman, *supra* note 138, at 256–61 (surveying the different calls for reform in the aftermath of *Kelo*).

223. *See supra* text accompanying notes 181–84.

administered, and collected by local governments, even within the group of market economies. In Britain, for example, the national government has taken over decision making and administration of nondomestic property taxation and distributes the proceeds to localities according to their population.[224] It also sets the assessment and rate principles for the new residential property tax (the Council Tax).[225] Property taxation is also very much centrally coordinated in other Western countries such as in Germany, in which the principles of the federal land-tax law govern this otherwise municipal tax.[226]

The overriding localization of the property tax in the United States thus has to do with much more than administrative feasibility. It represents the pillars of intergovernmental organization, politics, economy, and ideology. As William Fischel argues: "[P]eople view their property taxes as different from other taxes. They are part of their own city's or town's property."[227] Although this popular conception has its downside in that residents of wealthy localities are often hostile to interlocal transfers of tax revenues through regional or state mechanisms,[228] it nevertheless captures the unique way in which owning real property and paying real-property taxes combine to create a collective mechanism that is viewed as entrenching both individual and collective control of resources. The very same polls that point to the alleged "unpopularity" of property taxes also indicate that the American public regards their local governments as providing them with the most for their tax money, as compared to other levels of government.[229] One need not go here into a debate of whether incorporated municipalities are or should be considered genuine corporations in the sense that local property owners may be viewed as shareholders, although municipalities are formally located nowadays at the public end of the public/private corporate spectrum.[230] Even if we wholly deny any formal property or property-like right to individual residents in the collective assets of their local government, property taxation

224. *See* The Non-Domestic Rating (Transitional Period) Regulations, 1990, S.I. 1990/608 (U.K.) (establishing centralized national administration of local-business property taxes, known as National Non-Domestic Rates, or NNDRs).

225. Youngman, *supra* note 213, at 20–23.

226. Paul Bernd Spahn, *Land Taxation in Germany*, *in* INTERNATIONAL HANDBOOK OF LAND PROPERTY TAXATION, *supra* note 201, at 98, 98–99. German land-tax law is also peculiar in that a "standard tax" is set by the state tax administration and applies uniformly to all local governments within the state. *Id.*

227. WILLIAM A. FISCHEL, THE HOMEVOTER HYPOTHESIS: HOW HOME VALUES INFLUENCE LOCAL GOVERNMENT TAXATION, SCHOOL FINANCE, AND LAND-USE POLICIES 120 (2001).

228. *See id.* at 121–22 (citing Maine, a state in which voters in wealthy areas disproportionally supported a referendum to repeal a statewide property tax and school-funding reform, as an example of wealthy localities rejecting such transfers).

229. *See* Cole & Kincaid, *supra* note 189, at 199 ("Local government entered the new millennium being the most favored government, except for the property tax—a continuing tax thorn for many Americans, especially those living in the eastern half of the United States.").

230. *See* GERALD E. FRUG, CITY MAKING: BUILDING COMMUNITIES WITHOUT BUILDING WALLS 39–45 (1999) (describing the history of "the adoption of the private/public distinction").

is still considered to be an uncontestable part and parcel of what it means to be a "homevoter," as Fischel puts it.[231]

The central point I wish to make is that although the assessed value of properties, tax rate, internal-tax-burden sharing, or interlocal-tax-equalization schemes may all be a source of harsh political and legal conflict, the act of paying property taxes is not considered in itself a taking of property but rather a mechanism that inheres in property ownership and in local governance and control of resources. Property taxation is indeed conceptually different from other types of governmental interventions with property that are suspected, both intuitively and doctrinally, as a deprivation of property rights and are accordingly reviewed under the various tests of the Takings Clause.

The conflicts that do arise with respect to property taxes are therefore more appropriately reviewed under other sets of potential claims that more closely capture the essence of property taxation. Thus, for example, to the extent that one group of taxpayers (e.g., new homebuyers in California) claims that the assessment process or the tax rate levied on it is unequal to that of similarly situated groups, the Equal Protection Clause is indeed the jurisprudential route to follow, as was the case in *Nordlinger*.[232] The same goes for due process claims, including substantive due process ones that focus on the reasonableness of the public decision making, i.e., whether government power was exercised "without any reasonable justification in the service of a legitimate governmental objective,"[233] so as to root out arbitrary and capricious decisions that abuse what is otherwise a legitimate power that is not in itself considered an infringement of independently protected rights.[234] In other words, nothing in the conceptual distinction between taxings and takings means that taxation must always be awarded unconditional deference. What it does mean, however, is that the way in which taxation is constructed and embedded within the broader system of property calls for a different kind of constitutional review that more adequately addresses the typical concerns over taxation.

---

231. *See generally* FISCHEL, *supra* note 227 (developing a full description of homevoters as participants in the property system).

232. *See* Nordlinger v. Hahn, 505 U.S. 1, 11 (1992) ("The appropriate standard of review is whether the difference in treatment between newer and older owners rationally furthers a legitimate state interest. In general, the Equal Protection Clause is satisfied so long as there is a plausible policy reason for the classification . . . .").

233. County of Sacramento v. Lewis, 523 U.S. 833, 846 (1988).

234. This distinction that was made in *Lingle v. Chevron U.S.A., Inc.*, 544 U.S. 528 (2005), according to which the due-process-based "substantially advances" formula asks "whether a regulation of private property is *effective* in achieving some legitimate public purpose" and is thus conceptually and doctrinally different from the question of whether private property has been "taken." *Id.* at 542.

IV. The Virtues and Vices of Taxonomy in Property

The discussion of the taking/taxing taxonomy is evidently embedded in the broader enterprise of mapping out the different legal fields pertaining to property rights and interests. Although a comprehensive analysis is outside the scope of this Article, I wish to make here two essential points: one about the general nature of legal taxonomies in law and property law in particular, and the other about the tradeoff in adhering to the current doctrinal delineation of governmental interventions with private property.

To start with, the enterprise of legal taxonomy need not be understood as necessarily yielding to formalist or positivist conceptions of law, in which law purports to be capable of dividing the legal world into neat distinctive categories that simply reflect objective legal reality.[235] Taxonomies and legal categories are analytically and jurisprudentially essential to maintaining a reasonable level of clarity and certainty in organizing the world around us, developing legal expectations, and understanding the normative and policy considerations with respect to different actors, resources, and legal relationships.[236] And yet, no legal taxonomy can be portrayed as wholly detached from the institutional and normative foundations that stand as its basis.[237] Even the allegedly most basic distinctions in law, such as between private and public law, are not "natural" in the sense that these must follow a single formula or that they run across different legal systems irrespective of the governing normative and institutional principles in each one of them. The challenge that a legal system thus faces is to find the appropriate balance between the essentiality of creating a comprehensive taxonomy of legal orderings, while at the same time avoiding the pitfalls of enshrining legal categories as inherently superior to the underlying institutional and normative tenets of the legal system as a whole.[238]

Property law faces particularly intriguing challenges in creating and maintaining such a workable division. As a field of law that sets up the ways in which society orders resources and the human relationships around them, property is typified by the fact that entitlements and obligations in regard to resources regularly implicate numerous parties not only as a matter of

---

235. *See* Hanoch Dagan, *Legal Realism and the Taxonomy of Private Law*, *in* STRUCTURE AND JUSTIFICATION IN PRIVATE LAW: ESSAYS FOR PETERSON BIRKS 147, 147–49 (Charles Rickett & Ross Grantham eds., 2008) (criticizing formalism and positivism, and offering a legal realist conception of taxonomy).

236. *Id.* at 147–48.

237. *See id.* at 154–60 (arguing for a realist conception of taxonomies as inextricably connected to their larger social context).

238. *See* Lehavi, *supra* note 84, at 209–11 ("[T]he law has broader lessons to learn from the constant efforts in other fields to define and redefine the borders between the different spheres of life, recognizing the essentiality of line-drawing but at the same time understanding that these distinctions are uncertain, changeable, and often misleading.").

abstract analysis but also in social and economic practice.[239] Thus, although property is so laden with values and constant moral, political, and societal inquiries,[240] excessive ad hocery aimed at attaining resource-specific efficiency, justice, or some other underlying normative goal comes with its own high price tag because it undermines the broad and relatively straightforward signals that property should send about its core attributes to the large numbers of legal actors implicated by its rules.[241]

One more note is needed here about the nature of legal taxonomy, particularly in property law. A point that is often overlooked in the jurisprudential debate over the enterprise of legal taxonomy is that the link between the number of legal categories and the simplicity of the legal system is not straightforward. The question is not only how many different types of legal categories we have, and how easy it is for us to classify a particular event or situation as falling within a specific category, but also what is the type of legal norm that applies to each category—i.e., whether the norm is designed as a clear-cut rule that sets out a straightforward, relatively rigid decree, or rather, as a broadly phrased standard that requires further, later-stage crystallization.[242] This means that even what might seem at first glance to be a very orderly division of the world into legal categories can turn out to be quite different if each legal category is governed by a broad and vague standard that may more than offset the alleged tidiness of having carved out distinctive categories for different types of disputes. This is an issue of tremendous importance in takings law, within which the different categories can be governed by either per se rules or by highly complicated and "muddy" standards, mostly in the case of regulatory interventions with property.[243] It is thus essential to realize that "taxonomy" is not synonymous with "simplicity" or "rigidity."

I now briefly evaluate the taking/taxing taxonomy as part of the overall arc of American law that deals with the various kinds of governmental intervention with private property. My purpose is to demonstrate that although no taxonomy, carefully designed as it is, can ever be perfect in the sense that it

---

239. *See* Lehavi, *supra* note 21, at 2000–07 (arguing extensively and concluding that "property should thus be understood as a jurisprudential framework the primary purpose of which is to delineate basic kinds of entitlements and obligations in regard to certain types of resources").

240. *See generally* MERRILL & SMITH, *supra* note 33, at 243–392 (discussing the various "values subject to ownership").

241. *See* Lehavi, *supra* note 21, at 2014–21 (expanding on the core argument that "since property implicates a large—technically infinite—number of parties with certain entitlements or obligations for a specific resource (and resources in general), the legal regime has to rely on some core principles that are broadly understood and communicated, thus enabling a baseline of clarity, stability, and mutual respect").

242. *See supra* text accompanying note 52.

243. *See, e.g.*, John J. Costonis, *Presumptive and Per Se Takings: A Decisional Model for the Takings Issue*, 58 N.Y.U. L. REV. 465, 488 (1985) ("[T]he use-dependency test is more narrowly focused and often is considerably more demanding of government than is the reasonable relationship test.").

would create hermetic factual and normative borders that would never encounter some level of intra- or intercategory difficulty, overlapping, or inconsistency; such a delineation should be weighed for its overall systematic efficacy in giving a substantially coherent sense to what is undoubtedly a very muddy world—one in which there are so many different instances in which government affects in some way or another the rights, interests, and expectations of persons. Moreover, such an evaluation should be based not only on discerning an internal logic between different instances of government deprivations but also on the interrelationship between the takings-law taxonomy and the broader founding principles of property law.

Hence, in the example of land, beyond the taking/taxing division, one may observe quite a few other categorical delineations that typify American takings law.

Consider, first, the distinction between permanent physical invasion and regulatory intervention. As *Loretto* demonstrates, the per se taking rule in the former instance applies not only to full-scale eminent domain, but also to an allegedly trivial invasion such as the laying of television cable, despite the fact that the pure economic effect was at worst negligible and potentially was even one of positive value.[244] Needless to say, a critique evaluating the legal implications of a governmental act based on the scope and distribution of economic consequences could argue there is little sense in such line drawing, especially as compared to economically more significant instances of regulatory takings that are nevertheless governed by other legal rules.[245]

Second, for regulatory takings, the Court in *Lucas* famously applied a per se takings rule for a regulation that deprived owners of "all economically beneficial or productive use of the land."[246] This is opposed to the three-prong ad hoc test that generally applies to regulations that impose adverse effects, as developed in *Penn Central Transportation Co. v. City of New York*.[247] This distinction too has been criticized. Thus, for example, as Justice Stevens famously stated in his dissent in *Lucas*: "[T]he Court's new rule is wholly arbitrary. A landowner whose property is diminished in value 95% recovers nothing, while an owner whose property is diminished 100% recovers the land's full value."[248]

---

244. *See* Loretto v. Teleprompter Manhattan CATV Corp., 458 U.S. 419, 452 (1982) (Blackmun, J., dissenting) (noting that the majority decided that a taking had occurred even though "as a practical matter, the [occupation] . . . likely increases both the building's resale value and its attractiveness on the rental market").

245. *See, e.g.*, Costonis, *supra* note 243, at 529 (viewing the *Loretto* rule as "anachronistic" and "aberrational").

246. 505 U.S. 1003, 1015–16 (1992).

247. *See* 438 U.S. 104, 124 (1978) (identifying three relevant factors: the economic impact on the claimant, the extent of interference with distinct investment-backed expectations, and the character of the government action).

248. *Lucas*, 505 U.S. at 1064 (Stevens, J., dissenting).

Third, consider the opposite per se nontaking rule in the regulatory-takings context: the nuisance exception, according to which diminution in property value caused by nuisance-control measures never requires compensation.[249] This rule too did not evade criticism. Michael Heller and James Krier argue, for example, that this distinction unjustly burdens property owners who undertook certain activities that had been previously legally allowed and that generated a stream of revenues, but later came to be considered "noxious" due to a change of taste by government.[250]

Fourth, land-use law seems to draw a broad distinction between existing uses, which enjoy nearly complete immunity against new land-use regulation, and future uses, which, even if financially more significant at times, can be frustrated by regulation that receives substantial deference by reviewing courts applying the *Penn Central* test. Thus, for example, the vested-rights doctrine examines the extent to which the implementation of a real-estate project is sufficiently far along, and in such case awards the owner protection against a subsequently enacted regulation as if the existing use were already intact.[251] Criticizing this distinction, Chris Serkin argues that neither the Takings Clause nor the Due Process Clause provides for such a categorical differentiation and that, normatively speaking, one cannot view existing uses as worthy of a categorically stronger protection over other types of uses and property interests that are adversely affected by land-use regulation.[252]

In point-blank inspection, all of these distinctions may indeed be debatable. The real question, however, is whether the carving out of a certain category and a corresponding legal norm is deemed to be overall efficient and fair when considering both the entire spectrum of instances involving governmental intervention with private property and the broader

---

249. *See supra* text accompanying notes 171–75.

250. *See* Michael A. Heller & James E. Krier, *Deterrence and Distribution in the Law of Takings*, 112 HARV. L. REV. 997, 1009–13 (1999) ("[T]he conventional takings law treatment of nuisances does not necessarily promote fair results.").

251. *See* ROBERT C. ELLICKSON & VICKI L. BEEN, LAND USE CONTROLS: CASES AND MATERIALS 231–33 (2d ed. 2000) (articulating the generic rule that "[m]ost courts will recognize vested rights only if the owner has made substantial expenditures in good faith reliance upon the issuance of a building permit or other approval"). A related doctrine is that of "amortization," under which the government can eliminate a preexisting use that does not conform to a new zoning scheme without having to pay compensation by allowing the affected owner to continue use for long enough to amortize her investment. *See id.* at 222–26 (surveying the variations and nuances of the amortization doctrine as applied to nonconforming uses).

252. *See* Christopher Serkin, *Existing Uses and the Limits of Land Use Regulations*, 84 N.Y.U. L. REV. 1222, 1261 (2009) (observing that "current constitutional doctrine does not compel categorical protection for existing uses" and arguing that "none of the possible justifications for the categorical protection of existing uses withstands serious scrutiny"). Serkin thus calls for applying to existing uses the broader tests deriving from substantive-due-process jurisprudence or from the *Penn Central* regulatory-takings three-prong test. *See id.* at 1288–90 ("Once existing uses are dismantled as a specially protected category, it is possible to address whether a particular existing use should be protected with more precision and sophistication about the costs and benefits actually at stake").

institutional and normative principles of American property law. This is not merely a theoretical inquiry but one that has clear empirical and practical implications: How much do we actually stand to lose from having such differentiations within takings law, and do better mechanisms for delineating and maintaining a workable taxonomy while refraining from absurdities exist?

A detailed analysis of the overall taxonomy in takings law cannot be made here, but an illustrative example for each one of the realms of analysis (i.e., the existence of legal mechanisms to maintain the taxonomy while avoiding absurdities, and the conformity of the taxonomy to the broader principles of American property law—rights formalism and market propensity) may clarify why I consider the overall taxonomic enterprise of taking law to be sustainable.

First, the application of per se rules—aimed primarily at casting several anchors in the stormy waters of takings jurisprudence—has been criticized for inappropriately tackling borderline cases,[253] but I submit that these rules are generally able to resort to mitigating mechanisms to deal with such peculiarities. Thus, the *Loretto* per se rule seems to efficiently cater to dominant perceptions about the special gravity of physically invading someone's land on a permanent basis, while at the same time seems to avoid the absurdities of what is truly an idiosyncratic case in terms of pure economic impact by setting the amount of compensation at the nominal rate of $1.[254] This way of handling idiosyncrasy thus helps to preserve the category's approach to more paradigmatic situations, i.e., eminent domain cases that involve both permanent physical invasion and substantial economic damage.

Second, the broad differentiation in the legal protection against the frustration of existing uses vis-à-vis future uses, which may be criticized when one merely views the pure economic consequences of allegedly comparable land-use regulations, has independent merits that touch on the formal rights orientation of American property law. It aims at balancing the centrality of the liberty to use one's own property with the justification for limiting it to mitigate potential use conflicts and promote broader public needs. The line that has been drawn legally prioritizes the protection of an already existing use as a more firmly recognized strand of property.[255] And as is the case with other categories, the way in which such lines are drawn nevertheless tries to avoid potential absurdities. The vested-rights tests or amortization periods set for existing uses respect such rights without unduly inhibiting societal progress,[256] whereas on the other hand, under the *Penn*

253. *See supra* text accompany notes 245–48.

254. Loretto v. Teleprompter Manhattan CATV Corp., 458 U.S. 419, 435–38 (1982); *see supra* text accompanying notes 58–60.

255. For retroactivity in property law, see MERRILL & SMITH, *supra* note 33, at 1197–98.

256. *See supra* note 251 and accompanying text.

Central test, government cannot wholly disregard "investment-backed expectations" regarding future uses whenever these have established themselves objectively—well beyond mere anticipation—and are thus more appropriately viewed as being normatively and jurisprudentially linked to formal rights, not merely value.[257]

The bottom line, therefore, is that the enterprise of legally delineating the field of governmental interventions with private property, and the taking/taxing taxonomy within it, is instrumental not only for creating a reasonable level of order and security in an otherwise highly complicated area of law but also as a vehicle to implement and at the same time further develop the fundamental principles of American property law.

V. Conclusion

American property law is currently located at a crucial crossroads. Its longtime foundational premises and convictions are now being vigorously reexamined in the face of the domestic and global economic crisis. Governmental action that might have sounded farfetched only a few years ago—such as the partial or full nationalization of financial institutions; dramatically harsher regulation on securities, credit, or mortgage markets; initiatives for different forms of a mortgage moratorium;[258] or provision of a safety net for private savings[259]—questions the underlying features of American property law, including the protection of rights rather than mere economic value and the strong propensity toward markets as the chief vehicle to implement property rights. Whether the government measures taken in the months and years to come will create a major, long-standing upheaval in American property law remains to be seen.

At this point in time, however, it is important to identify the intricate and subtle ways in which the systematic features of American property law implicate, even if often in somewhat implicit ways, the various fields and doctrines in property and property-related matters. As this Article made

---

257. Penn Cent. Transp. Co. v. New York City, 438 U.S. 104, 130 (1978) ("[T]he submission that appellants may establish a 'taking' simply by showing that they have been denied the ability to exploit a property interest that they heretofore had believed was available for development is quite simply untenable.").

258. *See, e.g.*, David Leonhardt, *Life Preservers for Owners Under Water*, N.Y. TIMES, Oct. 22, 2008, at B1 (examining several contemporary propositions for mortgage foreclosure relief, including a ninety-day moratorium on foreclosures).

259. In the United States, individual money-market deposit accounts were insured at up to $100,000 by the Federal Deposit Insurance Corporation, but money-market mutual funds were not. Tara Siegel Bernard, *Money Market Funds Enter a World of Surprising Risks*, N.Y. TIMES, Sept. 18, 2008, at C13. The Emergency Economic Stabilization Act of 2008 raised the limit to $250,000. Pub. L. No. 110-343, tit. 1, sec. 115, 122 Stat. 3766, 3780 (to be codified at 12 U.S.C. § 5225). Other countries took even more sweeping measures in the face of the financial crisis. In October 2008, the German federal government announced it would guarantee all private savings accounts to address growing fears of wholesale collapse of banks. Carter Dougherty et al., *Financial Crises Spread in Europe*, N.Y. TIMES, Oct. 6, 2008, at A1.

clear, the taking/taxing taxonomy, or any other delineation drawn in regard to governmental intervention in property rights, is not carved in stone as an essentialist conclusion of property law. It is part and parcel of the understanding of a given legal system about what constitutes the bundle of rights in certain resources and what elements of it deserve constitutional protection, so that a change in the paradigms of property law will undoubtedly affect the taxonomy of property law and takings law in particular. As we await future developments in property law, we thus must keep in mind both the innumerable ways in which the currently prevailing convictions shape the landscapes of property law, and how this entire array of doctrines could change upon a societal reconstruction of the institution of property.

Book Review

Some Moving Parts of Jurisprudence

BETWEEN AUTHORITY AND INTERPRETATION: ON THE THEORY OF LAW AND PRACTICAL REASON. By Joseph Raz. New York, New York: Oxford University Press Inc., 2009. 396 pages. $85.00.

Kevin Toh*

I. Introduction

Jane Stapleton related to me something a scholar once told her. According to that scholar, the main difference between British and Australian book reviews on the one hand and American book reviews on the other is that Britons and Australians try to engage with the overall general arguments of the books that they review, whereas Americans pick and choose among the topics and arguments in the books mainly to push their own intellectual agenda.

I have some good reasons for going unabashedly American in reviewing Joseph Raz's *Between Authority and Interpretation*.[1] First, the book is a collection of articles, and unlike the articles collected in an earlier volume, *The Authority of Law*,[2] the ones in the new collection are not meant to, and have not been newly revised to, offer cumulative arguments. Sure enough, there are many intricate connections among the themes and arguments contained in the articles of the new volume. But they do not build on each other the way those in *The Authority of Law* do. Second, I do not think it would be wise to try to assess, or even get a good grasp of, the arguments contained in the new book without having quite an extensive acquaintance with Raz's earlier works. As a matter of fact, many of the arguments and assumptions in these articles are at times articulated quite impressionistically. (The book could

* For their helpful and challenging questions and comments, I thank the members of the Fall 2009 Monday afternoon discussion group in law and philosophy at the University of Texas Law School. I am grateful also to Peter Cane for a set of very helpful written comments, Mitch Berman for spending almost a full afternoon discussing this Review with me and recommending numerous changes, many of which I took up, and David Sosa and Nicos Stavropoulos for helping me to understand anti-individualism better. Most of all, I am indebted to David Enoch for raising numerous questions and challenges in a set of very detailed written comments and also during discussions, and for generally playing the role of "the resident Raz." The contents of Parts V and VI would have been much different and inferior without his generous help. Correspondences with Brian Leiter and Scott Shapiro, ostensibly about their works, also helped me formulate some of the thoughts herein.

1. JOSEPH RAZ, BETWEEN AUTHORITY AND INTERPRETATION: ON THE THEORY OF LAW AND PRACTICAL REASON (2009).

2. JOSEPH RAZ, THE AUTHORITY OF LAW (1979).

not be recommended to a newcomer to Raz's writings about the nature of law.) In effect, the articles in the new collection are like a set of appliances with disparate functions. It turns out, perhaps unsurprisingly given their common manufacturer, that the mechanical parts in many of the appliances, the ones that make the appliances do the respective works that they do, are the same. But those mechanical parts are often hidden from the consumer's view.

Instead of cataloguing and testing each appliance, I want to take some concentrated looks at the common mechanical parts. In the following pages, I will take up and examine a number of arguments and assumptions that Raz relies on in these articles—the arguments and assumptions that I believe do the real work in running these appliances—even when Raz does not explicitly or clearly invoke such arguments and assumptions in the pages of the new volume. I will take advantage of some references to and quotations from some of his older works when doing so will enable me to get a better look at the mechanical parts. I have chosen this way of going about things partly because for quite some time now I have been puzzled at some of the key features of Raz's approach to thinking about the nature of law, and I also have been astonished at how many and how often people take those features for granted in their own legal theorizing. The main disadvantage of my approach will be that I will not be giving the reader a good idea of what is contained in some of the articles in *Between Authority and Interpretation*. In particular, for space considerations, I have reluctantly decided not to include a part on the topic of legal interpretation, the topic that occupies Raz for five of the fourteen articles in this book.[3] I believe that the central argument that Raz makes on this issue, contained in "Intention in Interpretation,"[4] is a not-so-unpredictable application of the considerations of authority, which I will spend much time discussing in the following pages. I hope to have an opportunity to discuss Raz's ideas about interpretation more directly elsewhere.

Finally, a warning and a request before I proceed. Raz's writings are not kind on their readers. Anyone who has given a serious attempt to engage with them is bound to feel quite unsure of his steps and bearings. I am no exception, and I am less than fully confident that the positions that I attribute to him and go on to criticize in the following pages are really his. My guess is that at least some knowledgeable readers will examine my regimentations of his views and come away thinking that I have regimented them out of recognition, and that for this reason my criticisms miss their target. I have one request to make to such a reader: Try to formulate Raz's views and try to do so in a way that makes them reactive to and critical of H.L.A. Hart's benchmark views, as Raz clearly intended them to be. My guess and hope is that the reader will come away from such an exercise thinking that Raz's

3. RAZ, *supra* note 1, chs. 9–13.

4. *Id.* ch. 11.

views are very much like the ones I attribute to him. With some amount of trepidation then, I proceed.

II. Preliminaries I: Theories of Law

I begin with a couple of preliminary Parts. At one point in the book, Raz rather modestly situates his own theorizing about the nature of law as follows:

> Hart was right to claim that a combination of what he called primary and secondary rules is essential to law. But he was wrong to intimate that therein lay the key to the study of jurisprudence. There are other essential properties to the law. I have argued, for example, that it is essential to the law that it claims to have legitimate, moral, authority, and that it is source-based, and that it claims to have peremptory force, etc. These claims do not conflict with Hart's view that necessarily the law combines primary and secondary rules, that it includes a rule of recognition, and that it is accepted by the legal officials, and normally also by many others in the population it applies to.[5]

I myself find it helpful to gain entry into Raz's views on the nature of law by trying to figure out where and why he departs from the benchmark views that Hart developed in *The Concept of Law*[6] and elsewhere.[7] And in the following pages, I will take up and examine each of the new elements that Raz names in the above passage, before taking up an additional one that is argued for in *Between Authority and Interpretation*. In this first preliminary Part, I will quickly outline the nature of legal philosophical theorizing as it has been pursued by Hart and those who have been influenced by him, and then outline Hart's own particular legal theory. In the next Part, the second of the two preliminary Parts, I will propose a way to translate or regiment Raz's personalized characterizations about law, clear examples of which are provided in the passage quoted above—e.g., about what *the law* claims.

In devising a philosophical theory of law, we are trying to explain, or account for, certain facts that obviously or uncontroversially seem to be features of communities of people governed or regulated by laws. Such facts together amount to the pretheoretical data, and our goal is to come up with a hypothesis about what law is that best explains those facts.

Here are some pretheoretical facts that have been important in the development of legal philosophy in the past half century:

(F1) Some laws are power-conferring rules (as opposed to duty-imposing rules).[8]

---

5. *Id.* at 97.

6. H.L.A. HART, THE CONCEPT OF LAW (Penelope A. Bulloch & Joseph Raz eds., 2d ed. 1994).

7. *See generally* H.L.A. HART, ESSAYS ON BENTHAM (1982); H.L.A. HART, ESSAYS IN JURISPRUDENCE AND PHILOSOPHY (1983).

8. HART, *supra* note 6, at 27–42.

(F2) Some customary rules are laws even before being recognized as such by legal officials.[9]

(F3) Some communities governed by laws are without legislators whose legal powers are unlimited.[10]

(F4) Some laws retain their binding force even after the deaths of the lawmakers who enacted those particular laws.[11]

(F5) A thriving legal system may exist in a community even in the absence of a law-identifying rule that is commonly or jointly accepted by the community's officials.[12]

(F6) Even in hard cases—i.e., those cases in which decisions are not clearly determined by clearly valid laws—judges often consider their decisions as completely constrained by preexisting laws.[13]

(F7) Laws create genuine reasons (and even duties or obligations) to behave as the laws say.[14]

(F8) It is possible for unconfused people to have legal disagreements (as opposed to disagreements about what the law should be) that persist despite their agreements on all factual issues.[15]

(F9) There can be legal systems whose laws are so immoral or so evil that no set of moral principles could justify the laws of such legal systems.[16]

I shall be referring to some of these facts now and then in the following pages. Hart famously argued in the early chapters of *The Concept of Law* that the theories, proposed by Jeremy Bentham and John Austin, that characterize laws as commands of sovereigns fail to explain satisfactorily facts like (F1)–(F4).[17] Ronald Dworkin argued in turn that the alternative

---

9. *Id.* at 44–49.

10. *Id.* at 66–78.

11. *Id.* at 61–66.

12. *See generally* RONALD DWORKIN, LAW'S EMPIRE chs. 1, 4 (1986) [hereinafter DWORKIN, LAW'S EMPIRE]; RONALD DWORKIN, *The Model of Rules I*, *in* TAKING RIGHTS SERIOUSLY 14 (1977) [hereinafter DWORKIN, *Rules I*]; RONALD DWORKIN, *The Model of Rules II*, *in* TAKING RIGHTS SERIOUSLY, *supra*, at 46 [hereinafter DWORKIN, *Rules II*]; *cf.* HART, *supra* note 6, ch. 10.

13. *See generally* RONALD DWORKIN, *Hard Cases*, *in* TAKING RIGHTS SERIOUSLY, *supra* note 12, at 81 [hereinafter DWORKIN, *Hard Cases*]; DWORKIN, *Rules I*, *supra* note 12; DWORKIN, *Rules II*, *supra* note 12.

14. DWORKIN, LAW'S EMPIRE, *supra* note 12, at 108–13; DWORKIN, *Rules II*, *supra* note 12, at 48–49; JOSEPH RAZ, PRACTICAL REASON AND NORMS 56–58 (2d ed. 1990).

15. DWORKIN, LAW'S EMPIRE, *supra* note 12, ch. 1.

16. H.L.A. HART, *Legal Duty and Obligation*, *in* ESSAYS ON BENTHAM, *supra* note 7, at 127, 150 [hereinafter HART, *Legal Duty*]; Ronald Dworkin, *The Law of Slave Catchers*, TIMES LITERARY SUPP. (London), Dec. 5, 1975, at 1437; H.L.A. Hart, *Comment*, *in* ISSUES IN CONTEMPORARY LEGAL PHILOSOPHY 35, 40–42 (Ruth Gavison ed., 1987) [hereinafter Hart, *Comment*].

17. HART, *supra* note 6, chs. 2–4.

theory that Hart developed fails to explain facts like (F5)–(F8).[18] And Hart countered that Dworkin's legal theory fails to account for (F9).[19]

Of course, a proponent of a legal theory can dispute that certain putative facts are genuine data that need to be accounted for, as Hart did with respect to (F6) and (F7).[20] And a theorist can also try to show that his theory has resources to explain the proffered facts. That is what Hans Kelsen did, on Bentham and Austin's behalf, with respect to (F1) by characterizing power-conferring laws as fragments of duty-imposing laws addressed to legal officials.[21] Like the Ptolemaic attempts to explain planetary trajectories by positing epicycles, success of any such attempts depends very much on how cumbersome they get, and also on what alternative explanations are available.

Hart argued that most (if not all)[22] of the pretheoretical data that need to be explained, including (F1)–(F4), can be explained by conceiving existence of laws in the following way. Laws (or *legal* rules) come in hierarchically structured packages, with some specific kinds of second-order rules, or rules governing the operation of the rules within the package. These second-order rules—which Hart calls *secondary rules* to distinguish them from *primary rules* that directly govern conduct—include: the *rules of change*, which regulate any modification in the rules of the system; the *rules of adjudication*, which regulate settling of disputes about the content and application of the rules of the system; and the *rule of recognition*, which regulates the identification of the rules of the system. A legal system exists or prevails in a community if some powerful subset of the members of the community—call them the "officials" of that community—*accept* the secondary rules (in the sense to be specified presently) and follow them as the result, and the rest of the community at least follow the rules that are valid according to the rule of recognition prevailing in that community. A person accepts a rule, or takes an "internal point of view" toward that rule, according to Hart, when he believes there to be reasons to follow it.[23] Such acceptance is constituted by the person's dispositions to regulate his own conduct in accordance with the rule, to justify his own and others' conduct by appeals to that rule, and to criticize his own and others' deviance by appeals to that rule.[24] Presumably, to treat a rule of recognition as reason giving is to treat

18. *See* sources cited *supra* notes 12–15.

19. *See* sources cited *supra* note 16.

20. H.L.A. HART, *1776–1976: Law in the Perspective of Philosophy*, *in* ESSAYS IN JURISPRUDENCE AND PHILOSOPHY, *supra* note 7, at 145, 156–58; HART, *Legal Duty*, *supra* note 16, at 161; Hart, *Comment*, *supra* note 16, at 35–40.

21. HANS KELSEN, GENERAL THEORY OF LAW AND STATE 143–45 (Anders Wedberg trans., 1945).

22. See Hart's qualification on this point in HART, *supra* note 6, at 99.

23. *Id.* at 56–57, 89–91, 114–16.

24. Raz misleadingly characterizes Hart's theory when he at one point says that Hart gave an account "from the internal point of view." RAZ, *supra* note 1, at 93. Notice that the account I

generally the rules that are valid according to that rule of recognition as reason giving. Hart does not go into how thoroughly a person must be disposed to treat each valid law as reason giving to count as accepting a rule of recognition, but we can assume that some general tendency must be there. (Notice the difference between on the one hand the hypothesis of rule-acceptance that Hart employs to explain facts like (F1)–(F4), and on the other (F7).)

Those are the facts that would amount to the existence of a legal system in a community of people, according to Hart. A law with a particular content exists or prevails in a community if a legal system exists or prevails in that community and a rule with that particular content is valid according to the rule of recognition that is accepted by the officials of that community. But Hart was at pains to point out that not all legal statements are aimed at describing or representing such facts. Some legal statements, which Hart called "external legal statements," do actually purport to describe such facts.[25] These are analogous to the statements of sociologists or anthropologists whose goal is to describe the mores of a community without thereby expressing their own commitments to those mores. Other legal statements are more akin to the moral statements that speakers make in their morally committed hours. In uttering such "internal legal statements," as Hart called them, speakers do not *describe* or *represent* people's acceptances of a set of rules that constitute a legal system, but instead they *express* their own acceptances of such rules.[26] In other words, they display their commitments to be guided by the relevant rules, and to assess their own and others' conduct by appeals to such rules. Whereas external legal statements are descriptive or factual statements, internal legal statements are normative statements. In arguing that the existence of a legal system consists partly of at least some people in the relevant community accepting some secondary rules, Hart was arguing that where a legal system exists, at least some people are disposed to regulate their own conduct and to try to influence others' by way of the normative assessments that internal legal statements express.

This theory of law forms arguably the most important part of the backdrop for Raz's own theorizing about the nature of law. Moreover, I believe, we can domesticate many of the sometimes striking and surprising claims that Raz makes by determining how exactly such claims mark Raz's departures from the benchmarks provided by Hart's legal theory.

---

outline in the text is one that tries to depict or describe accurately the workings of legal systems, including those having to do with the attitudes and behavior of the participants who take the internal point of view towards their laws. Hart tried to give an account that depicts accurately people's internal point of view; but the account itself is not from the internal point of view.

25. HART, *supra* note 6, at 102–03.

26. *Id.*

III. Preliminaries II: Translating the Law's Claims

In the passage that I quoted at the beginning of Part II, Raz says that one of the important necessary features of law that Hart failed to notice is that the law "claims to have legitimate, moral, authority."[27] One characteristic feature of Raz's writings over the years that many have found suspect is his tendency to personalize the law in just this way. Raz explicitly addresses such qualms at a couple of places in the new volume. He says that he finds nothing amiss in such personalization, and goes on to say:

> The law's actions, expectations, and intentions are its in virtue of the actions, expectations, and intentions of the people who hold legal office according to law, that is we know when and how the actions, intentions, and attitudes of judges, legislators, and other legal officials, when acting as legal officials, are to be seen as the actions, intentions, and expectations of the law.[28]

This is consistent with what Raz says at one point in *The Morality of Freedom*[29] when he explains the usefulness of availing ourselves of the personalized claims as follows: "The law requires, permits and claims what the organs of government, acting lawfully, and in particular the courts, say that it does."[30] And the officials of a community who are presumed to accept the law are also then in turn presumed to accept the law's claims.[31]

These explanations, along with some other things that Raz says in the new volume and in some of his older works, I believe, warrant the following reconstruction of what Raz means when he says that law claims to have legitimate, moral, authority. Legal officials, when they are acting as officials and are enacting, enunciating, and interpreting the laws of their legal system, treat the laws of their legal system as legitimately, and morally, authoritative. Such treatment consists of the particular kind of normative attitudes that officials take towards the laws of their legal system, and the particular kind of normative force that the normative statements meant to enunciate and interpret the laws of their legal system purport to have. In the preceding Part, I described Hart as espousing the view that where a community is governed by laws, its officials accept the rules that make up their legal system, and the internal legal statements they make express such acceptance. And such acceptance amounts to treating the relevant rules as furnishing reasons to behave as those rules say. If this part of Hart's legal theory could be called "Acceptance Rationalism," then Raz can be seen as espousing a particular version of it that could be called "Acceptance Moralism." According to Raz, when a community is governed by laws, its officials treat the rules that make

27. RAZ, *supra* note 24, at 97.
28. RAZ, *supra* note 1, at 38, 93.
29. JOSEPH RAZ, THE MORALITY OF FREEDOM (1986).
30. *Id.* at 70.
31. RAZ, *supra* note 1, at 277.

up their legal system as furnishing *moral* reasons to behave as those rules say, and the legal statements they make express this particular kind of normative attitude towards the rules of the legal system.

Actually, in invoking the notion of authority, Raz is offering an even more specific version of Acceptance Moralism.[32] To treat a rule as legitimately authoritative is to treat it as furnishing what could be called "protected reasons." In addressing to someone a normative statement that purports to furnish such protected reasons, the addressor intends to furnish reasons for the addressee to act in some ways and also second-order reasons that exclude or preempt contrary first-order reasons. In saying that law claims to have legitimate, moral, authority then, Raz seems to mean that, where a community is governed by laws, the community's officials treat laws as furnishing protected moral reasons to behave as the laws say. And such treatment consists of the normative attitudes they take towards the laws of their legal system and the normative statements they make in enunciating and interpreting the laws. Actually, there are indications that Raz means something even more specific than that, and I will discuss these specificities in due course.

There is an all-important wrinkle in Raz's view that I have not yet pointed out. Raz is *almost*, but *not quite*, of the view that where laws exist legal officials of the relevant community necessarily take laws as furnishing protected moral reasons. Instead, he is of the view that legal officials either genuinely take *or only pretend to take* laws as furnishing protected moral reasons. They behave *as if* they take the laws as furnishing such reasons, and their internal legal statements often state the existence of protected moral reasons only in the pretended or simulated sense. These last kinds of legal statements are what Raz calls "detached legal statements," to be distinguished from "committed legal statements." The fact that according to Raz's view legal officials can take the laws of their community as real laws without thereby necessarily thinking that such laws furnish protected moral reasons is one of the features of his theory that enables Raz to claim to be a legal positivist.[33]

In sum, whereas Hart argues that in a community governed by laws, at least some members of the community—especially those who are in charge of the workings of the legal system—treat the rules that make up the legal system as furnishing reasons for action, Raz is of the view that in such a community the same kind of members of the community treat, or pretend to treat, the rules of the legal system as furnishing protected moral reasons to do as those rules say. That is what I take Raz to mean when he says that the law claims legitimate, moral, authority. I want to assess this aspect of Raz's thinking about the nature of law. I shall do so by breaking Raz's view into

32. The necessary textual support for my characterizations in this paragraph will be given in Part V, *infra*.

33. RAZ, *supra* note 1, at 100, 113–14.

two components and assessing them in separate steps: First, in Part IV, I address his view that legal officials take laws as furnishing moral reasons. Second, in Part V, I address his view that legal officials take laws as furnishing protected reasons (or protected reasons of some very specific variety that I will specify in due course). To facilitate my exposition, I would also like to dispense with the above-named wrinkle in the following pages, unless I specifically flag otherwise. When I talk about officials' and others' treatment of laws as furnishing reasons for action in outlining Raz's view, I should be taken as talking about their *sincere or pretended* treatment of laws as furnishing reasons for action.

IV. Acceptance Moralism

Hart adopted Acceptance Rationalism in order to explain facts like (F1)–(F4), listed in Part II above, which were not readily explained by command theories of law of the sort proposed by Bentham and Austin. In order to explain the existence of customary laws, laws that confer powers (as opposed to those imposing duties), laws that limit the powers of legislators, etc., Hart argued that we need to characterize the members of a community governed by laws as constrained in their behavior and practical thoughts by their allegiance to the rules themselves, rather than by their habit of obeying the persons who impose the rules.

I am not sure that Hart actually needed to adopt something as strong as Acceptance Rationalism to account for facts like (F1)–(F4). It seems to me that a compelling case could be made that so long as a sufficient number of the members of a community or its officials take themselves to have strong enough extraneous or extralegal reasons to behave as their laws demand, a legal system, very much with the features explanatory of (F1)–(F4), could prevail in that community without the members or their officials treating the laws themselves as furnishing reasons to act as the laws say. I do think, however, that the need to account for facts like (F6) and (F8) gives Hart compelling reasons to accept Acceptance Rationalism—which is a bit ironic in retrospect, given that it is Dworkin's view that the inadequacy of Hart's legal theory is revealed in its inability to explain facts like (F6) and (F8). (More will be said on this issue in Part VII below.)

What is important here is that Hart adopted Acceptance Rationalism, and in this Raz has followed him, for explanatory purposes. A question that naturally arises is what the explanatory payoffs are for adopting Acceptance Moralism, a particular version of Acceptance Rationalism that Raz has adopted and Hart has rejected. Raz seems to think that Acceptance Moralism is necessary to explain something like the following:

> Many people treat the laws of their community as furnishing reasons, and even duties or obligations, to act as the laws say, and they consider such reasons and duties/obligations as applicable to or binding on individual persons—even in some central areas of life—

> whether or not those individuals are personally committed to the relevant laws or the legal system as a whole.

This may seem a mere restatement of Acceptance Rationalism, albeit one that explicitly characterizes the reasons supposedly furnished by laws as categorical reasons, and that explicitly mentions that some of the reasons may amount to duties or obligations. But Raz seems to think that once the thesis of Acceptance Rationalism is put explicitly like this, there is no way to avoid ending up with Acceptance Moralism. That was what he asserted in some of his older works,[34] and he repeats them in the chapter called "Incorporation by Law": "[W]hatever else we grace with the title 'moral,' principles that impose, or give people power to impose on others, duties affecting central areas of life are moral principles. That much about the nature of morality is clear."[35] Raz's reasoning then seems to be that any plausible explanation of the nature of law must explain how people can believe that the laws of their community can impose categorical reasons and even duties to behave as the laws say,[36] and the only way to discharge that explanatory burden is to characterize people as treating legal rules as moral rules—that is, as furnishing moral reasons.

It appears here that there is a reductionist assumption running in Raz's line of thinking. He seems to think that all rules that furnish categorical reasons and obligations are moral rules. But why should we think this exactly? Why should we not think that there may be different types of rules that can be taken to give rise to categorical reasons and obligations? I do not think that Raz has ever really given an adequate explanation of his position on this point,[37] and as far as I can see, there is nothing new in the new volume that enhances the credibility of his position on this issue. Raz may be of the view that normative commitments on certain subject matters, or deontic commitments in general, are necessarily moral commitments. But in this dialectical context, such positions, without further arguments in support, would be question begging.

Richard Holton also argues for Acceptance Moralism and seems to offer a sharp articulation of what may be motivating Raz as well.[38] Holton takes

---

34. Joseph Raz, *Hart on Moral Rights and Legal Duties*, 4 OXFORD J. LEGAL STUD. 123, 130–31 (1984); Joseph Raz, *The Purity of the Pure Theory*, 35 REVUE INTERNATIONALE DE PHILOSOPHIE [R.I. PHIL.], 448, 454–55 (1981) (Belg.) [hereinafter Raz, *Purity*].

35. RAZ, *supra* note 1, at 188; *cf. id.* at 278 ("Why must those involved in making or applying the law believe in its moral acceptability? Because the law purports to determine or reflect (moral) rights and duties of its subjects.").

36. Christine Korsgaard places an analogous explanatory burden on any plausible explanation of the nature of morality. CHRISTINE M. KORSGAARD, THE SOURCES OF NORMATIVITY 13 (1996).

37. I have elsewhere discussed Raz's earlier arguments. Kevin Toh, *Raz on Detachment, Acceptance and Describability*, 27 OXFORD J. LEGAL STUD. 403, 414–20 (2007) (criticizing Raz's argument that the acceptance needed for an adequate portrayal of internal legal statements is necessarily a moral endorsement).

38. Richard Holton, *Positivism and the Internal Point of View*, 17 LAW & PHIL. 597, 600–06 (1998).

note that according to Hart a person can accept a law, in the sense of thinking that there are reasons to act as that law says, while thinking that morally speaking he ought not to accept that law. Holton then wonders: "How can one both think that one morally ought not accept the law, and yet still think that its demands are justified?"[39] But there is no mystery here. People like Friedrich Nietzsche and Bernard Williams had no problem noticing that morality points us in one direction, while arguing that we ought to think and act differently;[40] and we have no problem understanding what they meant. Nor do we have to resort to such extreme examples. To borrow Raz's own example from another article, a woman may recognize the moral reasons in favor of continuing her job of helping the homeless but still also recognize and go on to act on the reasons to give up her job in order to have a baby.[41] It may be thought here that this example does not really counter Holton's point because, although there are moral reasons for the woman to keep her job helping the homeless, she is not morally obligated to do so, and she is morally permitted to quit her job and have a baby. Holton's point may be that a person cannot think that he has reasons or obligations to do what he thinks he has moral obligation not to do.[42] But that seems to assume that a person can be thought to accept only one consistent set of obligations, or at least that a person can be thought to accept a set of moral rules only if he takes the obligations such rules issue as always overriding other considerations. Should not our moral psychology allow for greater complexity than what such an assumption entails?[43]

It seems that a sensible approach on this issue would be to come up with the best theory of morality and the best theory of law we can muster,[44] and then to see whether legal rules are necessarily moral rules—that is, whether our treatment of the rules that make up legal systems necessarily amounts to

---

39. *Id.* at 603. Although I disagree with Holton's Acceptance Moralism, I completely agree with him that Hart's exposition of the considerations against that view is marred by his conflation of several senses of "reason." For Holton's extremely helpful clarification, see *id.* at 603–04.

40. *See generally* FRIEDRICH NIETZSCHE, ON THE GENEALOGY OF MORALITY (Keith Ansell-Pearson ed., Carol Diethe trans., Cambridge Univ. Press 2006) (1887); BERNARD WILLIAMS, ETHICS AND THE LIMITS OF PHILOSOPHY (1985).

41. Joseph Raz, *On the Moral Point of View*, *in* REASON, ETHICS, AND SOCIETY: THEMES FROM KURT BAIER, WITH HIS RESPONSES 58 (J.B. Schneewind ed., 1996), *reprinted in* JOSEPH RAZ, ENGAGING REASON 247, 268–69 (1999).

42. In his discussion of the example, Raz himself resists imputing such decisive importance to the distinction between moral obligations on the one hand and moral reasons in general on the other to explain the seeming justifiability of the woman's decision. *Id.* at 269–70.

43. According to one plausible interpretation of the oft-mentioned example of Huckleberry Finn, Huck thought he was morally obligated to turn Jim in, yet also thought he had more compelling contrary reasons, upon which he eventually acted. Obviously, what is important here is the plausibility of the thought that someone could find himself in such a situation, not the faithfulness of my interpretation to Mark Twain's novel. *See generally* MARK TWAIN, ADVENTURES OF HUCKLEBERRY FINN (Oxford Univ. Press 1999) (1884).

44. To clarify, the theories I am talking about here are metatheories—i.e., theories about the nature of law and about the nature of morality—not first-order normative theories about how to act or what to pursue.

our treatment of them as furnishing moral reasons. In chapters 5 and 6 of *The Concept of Law*, Hart develops what he considers the best legal theory in the shape I outlined in Part II. And in chapter 8 of that book, for the purpose of assessing natural law theories, Hart specifies the rudiments of a theory of morality. The four features of moral rules that he names there—purported importance, immunity from deliberate change, voluntary nature of the offenses they address, and specific types of emotional pressures that they involve—seem to distinguish moral rules from legal rules.[45] But whatever the eventual fortunes of a theory of morality along such lines, what is important is that Raz needs a worked-out theory to ground his claim that legal rules are necessarily moral rules. In the absence of such a theory, it seems, the fact that legal rules can be deliberately modified whereas moral rules in general cannot, as well as the fact that violations of moral rules are thought to warrant necessarily the emotion of guilt on the part of the violator and that of resentment on others' whereas violations of legal rules do not, seem enough to create a fairly strong presumption against Acceptance Moralism.

In *On the Moral Point of View* and a number of related writings, Raz has argued against the idea that moral reasons form a distinct set of reasons that differ in a theoretically significant way from those involved in other areas of practical thought.[46] Perhaps these arguments amount to a response to a demand like mine for a worked-out theory of morality. Perhaps Raz can resist that demand by showing that such a theory would be poorly motivated. But such a move, with respect to the issue of Acceptance Moralism, would cut both ways. At least on first blush, there seems a difficulty in arguing both for Acceptance Moralism as a significant thesis and for the rejection of the conception of moral reasons as a significantly distinct set of reasons. Raz allows that there could be nonproblematic context-sensitive deployments of the distinction between moral reasons and other kinds of reasons,[47] and this is what enables him to formulate the above-described example. And perhaps in arguing for Acceptance Moralism, Raz intends to make a context-sensitive distinction of the sort that he finds unobjectionable.[48] But this suggestive

45. HART, *supra* note 6, at 168, 173–80. For similar conceptions of morality, see generally R.B. BRANDT, A THEORY OF THE GOOD AND THE RIGHT (1979); ALLAN GIBBARD, WISE CHOICES, APT FEELINGS (1990); and John Skorupski, *The Definition of Morality*, *in* ETHICS: ROYAL INSTITUTE OF PHILOSOPHY SUPPLEMENT 35, 121 (A. Phillips Griffiths ed., 1993), *reprinted in* JOHN SKORUPSKI, ETHICAL EXPLORATIONS 137 (1999).

46. RAZ, *supra* note 29, ch. 12; Raz, *supra* note 41, at 247. *See generally* Joseph Raz, *The Amoralist*, *in* ETHICS AND PRACTICAL REASON 369 (Garrett Cullity & Berys Gaut eds., 1997), *reprinted in* ENGAGING REASON, *supra* note 41, at 273; Joseph Raz, *The Central Conflict: Morality and Self-interest*, *in* WELL-BEING AND MORALITY: ESSAYS IN HONOUR OF JAMES GRIFFIN 209 (Roger Crisp & Brad Hooker eds., 2000), *reprinted in* ENGAGING REASON, *supra* note 41, at 303.

47. Raz, *supra* note 41, at 250–51.

48. Raz may be relying on a similarly context-sensitive distinction when he says that judgments about the functions of law are based on "evaluative," though nonmoral, considerations. Joseph Raz, *Authority, Law, and Morality*, 68 THE MONIST 295 (1985), *reprinted in* JOSEPH RAZ, ETHICS IN THE PUBLIC DOMAIN: ESSAYS IN THE MORALITY OF LAW AND POLITICS 194, 204–14 (1994) [hereinafter Raz, *Authority*]; Joseph Raz, *The Problem About the Nature of Law*, *in* 3

idea of context-sensitive distinctions is not sufficiently developed or even articulated by Raz for us to assess its viability. More important, Raz has insufficient bases for thinking that there cannot be a theoretically robust distinction between moral and other practical considerations. The conceptions of morality that he enumerates in the relevant writings, for the purpose of rejecting them thereafter, are quite specific and for this reason severely limit the implications of Raz's conclusions. In *On the Moral Point of View*, for example, Raz conceives moral values and reasons as aggregations of relational values and reasons.[49] In *The Amoralist*, moral reasons are conceived as reasons stemming from people's status as values in themselves.[50] Because of the very specific nature of Raz's conceptions of morality in such writings, even if Raz's arguments meant to cast doubt on the theoretically distinct status of morality as thus conceived were cogent and wholly successful,[51] that would not be enough to cast a *general* doubt on attempts by Hart and other like-minded philosophers in specifying the features of moral thought and practice which distinguish them from other kinds of practical thought and practice in a theoretically robust way. And if the distinction between morality and law, made partly by way of those features, were explanatorily useful, then that should be enough to deflect Raz's skepticism.

In sum, I find the considerations to which Raz can be seen to have appealed to argue for Acceptance Moralism quite insufficient for their purposes.[52]

CONTEMPORARY PHILOSOPHY: A NEW SURVEY 107 (Guttorm Fløistad ed., 1982), *reprinted in* ETHICS IN THE PUBLIC DOMAIN: ESSAYS IN THE MORALITY OF LAW AND POLITICS, *supra*, at 179, 193 [hereinafter Raz, *Problem*].

49. Raz, *supra* note 41, at 257.

50. JOSEPH RAZ, *The Amoralist*, *in* ENGAGING REASON, *supra* note 41, at 274.

51. This is a big "if." In particular, at key junctures in his arguments, Raz relies on a quite controversial conception of practical reason. According to this conception, practical reasons are necessarily considerations in light of which the option an agent pursues is seen by him as valuable or good. *See generally* JOSEPH RAZ, *Agency, Reason, and the Good*, *in* ENGAGING REASON, *supra* note 41, at 22; Joseph Raz, *Incommensurability and Agency*, *in* INCOMMENSURABILITY, INCOMPARABILITY, AND PRACTICAL REASON 110 (Ruth Chang ed., 1997), *reprinted in* ENGAGING REASON, *supra* note 41, at 46; Joseph Raz, *When We Are Ourselves*, *in* 71 PROC. ARISTOTELIAN SOC'Y 211 (1997), *reprinted in* ENGAGING REASON, *supra* note 41, at 5. This is not the place to try to evaluate this conception of practical reason. But it should be noted that many philosophers reject it. *See generally* Michael Stocker, *Desiring the Bad*, 76 J. PHIL. 738 (1979); Michael Stocker, *Raz on the Intelligibility of Bad Acts*, *in* REASON AND VALUE: THEMES FROM THE MORAL PHILOSOPHY OF JOSEPH RAZ 303 (R. Jay Wallace et al. eds., 2004); J. David Velleman, *The Guise of the Good*, *in* 26 NOÛS 3 (1992), *reprinted in* J. DAVID VELLEMAN, THE POSSIBILITY OF PRACTICAL REASON 99 (1999).

52. There is one particular sense of *moral* that Raz could be employing in arguing for Acceptance Moralism that I have intentionally avoided considering. Following Kelsen, Raz has taken the position that in order to characterize in an adequate way people's normative attitudes, we (the theorists) have to deploy first-order normative statements ourselves. JOSEPH RAZ, *Legal Validity*, *in* THE AUTHORITY OF LAW, *supra* note 2, at 146 [hereinafter RAZ, *Legal Validity*]; Raz, *Purity*, *supra* note 34. And Raz has followed Kelsen also in thinking that any such first-order normative statements are moral statements. Here, the contrast between "moral" and "nonmoral" is the contrast between first-order normative and metanormative. There are many moments in

V. Acceptance Protectionism and Acceptance Autonomism

At the heart of Raz's thinking about the authoritative nature of law is the notion of exclusionary reason. In a 1979 article, Raz says: "The law's claim to legitimate authority is not merely a claim that legal rules are reasons. It includes the claim that they are exclusionary reasons for disregarding reasons for non-conformity."[53] Raz introduced the notion of exclusionary reason in his seminal book *Practical Reason and Norms*.[54] In our practical reasoning, we do not treat all reasons for or against each of our available options as of a same kind, so that our job is merely that of balancing all known applicable reasons. Instead, some of the reasons that we respond to are those that call for or against our acting on other reasons. Those among these "second-order reasons" that call for the agent to refrain from acting on certain first-order reasons, and thereby *exclude* those first-order reasons from the agent's deliberative consideration, are what Raz calls "exclusionary reasons."[55]

According to Raz's conception of the law as "claiming legitimate authority" then, the officials of a community governed by laws treat their laws as authoritative. That means that they accept the rules that make up their legal system, in the sense not only of treating those rules as generating reasons to behave as those rules say but also treating those rules as generating exclusionary reasons to refrain from acting on reasons against acting as those rules say. And the legal statements, by way of which the officials enunciate and interpret laws, purport to give both first-order reasons and exclusionary reasons that exclude contrary first-order reasons. Raz uses the term "protected reason" to refer to this combination of a first-order reason and an attendant exclusionary reason that excludes contrary first-order reasons.[56] "Acceptance Protectionism" is my not-so-lovely neologism to refer to the aspect of Raz's view of the nature of law that I have just outlined. In stating that law claims "legitimate, moral, authority," Raz is endorsing both Acceptance Moralism and Acceptance Protectionism.

Between Authority and Interpretation where this conception of morality seems to surface. But I have intentionally avoided discussing this aspect of Raz's thinking for a couple of reasons. First, Raz's reasoning for adopting Kelsen's positions on these issues strikes me as quite inadequate, as I have tried to show elsewhere. Toh, *supra* note 37, at 421–26. And there is nothing new in the new volume that adds credibility to the Kelsenian positions. Second, given this inadequacy, we have reasons to figure out whether what Raz says about the moral nature of legal rules can be construed in alternative ways. And further given that Raz says many things that do not tie him to the Kelsenian lines just described, we should follow up on those leads. That is what I have tried to do in Part IV *supra*.

53. JOSEPH RAZ, *The Claims of Law*, *in* THE AUTHORITY OF LAW, *supra* note 2, at 28, 30; *see also* JOHN FINNIS, NATURAL LAW AND NATURAL RIGHTS 233–34 (1980).

54. RAZ, *supra* note 14, § 1.2.

55. *Id.*

56. RAZ, *supra* note 1, at 216; RAZ, *supra* note 14, at 191.

So, why should we accept Acceptance Protectionism? In *Practical Reasons and Norms*, Raz argued that the way mandatory rules[57] figure in our practical thinking should be modeled in terms of exclusionary reasons.[58] Mandatory rules are those that we think of as imposing obligations or duties, and perhaps more widely those that are formulated in terms of "must" and "ought." Raz proposes to conceive each such rule as furnishing a protected reason—i.e., a combination of a first-order reason to do something and an exclusionary reason that demands refraining from acting on contrary first-order reasons. This quite plausible conception of mandatory rules has obvious explanatory payoffs for theorizing about the nature of law. Undeniably, many legal rules purport to impose obligations or duties, and Raz's analysis provides us a compelling way of explaining how such laws figure in our practical thinking. Hart, for one, in one of his last articles, endorses and makes use of Raz's analysis in explaining legal obligations.[59] And it is not just duties or obligations that can be explained in terms of exclusionary reasons. Raz has also offered quite compelling arguments to show that significant classes of permissions and normative powers should be explained in terms of exclusionary reasons.[60] What he has shown, in my opinion, is that at least a significant portion of our practical thinking involving deontic concepts, including legal thinking, can be explained in terms of exclusionary reasons.

What we have in the resulting Acceptance Protectionism is a relatively thin conception of what it is to treat some rules as authoritative, and what I consider a well-motivated rationale for it based on its explanatory payoffs. We find in the pages of *Between Authority and Interpretation*, and in Raz's recent works more generally, some thicker or more inflated conceptions of the attitudes that officials and others take towards their laws in treating them as authoritative. I want to spend the rest of this Part outlining and assessing a number of such thicker conceptions. All of the ones that I will take up are variations on the idea that in treating some rules as authoritative, people treat them as furnishing not merely protected reasons, but *content-independent* protected reasons, in the sense to be specified presently. Let me call such a conception of the attitudes of acceptance of laws "Acceptance Autonomism." There is some evidence for thinking that Raz is of the view that, in treating some rules as authoritative in the sense of treating them as furnishing

57. Raz uses "mandatory norms." RAZ, *supra* note 14, at 49. By "norm," Raz intends to refer to "rules" and "principles" that are "mandatory" in the sense specified in the text. *Id.* at 9, 49. My guess is that Raz saw a need to mention both rules and principles given the significance that Dworkin attached to the existence of principles in his early criticism of Hart. *See generally* DWORKIN, *Rules I*, *supra* note 12. Like Hart, and unlike Dworkin and Raz, I use the term "rule" to refer to all kinds of practical standards, including what Dworkin calls principles.

58. RAZ, *supra* note 14, at 73–84.

59. *See generally* H.L.A. HART, *Commands and Authoritative Legal Reasons*, *in* ESSAYS ON BENTHAM, *supra* note 7, at 243.

60. RAZ, *supra* note 14, at 85–106.

content-independent protected reasons, people have to posit or presuppose the existence of some rule makers or "authors" of the rules. I will first examine the strong and weak versions of such "Authorial Acceptance Autonomism" before taking up the strong and weak versions of "Non-authorial Acceptance Autonomism." I am chagrined by the proliferation of such unattractive labels, but I adopt them because they will much facilitate my subsequent exposition.[61]

It is a striking fact about Raz's discussions of the authoritative nature of law that he (almost) invariably speaks of laws as issued by some authoritative persons or institutions. And that is certainly the case in the pages of *Between Authority and Interpretation*. It is not clear to me whether in such instances Raz is displaying a commitment to the view that in treating some set of rules as authoritative, a group of people have to see such rules as issued by some authoritative persons. Raz seems to disavow such a commitment in the 1979 article *The Claims of Law*, when he says:

> Authority was analysed in the context of a person in authority and his authoritative utterances. Such an analysis could in principle apply to a legislator and his acts of enactment. But not all law is enacted. Customary rules can be legally binding. Can they be authoritative despite the fact that they are not issued by authority? It is possible to talk directly of the authority of the law itself. A person's authority was explained by reference to his utterances: he has authority if his utterances are protected reasons for action, i.e. reasons for taking action they indicate and for disregarding (certain) conflicting considerations. The law has authority if the existence of a law requiring a certain action is a protected reason for performing that action . . . .[62]

Regardless of this earlier disavowal, however, it may be that Raz now subscribes to Authorial Acceptance Autonomism. That is what Andrei Marmor has argued in a recent article.[63] In particular, Marmor argues that the positing of authoritative figures (or "authors," as he puts it) is necessitated by the fact that people who treat some rules as authoritative treat them as generating content-independent protected reasons.[64] Some fact is supposed to constitute a content-independent reason to perform some action

---

61. I am also unhappy about the fact that nothing snappier than "Acceptance Autonomism" seems available as a label for the relevant conception of rule-acceptances.

62. RAZ, *supra* note 53, at 29.

63. *See generally* Andrei Marmor, *Authorities and Persons*, 1 LEGAL THEORY 337 (1995), *revised and reprinted in* ANDREI MARMOR, POSITIVE LAW AND OBJECTIVE VALUES 89 (2001). And that is also what David Enoch has argued in conversations. Thanks to Enoch for pointing me to Marmor's article. In his article, Marmor does not mention the above-quoted passage from *The Claims of Law*, and more generally does not mention any change in Raz's position on this issue. In personal communication, Les Green told me that he too suspects that Raz is committed to an authorial conception of authority, but Green also expressed a puzzlement as to how such a conception can be squared with the goal of explaining customary laws.

64. MARMOR, *supra* note 63, at 97–98.

if that fact is a reason for performing such an action, and that fact is not directly related to the nature or character of the called-for action.[65] According to Raz and Marmor, in treating certain rules as authoritative, people treat the fact that an authoritative figure has issued or instituted those rules, instead of the facts having directly to do with the nature or character of the actions those rules call for, as the reason to adhere to those rules.[66]

It seems undeniable that people often treat the laws of their community as furnishing content-independent reasons. The fact, noted in the preceding Part, that laws are amenable to modification at will seems to imply that at least in some cases people will treat their laws as furnishing content-independent reasons. The question is whether such treatment really requires, as a conceptual matter, authoritative figures as the authors of those laws. Much here depends on how strongly or literally we interpret the alleged requirement of authoritative figures. If we were to require persons or institutions that deliberately make laws, then we would be ill-equipped to explain people's treatment of customary laws as authoritative. That was what motivated Raz earlier to disavow the authorial conception of authority. Even Marmor admits that customary laws present a problem for thinking that authoritative figures are necessary.[67]

Marmor, however, seems amenable to adopting a weaker version of Authorial Acceptance Autonomism when he highlights what he deems "the crucially important point" that "customs do reflect the view of the bulk of the community about how they themselves ought to behave."[68] This statement seems intended to reflect a view of customary laws articulated earlier by Raz.[69] Also relevant here is the fact that when Raz appeals to the considerations having to do with authority to argue for what he calls the "Authoritative Intention Thesis" in the chapter entitled "Intention in Interpretation"[70]—according to which, the intentions of lawmakers should be treated as the controlling considerations in interpretations of laws—he explicitly limits the scope of that thesis to "deliberately made" laws, and leaves out "practice-based" laws.[71] In such moments, both Raz and Marmor seem to recognize that if their goal is to provide a general explanation of people's treatment of laws as authoritative—i.e., an explanation that applies to people's treatment of customary laws as authoritative as well as to such treatment of deliberately

65. HART, *supra* note 59, at 254; RAZ, *supra* note 1, at 210; RAZ, *supra* note 29, at 35. The possibility that Raz may have changed his mind after 1979 on the need for authoritative figures gains some plausibility once we notice that he was probably not fully mindful of the content-independent nature of the reasons stemming from authoritative directives until Hart's 1982 article *Legal Duty and Obligation* appeared.

66. RAZ, *supra* note 1, at 210; MARMOR, *supra* note 63, at 97–98.

67. MARMOR, *supra* note 63, at 110.

68. *Id.*

69. Raz, *Authority*, *supra* note 48, at 205.

70. RAZ, *supra* note 1, ch. 11.

71. *Id.* at 275, 289.

enacted laws—then they cannot stick to the strong or literal version of the personal conception of authority. People do not necessarily posit deliberate lawmakers in treating laws as authoritative. At most, people need to think that their laws reflect the judgments of some persons, who may or may not have deliberately enacted those laws, as to how the people ought to behave.

But it is unclear that even this weaker form of "authorship" is required. What Marmor emphasizes is the fact that people often treat their laws as furnishing content-independent reasons. What exactly is it to treat some rules as furnishing content-independent reasons? It seems to me that Raz has something like the following in mind: in accepting a set of rules as furnishing content-independent reasons, a person is motivated to act as the rules demand out of his desire to do *what those rules demand*, where this italicized phrase is read *de dicto* rather than *de re*. In other words, it is not that there are certain things that the relevant rules demand of the person, and the person desires to do those things (*de re* reading); instead, the person desires to do whatever things are required of him by those rules (*de dicto* reading). A person motivated by the second type of desire is someone who treats the rules as furnishing content-independent reasons. Even if the relevant rules were to require him to do things quite at variance with what they actually require, the person would be motivated to act in those alternative ways. Now, it seems to me that a person can treat some laws of his community as furnishing content-independent reasons in this sense without having to think that those laws reflect the judgments of some persons as to how the members of his community ought to behave. It may be wondered how anyone could develop a trust or confidence of a requisite kind in a set of rules without thinking that those rules reflect some trusted person's judgment. Here is one possibility. One may think that heavy selection pressures (dissociated from people's normative judgments) over time favor adherents of good rules and weed out adherents of bad rules, so that the rules that survive and prevail in one's community are more likely than not good, trustworthy rules. I am not sure how plausible a hypothesis like this is, and when if ever we are warranted in taking such a view of some set of rules. But clearly, the hypothesis is not a conceptual impossibility, and at least some influential writers have taken seriously a similar view of the laws prevailing in their communities.[72] And we can cook up other hypotheses. For example, a group of people could treat some rule as generating content-independent reasons without thinking that the rule represents someone's normative view; instead, they may simply think, for example, that that is how things have been done over the years.[73] My example here resembles an example that Raz himself uses at one point in

72. *See generally* EDMUND BURKE, REFLECTIONS ON THE REVOLUTION IN FRANCE (Conor Cruise O'Brien ed., 1976) (1790); MONTESQUIEU, THE SPIRIT OF THE LAWS (Anne M. Cohler et al. eds., 1989) (1748).

73. If I am not mistaken, I have met at least half a dozen English people who have family traditions of exactly that sort—e.g., what kind of pudding to eat on Christmas.

Reasoning with Rules. A person may have a personal routine of taking a particular route in walking to and from work, and the existence of that routine, Raz says, could be a content-independent reason to walk along that route on some particular day.[74] Such a routine does not have to be seen by the person as a normative opinion of anyone, not even his earlier self, as to how he ought to behave. These considerations seem enough to show that a person can treat some rules as furnishing content-independent reasons without thereby thinking that those rules reflect normative judgments of anyone. In the first hypothetical case I discussed, for example, what a person defers to is not someone's normative judgments, but instead the mechanism that applies the selection pressures. Analogous things could be said about the other examples.[75]

I am then inclined to reject both versions of Authorial Acceptance Autonomism, to one or the other of which Raz *may* be committed. The fact that laws are often treated as furnishing content-independent protected reasons does not seem to require positing authors in either the strong or weak sense we have just discussed. The second pair of Acceptance Autonomism I want to consider are those according to which people treat the laws of their community as furnishing content-independent protected reasons, without thinking thereby that there are deliberate lawmakers or tacit endorsers. Once again, there are strong and weak versions of such Non-authorial Acceptance Autonomism. I believe that the strong version is problematic, whereas the weak version is probably correct.

As noticed above, people often treat the laws of their community as furnishing content-independent reasons. But how thoroughly or extensively can they do so? When it comes to nonlegal rules, there are often clear limits as to how thoroughly content independent the output reasons can be so treated. Raz says the following at one point in *Reasoning with Rules*:

> It is important not to confuse content-independence with unlimited jurisdiction. A justification can be, and typically will be, both content-independent and limited. The club committee [authorized e.g., to decide how many guests could be invited by members] cannot

---

74. RAZ, *supra* note 1, at 213. Raz comments that the justification of such a decision would not be "entirely content-independent" since the appeal to the routine would have been insufficient if alternative routes were much more attractive. But that alone does not undermine the point I am making. The non-existence of clearly more attractive alternative routes underdetermines the choice of the route picked out by the routine. The existence of the routine is necessary to pick out that particular route, and therefore amounts to a content-independent, albeit content-sensitive, reason for taking that route.

75. Both Raz and Marmor distinguish practical authorities from theoretical authorities by saying that the former make, whereas the latter do not, "practical differences" of the following type. MARMOR, *supra* note 63, at 100–01; RAZ, *supra* note 29, at 30. Even when a person judges that the particular directives of a practical authority were mistaken, he could treat those directives as furnishing reasons. Notice that this requirement could be satisfied in the hypothetical scenarios I am imagining. For example, even when a person judges that some particular rule generated by a selective mechanism is a bad rule, he may abide by it out of concerns about societal coordination or some such.

> be authorized to commit or order others to commit murder, etc. It is, if you like, content-sensitive in that it does not allow for any content whatsoever, while being content-independent, in not being specific to one rule. What makes a justification content-independent is not whether it can justify more or less possible rules . . . . That there are other rules which, because of their content, the justification does not show to be binding is immaterial.[76]

It may be thought that legal authorities and rules are no different. There are invariably limitations on the powers of legal authorities, and also limitations on the kind of contents that legal rules can have. Such limitations are imposed for example by constitutional provisions, and people's commitments to such constitutional provisions do not usually amount to treating them as furnishing content-independent reasons. In other words, people's commitments to those provisions are conditional on what those provisions actually say. And that would mean that ultimately, people's acceptances of their laws usually are not thoroughly content independent.

Raz has, however, argued that laws can and must ultimately be treated as furnishing content-independent reasons in a completely thoroughgoing way. At one point in *The Morality of Freedom*, he says:

> In most contemporary societies the law is the only human institution claiming unlimited authority. . . . Parliament, according to English constitutional theory, can make and unmake any law, on any matter, and to any effect whatsoever. . . . [And] things are much the same in countries with strong constitutional traditions. The American Congress's power to legislate may be limited by the Constitution, but the Constitution itself may be changed by law. Hence, even in the USA, the law claims unlimited authority.[77]

Raz goes on to add:

> Any conditional or qualified recognition of legitimacy will deny the law the authority it claims for itself. . . . [T]he law provides ways of changing the law and of adopting any law whatsoever, and it always claims authority for itself. That is, it claims unlimited authority, it claims that there is an obligation to obey it whatever its content may be.[78]

According to Raz's reasoning, in recognizing that any existing legal limitations on a community's laws could be changed by the law itself, people conceive their laws as generating reasons that are *ultimately* thoroughly content independent.

There is, however, a tension between such a conception of laws and what Raz calls the "Normal Justification Thesis." According to that thesis, a person's treatment of some rule as authoritative is normally or primarily jus-

---

76. RAZ, *supra* note 1, at 211 n.13.

77. RAZ, *supra* note 29, at 76.

78. *Id.* at 76–77.

tified only if that person, by being guided by that rule, would better conform to the reasons that apply to him anyway than he would otherwise.[79] The thesis, or more generally Raz's *Service Conception of Authority* of which the thesis is a central component, has been a topic of much philosophical controversy, and Raz spends one of the longest articles in the new volume responding to a number of the criticisms that that conception has elicited.[80] But I do not think that these criticisms bear on the point that I here want to make in relation to the Normal Justification Thesis; I think we can assume here that the thesis is roughly true. According to the thesis, in treating certain rules as authoritative, people would be treating those rules in a way that would be warranted normally or primarily only if such treatment would enable them to better conform to the reasons that apply to them anyway. We may or may not be warranted in assuming that people are at least implicitly cognizant of the (rough)[81] truth of the Normal Justification Thesis (although obviously not by that name). Raz has observed that people cannot accept advice without some understanding of the reason for which they should normally take advice—i.e., that the advice is likely to be sound.[82] Analogously, we might assume that people cannot treat some set of rules as authoritative without some understanding of the reasons for which those rules are to be so treated. This would mean that people could not normally treat any set of rules as furnishing protected reasons that are as thoroughly content independent as Raz seems to think that those furnished by laws are. Presumably, people have opinions and commitments about what ground-level reasons apply to them anyway. And assuming that they are cognizant of the truth of the Normal Justification Thesis, people could not in most cases coherently treat a set of rules as authoritative—in the sense at least of treating them as generating protected reasons—unless they believed that those rules, generally speaking, effectively mediate between the applicable reasons and their actions. Or at the least, people could not believe that the rules they treat as authoritative in fact undermine or adversely affect, in a systematic way, their abilities to conform to the applicable reasons.

Alternatively, we might assume that people are not generally cognizant of the truth of the Normal Justification Thesis. Even in that case, if we were

79. RAZ, *supra* note 1, at 136–37, 216; RAZ, *supra* note 29, at 53.

80. See the chapter entitled "The Problem of Authority: Revisiting the Service Conception" for a treatment of this issue. RAZ, *supra* note 1, ch. 5. In this chapter, Raz actually introduces a condition for justification of authoritative treatment additional to the one that the Normal Justification Thesis designates. He says that not only must the relevant person do better in conforming to applicable reasons by being guided by the relevant rules than otherwise but also that on the issue at hand it is more important for the person to conform to reason than to decide for himself without deferring to the rule issuer. *Id.* at 136–37. I doubt, however, that this is an improvement over the original self-standing Normal Justification Thesis. Why should we not see the set of reasons that apply to the person anyway as including the reasons having to do with autonomy or independence?

81. I will dispense with this qualifier in the sequel.

82. RAZ, *supra* note 29, at 54.

to take Raz's position that people conceive of their laws as furnishing reasons that are ultimately thoroughly content independent, and at the same time assume as we should that they have opinions and commitments about what ground-level reasons apply to them, then we would be characterizing them as having normative attitudes that are systematically incorrect. In effect, people would be seeing themselves as having reasons to behave as the laws demand no matter what those laws say, and according to the Normal Justification Thesis, such attitudes would be normatively defective. And given the systematic incorrectness of such attitudes, we should attribute them to people only if such attribution is really unavoidable because no comparable alternative explanation is in the offing.[83] In the current case, we actually have a plausible alternative explanation. We can distinguish between two versions of Non-authorial Acceptance Autonomism, both of which recognize that people treat laws as furnishing content-independent protected reasons, and neither of which necessarily posits some lawmaker. According to one, the one which Raz seems to support, the furnished reasons can and must be thoroughly or unlimitedly content independent. According to the other version, which amounts to the alternative we want, the reasons can be content independent only to some limited extent (which we need not specify). As far as I can see, the stronger version of Non-authorial Acceptance Autonomism does not offer any additional explanatory payoffs, and we can reject it without regret. In sum, whether or not we assume that people are cognizant of the truth of the Normal Justification Thesis, we should reject the strong version of Non-authorial Acceptance Autonomism in favor of the weak version. Such rejection is required by the Normal Justification Thesis, which builds in a certain amount of content sensitivity to people's treatment of rules as authoritative.

To recapitulate this relatively long Part, we began with the simple Acceptance Protectionism, according to which people treat their laws as generating protected reasons. We then took up and assessed thicker or more inflated conceptions of the attitude of rule-acceptance—the four versions of Acceptance Autonomism—each of which recognizes the fact that people treat laws as furnishing content-independent protected reasons. We rejected the two authorial versions of Acceptance Autonomism. And as for the non-authorial versions, we accepted the version that recognizes some content sensitivity, while rejecting the version that does not. In effect, according to the version of Acceptance Autonomism that is left standing, the one that Raz is entitled to in saying that the law "claims legitimate authority," people treat the laws of their community as furnishing limitedly content-independent

83. For similar reactions to J.L. Mackie's "error theory" in ethics, see Simon Blackburn, *Errors and the Phenomenology of Value*, *in* ETHICS AND OBJECTIVITY (Ted Honderich ed., 1985), *reprinted in* SIMON BLACKBURN, ESSAYS IN QUASI-REALISM 149, 149 (1993); David Lewis, *Dispositional Theories of Value*, 63 PROC. ARISTOTELIAN SOC'Y 113 (1989), *reprinted in* DAVID LEWIS, PAPERS IN ETHICS AND SOCIAL PHILOSOPHY 68, 92–93 (2000). Mackie proposed his error theory in J.L. MACKIE, ETHICS: INVENTING RIGHT AND WRONG ch. 1 (1977).

protected reasons. This conclusion will prove useful in assessing Raz's argument for the Sources Thesis.

VI. Sources Thesis

For some time now, legal philosophers have been preoccupied with an increasingly baroque—some would say "scholastic"—debate about whether communities' ultimate criteria of legal validity can include some normative, nonfactual criteria. Hart took note that constitutions of countries like the United States explicitly incorporate moral standards.[84] But it is Dworkin who gave the debate its current contours by arguing that the nature of adjudication indicates that ultimate criteria of legal validity can include normative tests, and that legal positivism is erroneously committed to there being only purely factual criteria of legal validity.[85] The question of whether moral criteria can be parts of the ultimate criteria of legal validity thereby became the question of whether nonfactual criteria can be parts of the ultimate criteria of legal validity.

Some legal positivists took the "Oh yeah?" approach in reacting to Dworkin and argued that legal positivism can allow ultimate criteria of legal validity to include normative criteria.[86] Others have taken the "So what?" approach instead, and have argued that legal positivism is indeed committed to there being only factual criteria of legal validity, but that this is not a problem.[87] The two positions often go by the labels "inclusive legal positivism" and "exclusive legal positivism" respectively, and Raz has been the leader among exclusive legal positivists.

The particular version of exclusive legal positivism that Raz has defended over the years is called the "Sources Thesis." A law has a "source," and therefore is "source based," if its existence and content can be determined by appeals to what Raz calls "social facts"—i.e., behavioral and

---

84. HART, *supra* note 6, at 204.

85. *See generally* DWORKIN, *Hard Cases*, *supra* note 13, at 81; DWORKIN, *Rules I*, *supra* note 12, at 14; DWORKIN, *Rules II*, *supra* note 12, at 46. A plausible reading of these articles is to see Dworkin as arguing that the best explanation of the fact that judges often appeal to normative considerations in deciding cases, and the fact that when they do so judges do not see themselves as exercising discretion but instead as legally constrained (what I call "(F6)" in Part II *supra*), is provided by the hypothesis that criteria of legal validity can include normative tests.

86. *E.g.*, JULES L. COLEMAN, THE PRACTICE OF PRINCIPLE 107 (2001); W.J. WALUCHOW, INCLUSIVE LEGAL POSITIVISM ch. 5 (1994); Jules L. Coleman & Brian Leiter, *Legal Positivism*, *in* A COMPANION TO PHILOSOPHY OF LAW AND LEGAL THEORY 241, 243 (Dennis Patterson ed., 1996). It is not clear whether such philosophers are really entitled to say that the ultimate criteria of legal validity can be partly normative, for they believe that the inclusion of such normative criteria occurs, if it does, only if certain facts underwrite such inclusion. Because of this qualification, it is far from clear how the inclusive legal positivist answer provides an adequate reply to Dworkin's question.

87. In the text, I am taking advantage of the common lore among philosophers that all reactions to arguments could be divided into the "Oh yeah?" and "So what?" varieties. My first exposure to this lore was from Nicholas Sturgeon, *What Difference Does It Make Whether Moral Realism Is True?*, 24 S. J. PHIL. 115, 115 (1986).

psychological facts that constitute social practices, legislative enactments, judicial decisions, etc.—and without an appeal to any normative considerations.[88] The Sources Thesis is the view that all laws are source based. It is meant to tell us what legal reasoning *in the strict sense* is like. According to the Sources Thesis, legal reasoning, in the strict sense of reasoning aimed at determining the existence and content of existing laws, is a purely factual or nonnormative reasoning. The modifier *in the strict sense* is necessary because Raz does not think that legal reasoning *broadly conceived* is limited to that which is aimed at determining the existence and contents of existing laws. When the reasons not represented by or depended upon by legal rules call for it and outweigh the reasons represented by legal rules, judges and other officials must exercise discretion and reason contrary to what the existing laws say.[89] In some such cases, they create new laws that control future legal reasoning in the strict sense.

Raz has influentially argued that the authoritative nature of law provides very strong grounds for thinking that the Sources Thesis is true. A version of that authority-based argument is given in "On the Nature of Law," but that version is quite incomplete.[90] What Raz seems to be relying on there, and elsewhere in *Between Authority and Interpretation*,[91] is what many consider the canonical version of the argument in the article *Authority, Law, and Morality*.[92] Given the importance of this argument for Raz's overall thinking about the nature of law, let me first outline it in the personalese that he favors. Throughout, authority means "legitimate authority."

(6.1) The law claims authority.

(6.2) Since the law claims authority, it normally or typically is capable of possessing authority.

(6.3) To be capable of possessing authority, the law must have at least the following two nonnormative features: (i) The law's directives reflect someone's views as to how the subjects ought

88. RAZ, *supra* note 1, at 111, 114–16. In these pages, Raz does not use the term "the Sources Thesis" to refer to the view that he summarizes and endorses. He does at page 380, but there he is quoting from Gerald Postema's formulation of the thesis. *Id.* at 380 (quoting Gerald J. Postema, *Law's Autonomy and Public Practical Reason*, THE AUTONOMY OF LAW 79, 92 (Robert P. George ed., 1996)). For more canonical formulations of the Sources Thesis, where Raz also identifies the thesis by its name, see, for example, RAZ, *supra* note 2, at vi–vii; JOSEPH RAZ, *Legal Positivism and the Sources of Law*, *in* THE AUTHORITY OF LAW, *supra* note 2, at 37, 47–48; J. RAZ, *Legal Reasons, Sources and Gaps*, 11 ARCHIV FÜR RECHTS- UND SOZIALPHILOSOPHIE 197, 205 (1979) (F.R.G.), *reprinted in* THE AUTHORITY OF LAW, *supra* note 2, at 53, 63; and Raz, *Authority*, *supra* note 48, at 194.

89. RAZ, *supra* note 1, at 116, 333, 376–79.

90. *Id.* at 106–15. It is also marred by its reliance on a conception of laws as literally representing decisions of the community acting as an agent on behalf of its members. *Id.* at 101–04. In *Reasoning with Rules*, a chronologically later article, Raz explicitly repudiates this conception. *Id.* at 217–18.

91. *E.g.*, *id.* at 381.

92. Raz, *Authority*, *supra* note 48, at 194.

to behave; and (ii) both these directives' status as issues of the law and their contents can be ascertained without relying on or delving into the reasons on which the directives are meant to be based and that the directives are meant to exclude in the subjects' practical thinking.

(6.4) It follows that the law normally or typically possesses the two nonnormative features.

(6.5) If the Sources Thesis were true, then the law would have the two nonnormative features, whereas the truths of the competing leading theses about the nature of law would not ensure the law's having those features.

(6.6) By inference to the best explanation, the Sources Thesis is probably true.

I believe that this argument can be translated without distortion into the following nonpersonalized form:

(6.7) Necessarily, many members of a community governed by laws treat the laws of their community as authoritative.

(6.8) Necessarily, if many people treat some rules as authoritative, then those rules must normally or typically be appropriate or fitting objects of being treated as authoritative.

(6.9) Necessarily, if some rules are appropriate or fitting objects of being treated as authoritative, then those rules have at least the following two nonnormative features: (i) They reflect someone's views as to how the members ought to behave; and (ii) their existence and contents can be ascertained without relying on or delving into the reasons on which the rules are meant to be based and that the directives are meant to exclude in the members' practical thinking.

(6.10) It follows that normally or typically the laws of a community have the two nonnormative features.

(6.11) If the Sources Thesis were true, then the laws would have the two nonnormative features, whereas the truths of competing leading theses about the nature of law would not ensure the laws having those features.

(6.12) By inference to the best explanation, the Sources Thesis is probably true.

I want to assess Raz's argument for the Sources Thesis in this last form. I am mindful that there are other versions of the argument, and also alternative arguments, for the Sources Thesis that Raz has relied on in various places.[93] But given the canonical status that the argument in *Authority, Law, and Morality* seems to enjoy, an assessment of it alone would be very much

93. As a matter of fact, Raz quickly outlines an alternative argument for the thesis in *Authority, Law, and Morality* itself. *Id.* at 204.

worth our while. I will not speculate here as to how widely the considerations that I will be using to criticize that argument will be applicable to the various variations of and alternatives to this argument.

Clearly, (6.9) is the key move in the argument. But let me take up (6.8) first. Here, Raz is claiming that certain conceptual mistakes are normally impossible or at least extremely unlikely. Obviously, how strong or plausible a claim that Raz is making here depends on what he thinks is involved in thinking that certain rules are fitting objects of being treated as authoritative (or in the personalized version of the argument, what he thinks is involved in being capable of possessing authority)—in other words, on the specific non-normative features that (6.9) enumerates. But bracketing for the moment the specific content of (6.9), we may wonder why Raz would rule out *any* kind of conceptual mistake on the part of those who take their laws as authoritative. Raz explains as follows:

> Since the claim [that the law possesses legitimate authority] is made by legal officials wherever a legal system is in force, the possibility that it is normally insincere or based on a conceptual mistake is ruled out. It may, of course, be sometimes insincere or based on conceptual mistakes. But at the very least in the normal case the fact that the law claims authority for itself shows that it is capable of having authority.
>
> Why cannot legal officials and institutions be conceptually confused? One answer is that while they can be occasionally they cannot be systematically confused. For given the centrality of legal institutions in our structures of authority, their claims and conceptions are formed by and contribute to our concept of authority. It is what it is in part as a result of the claims and conceptions of legal institutions.[94]

In other words, according to Raz, our understanding of the nature of authority has been fundamentally shaped by and intertwined with our understanding of the nature of law. The law is paradigmatic of what can be authoritative. And for this reason, Raz seems to think, people are quite unlikely to make certain *elementary* mistakes, or at least to make such elementary mistakes systematically, in attributing authoritativeness to laws.

I am not sure what exactly to think about this hypothesis of conceptual safety net, so to speak, but if pressed I would be inclined to reject it. The development of anti-individualist semantics (which will be discussed in the next Part) indicates that people can be quite wrong (and systematically so) about what at one time was considered a paradigmatic or archetypal member of the extension of a concept. But putting this off-the-cuff skepticism to one side, let me make one other observation about Raz's argument here before moving on to a consideration of (6.9), which will enable us to get a real measure of (6.8). If, as Raz believes, our concept of authority has been

94. *Id.* at 201.

shaped fundamentally by our understanding of law, then that would be an additional reason to reject Authorial Acceptance Autonomism, at least in its strong version. Customary laws traditionally have been a large and important part of our legal practice. Assuming Raz's safety-net hypothesis, if in fact our concept of authority has been fundamentally shaped by our understanding of the nature of law, then our laws cannot be authoritative only in the sense of requiring deliberate lawmakers.

According to clause (i) of (6.9), only those rules that are reflective of someone's normative judgments could be treated as authoritative. But my arguments in the preceding Part against both the strong and weak versions of Authorial Acceptance Autonomism show that this is not the case. People can treat some rules as content-independent protected reasons even when they do not view the rules as reflecting someone's normative judgments or views as to how people should behave. Turning now to clause (ii), clearly the central move in the argument for the Sources Thesis, we should ask: Does treating some rules as authoritative really necessarily involve thinking that whether they prevail and what they say can be determined without resorting to the underlying reasons on which the rules are supposed to be based? What is all-important here is the sense of "authoritative" that is in play. And that in turn depends on which among the various conceptions of the psychological attitude of rule-acceptance that we have enumerated and considered in the last Part is correct. If "authoritative" were defined by way of the simple Acceptance Protectionism, the one according to which people treat laws as furnishing protected but not necessarily content-independent reasons, then Raz's argument for the Sources Thesis would have a very unwelcome implication. Remember that there are moral rules that we also treat as authoritative in the simple sense of furnishing protected reasons to act as the rules say.[95] Some moral rules can be characterized as source based—e.g., those having to do with promises, agreements, and also perhaps individual volitional commitments such as decisions and intentions. But Raz would not want to characterize all moral rules, or even all deontic moral rules, as source based. Again and again, he has said that legal rules are distinct from moral rules (or given Raz's Acceptance Moralism, nonlegal moral rules) in being source based.[96] And rightly so.

What would enable Raz to steer clear of this unwanted implication is a conception of the attitude of rule-acceptance that is thicker or more inflated than Acceptance Protectionism, and more particularly one that recognizes the content-independent nature of the reasons that are putatively furnished by legal rules. Both versions of Authorial Acceptance Autonomism are unacceptable for the reasons that we have already discussed. Non-authorial Acceptance Autonomism is more promising, but here I have argued that the strong or unconstrained version, according to which laws are taken as gener-

95. RAZ, *supra* note 1, at 382.
96. *Id.* at 115–16.

ating protected reasons that are ultimately completely or unlimitedly content independent, cannot be correct. Such a version is inconsistent with the Normal Justification Thesis. The only version of Non-authorial Acceptance Autonomism that was left standing at the end of the preceding Part is the weak version, according to which people see their laws as furnishing ultimately only limitedly content-independent protected reasons. But what is crucial here is that Raz could maintain clause (ii) of (6.9) only if the strong version of Non-authorial Acceptance Autonomism were true. If people were to see their laws as furnishing protected reasons that are only limitedly content independent, then that would mean that they could not appeal merely to social–factual considerations to determine the existence and contents of their laws; they would ultimately have to delve into the normative character or nature of the actions that are called for by their laws. Or at the least, even when they are not attending to such normative considerations, people would have to be disposed to react appropriately to any appearances of great or systematic discrepancies between what their laws demand of them and what the ground-level reasons call for in their behavior. The fact that the reasons are content independent (to some extent) would mean that the relevant normative considerations by themselves are not enough to determine the existence and contents of the laws. But the fact that the reasons are also content sensitive would mean that the social–factual considerations by themselves are also not enough for those purposes. In sum, if I am right in thinking that only the weak version of Non-authorial Acceptance Autonomism is plausible, then Raz's argument for the Sources Thesis hits a serious snag at clause (ii) of (6.9).

Let me try to mitigate the quite abstract and even scholastic appearance of the foregoing discussion. We can think of a scenario typical of those that Raz works with, in which a legal authority in a community enacts laws that are meant to govern the behavior of the community members. The officials and others of the community who accept their laws treat such laws as furnishing content-independent protected reasons to act as those laws say. The question has been whether the nature of such an attitude towards the laws makes it the case that the existence and contents of all laws of the relevant community can be determined by appealing only to the social–factual considerations completely divorced from the considerations having to do with the normative character or nature of the actions that the laws call for. My argument has been that the answer is "no." Given the Normal Justification Thesis, the members of the community cannot treat their laws as furnishing reasons that are content independent in the unlimited sense. And it follows that the nature of the attitude of acceptance—or to put things in terms that Raz himself favors, the nature of the law's claim of authority—does not warrant the conclusion that the existence and contents of all laws can be determined by consulting only social–factual considerations. To put the point slightly more concretely, the legal authority in our scenario may or may not be explicitly limited in its authority by the power-conferring laws that

confer the lawmaking authority on it. But even if the legal authority were not so explicitly limited, and even if the particular legal authority were the ultimate legal authority in the relevant community so that it—by itself or together with some other legal authorities—could change the legal limitations on its lawmaking powers, the community's members cannot treat the enacted laws as furnishing unlimitedly content-independent reasons. For such an attitude towards the laws would imply their willingness to comply with the laws no matter how adversely they believe such compliance would affect their abilities to conform to the ground-level reasons that apply to them anyway. And that would go against the very purpose of treating any rules as authoritative—that is, as furnishing content-independent protected reasons. In effect, the sense of "authoritative," in which we are warranted in thinking of the members of a community as treating their laws as authoritative, is one according to which the members treat the laws as furnishing only limitedly content-independent protected reasons. And if we insert this particular sense of "authoritative" in Raz's argument for the Sources Thesis, then the argument does not go through. This is not to say that there could not be considerations other than the nature of the attitude of rule-acceptance that could be appealed to to argue for the Sources Thesis. But the canonical version of the argument for the Thesis in *Autonomy, Law, and Morality* seems to fail.[97]

VII. Theoretical Disagreements

Among the articles contained in *Between Authority and Interpretation*, the one titled *Two Views of the Nature of the Theory of Law: A Partial Comparison* seems to me to contain more of Raz's new thinking about the nature of law than any other article in the collection.[98] The article consists in the main of an attempt to respond to Dworkin's argument, made in the first chapter of *Law's Empire*, that legal positivism, and more particularly Hart's version of it, fails to account for what Dworkin calls "theoretical legal disagreements"—that is, for (F8) mentioned in Part II above.[99] I am afraid

97. What about the apparently obvious counterexamples provided by the various moral standards embedded in, for example, the U.S. Constitution? Here, I believe, Raz has succeeded in blunting the force of such apparent counterexamples by likening them to the cases in which terms of private contracts or company regulations are given legal effect, or those in which foreign laws are given legal effect through domestic laws governing choice of laws. *Id.* at 193–94. I am not sure what exactly would be wrong in thinking that the relevant terms of private contracts, company regulations, and foreign laws are parts of the law of the land. And Raz also says that "the distinction between what is part of the law and what are standards binding according to law but not themselves part of the law is particularly vague," and that "much of the time the practical implications of a standard are the same either way." RAZ, *supra* note 1, at 195. But I also do not know of any considerations in favor of the more inclusive conception of the law of the land over Raz's more exclusive conception.

98. *Id.* ch. 3.

99. DWORKIN, LAW'S EMPIRE, *supra* note 12, ch. 1. Dworkin's argument is meant to apply to all traditional legal theories, not just legal positivist ones. But in order to simplify my discussion, I will concentrate mainly on his argument as it applies to legal positivist ones.

that quite a bit of stage setting is necessary before we can discuss and assess Raz's arguments in *Two Views*.[100]

What Dworkin calls "theoretical disagreements" are legal disagreements that persist despite discussants' agreement on all issues having to do with the so-called social facts—i.e., the behavioral and psychological facts that constitute social practices, legislative enactments, judicial decisions, etc. They are to be distinguished from on the one hand social–factual disagreements (which Dworkin calls "empirical disagreements") and on the other disagreements about what the law should be. Dworkin spends a large chunk of chapter 1 of *Law's Empire* characterizing the judicial opinions of some well-known American and British appellate cases as instancing such theoretical disagreements among judges (and commentators)[101] and argues that legal positivist theories, and more particularly Hart's theory, cannot explain occurrences and even prevalence of theoretical legal disagreements in actual legal settings.

Exactly what the relationship is between Dworkin's earlier arguments and this new argument based on the possibility of theoretical disagreements is a vexing issue that I cannot pursue here. In any case, in *Law's Empire*, Dworkin sought to offer a deeper diagnosis of legal positivism's ills than what he offered in his earlier articles. He argued, quite surprisingly to many, that legal positivist theories' inability to account for theoretical disagreements can be traced to their being a species of "semantic theories of law," according to which all competent speakers, or at least all competently trained lawyers, know and agree on the meaning or concept of "law," and that meaning or concept determines the ultimate criteria of legal validity in each legal system.[102] What is distinctive of the semantic legal theories that legal positivists espouse is that, according to them, the meaning or concept of "law" designates only factual criteria, and more specifically social–factual criteria, as the ultimate criteria of legal validity, the criteria that everyone who properly understands the meaning or concept of "law" accepts.[103] As replacements for semantic theories of law, Dworkin offered what he called "interpretive theories" of law, according to which people in legal disputes offer competing "constructive interpretations" of the existing societal standards and practices.[104] And given that each constructive interpretation is a normative conclusion inferred from a set of normative standards that best fit

100. Recently, Scott Shapiro, *The 'Hart-Dworkin' Debate: A Short Guide for the Perplexed*, *in* RONALD DWORKIN 22 (Arthur Ripstein ed., 2007) and Brian Leiter, *Explaining Theoretical Disagreements*, 76 U. CHI. L. REV. 1215 (2009), have revived to some extent the interest in theoretical disagreements. While my disagreements with many aspects of these two discussions will be implicit in my subsequent discussion, I will not have space here to make my disagreements explicit.

101. DWORKIN, LAW'S EMPIRE, *supra* note 12, at 15–30.

102. *Id.* at 31–46.

103. *Id.* at 7, 32–35.

104. *Id.* ch. 2.

and morally justify some paradigms of the prevailing legal standards and of the existing societal practices, those who issue such interpretations are not constrained by any existing agreement among them as to what the ultimate criteria of legal validity are, whether such agreement is determined by their shared linguistic or conceptual understanding or anything else.

Many legal philosophers found Dworkin's characterization of legal positivist theories, and of Hart's theory in particular, as semantic theories of law puzzling and even perverse. But some have read in Dworkin's dissatisfaction with Hart's legal theory the same kind of dissatisfaction that philosophers like Hilary Putnam and Tyler Burge slightly earlier had with the traditional individualist semantics and philosophy of mind.[105] Nicos Stavropoulos, in particular, has argued that what Dworkin should have distinguished between are not semantic theories of law and interpretive theories, but instead legal theories based on the traditional criterial semantics and those based on the newer anti-individualistic semantics.[106] Apparently, Dworkin has come to accept Stavropoulos's revisionary reading of the argument in *Law's Empire*, and Raz's arguments in *Two Views* are aimed at responding to that revisionary reading of Dworkin's argument.[107]

Now, what is criterial semantics? The basic idea of criterial semantics is that for each subject matter, there are certain truths about it that are *criterial* in the following sense: (i) they are true purely in virtue of meaning or are conceptual truths, which are a species of truths by convention or stipulation; (ii) an understanding necessary for any genuine talking or thinking about the subject matter requires grasping these truths; and (iii) the necessary understanding also requires an understanding that these are conventional or stipulative truths, and hence both impossible and pointless to doubt. An example of this type of allegedly criterial truths is the following: "Sofas are pieces of furniture made or meant for sitting." According to criterial semantics, we would not credit someone as genuinely talking or thinking

105. *See generally* NICOS STAVROPOULOS, OBJECTIVITY IN LAW (1996); David O. Brink, *Legal Theory, Legal Interpretation, and Judicial Review*, 17 PHIL. & PUB. AFF. 105 (1988).

106. David Sosa has pointed out in a conversation with me that philosophers of language and of mind use the terms "criterial semantics," "anti-criterial semantics," "individualism," and "anti-individualism" in a wide variety of ways. So some philosophers may complain that, as they understand these terms, criterial semantics and individualist semantics are not necessarily the same, and that there is even a possibility of being a criterialist anti-individualist. I should point out that I am using the terms "criterial semantics" and "anti-individualist semantics" in the senses that Stavropoulos uses them, and that he in turn intends to use these terms to refer to what Burge calls individualism and anti-individualism, respectively. STAVROPOULOS, *supra* note 105, at 34–37; Tyler Burge, *Intellectual Norms and Foundations of Mind*, 83 J. PHIL. 697 (1986). The bottom line is that I use the relevant terms only to mean the views that I am about to specify in the text, while recognizing that there have been in the philosophical literature somewhat different usages of the terms. No matter how representative or unrepresentative my uses of these labels are, if my descriptions of the two views allow us to get at the Dworkinian argument that Stavropoulos has constructed for us, then that should be good enough for my purposes.

107. RAZ, *supra* note 1, at 53 & n.16, 59.

about sofas unless he grasped this truth, and that in turn means that no one could coherently question this truth.

Why would anyone think that Hart was guilty of criterial semantics? The thinking goes as follows.[108] According to Hart, whether a legal statement or judgment is true or correct depends ultimately on whether there exists or prevails among the officials of the relevant community a convergent practice of adhering to some rules of validity that validate the statement or judgment. This is meant to be a criterial truth in the sense that no competent lawyer could fail to know it, and no competent lawyer would question it. And it could be further said that this truth follows from the meaning or concept of "law." It follows from this characterization of Hart's legal theory that two competent lawyers could have empirical disagreements, disagreements about how the law should be changed, or merely apparent disagreements in which they talk at cross purposes. But they could not have theoretical disagreements.

Anti-individualist semantics attempts to provide a more accurate characterization of our cognitive practices than does criterial semantics. According to it, the understanding that is necessary for being credited with genuine talk or thought about some subject matters is not the grasping of some conventional or stipulative truths. Instead, what is necessary is recognition of a set of some examples as paradigms of the relevant subject matter (the membership in which set could be revised upon further inquiry), and a certain degree of susceptibility to regulation by the norms of intellectual inquiry. ("After you showed me those chairs bolted to the lunch counter, I stand corrected, and I no longer believe that chairs must have four legs.") An implication is that a person may have wildly unorthodox and incorrect beliefs about a subject matter and still count as genuinely talking and thinking about it. These features of our cognitive practices reflect our recognition that what is true or correct about many subject matters is not ultimately a matter of convention or stipulation, or even a matter of expert consensus achieved through reflective equilibrium, but instead a matter of the nature of the subject matter, to which we gain access through our investigation of the paradigms. What we hold ourselves and each other responsible to in our intellectual inquiry about many a subject matter is not some convention or consensus opinion about it, but instead the norms of intellectual inquiry and ultimately the nature of our subject matter.

It is not a huge stretch to think that Dworkin's criticism of legal positivism, especially of Hart's legal theory, and his advocacy of interpretive theories of law were born out of his implicit anti-individualism. A person

108. I will eventually be questioning whether what follows is a proper understanding of Hart's legal theory. But that is not to say that there are not passages in *The Concept of Law* that invite (or at least do not exclude) such understanding. Kevin Toh, *An Argument Against the Social Fact Thesis (And Some Additional Preliminary Steps Towards a New Conception of Legal Positivism)*, 27 LAW & PHIL. 445, 482–86 (2008).

who offers a constructive interpretation, according to Dworkin, offers a legal opinion that is meant to be a part of the best moral justification of the paradigms of valid laws and legally valid practices in his community, and such an opinion can be quite at odds with the legal consensus that prevails among the officials of the person's community. Legal truths or correctness, according to Dworkin, are ultimately not a matter of convention or even expert consensus, or at least not of them alone. What is the law in a community can transcend or outstrip the practices that prevail in that community. This is why, according to Dworkin, two lawyers can agree on all the relevant social facts but still disagree about what the law is.

Raz's reaction to this recasting of Dworkin's argument is to stick to criterial semantics while incorporating what he considers the insights of anti-individualism.[109] Raz has (mostly)[110] adhered to the view, which he attributes to Hart, that whether a legal statement or judgment is true or correct depends ultimately on whether there exists or prevails among the officials of the relevant community a convergent practice of adhering to some rules of validity that validate the statement or judgment. In the important article *Legal Validity*, for example, he offered the following as a slight variation and improvement on what Hart was getting at:

> The legal validity of a rule is established not by arguments concerning its value and justification but rather by showing that it conforms to tests of validity laid down by some other rules of the system which can be called rules of recognition. These tests normally concern the way the rule was enacted or laid down by a judicial authority. The legal validity of rules of recognition is determined in a similar way except for the validity of the ultimate rules of recognition which is a matter of social fact, namely those ultimate rules of recognition are binding which are actually practised and followed by the courts.[111]

I assume that the "validity" that Raz is talking about in the emphasized portion of this passage is legal validity. What Raz and like-minded legal philosophers believe is that this view of what ultimately determines the legal validity of rules is something that Hart has come to discover through his

---

109. I should warn that in the course of his discussion in *Two Views* Raz offers somewhat different characterizations of the dialectic between criterial semantics and anti-individualism, and of their relevance for Dworkin's argument than what I have offered above. As I said above, *supra* note 106, the characterizations I offered were meant to duplicate what I believe are Burge's understanding of the two semantic views, and Stavropoulos's (and presumably also Dworkin's) understanding of their relevance for legal philosophy. Raz offers somewhat different characterizations. But in my subsequent discussion I shall ignore most of the differences as I do not think that they have any repercussions for the central issues at stake; I will mention only the ones that I think do matter.

110. The significance of this qualification will be clarified in due course.

111. RAZ, *Legal Validity*, *supra* note 52, at 150–51 (emphasis added); *see also* RAZ, *supra* note 1, at 77–78. Raz's talk in the quoted passage of multiple rules of recognition and a hierarchy among them is explained by his view that Hart was wrong to assume that a legal system can have only one rule of recognition. JOSEPH RAZ, *The Identity of Legal Systems*, *in* THE AUTHORITY OF LAW, *supra* note 2, at 78, 95–96.

philosophical investigation of the concept of law. Raz's Sources Thesis and his authority-based argument for it were a further elaboration of this allegedly criterial truth.

For all I have said so far in the last paragraph, Hart's legal theory, as Raz conceives it, and Raz's own theory could be consistent with the characterization of criterial semantics that I offered above. In *Two Views*, Raz explicitly departs from that characterization when he denies that all competent lawyers or speakers grasp this criterial truth, and when he grants that competent lawyers or speakers could coherently doubt this truth. Raz says: "Having a concept . . . is compatible with a shallow and defective understanding of its essential features, and of the nature of what it is a concept of."[112] This he takes to be a core insight of anti-individualism, and one that criterial semantics can incorporate without much ado. Raz says:

> Where the individualistic approach goes wrong is in thinking that the criteria set by each person's personal rule for the correct use of terms and concepts are fully specified. In fact, their personal rules are not specified. Each person takes his use of terms and concepts to be governed by the common criteria for their use. That is all their personal rule says. . . . It is part of each person's rule for the use of the term or concept that mistakes can occur, for the rule refers to the criteria as they are, rather than to what that person thinks they are. What they are, however, does depend on what people think they are. The correct criteria are those that people who think they understand the concept or term generally share, i[.]e[.] those that are generally believed to be the correct criteria are the correct criteria.[113]

It is important to keep apart two kinds of criteria that are in play here. One has to do with the criterial truth that Hart is taken to have discovered, according to which the truth or correctness of legal statements or judgments is ultimately a matter of social facts concerning which rules of legal validity a community or its officials converge upon in their practices. The other kind of criteria are those ultimate rules of legal validity that a community or its officials converge upon in their practices. The point that Raz is making in the above passage is about conceptual criteria in the first sense. Two competent people can differ on the criteria for application of a single concept—in this case the concept of law—but still deploy that concept in their disagreement with each other. The important implication for Raz is that such criterial disagreement can produce legal criterial disagreements of the second kind, which are not empirical disagreements, disagreements about what the law should be, or merely apparent disagreements. They could be genuine disagreements even about the ultimate criteria of legal validity. Could not these be the kind of disagreements that Dworkin was gesturing at by "theoretical disagreements"?

---

112. RAZ, *supra* note 1, at 55.

113. *Id.* at 64.

Not quite. What Raz adopts is not anti-individualism (at least as Stavropoulos and Dworkin seem to understand that view and its implications for legal philosophy);[114] it is at best only a halfway house to anti-individualism. In a way, the term "anti-individualism" is a bit of a misnomer. For the view not only denies that the crucial sort of understanding necessary for genuine talk and thought about a subject matter is an understanding of truths of meaning (or conceptual truths) that are stipulated in an individual's idiolect, but it also denies that the relevant kind of understanding is an understanding of truths of meaning (or conceptual truths) that are conventionally established among a community's members. In accepting what he deems anti-individualism, Raz rejects the idiolectical version of individualism, but he adheres to the communal version. As I explained above, what anti-individualism says is that for the relevant subject matters, we hold ourselves responsible to the norms of intellectual inquiry and ultimately to the nature of those subject matters, and not to the consensus, even expert consensus, prevailing in our community. According to the position that Raz adopts, in legal matters, what we hold ourselves responsible to ultimately is the community consensus, or at least expert consensus. Of course, we know why he opts for this halfway measure. He cannot be a thoroughgoing anti-individualist without abandoning his Sources Thesis.

Perhaps it may be thought that Raz gets law right, even if he gets anti-individualism wrong. There are some subject matters about which we do actually think that the community consensus or expert consensus is the final arbiter, and perhaps law or legal validity is one such subject matter. Anti-individualism was never meant to be a global semantic theory. This is a large and complicated issue, and I cannot do justice to it, especially near the end of an already very long review essay. But in closing, I want to offer two admittedly inconclusive considerations against viewing law that way, just to indicate that Raz is not clearly right.

First, think of what Justice Scalia takes himself to be doing when he makes his originalist pitch. He takes himself to be making a claim about what the law is, not a claim about what the law should be. However, presumably he fully realizes that the prevailing current practice or consensus among legal experts and officials does not support originalism. At least Justice Scalia and many like him—not just originalists but many who make fairly unorthodox fundamental legal claims—do not view themselves as taking part in a cognitive practice in which the final arbiter of correctness is community or expert consensus. No amount of sociological, psychological, and historical evidence indicating actual judicial deviance from originalism is likely to persuade Justice Scalia that he is wrong. Nor can it plausibly be said that Justice Scalia's "intransigence" is due to conceptual ignorance or confusion. We could have him study *The Concept of Law* thoroughly and

114. On the need for this parenthetical, see *supra* note 106.

have him spend weeks on end with Raz himself explaining the book to him, and with Frank Jackson explaining to him the virtues of conceptual analysis.[115] Raz may be able to convince him that conceptual investigations clearly show that whether a legal statement or judgment is true or correct depends ultimately on whether there exists or prevails among the officials of the relevant community a convergent practice of adhering to some rules of validity that validate the statement or judgment. I suspect that Justice Scalia's reaction then would resemble many philosophers' reaction to R.M. Hare's argument in *Moral Thinking* that our moral concepts or the meanings of our moral terms commit us to a form of act-utilitarianism.[116] Many philosophers reacted by saying that if that really were what our moral concepts or meanings commit us to, then henceforth we take ourselves to be deploying different concepts and meanings when we engage in what seem like moral thinking and talking.[117] Any protests, on Hare's behalf, that such reactions are irrational, incoherent, or downright immoral are apt to seem superficial and question begging. I believe that analogous protests to what I presume will be Justice Scalia's reactions would be similarly superficial and question begging.

I doubt, for another reason, that *The Concept of Law* could be used to persuade Justice Scalia. And here I come to the second of the two considerations. I do not think that legal philosophers like Raz and Dworkin have been right in thinking the following: according to Hart it is a criterial truth, which can be discerned in the meaning or concept of "law," that whether a legal statement or judgment is true or correct depends ultimately on whether there exists or prevails among the officials of the relevant community a convergent practice of adhering to some rules of validity that validate the statement or judgment. To be contrasted with that characterization of Hart's position, which can be found in the above-quoted passage from Raz's *Legal Validity*,[118] is the following characterization that Raz gives in his response to Gerald Postema:

> [Postema] alleges that Hart regarded the validity of rules of recognition as resting solely on their social acceptance. But as you

115. At least in North America, philosophers have for many years thought that conceptual analysis as a philosophical method had been thoroughly discredited by the criticisms of W.V. Quine and others. Jackson has been one of the main instigators of the revitalization (or at least reconsideration) of conceptual analysis as a viable philosophical method in recent years. *See generally* FRANK JACKSON, FROM METAPHYSICS TO ETHICS (1998). To the extent that my argument in this Part questions the supposed indubitability of certain legal claims based on their conceptual or semantic (or criterial) nature, my argument is Quinean in spirit. For wide-ranging Quinean reflections on the use of conceptual analysis in legal philosophy, see generally the articles collected in BRIAN LEITER, NATURALIZING JURISPRUDENCE (2007).

116. R.M. HARE, MORAL THINKING 35–43 (1981).

117. *E.g.*, Thomas Nagel, *The Foundations of Impartiality*, *in* HARE AND CRITICS: ESSAYS ON MORAL THINKING 101 (Douglas Seanor & N. Fotion eds., 1988); Bernard Williams, *The Structure of Hare's Theory*, *in* HARE AND CRITICS, *supra*, at 185.

118. *See supra* note 111 and accompanying text.

> will know Hart denied that it is right to talk of the validity of rules of recognition. They merely exist, and to say that they do is no more than to say that they are accepted. In The Concept of Law he never said that the fact that they were accepted gives anyone reason to follow them. That they are accepted entails, of course, that those who accept them think that they are binding, and take themselves to have reason to follow them. But it does not follow that that reason is the fact of their acceptance.[119]

I think we can assume here that the "validity" Raz is talking about in the first part of the quote is legal validity, and that in talking of "reasons to follow" rules in the latter half he is again speaking of legal validity. In this passage, unlike in the passage from *Legal Validity*, Raz does not say that the legal validity of rules of recognition is a matter of social facts. Here, Raz seems to say, the truth or correctness of a legal statement or judgment is not ultimately a matter of certain convergent practices and the attendant attitudes of acceptance on the part of the community members existing. I think this is right and more accurate of Hart's commitments, whereas the position that Raz (most of the time), Dworkin, and others have attributed to Hart, the one that gave initial impetus to the characterization of his theory as committed to criterial semantics, is not really Hart's.

What Hart argued was that the existence of law in a community requires the existence in that community of a system of rules with some very specific kind of secondary rules and that the existence of such a system of rules requires the existence of convergent practice and the attendant attitudes of acceptance among the members of that community. Analogous things could have been said about the existence of a set of mores in a community or about the existence of numerous other kinds of norms such as those meant to regulate intellectual inquiries, artistic endeavors, etiquette, games, and rituals, etc. As far as I can see, what Hart said about the existence of law, and what analogous things that could be said about the existence of mores or some other kinds of norms, does not have any implication that the truth or correctness of the relevant normative statements is ultimately a matter of certain convergent practices and attitudes existing. In order to draw that inference, we would have to presuppose a normative bridge principle that says something like: "Do what your fellows do!" or "Do what the experts among your fellows do!" And I believe it is very difficult to see in *The Concept of Law* an argument, let alone a conceptual argument, for such a bridge principle. I want to be quite upfront here. There are a handful of passages in *The Concept of Law* that invite, or at least do not exclude, interpretations like Raz's and Dworkin's. But as I have argued elsewhere,[120] they fail to sustain those interpretations when read in the context of other relevant passages.

---

119. RAZ, *supra* note 1, at 381.

120. Toh, *supra* note 108, at 482–90.

As for theoretical legal disagreements, we should remember from Part II that Hart conceived of *internal legal statements*, the statements that participants in legal systems use to enunciate and interpret laws, as normative statements. If we had remembered this feature of Hart's legal theory, we would not have anticipated any problem for Hart in accounting for post-social–factual disagreements of the sort that Dworkin had in mind. What I suspect has caused much of the confusion is the failure to keep apart internal and external legal statements, as Hart urged us to do.[121] According to Hart, "what the law is" could be considered from two different points of view—from the point of view of a theorist, and from the point of view of a participant. Or better yet, there are two distinct concepts of what the law is, which play different explanatory roles. What the law is from the point of view of a theorist is ultimately a matter of what convergent practices prevail in a community. What the law is from the point of view of a participant, on the other hand, may exceed or transcend the prevailing convergent practices. In case that appears puzzling, think of how things are in morality. What is morally right from the point of view of a theorist (who is trying to determine the mores of a community) is ultimately a matter of what practices exist in a community, whereas what is morally right from the point of view of a participant can exceed or transcend prevailing convergent practices. If we keep in mind that there are two distinct concepts of "what is morally right" that play different explanatory roles, then we see that nothing is puzzling here. That is the case in law as well. In the case of morality, we have the terminological distinction between "mores" and "morality." Unfortunately, we do not have such a terminological distinction in law. That may be a reflection of a deep fact about law, and that would be the case if Raz were right. But I do not think that that is the case. Notice also that the distinction between internal and external legal statements enables us to make good sense of Justice Scalia's (projected) conception of what he is doing in advocating originalism.

What is undeniable is that in many cases, though not all, facts of the sort that a theorist of law observes from the external point of view in reaching his theoretical conclusions, what Raz calls "social facts," constitute the grounds in virtue of which participants' internal legal statements or judgments are true or correct. This is something that a legal theory should try to explain. Hart's legal theory, as I have interpreted it, does not offer such an explanation, whereas Raz's theory (and also Hart's theory as Raz interprets it) does. But Raz's is not the only explanation that we have available. Dworkin also has offered us an explanation of this fact. According to him,

---

121. HART, *supra* note 6, at 102–03; Eugenio Bulygin, *Norms, Normative Propositions, and Legal Statements*, *in* 3 CONTEMPORARY PHILOSOPHY: A NEW SURVEY 127, 136–48 (Guttorm Fløistad ed., 1982). Raz has in the past displayed some skepticism about Hart's distinction. RAZ, *Legal Validity*, *supra* note 52; Raz, *Purity*, *supra* note 34. I have argued elsewhere that Raz did not have good reasons for his skepticism. Toh, *supra* note 37, at 407–14.

legal judgments are constructive interpretations, which are roughly inferences from the best moral justifications of the existing societal standards and practices. I am not completely won over by Dworkin's explanation, partly because of its commitment to Acceptance Moralism. But in many ways, I believe, it is more accurate than Raz's explanation, and I suspect that it or something very much like it can be readily combined with Hart's theory about when a legal system exists or prevails among a community of people. Dworkin's explanation, unlike Raz's, can also explain the existence of genuine theoretical legal disagreements, and part of this latter explanation indicates that social facts cannot be the sole grounds in virtue of which internal legal statements or judgments are taken to be true or correct. And that is what we would have expected from Hart's conception of internal legal statements and judgments as normative statements and judgments.[122]

VIII. Conclusion

I am told by those who are more knowledgeable than I about these matters that the durability and the resulting popularity of products like the old Volkswagen Beetle, Kalashnikov rifles, and CD players manufactured in the 1980s stem mainly from the simplicity of their design and the small number of moving mechanical parts within them. I have a similar take on Hart's legal theory. It explains a lot with a small number of relatively durable philosophical components—which is not to say that such components could not be improved upon or supplemented. In the preceding pages, I have been somewhat skeptical of some of the different and additional moving parts that Raz has employed in his own theorizing about the nature of law. I have found many of these new parts less than well made and without obvious benefits. It is not clear to me how much of my skepticism has been driven by my desire to keep legal theorizing simple and uncluttered. Come to think of it, I have never cared for those CD players with five-CD changers.

122. Thanks to David Enoch for urging me to address the issues discussed in this paragraph. I have elsewhere tried to take some first steps in devising a conception of internal legal judgments that offers an explanation of why social facts play such an important role as the grounds of such judgments. Kevin Toh, *Legal Judgments as Plural Acceptances of Norms*, 1 OXFORD STUD. PHIL. L. (forthcoming 2010). My conception is intended to be very much Hartian in its foundations.

* * *

Notes

Post-9/11 Anti-terrorism Policy Regarding Noncitizens and the Constitutional Idea of Equal Protection Under the Laws*

> If you don't stick to your values when they're being tested, they're not values; they're hobbies.
>
> —Jon Stewart[1]

I. Introduction

The Bush Administration frequently justified its policies by arguing that September 11th changed everything.[2] To be sure, the terrorist attacks on the United States commanded a great deal of attention and brought concern for national security to the fore of both the government's policy-making agenda and America's shared consciousness. Against the psychological backdrop of the smoking Twin Towers, the country plunged headlong into two wars,[3] enacted sweeping legislation aimed at securing the homeland,[4] reorganized the government's administrative landscape,[5] and constructed an apparatus for detaining individuals who might pose a prospective threat to Americans' safety.[6] As this Note will detail, many of the policy changes that were made in response to the September 11th attacks affected noncitizens far more adversely than they did citizens. The primary focus of this Note will be to examine post-9/11 law as it pertains to noncitizens and comment on what

---

* I would like to thank Professor Robert Chesney for his guidance in developing this Note, and Professors Patrick Woolley and H.W. Perry for their invaluable comments and suggestions. Additionally, I would like to thank the Texas Law Review Editorial Board, particularly Jessica Miller, Nick Jackson, Brice Wilkinson, and Kat Hacker for their hard work in preparing the Note for publication.

1. *The Daily Show* (Comedy Central television broadcast Jan. 22, 2009).

2. *See, e.g.*, *Meet the Press* (MSNBC television broadcast Sept. 14, 2003) (broadcasting Vice President Dick Cheney's statement, "[T]he theme that comes through repeatedly for me is that 9/11 changed everything.").

3. *See* Bob Herbert, Editorial, *Wars, Endless Wars*, N.Y. TIMES, Mar. 3, 2009, at A27 (questioning the wisdom of American military policy and the wars in Afghanistan and Iraq).

4. *See, e.g.*, USA PATRIOT Act of 2001, Pub. L. No. 107-56, 115 Stat. 272 (codified as amended in scattered titles of U.S.C.) (enacting numerous reforms, including enhancing the government's surveillance abilities, expanding the Secretary of the Treasury's authority to regulate financial transactions, and refining the legal definition of "terrorism").

5. *See, e.g.*, Homeland Security Act of 2002, Pub. L. No. 107-296, 116 Stat. 2135 (codified as amended in scattered titles of U.S.C.) (consolidating several executive organizations, such as the Secret Service and the Coast Guard, under the umbrella of a single Cabinet-level agency).

6. *See* SARAH E. MENDELSON, CTR. FOR STRATEGIC INT'L STUDIES, CLOSING GUANTANAMO: FROM BUMPER STICKER TO BLUEPRINT 3–4 (2008) (reviewing the history of the Bush Administration's use of the U.S. base at Guantanamo Bay, Cuba as a detention facility).

these laws—and their reception in the legal world—say about the constitutional principle of equal protection: namely, that it carries relatively little weight.

The laws and policies of a democratic government necessarily reflect the values of that society. Accordingly, because the Constitution is the foundational legal document of the United States, it should embody the country's most deeply felt societal beliefs. As such, because the notion of equal protection is enshrined in the Constitution,[7] it might be expected that the United States places a high value on equality. Of course, before the Fourteenth Amendment was ratified, Thomas Jefferson famously invoked equality as a primary justification for American independence.[8] And since then the mantra of equality as an important American ideal has been taken up time and time again throughout the country's history.[9] However, the national-security-related policies, statutes, and case law adopted since September 11th that treat citizens differently from noncitizens betray a marked lack of commitment to equality norms. This differential treatment cannot even meaningfully be justified on national-security grounds, suggesting that America's professed devotion to equality rings hollow. Perhaps, then, with respect to the sanctity of the Equal Protection Clause and its broader meaning to the country, September 11th did not change anything. Rather, our response to the attacks simply serves to highlight the relative lack of importance placed on the constitutional principle of equality.

This Note will proceed in five substantive Parts. Part II will argue that there are some constitutional norms held in the highest regard, using Philip Bobbitt's *Constitutional Fate* as a jumping-off point for examining Supreme Court decisions that craft awkward doctrine in order to preserve certain constitutional ideals. These cases and norms will serve as points of contrast to the discussion of equal protection later in the Note. Part III will outline equal protection doctrine and theory in general, serving as the backdrop against which post-9/11 policy pertaining to noncitizens will be viewed in Part IV. Part V will examine those policies' constitutional merits under the Equal Protection Clause, focusing on the structure of the Fourteenth Amendment and its ratification history. And that Part will propose that Fourteenth Amendment jurisprudence in the context of anti-terrorism policies demonstrates a stark contrast between the Equal Protection Clause and material addressed in Part II. Accordingly, Part V will argue that the post-9/11 noncitizen experience casts a negative light on equal protection norms, showing that the notion of equality might not be as important an ideal as it is often claimed to be. Part VI will take a second look at particular

7. U.S. CONST. amend. XIV, § 1.

8. THE DECLARATION OF INDEPENDENCE para. 2 (U.S. 1776).

9. *See, e.g.*, Martin Luther King, Jr., Address at March on Washington: I Have a Dream (Aug. 28, 1963) (referring to the Declaration of Independence's call for equality as America's "creed").

terrorism policies and argue that they do not even appear to be "reasonable" in any meaningful sense. The Note will conclude by offering a brief normative assessment of the implications of its argument.

II. Constitutional Ideas and the Ways in Which the Supreme Court Stretches Doctrine to Preserve Them

The U.S. Constitution is more than the foundational body of laws that established American government and outlined the boundaries of its power. Its provisions not only set out the most basic guiding principles for the government but also ratify a set of fundamental values, some overtly stated[10] and some implicitly embedded,[11] that help make up the moral fabric of American society. Indeed, Ronald Dworkin claims that "the American ideal of government not only under law but under principle as well is the most important contribution our history has given to political theory."[12] Among these principles are popular participation in government,[13] various individual liberties,[14] and equality—given a voice in the Equal Protection Clause of the Fourteenth Amendment.[15] The mere fact of their inscription in the Constitution, in itself, indicates the reverence that American society has for those values—or, at least, has had for them in the past. Not all constitutional norms are created equal, however. Some values are so fundamental to the American conception of government and the boundaries of its power that the Court will tie itself up in knots in order to maintain them. Others, by contrast, garner far less judicial fidelity.

In his seminal work, *Constitutional Fate*, Philip Bobbitt highlights several constitutional principles of primary importance by noting that some of the Court's most significant decisions have been made based on reasoning that does not fall neatly into any of the conventional canons of constitutional interpretation.[16] Bobbitt identifies and describes the five commonly recognized modes of constitutional argument[17] before suggesting that a sixth mode, which he coins the "ethical argument,"[18] drives many Supreme Court decisions that seem unsatisfyingly reasoned.[19] By "ethical" he does *not* mean

---

10. *See, e.g.*, U.S. CONST. amend. I (establishing unambiguously freedoms of religion, speech, the press, the right to assemble, and the right to petition the government).

11. *See, e.g.*, Griswold v. Connecticut, 381 U.S. 479, 484 (1965) (finding the right to privacy to be implicitly protected by the Bill of Rights).

12. RONALD DWORKIN, FREEDOM'S LAW: THE MORAL READING OF THE AMERICAN CONSTITUTION 6 (1996).

13. U.S. CONST. amends. I, XV, XIX, XXVI.

14. *See, e.g., id.* amend. II (recognizing the right of the people to bear arms).

15. *Id.* amend. XIV, § 1.

16. PHILIP BOBBITT, CONSTITUTIONAL FATE (1982).

17. *See id.* at 1–92 (examining the textual, historical, doctrinal, prudential, and structural modes of constitutional argument).

18. *Id.* at 94.

19. *See, e.g., id.* at 137 ("Even the most gullible student is reluctant to accept the doctrinal justification in, say, *Shapiro v. Thompson* that welfare residency requirements are unconstitutional

"moral," but rather "constitutional argument whose force relies on a characterization of American institutions and the role within them of the American people. It advances the character, or *ethos*, of the American polity."[20] The foundation of that *ethos*, he contends, and thus the force behind the ethical argument, is the notion that ours is a government of limited powers, an idea that is inherent in and crucial to the American Constitution.[21]

In this way Bobbitt explains a whole host of Supreme Court decisions that, despite being of tremendous constitutional importance, appear to be awkwardly reasoned at best. Consider, for example, one of the Court's most controversial decisions of the twentieth century: *Roe v. Wade*.[22] Although the Constitution does not explicitly protect a right to privacy,[23] the Court in *Roe*, ostensibly relying on precedent, held a woman's right to choose to have an abortion to be encompassed by such a constitutional right.[24] Bobbitt argues that because "[t]he two principal propositions on which [the decision] rests are neither derived from precedent nor elaborated from larger policies that may be thought to underl[ie] such precedent," the decision "is a doctrinal fiasco."[25] Bobbitt does, however, suggest that by looking through the lens of the "ethical" approach, the decision can be easily explained. He argues that the Court, in reality, likely based its decision on reasoning along the following lines: "Government may not coerce intimate acts,"[26] a notion that flows naturally from the principle of limited government.[27] As is typically the case with ethical reasoning, the Court was unwilling to admit to such a basis for its decision,[28] thus leading to the problematic reasoning put forth in the majority opinion.

Bobbitt's evaluation of *Roe* is elegantly simple, and (to this reader) convincing. But, regardless of whether the ethical argument provides a satisfactory explanation for it, it is plain that in *Roe* the Court stretched to apply a constitutional principle that finds little clear support in the

---

because they interfere with the equal protection to be afforded *travel*." (citing 394 U.S. 618 (1969))).

20. *Id.* at 94.

21. *See id.* at 118, 150 (discussing the role of the idea of a limited federal government in shaping the ethical argument).

22. 410 U.S. 113 (1973).

23. *See, e.g.*, Griswold v. Connecticut, 381 U.S. 479, 508 (1965) (Black, J., dissenting) ("The Court talks about a constitutional 'right of privacy' as though there is some constitutional provision forbidding any law to be passed which might abridge the 'privacy' of individuals. But there is not.").

24. *Roe*, 410 U.S. at 154.

25. BOBBITT, *supra* note 16, at 159.

26. *Id.*

27. *Id.* at 160.

28. *See id.* at 163 (explaining that the Court was so specific in its remedies in *Roe* because such an approach is "virtually forced on a court that is unwilling to make explicit the constitutional rules by which it arrives at a decision").

Constitution itself. Accordingly, it may be said that the right to privacy (manifested, in this context, in a right to reproductive autonomy) is a constitutional principle that garners a deep fidelity. *Roe*'s awkward reasoning suggests that this is likely true whether or not that principle has its roots in a largely unrecognized—if compelling—mode of constitutional interpretation based on the Constitution's appeal to the principle of limited government. And so, despite the fact that the result hardly put the issue of abortion to rest,[29] *Roe v. Wade* demonstrates that privacy is a constitutional ethic of the utmost importance.[30]

Not all poorly reasoned—yet crucially important—Supreme Court decisions are controversial. Despite being fully sympathetic to the result, Bobbitt criticizes the Court's reasoning in *O'Connor v. Donaldson*[31] as being "elliptical."[32] In that case, Kenneth Donaldson, who had been committed for almost fifteen years in a mental facility, challenged the conditions of his confinement.[33] The Supreme Court held that because Donaldson was not dangerous, the Constitution did not allow him to be held against his will, despite his mental illness.[34] Bobbitt argues that to the extent that there was any real reasoning underlying the Court's decision at all, it was wrongly based on Fourteenth Amendment procedural-due-process grounds.[35] He sees the Court's opinion as an unartful conflation of the deprivation of liberty without due process of law with the deprivation of liberty in itself (regardless of the process afforded).[36] He contends that the holding can be much more easily justified on the basis of the simple principle that the government may

29. Quite the contrary, in fact. *See* KRISTIN LUKER, ABORTION AND THE POLITICS OF MOTHERHOOD 144 (1984) ("[T]he *Roe* decision would mobilize a new and much stronger opposition to abortion reform.").

30. It is true that if *Roe v. Wade* were decided by the current Roberts Court, instead of a Burger Court full of Warren Court holdovers, the result might well have been different. As such, one may protest that the case is a poor vehicle for the argument that there are some constitutional principles garnering deep judicial fidelity despite having scant clear foundation in the Constitution itself. But seven Justices, with only two dissenting, joined the opinion of the Court in *Roe*. Roe v. Wade, 410 U.S. 113, 115 (1973). That the reasoning got *any* support—let alone that of a significant majority—despite failing to meaningfully anchor itself in an apparent constitutional principle—bolsters the argument of the importance of privacy. If *Roe* came before the Court today as a case of first impression and only four Justices were to discover a constitutional principle in the manner that the actual *Roe* Court did, it would still demonstrate the importance placed on that principle, even if those Justices could not muster a majority. Furthermore, *Roe* survived an attack in *Planned Parenthood of Southeastern Pennsylvania v. Casey*, 505 U.S. 833 (1992), and will likely continue to be good law in spite of its problematic reasoning. *Roe* is established in American jurisprudence. Thus, it seems fair to use the case to support the argument advanced herein, in spite of the fact that Supreme Court personnel have changed in such a way as to call into question how it would have been treated today.

31. 422 U.S. 563 (1974).

32. BOBBITT, *supra* note 16, at 165.

33. *Donaldson*, 422 U.S. at 564–65.

34. *Id.* at 576.

35. BOBBITT, *supra* note 16, at 165–66.

36. *Id.* at 165.

not incarcerate someone who poses no danger to himself or to others.[37] This principle has its origins (of course) in the broader constitutional principle that the government is one of limited, specifically enumerated powers.[38]

Here again, Bobbitt's interpretation of the case seems not only plausible but also satisfying. But, even if he is wrong that the ethical argument served as the true basis for the Court's decision, his analysis sheds light on the fact that there are some principles that the Court finds in the Constitution but must struggle to justify having found. Just as *Roe* demonstrates a fidelity to an inherent constitutional right of privacy, *Donaldson* shows us that the constitutional norm of liberty extends beyond the textual boundaries of the Fifth and Fourteenth Amendments. In order to reach a decision consistent with that constitutional norm, the Court was forced to rely on reasoning that was somewhat dubious, even though the result was probably never in doubt.[39] The Court's positions in *Roe* and *Donaldson* suggest that some constitutional principles, such as privacy and liberty, are held in the highest regard.[40] Parts IV–VI will attempt to show that the principle of equality does not enjoy that level of reverence.

III. Equal Protection Doctrine and Theory

The constitutional rights of privacy and liberty (over and above the protections of the Due Process Clause) have no clear textual support in the Constitution itself. As such, Court decisions that uphold these rights indicate their paramount importance. At least by comparison, the right to equal protection is apparent and unambiguous: "No state shall deny to any person within its jurisdiction the equal protection of the laws."[41] Thus, unlike the right to privacy, the Court need not reach to find the principle of equality in the Constitution. Yet, in spite of the relative clarity of the clause's purpose, equal protection doctrine has not evolved to be as simple as an initial reading of the Constitution might make it seem.

---

37. *Id.* at 166.

38. *Id.*

39. The Court was unanimous in its decision. *Donaldson*, 422 U.S. at 573–75.

40. That list is not exhaustive. There are several constitutional norms that command such importance that the Court will uphold them without feeling particularly bound by typical constraints, such as plain-meaning textualism. First Amendment norms provide another such example. *See, e.g.*, BOBBITT, *supra* note 16, at 101 (noting that in New York Times Co. v. United States, 403 U.S. 713 (1971), in which the *Times* contested the press-hindering actions of the *Executive* Branch, "the Court applied conventional First Amendment analysis, despite the clear terms of that Amendment limiting the powers of *Congress*").

41. U.S. CONST. amend. XIV, § 1. Although by its terms the Fourteenth Amendment restricts states only, it is a matter of well-established jurisprudence that the federal government is also bound by Fourteenth Amendment requirements by way of that Amendment's incorporation into the Fifth Amendment's Due Process Clause. *See* Bolling v. Sharpe, 347 U.S. 497, 499–500 (1954) (applying Fourteenth Amendment equal protection doctrine via the Fifth Amendment to hold that the District of Columbia must integrate its public schools); *see also* Adarand Constructors, Inc. v. Peña, 515 U.S. 200, 213, 235 (1995) (applying, in a claim arising under the Fifth Amendment, the same standard of review to federal and state government racial classifications).

Instead, the Court has developed a three-tiered method of review that it employs when evaluating the constitutionality of a law under the Equal Protection Clause. In *United States v. Carolene Products Co.*,[42] the Court announced that it would analyze certain types of classifications more closely than others.[43] Out of *Carolene Products*' Footnote Four the Court crafted a body of equal protection doctrine that submits laws that adversely affect certain groups—or infringe on a "fundamental right"—to "strict scrutiny";[44] laws adversely affecting other groups to "intermediate scrutiny";[45] and all other laws to "rational-basis review."[46] Strict scrutiny is a very difficult standard to meet, requiring a showing that the law at issue be "narrowly tailored" to satisfy a "compelling" governmental interest.[47] Intermediate scrutiny is somewhat more deferential, requiring a showing that the law or policy is "substantially related" to an "important" governmental interest.[48] And rational-basis review is highly deferential: laws are almost never overturned when put to this test.[49] Under rational-basis review a law merely must be shown to be "rationally related" to a "legitimate" governmental interest.[50] It is apparent, then, that the survival of a law under an equal protection analysis will depend heavily on which level of scrutiny a court employs.

Thus, equal protection doctrine makes fairly complicated what seems at first blush to be a simple principle, and it does so in a way that may be unfair to plaintiffs relegated to the rational-basis review category because it

---

42. 304 U.S. 144 (1938).

43. *See id.* at 153 n.4 (leaving open the possibility that the Court would submit some types of legislation to more exacting scrutiny than it would others, depending on the group affected by it).

44. *See* Maher v. Roe, 432 U.S. 464, 470 (1977) ("'We must decide, first, whether [state legislation] operates to the disadvantage of some suspect class or impinges upon a fundamental right explicitly or implicitly protected by the Constitution, thereby requiring strict judicial scrutiny.'" (quoting San Antonio Indep. Sch. Dist. v. Rodriguez, 411 U.S. 1, 17 (1973))).

45. *See, e.g.*, United States v. Virginia, 518 U.S. 515, 575 (1996) (describing intermediate scrutiny for gender-based classifications).

46. *See, e.g.*, United States v. Lopez, 514 U.S. 549, 606 (1995) (discussing the emergence of the rational-basis test).

47. *See, e.g.*, Parents Involved in Comm. Sch. v. Seattle Pub. Sch. Dist. No. 1, 551 U.S. 701, 704–05 (2007) (asserting that government action dividing people by race is inherently suspect and that the allegedly compelling state interest of diversity in education could not justify the school districts' use of racial classifications in student assignment plans).

48. *See, e.g.*, United States v. Virginia, 518 U.S. at 516 (noting that for a gender-based classification to withstand equal protection scrutiny, it must be established "at least that the [challenged] classification serves important governmental objectives and that the discriminatory means employed are substantially related to the achievement of those objectives" (internal quotations omitted)).

49. *See, e.g.*, FCC v. Beach Commc'ns, Inc., 508 U.S. 307, 314–15 (1993) (characterizing rational-basis review as "a paradigm of judicial restraint"); *Developments in the Law—Legal Responses to Domestic Violence*, 106 HARV. L. REV. 1498, 1570 (1993) (noting that rational-basis review is generally "something of a rubber stamp on governmental action").

50. *See, e.g.*, Vance v. Bradley, 440 U.S. 93, 97 (1979) (emphasizing that the Court will overturn a legislative classification only if it "is so unrelated to the achievement of any combination of legitimate purposes that [the Court] can only conclude that the [government's] actions were irrational").

severely limits their opportunity to prove out an equal protection violation. Certainly, then, the Equal Protection Clause does not stand for an absolute grant of equality—a prohibition of any classification by the states or the federal government. Nor could it. After all, a government must be able to classify and differentiate in order to govern, and so it must be given some deference in constitutional review in order to do so effectively.[51] Indeed, it would be nonsensical for the Equal Protection Clause to guarantee a six-year-old child the right to vote and drink alcohol just because adults may do so, for example. Accordingly, the Equal Protection Clause can only proscribe unfair discrimination between those who are "similarly situated."[52]

So, even if somewhat complicated, the *Carolene Products* framework makes sense. It grants leeway to government so that it can function properly, while at the same time it allows courts to take a hard look at policies that target historically disadvantaged groups.[53] Furthermore, in this second respect, the *Carolene Products* framework represents a vast improvement over prior equal protection doctrine, which left much to be desired in its ability to ensure the civil rights of African Americans and other minorities.[54]

Perhaps recognizing the Equal Protection Clause's ineffectual sway over such issues at the time, Justice Oliver Wendell Holmes once labeled equal protection the "last resort of constitutional arguments."[55] However, in part because of the *Carolene Products* framework, equal protection has not only become a viable claim but also a major mechanism for challenging discriminatory laws and policies.[56] In particular, the Equal Protection Clause was a powerful tool in ensuring equal rights for African Americans in the second half of the twentieth century. For example, in *Brown v. Board of Education*[57] the Court reversed course on *Plessy*'s "separate but equal" approach and relied on the Equal Protection Clause to compel integration in public schools;[58] and in *Loving v. Virginia*[59] the Court relied on the Equal

51. *See, e.g.*, DWORKIN, *supra* note 12, at 159 ("If the equal protection clause forbade *any* classification by groups, all such legislation would be unconstitutional, which is absurd.").

52. *See* Joseph Tussman & Jacobus tenBroek, *The Equal Protection of the Laws*, 37 CAL. L. REV. 341, 344 (1949) ("The measure of the reasonableness of a classification is the degree of its success in treating similarly those similarly situated.").

53. *See* United States v. Carolene Prods. Co., 304 U.S. 144, 153 n.4 (1938) (identifying classifications affecting "discrete and insular minorities" as deserving of heightened scrutiny).

54. *See, e.g.*, Plessy v. Ferguson, 163 U.S. 537, 548 (1896) (establishing the doctrine of separate but equal and holding that the principle did not deprive African Americans of "equal protection of the laws").

55. Buck v. Bell, 274 U.S. 201, 208 (1927).

56. *See* ALEXANDER TSESIS, WE SHALL OVERCOME: A HISTORY OF CIVIL RIGHTS AND THE LAW 251 (2008) (describing the *Carolene Products* Footnote Four analysis as enabling "the Warren Court, which sat from October 1953 to June 1969, to change American civil, social, and political dynamics").

57. 347 U.S. 483 (1954).

58. *Id.* at 483.

59. 388 U.S. 1 (1967).

Protection Clause to hold anti-miscegenation statutes unconstitutional.[60] In addition, the Equal Protection Clause has become a powerful tool in securing equal rights for women,[61] among other groups.[62]

In light of equal protection's influence on major issues of the twentieth century, it is unsurprising and understandable that it has come to be regarded as representing a sweeping and robust constitutional principle. Larry Sager, for example, has written that the norm of equal protection has meaning over and above the parameters of equal protection doctrine.[63] He argues that the doctrine has been bounded by institutional constraints—such as deference to other branches of government—rather than analytical ones.[64] Sager thus contends that the Equal Protection Clause "should be understood to be legally valid to [its] full conceptual limit[]," even though the doctrine does not reach that far.[65] Implied throughout Sager's piece is the notion that the "full conceptual limit" is expansive.[66] He argues that lawmakers are bound by the full force of equal protection, despite the Court's unwillingness to enforce it to that limit.[67]

Even more, Ronald Dworkin considers the Equal Protection Clause to represent an "abstract moral principle[]" with a hugely broad scope.[68] To Dworkin, the Equal Protection Clause not only extends beyond the boundaries of the existing doctrine it has shaped but also stands for a right to "equal concern and respect" for everyone within the purview of the Constitution.[69] He thus considers the scope of the Fourteenth Amendment to be "breathtaking."[70] He does not consider the Equal Protection Clause to have *established* this right, however. Rather, he sees the right to equal concern and respect as a bedrock principle of American government, implicit in the

60. *See id.* at 12 ("There can be no doubt that restricting the freedom to marry solely because of racial classifications violates the central meaning of the Equal Protection Clause.").

61. *See* United States v. Virginia, 518 U.S. 515, 515–16 (1996) (holding that the male-only admissions policy at the Virginia Military Institute violated the Equal Protection Clause).

62. *See, e.g.*, City of Cleburne v. Cleburne Living Ctr., Inc., 473 U.S. 432 (1985) (holding that the City of Cleburne discriminated against mentally handicapped persons in violation of the Equal Protection Clause when it denied a permit required for their group home on the basis of their mental disability).

63. *See* Lawrence Gene Sager, *Fair Measure: The Legal Status of Underenforced Constitutional Norms*, 91 HARV. L. REV. 1212, 1218 (1978) (noting the "disparity between [the current] construct and a true conception of equal protection," which indicates that "equal protection is an underenforced constitutional norm").

64. *See id.* at 1221 ("[A] federal judicial construct is found not to extend to certain official behavior because of institutional concerns rather than analytical perceptions . . . .").

65. *See id.*

66. *See id.* at 1264 ("I have offered . . . the view that constitutional norms which are significantly underenforced by the federal judiciary nevertheless ought to be regarded as legally valid to their conceptual boundaries.").

67. *Id.* at 1221.

68. DWORKIN, *supra* note 12, at 7.

69. *Id.* at 72.

70. *Id.*

Constitution.[71] The Equal Protection Clause, then, can be seen as reflecting and reinforcing this foundational principle—expressing it and giving weight to it in a way that is not the case for the extra-textual principle of privacy, for example.

If Sager and Dworkin are correct about the scope and meaning of the provision, it seems that the Equal Protection Clause—like the elusive emanations and penumbras that support the right to privacy[72]—stands for a vibrant constitutional ethic in its own right.

IV. Post-9/11 Anti-terrorism Policy Regarding Noncitizens

If the Equal Protection Clause embraces a fundamentally important constitutional principle in the way that Sager and Dworkin suggest,[73] and in the manner in which the Due Process Clauses of the Fifth and Fourteenth Amendments represent a substantive norm of liberty *in itself*,[74] then consideration for equal protection might be expected to pervade lawmaking as well as court decisions. But, despite equal protection's rise in prominence in the second half of the twentieth century, a question arises as to whether equal protection norms are as meaningful as those of privacy and liberty. The latter have been vindicated as critical by awkwardly reasoned Supreme Court opinions. So, does equal protection receive the same kind of deference accorded the principles of privacy and liberty? In order to investigate the relative vivacity of the equal protection ethic, the focus of this Note narrows to examine a particular area of law in which equal protection may be expected to be implicated: twenty-first-century counterterrorism policy.

September 11th precipitated a sea change in American national-security law. The terrorist attacks engendered a perception of vulnerability, which understandably spurred broadly sweeping legislation aimed at protecting the country from further harm.[75] However worthy their aims, though, running throughout many of the new policies that the United States adopted in the wake of 9/11 was a systematic differentiation between American citizens and aliens that appears to be unbounded by any concern for the aliens' constitutional right of equal protection under the laws. Certainly, any consideration of the noncitizens' right to "equal concern and respect"[76] seems conspicuously absent from the policy-making agenda.

---

71. *See id.* at 17 (arguing that the "constitutional conception of democracy" envisions a system wherein all members of the political community are treated with "equal concern and respect").

72. *See* Griswold v. Connecticut, 381 U.S. 479, 484 (1965) (arguing that the "specific guarantees in the Bill of Rights have penumbras, formed by emanations from those guarantees that help give them life and substance").

73. *See supra* notes 63–71 and accompanying text.

74. *See supra* notes 31–38 and accompanying text.

75. *See supra* notes 2–5 and accompanying text.

76. DWORKIN, *supra* note 12, at 17.

Consider, for example, the USA PATRIOT Act.[77] The giant bill is replete with unjustified distinctions between citizens and aliens. For example, "for citizens, terrorism has a limited definition that roughly corresponds to common understandings of the phenomenon, whereas for foreign nationals, Congress has labeled 'terrorist' wholly nonviolent activity and ordinary crimes of violence."[78] This broader definition of terrorism applied to noncitizens has monumental implications for foreign nationals under the law—aliens may be deported for engaging in activity that is ordinarily protected by the First Amendment.[79] In this way, under the USA PATRIOT Act noncitizens do not enjoy the benefits of the Bill of Rights to the same extent as do citizens.

The disparate treatment of noncitizens does not end there. The USA PATRIOT Act also allows the Executive Branch to preventatively, and indefinitely, detain aliens who fall under the expansive, alien-specific definition of "terrorist."[80] In order to prevent people who might pose a threat to America's national security from carrying out or aiding in a terrorist attack, the Attorney General's ability to detain aliens under immigration proceedings was expanded.[81] However, as David Cole notes, the USA PATRIOT Act vests in the Attorney General the power to continue to detain noncitizens even after they have prevailed in their immigration proceedings[82] and thus no longer seem to pose a threat.

While the immigration-related detention practices may seem unfair, by definition they *must* apply only to noncitizens. However, differential treatment regarding preventative detention has not been limited to the immigration context. Two months after the 9/11 attacks, President Bush issued a military order asserting the authority to detain aliens who presented a threat to the national security of the United States because of membership in al Qaeda or some other relevant affiliation, motive, or capacity to do harm.[83] The order was justified on the basis of the national emergency precipitated by the terrorist attacks of September 11th,[84] but there was no explanation given as to why it should apply only to *foreigners* who posed such a threat, rather than to *anyone* who posed such a threat, except for the

77. Pub. L. No. 107-56, 115 Stat. 272 (2001).

78. DAVID COLE, ENEMY ALIENS: DOUBLE STANDARDS AND CONSTITUTIONAL FREEDOMS IN THE WAR ON TERRORISM 58 (2003).

79. *See id.* ("[The Act] makes foreign nationals deportable for wholly innocent associational activity . . . .").

80. *See id.* at 65 (describing the circumstances under which the Attorney General may incarcerate noncitizens under the Act).

81. *Id.*

82. *Id.* at 66.

83. Military Order of Nov. 13, 2001, Detention, Treatment, and Trial of Certain Non-citizens in the War Against Terrorism, 66 Fed. Reg. 57,833, 57,834 (Nov. 16, 2001).

84. *Id.* at 57,833.

assertion that the perpetrators of the 9/11 attacks were "international terrorists."[85]

That distinction, fundamental to the order, proved to be as shortsighted as it was unexplained. In order to detain Jose Padilla—an American citizen who was associated with al Qaeda and allegedly planned to detonate a "dirty bomb" in Chicago[86]—without criminal charges, the government was forced to revise its policy to include citizens and aliens alike.[87] And, but for the Padilla wrinkle, the government would still have had to revise its policy after discovering that a Guantanamo detainee, Yaser Esam Hamdi, was a natural-born American citizen.[88] Although the government took a different tack with respect to preventative detention in light of the cases of Padilla and Hamdi, it remained steadfast in drawing a distinction between the rights of citizens and those of noncitizens. After all, despite being labeled "enemy combatants," Padilla and Hamdi were nonetheless afforded relatively speedy trials, whereas hundreds of noncitizen detainees were repeatedly denied any semblance of due process.[89]

Justice O'Connor, for a plurality of the Supreme Court in *Hamdi*, wrote that the government may detain citizens as unlawful enemy combatants.[90] The opinion goes on to say, however, that "a *citizen*-detainee seeking to challenge his classification as an enemy combatant must receive notice of the factual basis for his classification, and a fair opportunity to rebut the Government's factual assertions before a neutral decisionmaker."[91] Thus, the plurality's decision makes clear that the Court, like the other branches of the federal government, identified a meaningful constitutional difference between citizens and noncitizens, and was seemingly untroubled by any

85. *Id.*

86. *See* Padilla v. Hanft, 423 F.3d 386, 388 (4th Cir. 2005) (summarizing the background of the case); Philip Shenon & James Risen, *Traces of Terror: The Investigation; Terrorist Yields Clues to Plots, Officials Assert*, N.Y. TIMES, June 12, 2002, at A1 (describing the efforts to find and arrest Padilla in connection with a suspected "dirty bomb" plot).

87. *See* Order from President George W. Bush to Donald Rumsfeld, Sec'y of Def., to Detain Jose Padilla (June 9, 2002), *available at* http://news.findlaw.com/hdocs/docs/padilla/padillabush60902det.pdf (basing the authority to detain an American citizen on the Authorization for Use of Military Force, Pub. L. No. 107-40, 115 Stat. 224 (2001) (codified at 50 U.S.C. § 1541 note (2006))).

88. *See* Hamdi v. Rumsfeld, 542 U.S. 507, 510 (2004) (plurality opinion) (noting that Hamdi was born in Louisiana and was an American citizen when detained).

89. *See* Military Commissions Act of 2006, Pub. L. No. 109-366, sec. 3, 120 Stat. 2600, 2600–31 (codified at scattered sections of 10 U.S.C. (2006)) (establishing, several years after detentions began, a system by which alien detainees were to be tried that had few traditional protections resembling those in the American court system).

90. *Hamdi*, 542 U.S. at 519.

91. *Id.* at 533 (emphasis added). But, the Court did extend the right to habeas to *all* Guantanamo detainees a few years later in *Boumediene v. Bush*, 128 S. Ct. 2229, 2275 (2008).

notion that the Equal Protection Clause of the Fourteenth Amendment might prohibit the government from making such a distinction.[92]

V. Equal Protection Applied

Read in the context of the entire first section of the Fourteenth Amendment,[93] the Equal Protection Clause plainly applies to citizens and noncitizens alike. The amendment declares that all natural-born Americans are citizens, affording all such people the full set of "privileges and immunities" of citizenship.[94] The amendment goes on to guarantee due process and equal protection to all "persons" within the Constitution's jurisdiction, thereby expressly reserving these rights for more than just the American citizenry.[95] Indeed, the Supreme Court has explicitly held that the term "persons" in the Equal Protection Clause includes noncitizens.[96]

In this way, the structure of the Fourteenth Amendment could be read to imply that differentiation on the basis of citizenship status, without more, constitutes an inappropriate classification under the Equal Protection Clause. After all, if the framers intended to bring noncitizens under the purview of the Equal Protection Clause, then their intent regarding aliens' status in American society might be frustrated by policies that draw a distinction between citizens and aliens without meaningfully justifying that distinction. Alternatively—and somewhat more modestly—such a classification could be seen as presumptively invalid for the same reason. Or, more modestly still, the Fourteenth Amendment's structure might lead to the inference that laws discriminating on the basis of citizenship status should be subject to heightened judicial scrutiny. If this interpretation of the Equal Protection Clause were doctrinally salient, it would call into question the constitutionality of all of the laws and policies addressed in Part IV.

Courts do not appear to have drawn any such inferences from the Fourteenth Amendment, however; as a result, the policies addressed in Part IV would likely survive constitutional challenge. Instead, courts have held alienage not to be a suspect classification when a federal law is at issue,

---

92. For its part, the dissent drew an even clearer distinction, claiming that a detention such as Hamdi's was illegal without an explicit suspension of the writ of habeas corpus but maintaining that its position in this regard applied only to citizens. *Hamdi*, 542 U.S. at 577 (Scalia, J., dissenting).

93. The text of the amendment reads:

> All persons born or naturalized in the United States, and subject to the jurisdiction thereof, are citizens of the United States and of the State wherein they reside. No State shall make or enforce any law which shall abridge the privileges or immunities of citizens of the United States; nor shall any State deprive any person of life, liberty, or property, without due process of law; nor deny to any person within its jurisdiction the equal protection of the laws.

U.S. CONST. amend. XIV, § 1.

94. *Id.*

95. *Id.*

96. Plyler v. Doe, 457 U.S. 202, 210 (1982).

and thus federal laws discriminating against noncitizens are subject only to rational-basis review.[97] So, in order for any of the laws or policies described above to withstand a constitutional challenge, the United States would merely have to show that the law is rationally related to a legitimate governmental interest.[98] The government would certainly meet this requirement, as national security is undoubtedly a governmental interest of the highest order, and "rationally related" is a very low threshold to reach.[99] Indeed, simply asserting—without even showing—that a terrorist is more likely to be an alien than a citizen likely would be more than enough to satisfy this requirement.[100] Because rational-basis review is such a lenient check on governmental action, noncitizens effectively have no equal protection claim against the United States for their differential treatment under twenty-first-century counterterrorism law.

It is troubling that this differential treatment of noncitizens would survive constitutional challenge, not only because the structure of the Fourteenth Amendment seems to preclude such differential treatment, but also because its specific purpose was to invalidate Southern Reconstruction laws that bore a striking resemblance to some of America's current counterterrorism policies in one important respect. Although the meaning of the Equal Protection Clause was not widely agreed upon by the framers,[101] the central, immediate purpose of the Fourteenth Amendment and its guarantee of equal protection was clear. Above all, the framers of the Fourteenth Amendment intended to eliminate the Black Codes.[102] The Codes, adopted in the wake of emancipation and the ratification of the Thirteenth Amendment, drew the attention of the framers as especially pernicious policies.[103] Among other things, the Black Codes singled out newly freed

97. *See, e.g.*, United States v. Ferreira, 275 F.3d 1020, 1025 (11th Cir. 2001) ("Appellants' argument is grounded on the erroneous foundation that congressional classifications based on alienage are subject to strict scrutiny."). *But see* Graham v. Richardson, 403 U.S. 365, 371–72 (1971) (reviewing Supreme Court precedent to note that alienage is a suspect classification when a *state* law is at issue).

98. *See Ferreira*, 275 F.3d at 1025 ("In other words, Congress can pass laws regulating the conduct of non-citizens within the United States, and those laws do not violate equal protection so long as they are rationally related to a legitimate government interest.").

99. *See supra* note 49 and accompanying text.

100. *See* Mathews v. Diaz, 426 U.S. 67, 82 (1976) (noting that decisions made by Congress or the President in the area of immigration and naturalization are subject to "a narrow standard of review" equivalent to rational-basis scrutiny).

101. *See* ANDREW KULL, THE COLOR-BLIND CONSTITUTION 68 (1992) (observing that the framers of the Fourteenth Amendment chose the language of equality because its meaning was disputed in 1866, just as it remains today).

102. *See id.* at 81 (examining proposed drafts of what would become the Fourteenth Amendment and concluding that either draft would have accomplished the objective of eliminating the Black Codes).

103. *See, e.g.*, CONG. GLOBE, 39th Cong., 1st Sess. 475 (1866) (statement of Sen. Lyman Trumbull) (characterizing the Black Codes as being "in violation of the rights of a freeman").

slaves for special and unequal punishment for criminal violations.[104] For example, the Mississippi Black Codes provided for substantial monetary penalties and possible jail time for African-Americans who, for instance, "exercis[ed] the function of a minister of the Gospel without a license from some regularly organized church."[105] Such penalties were not imposed on whites for committing the same or similar offenses.[106]

In this respect, some of America's twenty-first-century counterterrorism laws resemble the Black Codes. An alien who commits an "ordinary crime of violence" may be labeled a terrorist, whereas a citizen would not, and thus, that alien may be subject to harsher penalties than would be a citizen who committed the same crime.[107] Further, at least until *Hamdi* and *Padilla* caused the Bush administration to modify its approach, only a noncitizen could be labeled an enemy combatant and be subject to indefinite detention.[108] Notwithstanding this important similarity between current policies affecting noncitizens and the policies that the Fourteenth Amendment was specifically designed to eliminate, the Equal Protection Clause provides little recourse for adversely affected aliens because—in this context—alienage is not a suspect classification.[109]

The post-9/11 terrorism policies' systematic differentiation between citizens and noncitizens, coupled with an unsympathetic equal protection doctrine that presents a nearly insurmountable obstacle to challenging federal laws on the basis of that differentiation, provides a stark contrast to the privacy and liberty case law addressed in Part II. This contrast offers a hint as to where equal protection norms fit into the fabric of American constitutionalism.

There are some constitutional ideas that are so engrained in the national consciousness that the Court maintains a fidelity to those ideas even when constitutional support is hard to come by.[110] On the other hand, unlike somewhat vague notions of privacy, for example, the principle of equal protection is explicitly written into the Constitution.[111] For that reason, the Court need

---

104. KULL, *supra* note 101, at 76.

105. Miss. Black Codes, Act of Nov. 25, 1865, ch. 4, § 2, 1865 Miss. Laws 82, *invalidated by* U.S. CONST. amend. XIV; *see also* Paul Finkelman, *"Let Justice Be Done, Though the Heavens May Fall": The Law of Freedom*, 70 CHI.-KENT L. REV. 325, 354–61 (1994) (detailing Mississippi's legislation and the subsequent federal response).

106. Miss. Black Codes, ch. 4, § 2 (limiting its application to "any freedman, free Negro, or mulatto").

107. *See supra* notes 78–79 and accompanying text.

108. *See supra* note 82 and accompanying text.

109. *See* United States v. Ferreira, 275 F.3d 1020, 1025 (11th Cir. 2001) (holding that because of Congress's broad power to regulate immigration, federal laws discriminating based on alienage do not violate the Equal Protection Clause if "they are rationally related to a legitimate government interest"); *see also* Matthews v. Diaz, 426 U.S. 67, 78 (1976) ("[A] legitimate distinction between citizens and aliens may justify attributes and benefits for one class not accorded to the other.").

110. *See supra* Part II.

111. U.S. CONST. amend. XIV, § 1.

not bend over backwards to preserve a national ethic of equality when the occasion arises. Sometimes, however, it appears to do just the opposite, turning a blind eye to issues that—at least arguably—raise equal protection concerns. Indeed, relying on either the structure of the Fourteenth Amendment or its specific purpose of eliminating the unequal punishment fostered by the Black Codes would provide the Court with a relatively clear path to strike down federal counterterrorism policies that differentiate between citizens and aliens. And yet, the Supreme Court seems largely sympathetic to the counterterrorism policies addressed in Part IV, notwithstanding the equal protection concerns that they might raise.[112]

Furthermore, the three-tiered framework of constitutional review that the Court has constructed regarding the Equal Protection Clause makes successfully challenging those policies on equal protection grounds all but out of the question. Therefore, despite its noble origins and frequent, impassioned invocation as a fundamental principle of American society, it seems that the constitutional norm of equal protection is not one on which the country places particularly high value.

VI. The Reasonableness of the Counterterrorism Policies

The structure and the specific purpose of the Fourteenth Amendment suggest that, at least in the terrorism context, federal classification on the basis of citizenship status should be subject to a more searching scrutiny under the Equal Protection Clause than rational-basis review. But there are compelling arguments of institutional competence that justify rational-basis review as appropriate. In *Mathews v. Diaz*,[113] for example, the Supreme Court explained that for reasons of foreign relations, politics, and economics, decisions regarding noncitizens "are frequently of a character more appropriate to either the Legislature or the Executive than to the Judiciary."[114] Sager might say that this concession to institutional competence should not be seen as an indictment of the equal protection principle because the Clause is still

112. *See supra* notes 90–92 and accompanying text. It is true that it is difficult to draw an inference as to the Court's position on equal protection for noncitizens in the context of anti-terrorism law from a case—*Hamdi v. Rumsfeld*—where the Equal Protection Clause was not in issue. *See* 542 U.S. 507, 525 (2004) (plurality opinion) (stating that resolution of the case required only an examination of the writ of habeas corpus and the Due Process Clause). But, for one, there are other examples of disparate treatment by the Court. *See, e.g.*, Boumediene v. Bush, 128 S. Ct. 2229, 2259 (2008) (holding that the citizenship status of a detainee is a factor to be considered in "determining the reach of the Suspension Clause"). And perhaps more importantly, there is basically no case law on the subject that would allow for better insight into the Court's position. There have been equal protection challenges to the government's post-9/11 counterterrorism policies brought in federal court. *See, e.g.*, Ashcroft v. Iqbal, 192 S. Ct. 1937, 1943–44 (2009) (challenging the constitutionality of immigration-related detention). But, although these challenges were brought by noncitizens, they were grounded in claims of differential treatment on the basis of national origin, ethnicity, and religion, rather than citizenship status. *Id.* at 1944.

113. 426 U.S. 67 (1976).

114. *Id.* at 81.

legally valid to its conceptual limits and because the Legislature and the Executive are still bound at those limits.[115] Maybe, then, the conclusions reached in Part V are premature.

But a second look at the laws addressed in Part IV reveals that the political branches do not in any way appear to feel bound beyond the limits the Supreme Court has drawn. Indeed, a closer analysis of those laws shows that they may not even fairly be seen as reasonable. The unreasonableness of the citizenship-based classification in counterterrorism law suggests that if the Equal Protection Clause stood for an important constitutional principle in the way Dworkin understands it to,[116] then rational-basis review would serve as more than a rubber stamp.

For the laws and policies addressed in Part IV to be reasonable in any true sense, there would have to be something about the distinction between aliens and citizens that makes codifying that distinction defensible in this context.[117] As a general matter, there is not. American citizens are at least as capable of carrying out a terrorist attack as noncitizens are. In fact, they are probably more so in some ways because aliens—at least those who arrive legally—have to go through customs when they come to the United States, and thus the government maintains some minimal ability to keep track of those aliens who might be dangerous or to keep them out of the country altogether. Even the Bush Administration admitted as much. When he was the Attorney General, Alberto Gonzales noted, "The threat of homegrown terrorist cells . . . may be as dangerous as groups like al Qaeda, if not more so."[118]

It might be argued that because noncitizens do not have a vested interest in American society the way that citizens do, they are more apt to commit acts of terrorism, even if they are no more practically capable of doing so. And, therefore, the distinction between citizens and aliens in counterterrorism law is appropriate. But this is plainly a false premise. Unfortunately, American citizens repeatedly have shown that the desire and willingness to

---

115. *See* Sager, *supra* note 63, at 1221 ("[C]onstitutional norms which are underenforced by the federal judiciary should be understood to be legally valid to their full conceptual limits, and federal judicial decisions which stop short of these limits should be understood as delineating only the boundaries of the federal courts' role in enforcing the norm . . . .").

116. *See* DWORKIN, *supra* note 12, at 9–10 (declaring that it has "been settled by unchallengeable precedent" that the Equal Protection Clause of the Fourteenth Amendment is a very "robust" principle).

117. *See* Tussman & tenBroek, *supra* note 52, at 346 ("A reasonable classification is one which includes all persons who are similarly situated with respect to the purpose of the law.").

118. Alberto Gonzales, U.S. Attorney Gen., Remarks at the World Affairs Council of Pittsburgh on Stopping Terrorists Before They Strike: The Justice Department's Power of Prevention (Aug. 16, 2006), *quoted in Implications of the Supreme Court's* Boumediene *Decision for Detainees at Guantanamo Bay, Cuba: Non-governmental Perspective, Hearing Before the H. Armed Services Comm.*, 110th Cong. 9 (2008) (statement of Neal K. Katyal, Professor of Law, Georgetown Law School), *available at* http://armedservices.house.gov/pdfs/FC073008/Katyal_Testimony073008.pdf.

resort to terrorism on American soil is not unique to foreigners. After all, the Oklahoma City bombing, the worst act of terrorism in America prior to September 11th, was the work of a citizen.[119] More recently, there have been citizen enemy combatants.[120] More recently still, a citizen-member of the U.S. military shot and killed thirteen Americans at the Fort Hood army base.[121]

Moreover, by their very terms, many of these specific laws and policies render any differentiation between citizens and aliens nonsensical and even counterintuitive. Consider, for example, provisions of two of the laws noted above: the USA PATRIOT Act and the first military order asserting authority to preventatively detain aliens. The former defined terrorism so broadly for noncitizens that it allowed the government to detain some aliens who were, at best, tangential to terrorism,[122] whereas the first incarnation of the Bush military order, on its face, did not even allow the Administration to detain citizen-members of al Qaeda.[123] Thus, the government asserted broad authority to detain arguably harmless noncitizens but did not—at first—claim the same authority over citizens who, all else being equal, were much more dangerous. It seems absurd for the government to claim that there is something intrinsic to noncitizen status that makes an alien with tenuous, indirect ties to al Qaeda more dangerous, and thereby more detainable, than a citizen who is a full-blown *member* of the terrorist organization. But that is the position the government took before it apprehended Padilla.

In spite of the unreasonableness of the counterterrorism laws' distinction between citizens and noncitizens, those laws remain unchallenged under equal protection. This is, perhaps, because rational-basis review is so deferential that the laws would survive a challenge anyway. Therefore, while equal protection may no longer be a constitutional argument of last resort, the clause nonetheless continues to stand for a second-class constitutional ethic. Unlike liberty norms and privacy norms, a full-fledged commitment to the constitutional ethic of equality still takes a backseat to countervailing considerations, even those that cannot be reasonably justified.

---

119. *See* Jo Thomas, *Verdict Is Cheered: Affixing Blame in Worst Terrorism Attack on United States Soil*, N.Y. TIMES, June 3, 1997, at A1 (detailing Timothy McVeigh's guilty verdict in the Oklahoma City bombing and describing the event as "the worst act of terrorism on American soil").

120. *See supra* notes 87–88 and accompanying text.

121. Eric Schmitt & Eric Lipton, *Focus on Internet Imams as Recruiters for Al Qaeda*, N.Y. TIMES, Jan. 1, 2010, at A14.

122. *See* USA PATRIOT Act of 2001, Pub. L. No. 107-56, sec. 802, 115 Stat. 272, 376 (codified at 18 U.S.C. § 2331 (2006)) (defining domestic terrorism to be "activities that involve acts dangerous to human life that are a violation of the criminal laws of the United States or of any State; [that] appear to be intended to intimidate or coerce a civilian population, to influence the policy of a government by intimidation or coercion, or to affect the conduct of a government by mass destruction, assassination, or kidnapping; and occur primarily within the territorial jurisdiction of the United States").

123. Military Order of Nov. 13, 2001, Detention, Treatment, and Trial of Certain Non-citizens in the War Against Terrorism, 66 Fed. Reg. 57,833, 57,834 (Nov. 16, 2001).

VII. Equal Protection Should Stand on the Same Footing as Liberty and Privacy

If one accepts the conclusions above, the next question might be whether it is necessarily a bad thing that equal protection norms are of lesser importance than other constitutional ideas. Maybe issues of national security are so critical that, even if certain policies are unfair to a group (and even if they are irrationally so), it is better to err on the side of safety than potentially to handicap terrorism prevention by worrying about the Fourteenth Amendment implications of those policies. After all, Marc Sageman argues that the problem of homegrown terrorism is significantly less pronounced in America than it is in other Western countries,[124] and so it might make some sense to distinguish between citizens and noncitizens in a way that strict scrutiny would not allow. Perhaps, despite some seemingly problematic aspects of post-9/11 terrorism policy, the interests of national security are best served when equal protection considerations do not inform such policy making. For example, with respect to administrative detention, one prominent Washington insider sees compelling reasons to make distinctions like those outlined above, but he would draw the line at residency rather than citizenship.[125] Benjamin Wittes argues that because "[c]itizens and permanent resident aliens are unambiguously part of the American social compact," and have "deep ties" to the country, they should be privy to a different system of incarceration and trial than non-resident aliens, because otherwise it would "injure[] the American body politic."[126] Perhaps, then, some people are just more deserving of the Constitution's protections than others, and a deep fidelity to equal protection norms would disrupt the natural order of things in this regard.

Such thinking is wrongheaded and misguided on a number of levels. For one thing, Sageman argues that the American promise of equality (however illusory that promise might be) is the very reason that the phenomenon of homegrown terrorism is insignificant in the United States.[127] For this reason, the broader principle of equal protection of the laws ought to be embraced, rather than disregarded or minimized. Strictly from a policy perspective, the United States is better off faithfully adhering to its professed principles. Doing so would be advantageous in the country's campaign for the hearts and minds of prospective terrorists, both at home and abroad. Indeed, the best way to prevent terrorism is not to detain terrorists but rather,

---

124. MARC SAGEMAN, LEADERLESS JIHAD: TERROR NETWORKS IN THE TWENTY-FIRST CENTURY 90 (2008).

125. BENJAMIN WITTES, LAW AND THE LONG WAR: THE FUTURE OF JUSTICE IN THE AGE OF TERROR 178 (2008).

126. *Id.*

127. *See* SAGEMAN, *supra* note 124, at 96–97 (arguing that the perception of equal access to the "American Dream" serves to reduce significantly the likelihood of Muslim radicalization, thereby reducing the terrorist threat).

where possible, to discourage people from becoming terrorists in the first place. In this way, America's apparent lack of commitment to equal protection norms is problematic.

More significantly, Wittes's conception of a "social compact," from which members of society derive their rights, is fundamentally at odds with the philosophy underlying the American constitutional system, especially as Bobbitt understands it. The Equal Protection Clause requires that American law apply equally to everyone within the Constitution's purview, regardless of her place in society. Thus, the Constitution prohibits differential treatment not only of aliens but also of political dissidents and social outcasts, despite the fact that they may be wholly without meaningful "ties" to the social order. But because of the "*ethos*[] of the American polity"[128]—that of a government with limited, specifically enumerated powers—the Equal Protection Clause is not even necessary to effect this prohibition. Indeed, American society does not operate on an opt-in basis, whereby one is granted rights when she signs onto an ethereal social compact. Instead, as Bobbitt makes clear, the Constitution grants the *government* certain limited rights and powers.[129] And so, Bobbitt might argue, since the federal government has not been bestowed with the power to do so, it is constitutionally prohibited from making some people more deserving of legal protections than others, regardless of the attractiveness of social-compact theory. In this way, the Equal Protection Clause can be seen as merely adding rhetorical force to a prohibition already implicit in the Constitution.[130] In fact, long before the ratification of the Fourteenth Amendment, the Court in *Martin v. Hunter's Lessee*[131] acknowledged that equality is a central principle behind the Constitution.[132]

Ultimately, then, the relatively low importance given to the principle of equal protection, as evidenced by post-9/11 terrorism policies, is not problematic solely from the practical perspective of combating terrorism. Even more significantly, it represents a troubling departure from the most basic foundational principle of the American system of government. If the broader constitutional ethic of limited government, which Bobbitt argues informs a wide range of judicial decision making, encompasses the principle of equal

128. BOBBITT, *supra* note 16, at 94.

129. *See* PHILIP BOBBITT, CONSTITUTIONAL INTERPRETATION 20 (1991) ("The fundamental American constitutional ethos is the idea of limited government . . . .").

130. But, admittedly, one that was buried and badly obscured by America's history of institutionalized racial injustice.

131. 14 U.S. (1 Wheat.) 304 (1816).

132. *See id.* at 348 ("The constitution of the United States was designed for the common and equal benefit of all the people of the United States.").

protection, then a disregard for equal protection cheapens the larger principle. As such, the Court should renew its commitment to the concept of equality and should enforce the Equal Protection Clause much closer to its conceptual boundaries.

—Daniel H. Cohen

* * *

The Conman and the Sheriff: SEC Jurisdiction and the Role of Offshore Financial Centers in Modern Securities Fraud*

Congress founded the Securities and Exchange Commission (SEC) in 1934 in the wake of the greatest financial collapse in history. More than seventy years later, the SEC, charged by Congress with a mandate to preserve the national public interest through fair and honest markets, remains a critical force in policing U.S. markets. But while the scope and authority of the SEC's power has expanded over time, so too have the crimes it seeks to prevent. Over the last five years, a carousel of large, international frauds by well-known and well-regarded financiers undermined the integrity of global securities markets and international cooperation in market enforcement. Specifically, these crimes, perpetrated by rogue financiers like Bernard Madoff and Sir Allen Stanford, have cost investors in the Americas and beyond billions of dollars. In light of these inventive frauds, no simple solution can prevent all forms of financial crime; however, this Note advocates an expansion of SEC authority to place a substantial hurdle in the way of these clandestine conmen in the hopes of stripping them of a primary tool—the use of offshore financial centers (OFCs).

To prevent conmen from avoiding SEC jurisdiction and capitalizing on self-interested local regulation in OFCs, this Note encourages Congress to grant the SEC the authority to initiate investigations on foreign soil—with prior consent from foreign regulators—when it perceives a substantial threat to investors in the United States. To support such an expansion of authority, this Note first provides a primer on OFCs and their use by international fraudsters, drawing from three recent case studies of securities fraud. Then, this Note will explain the expansion of SEC authority while addressing both the inefficiencies in the SEC's current cooperative model and the precedent for such an expansion of authority—Title III of the USA PATRIOT Act. Finally, this Note will apply the SEC's expanded jurisdiction to the three case studies discussed earlier to demonstrate how this new approach can prevent the types of frauds that threaten the integrity of U.S. markets and the national public interest.

---

* J.D. candidate, 2010, The University of Texas School of Law. The author would like to thank first and foremost the staff of the Texas Law Review for their assistance in editing this Note, as well as their constant friendship, integrity, and character. Henry Hu—the Director of the SEC's Division of Risk, Strategy, and Financial Innovation, and a law professor at the University of Texas—deserves particular credit for his constructive criticism and guidance in producing this Note. Finally, the author would like to thank his family for their love and support—responsible for any ounce of inspiration contained herein.

I. Enhancing SEC Jurisdiction to Combat International Fraud—An Introduction

In light of persistent technological innovation and the integration of global markets, the Securities and Exchange Commission (SEC) must adapt constantly to new challenges threatening its domain.[1] As the recent financial crisis again demonstrates, the effects of adverse market conditions in one nation can cause unforeseeable damage across the world.[2] Moreover, given the speed of modern transactions, where money deposited in Omaha before breakfast can be in London for tea and in Hong Kong by dinner, regulators often have trouble keeping pace with increasingly complex, rapid communications.[3] Amidst this changing economic and regulatory landscape, a new villain has emerged to challenge the SEC: the international fraudster or conman.[4] Recently, these conmen, like Sir Allen Stanford and Bernard Madoff, have dominated the press with the exploits of their multi-billion-dollar frauds.[5] To impede the proliferation of international conmen, this Note will investigate the role that offshore financial centers (OFCs) play in modern cons and offer a possible mechanism by which the SEC may be able to mitigate the damage done by fraudsters.

Specifically, the manner in which conmen utilize OFCs in their frauds often raises red flags that something is amiss, giving regulators an

---

1. *See* Luis A. Aguilar, Comm'r, SEC, Speech by SEC Commissioner: Empowering the Markets Watchdog to Effect Real Results (Jan. 10, 2009) (transcript available at http://sec.gov/news/speech/2009/spch011009laa.htm) (discussing the importance of developing "a more efficient and modern regulatory structure . . . to address new technologies, globalization, and new innovative financial products and services").

2. *See* David Reiss, *The Federal Government's Implied Guarantee of Fannie Mae and Freddie Mac's Obligations: Uncle Sam Will Pick Up the Tab*, 42 GA. L. REV. 1019, 1024 (2008) (addressing the argument that "if the federal government did not bail out Fannie or Freddie, it could lead to an international financial crisis that could be greater than those posed by the 1994 Mexican peso collapse, the 1997 East Asian 'flu' and the 1998 Russian bond default").

3. *See* Tom McGinty & Kara Scannell, *SEC Plays Keep-Up in High-Tech Race*, WALL ST. J., Aug. 20, 2009, at C1 ("[T]he [SEC] is outmatched by the traders and market venues with technology that is remaking the trading world."); *cf. Strengthening the S.E.C.'s Vital Enforcement Responsibilities: Hearing Before the Subcomm. on Securities, Insurance, and Investment of the S. Comm. on Banking, Housing, and Urban Affairs*, 111th Cong. 5 (2009) (statement of Richard J. Hillman, Managing Director, Fin. Mkts. & Cmty. Inv., Gov't Accountability Office), *available at* http://www.gao.gov/new.items/d09613t.pdf (admonishing the SEC's "burdensome system for internal case review" that slows enforcement actions).

4. *See, e.g.*, ARTHUR HERZOG, VESCO: FROM WALL STREET TO CASTRO'S CUBA, THE RISE, FALL, AND EXILE OF THE KING OF WHITE COLLAR CRIME, at xv (2003) (noting that in the 1970s, financier Robert Vesco was accused of stealing roughly $250 million, thus placing him "at the pinnacle of white collar thieves").

5. *See, e.g.*, Alyssa Abkowitz, *The Investment Scam-Artist's Playbook: Bernie Madoff and R. Allen Stanford's Tactics May Suggest a Formula on How to Get Mixed Up in a Massive Government Fraud Case*, CNNMONEY.COM, Feb. 25, 2009, http://money.cnn.com/2009/02/25/news/madoff_stanford_playbook.fortune/index.htm?postversion=2009022516 (discussing the similarities of the two cases).

opportunity to stop a fraud before it escalates.[6] However, despite the prevalence of these indicators, the SEC cannot act peremptorily to halt a growing fraud because it is unable to rely on foreign regulators in these OFCs,[7] and it lacks the authority to initiate unilateral investigations within foreign borders.[8] As a result, conmen engage in OFC-based activity to avoid SEC jurisdiction and capitalize on self-interested local regulation,[9] placing substantial hurdles in the way of the SEC's enforcement efforts. In order to overcome these obstacles and the SEC's reliance on foreign regulators—who often appear complicit in the frauds[10]—Congress should grant the SEC the authority to initiate investigations on foreign soil when it perceives a substantial threat to investors in the United States without waiting for approval from foreign regulators.

After a brief primer on OFCs in Part II, this Note will investigate three recent frauds employing OFCs in Part III, noting the existence of red flags in each instance. Then, in Part IV, this Note will present an argument for expanding the SEC's authority to foreign soil while addressing the inefficiencies of the SEC's current cooperative model and the notion that this expansion is not without precedent in light of Title III of the USA PATRIOT Act.[11] Finally, in Part V, this Note will apply the solution developed in the previous Part to the three case studies discussed in Part III, demonstrating the potential of an expanded investigative right to overcome the jurisdictional benefits of OFCs to conmen, before concluding.

---

6. *See infra* sections III(A)(3), III(B)(3), III(C)(3) (identifying the red flags that appeared in each case study).

7. *See infra* subsections III(A)(2)(b)–(c) (describing how the close relationship between the conman and the local regulatory authority made the latter an unreliable partner for SEC investigations).

8. *See* Securities Exchange Act of 1934 § 21(a)(2), 15 U.S.C. § 78u(a)(2) (2006) (granting the SEC authority to assist a foreign government in a securities investigation *if* that foreign government requests assistance).

9. *See infra* subpart II(B) (explaining the incentives that lead OFCs to avoid meaningful regulation).

10. *See Alleged Stanford Financial Group Fraud: Regulatory and Oversight Concerns and the Need for Reform: Hearing Before the S. Comm. on Banking, Housing, and Urban Affairs*, 111th Cong. 9 (2009) (statement of Onnig H. Dombalagian, George Denégre Assoc. Professor of Law, Tulane Univ.), *available at* http://banking.senate.gov/public/index.cfm?FuseAction=Files.View&FileStore_id=9f74be50-89f5-4e07-b9ef-f9aa542f9df6 (contending that evidence suggests that Stanford likely procured the complicity of the Antiguan regulator responsible for oversight of Stanford's operations to hinder the SEC's investigation).

11. USA PATRIOT Act, Pub. L. No. 107-56, tit. 3, 115 Stat. 272, 296–342 (2001) (codified as amended in scattered titles of U.S.C.).

II. A Primer on Offshore Financial Centers

A. An Operational Definition

Much of the current scholarship on offshore financial centers focuses on providing an operational definition of what constitutes an OFC.[12] From this literature, three common characteristics emerge. First, OFCs direct their business toward nonresidents.[13] In particular, laws in OFCs are designed to encourage the export of financial services to higher regulation nations in exchange for capital.[14] As a result, these regions "provide[] financial services to nonresidents on a scale that is incommensurate with the size and the financing of [their] domestic econom[ies]."[15]

Second, to encourage this exchange, OFCs present favorable regulatory environments relative to higher regulation nations like the United States.[16] These favorable conditions include flexible incorporation and licensing standards, low government supervision, robust secrecy laws, a lack of physical-presence requirements, and extensive use of special-purpose vehicles and trusts without government review.[17]

Third, OFCs offer "low- or zero-taxation schemes" on business and investment income for nonresidents.[18] This characteristic exemplifies the notion that OFCs systematically create an economic climate that favors foreign capital, as many of the advantages granted to foreigners are not extended to locals.[19]

---

12. *See, e.g.*, Ahmed Zoromé, *Concept of Offshore Financial Centers: In Search of an Operational Definition* 26 (Int'l Monetary Fund, Working Paper No. 07/87, 2007), *available at* http://www.imf.org/external/pubs/ft/wp/2007/wp0787.pdf (listing various definitions of OFCs found in a survey of the literature).

13. *See* GUNTER DUFEY & IAN H. GIDDY, THE INTERNATIONAL MONEY MARKET 37 (1978) (defining offshore banking as attracting nonresident borrowers and lenders with sparse or flexible regulation and taxation on the banking industry); R.B. JOHNSTON, THE ECONOMICS OF THE EURO-MARKET: HISTORY, THEORY AND POLICY 18 (1982) (adding that territories with active offshore banking industries are also distinguished by a level of banking business that far exceeds the needs of the local market); Ian McCarthy, *Offshore Banking Centers: Benefits and Costs*, FIN. & DEV., Dec. 1979, at 45, 45–46 (emphasizing ease of entry and low license fees, taxes, and levies as characteristic of successful offshore banking centers).

14. *See* Zoromé, *supra* note 12, at 6 (noting "the intrinsic feature of the OFC phenomenon, which is its raison d'être—the provision of financial services to nonresidents, namely, exports of financial services" (emphasis omitted)).

15. *Id.* at 7 (emphasis omitted) (citation omitted).

16. McCarthy, *supra* note 13, at 47.

17. *See infra* notes 31–32 and accompanying text. *See generally* Zoromé, *supra* note 12. It is worth noting that, while not always individually dangerous, these elements can be more potent when used collectively.

18. Zoromé, *supra* note 12, at 4.

19. *See* ORG. FOR ECON. CO-OPERATION & DEV., HARMFUL TAX COMPETITION: AN EMERGING GLOBAL ISSUE 26–28 (1998), *available at* http://www.oecd.org/dataoecd/33/1/1904184.pdf (describing "ring-fencing"—either preventing residents from benefitting from tax advantages or preventing advantaged nonresidents from participating in the domestic market—as a key factor in identifying harmful preferential tax regimes typical of tax havens).

Taking these three factors together, one can consider an OFC to be a region that offers favorable regulatory standards to nonresidents, including low tax and transparency burdens, in an effort to attract foreign capital. The key corollary to this definition is that OFCs willfully undercut foreign regulations to attract capital, creating an opportunity for residents in higher regulation countries to move operations to regions with lower standards.[20] To any conman, the advantages of doing business in an OFC become clear: low regulation and high secrecy decrease the likelihood that home-country regulators, like the SEC, will be able to detect suspicious behavior, much less investigate with enough diligence to bring an enforcement action.

Despite the commonality of certain OFC indicators, not all OFCs are the same; a particular dichotomy essential to understanding OFCs deserves attention. In the world of OFCs, two subgroups emerge: nations like Switzerland, which have high regulatory standards but evince high secrecy laws that frustrate foreign regulators,[21] and nations like the Seychelles, which function purely as a regulatory façade.[22] It is particularly telling that the latter group is the one more likely to be targeted by international fraudsters, as evidenced by the case studies.[23] As a result, this group will be subject to greater scrutiny under the model proposed in Part IV.

B. Inadequate Incentives to Self-Regulate

An alternative model to the one advocated in this Note centers on self-regulation by OFCs. Rather than granting the SEC unilateral authority to investigate beyond U.S. borders, the model rests on the notion that the SEC can cooperate with local regulators that police themselves. Given the definition of OFCs presented in the preceding subpart, this model is inherently nonfunctional. Two particularly compelling forces underlie the inefficacy of the self-regulation model.

---

20. *See* DUFEY & GIDDY, *supra* note 13, at 38 ("Thus a government can often legislate (or better, delegislate) its country into financial prominence, if it wishes to do so.").

21. Despite being listed on the Organisation for Economic Cooperation and Development's (OECD) "grey list" of substandard tax regulators, the Financial Stability Forum elects not to classify Switzerland as an OFC based on its cooperation with the SEC and other foreign regulators. FIN. STABILITY FORUM, REPORT OF THE WORKING GROUP ON OFFSHORE CENTRES 14 (2000), *available at* http://www.financialstabilityboard.org/publications/r_0004b.pdf [hereinafter FSF REPORT]. *But see* ORG. FOR ECON. CO-OPERATION & DEV., A PROGRESS REPORT ON THE JURISDICTIONS SURVEYED BY THE OECD GLOBAL FORUM IN IMPLEMENTING THE INTERNATIONALLY AGREED TAX STANDARD (2009), *available at* http://www.oecd.org/dataoecd/38/14/42497950.pdf ("grey" listing Switzerland); *Swiss Bank Refuses US Tax Request*, BBC NEWS, May 1, 2009, http://news.bbc.co.uk/2/hi/business/8028174.stm (discussing the Swiss bank UBS's efforts to hide behind its nation's secrecy laws to avoid furnishing evidence in conjunction with a pending tax-evasion investigation in the United States).

22. *See, e.g.*, INT'L MONETARY FUND, COUNTRY REPORT NO. 04/381, SEYCHELLES: REVIEW OF FINANCIAL SECTOR REGULATION AND SUPERVISION 14–16 (2004), *available at* http://www.imf.org/external/pubs/ft/scr/2004/cr04381.pdf (discussing the "urgent need" for the Seychelles to upgrade its banking regulations to meet the Basel standards).

23. *See infra* Part III.

First, at a national level, OFCs stand to lose far more than they can gain by enhancing regulations. The Cayman Islands, for example, relies heavily on its financial-services sector, which is predicated on low regulations of foreign capital:

> Forty years ago the Cayman Islands were on the verge of economic extinction. There was no room in the modern world for an island that made rope and caught turtles and whose wetlands prevented any agriculture. Today the standard of living is the highest in the Caribbean, surpassing the United States and Britain. There is virtually no unemployment. Only the beautiful alliance of capital and state has made this possible.[24]

Conversely, the only tangible benefit from complying with foreign regulation is avoiding public shaming.[25] Thus, at a national level, OFCs lack an adequate incentive or substantial threat to forgo the wealth benefit of low regulation.

This conflict is replicated at the individual-actor level. In particular, domestic political pressure tends not to incentivize endorsing greater regulation.[26] Moreover, given the sizable role that the financial-services industry plays in many local economies, it is unlikely that voters in an OFC will forgo increased domestic liquidity out of concern for the potential negative externalities caused by their regulations in other countries.[27] This bias ultimately manifests itself in local politicians who support low regulations to bring greater liquidity to their communities. Taken together, the national and private-actor interests in maintaining low regulations imply that efforts to control the fraudulent use of OFCs cannot occur through enhanced self-regulation by the OFCs themselves. Instead, if the SEC is to meaningfully monitor the effects of OFC-based behavior on the U.S. markets, it will have to do so on its own initiative.

---

24. WILLIAM BRITTAIN-CATLIN, OFFSHORE: THE DARK SIDE OF THE GLOBAL ECONOMY 8 (2005). Similarly, in the British Virgin Islands, "[t]he revenue from registering foreign companies has paid for a community college and a hospital." Ben Fox, *Islands Resent Crackdown of Tax Havens by G-20*, ABC NEWS, Apr. 3, 2009, http://abcnews.go.com/International/wireStory?id=7247748.

25. *Cf., e.g.*, Press Release, Org. for Econ. Co-operation & Dev., Four More Countries Commit to OECD Tax Standards (Apr. 7, 2009), *available at* http://www.oecd.org/documentprint/0,3455,en_2649_34487_42521280_1_1_1_1,00.html (announcing that four previously blacklisted countries—Costa Rica, Malaysia, Philippines, and Uruguay—had committed to improving their respective tax standards following a public shaming by the OECD).

26. *Cf., e.g.*, Pierre-Hugues Verdier, *Transnational Regulatory Networks and Their Limits*, 34 YALE J. INT'L L. 113, 127–28 (2009) (observing that administrative agencies, like financial regulators, are subject to domestic political pressures sometimes allowing "concerned parties [to] succeed in convincing legislatures to override agency rules . . . or even to radically restructure or consolidate agencies").

27. *See* BRITTAIN-CATLIN, *supra* note 24, at 8 (acknowledging the effect of importing financial services on the unemployment level and standard of living in the Cayman Islands).

C. *Current Examples of OFCs*

In 2000, the Financial Stability Forum (FSF) identified a list of OFCs based on the level of technical assistance offered by these regions in support of international securities-law enforcement and the extent of their financial regulations.[28] Nations that demonstrated an unwillingness to improve their low standards were more likely to be relegated to the FSF's OFC list, presented in Table 1 below.[29]

Table 1: Countries Classified as "Financial Centres with Significant Offshore Activities" by the FSF[30]

Andorra	British Virgin Islands	Lebanon	St. Kitts & Nevis
Anguilla	Cayman Islands	Liechtenstein	Niue
Antigua	Cook Islands	Macau	Panama
Aruba	Costa Rica	Malta	St. Lucia
Bahamas	Cyprus	Marshall Islands	St. Vincent
Bahrain	Gibraltar	Mauritius	Samoa
Barbados	Guernsey	Monaco	Seychelles
Belize	Isle of Man	Nauru	Turks & Caicos Islands
Bermuda	Jersey	Netherlands Antilles	Vanuatu

The FSF conducted various surveys of both OFCs and traditional, onshore financial centers to better determine the key characteristics, issues, and motivations associated with OFCs. The FSF determined that OFCs are used "to maximize profits in low tax regimes," to create special-purpose vehicles (SPVs) to issue securitized products like asset-backed securities, to hide assets from seizure and repatriation, to avoid tax-related disclosures, and to launder money.[31] Ultimately, the FSF concluded that regulatory exploitation in OFCs developed from inadequate due diligence in licensing and monitoring corporations and SPVs, inadequate disclosure rules, inadequate

28. FSF REPORT, *supra* note 21, at 9.

29. *Id.*

30. *Id.* at 14 tbl.1. Only twenty-five of the thirty-seven countries responded to the survey, but the FSF did not disclose which did or did not, so the full list is included here.

31. *Id.* at 10.

information sharing between the OFCs and investors, inadequate enforcement capital and personnel, excessive secrecy laws, and an "[a]bsence of political will to improve the quality of supervision."[32] Significantly, each of the nations implicated in the case studies discussed below appears on the FSF's list of OFCs, and each embodies the characteristics of high secrecy and low monitoring of corporate activity, making them attractive tools in any con.

III. The Conmen: Case Studies from the Post-Enron Era

Modern securities-fraud cases are as diverse as the personalities of those who perpetrate them. However, as the case studies will demonstrate, many frauds rely on the obscurity of OFCs to effect material elements of the crime. Capitalizing on the SEC's jurisdictional limitations and the lack of transparency promulgated by local regulators, fraudsters are able to engage in brazen techniques to defraud investors. The three case studies will demonstrate that, while the scale and techniques of fraud vary, certain red flags tied to activities in OFCs can be detected, giving the SEC the opportunity to shut down a fraud before it escalates. At a minimum, increased authority for the SEC would make it far more difficult for the fraudsters to effect their crimes by shedding greater transparency on their offshore activities, ultimately decreasing the ability of conmen to successfully engage in fraudulent practices.

A. Sir Robert Allen Stanford

> This tournament is unique in so many ways, not only because of the prizes up for grabs but also for the different elements that we are going to add to the game that will make it even more exciting. I don't want to reveal too much[,] but I hope everyone will come out and see what we have in store.[33]

1. The Con.—According to the SEC, Allen Stanford and Antigua-based[34] Stanford International Bank, Ltd. (SIB) engaged in a two-pronged effort to defraud investors of over $8 billion.[35] The first aspect involved the sale of allegedly low-risk certificates of deposit (CDs) to investors, professing returns at twice the market rate.[36] In particular, the SEC alleges that SIB

32. *Id.* at 12–13.

33. *Stanford Ups the Ante*, CRICINFO.COM, March 31, 2006, http://www.cricinfo.com/pakistan/content/story/242809.html.

34. Julie Creswell et al., *Fraud Parade: $8 Billion Case Is Next in Line*, N.Y. TIMES, Feb. 18, 2009, at A1.

35. Complaint at 4, SEC v. Stanford Int'l Bank, Ltd., No. 3:09-CV-298-N (N.D. Tex. Feb. 16, 2009) [hereinafter Stanford Complaint].

36. *See id.* at 1 (explaining that Stanford promised "high return rates that exceed[ed] those . . . offered by traditional banks"); *see also* Matthew Goldstein, *The Pressure Mounts on Stanford*, BUSINESSWEEK, Feb. 16, 2009, http://www.businessweek.com/investing/wall_street_

misled investors into believing the CDs were safer than they actually were in three ways.

First, SIB claimed to re-invest "client funds primarily in 'liquid' financial instruments."[37] However, "a substantial portion of the bank's portfolio was placed in illiquid investments, such as real estate and private equity," increasing the risk faced by investors without disclosure.[38] Second, SIB claimed to monitor "the portfolio through a team of 20-plus analysts."[39] But in actual fact, the SEC alleges that only "two people—Allen Stanford and [CIO] James Davis"—were monitoring the multibillion-dollar portfolio.[40] Finally, the SEC asserts that SIB, to pacify investor concerns, claimed that the firm was "subject to yearly audits by Antiguan regulators."[41] However, "the Antiguan regulator responsible for oversight of the bank's portfolio, the Financial Services Regulatory Commission, [did] not audit SIB's portfolio or verify the assets SIB claim[ed] in its financial statements."[42] The cumulative effect of these misrepresentations was the cultivation of an $8.4 billion portfolio predicated on erroneous information and overinflated estimations of investment returns.[43]

The second element of Stanford's massive fraud involved more than $1 billion in sales of a "proprietary mutual fund wrap program" based on materially false historical-performance data.[44] This program allegedly enabled Stanford to recruit financial advisers "to re-allocate their clients' assets to SIB's CD program."[45]

Collectively, Stanford's fraud centered on his ability to mislead investors into believing that the too-good-to-be-true returns on the CDs offered by SIB were both real and secure, while in fact investor proceeds were placed in speculative, unsecure investments.[46] Through his con, Stanford convinced his investors—most of whom lived in the United States,

---

news_blog/archives/2009/02/the_pressure_mo.html (commenting that the professed returns on Stanford's high-yield CDs were "twice the market average"); Bill Zielinski, *Stanford Financial Is Next Hedge Fund Investigated*, SEEKING ALPHA, Feb. 16, 2009, http://seekingalpha.com/article/120791-stanford-financial-is-next-hedge-fund-investigated (explaining that Stanford proffered "tantalizing" interest rates nearly twice the rates offered on average (4.5% instead of 2% for one-year CDs and 7.03% instead of 3.9% for five-year CDs)).

37. Stanford Complaint, *supra* note 35, at 3.

38. *Id.* at 3–4.

39. *Id.* at 3.

40. *Id.* at 4.

41. *Id.* at 3.

42. *Id.* at 4.

43. *Id.* at 8–9.

44. *Id.* at 4.

45. *Id.* at 5.

46. *See* Ray Hennessey, *Fraud at Stanford Was Suspected as Early as 2002*, FOXBUSINESS.COM, May 6, 2009, http://www.foxbusiness.com/story/markets/industries/government/fraud-stanford-suspected-early/ (quoting an unnamed insider who claimed that "investor proceeds [were] being directed into speculative investments like stocks, options, futures, currencies, real estate and unsecured loans").

Venezuela, Ecuador, and Mexico[47]—that his firm was capable of extraordinary returns, while stripping them of billions of dollars, much of which cannot be located.[48]

2. The OFC Element.—Antigua played a significant role in Stanford's fraud in three principal ways. First, the financier capitalized on the island's lack of transparency to institute and execute his suspicious business model with virtually no oversight.[49] Second, he cultivated close links with the local community, placing him beyond the reproach of local regulators and politicians.[50] Third, by basing his operation outside of the SEC's jurisdiction, Stanford was able to expand his fraud while the SEC was left clamoring for local assistance from Antiguan regulators.[51]

a. Low Transparency Regarding Investment Sources and Inadequate Verification by Local Auditors.—Stanford took advantage of Antigua's low level of supervision to implement two aspects of his fraud that were intended to deceive investors into believing that his investments were safe and successful: his business model and his local auditor. First, Stanford's fraud was predicated on his ability to raise constant investments from the public without having to prove that his alleged returns were real. By establishing a bank in Antigua and issuing CDs predominantly to Latin American investors,[52] who value low transparency,[53] Stanford was able to claim twice-market returns without having to provide proof to American investors and regulators.[54]

To the extent that proof of his model's fiscal viability was requested to satisfy investors, Stanford was able to take advantage of a local auditor that would not be vetted by American authorities. Since the institution of the Public Company Accounting Oversight Board (PCAOB),[55] which requires accounting firms that participate in audits of SEC-reporting companies to

47. Neil Roland, *Alleged "Massive Fraud" at Stanford Financial Exposes Regulatory Gaps—Again*, FIN. WK., Feb. 17, 2009, http://www.financialweek.com/apps/pbcs.dll/article?AID=/20090217/REG/902179991/1049/COMPLIANCE.

48. Marie Colvin, *$8bn "Missing" from Allen Stanford's Offshore Bank*, TIMES ONLINE, Feb. 22, 2009, http://business.timesonline.co.uk/tol/business/industry_sectors/banking_and_finance/article5780591.ece.

49. *See infra* subsection III(A)(2)(a).

50. *See infra* subsection III(A)(2)(b).

51. *Cf.* Hennessey, *supra* note 46 (explaining that the SEC had difficulty obtaining information due to Antiguan confidentiality laws).

52. Roland, *supra* note 47.

53. *See Latin Investors Could Be Big Madoff Losers*, MSNBC.COM, Dec. 29, 2008, http://www.msnbc.msn.com/id/28423227 (discussing the secrecy premium valued by Latin American investors).

54. *See* Roland, *supra* note 47 ("[T]he Stanford Group Co. investment advisers did not register with the SEC for many years, precluding potential inspection by regulators . . . .").

55. *See* Sarbanes–Oxley Act of 2002, Pub. L. No. 107-204, tit. 1, sec. 101(a), 116 Stat. 745, 750 (codified at 15 U.S.C. § 7211(a) (2006)) (establishing the PCAOB).

register with the PCAOB and adhere to its rules and inspections,[56] the SEC has been forced to rely significantly on foreign regulators to monitor local auditors.[57] When relying on foreign regulators, the PCAOB has employed a "'sliding scale approach': its rules permit varying degrees of reliance on the home country" depending on how independent that regulator is.[58] In determining independence, the PCAOB considers "the home country system's funding arrangements, transparency measures, and track record."[59] However, in the interest of administrative feasibility, the PCAOB must rely on the representations of local regulators, particularly with respect to determining transparency and funding.[60]

Therefore, to the extent that a fraudster would want to deceive the PCAOB and SEC, she mainly would have to mislead a local regulator into accepting the audit information as accurate. As will be demonstrated shortly, in light of the close relationship between Stanford and the Antiguan political community, the fraudster's "burden" of convincing local regulators becomes increasingly facile when the regulators are socially and financially integrated with the fraudster.

According to Stanford's SEC filings, SIB's auditor was the Antiguan branch of St. John's-based C.A.S. Hewlett & Co.[61] However, employees at Hewlett's office have no knowledge of working on any Stanford accounts.[62] The prevailing suspicion among the SEC and Hewlett employees is that the auditor's now-deceased former CEO Charlesworth Hewlett—if anyone—would be the only party with any "possible knowledge of a relationship to Stanford."[63] Moreover, the firm itself, a ten-person operation working out of a residential neighborhood, appeared to be "an unlikely operation to manage books for an $8 billion enterprise."[64] Insufficient auditing proves material in light of Stanford's gross financial misrepresentations. In one case, for instance, SIB purchased 1,587 acres in Antigua for $63.5 million.[65] Months later, this real estate was valued on SIB's books at $3.2 billion.[66] Clearly a

56. 15 U.S.C. §§ 7212–7214.

57. *See* Roberta S. Karmel, *The EU Challenge to the SEC*, 31 FORDHAM INT'L L.J. 1692, 1704 n.54, 1708 (2008) (discussing the SEC's reliance on foreign regulators when unable to subject foreign auditors to PCAOB's standards and reliance on foreign regulators as a key element in moves to harmonize American and foreign regulatory systems).

58. Stavros Gadinis, *The Politics of Competition in International Financial Regulation*, 49 HARV. INT'L L.J. 447, 487 (2008).

59. *Id.* at 488.

60. *Id.*

61. Jason Szep, *At Small Antigua Accounting Firm: Who's Stanford?*, REUTERS, Feb. 19, 2009, http://www.reuters.com/articlePrint?articleID=USN1951126620090219.

62. *Id.*

63. *Id.*

64. *Id.*

65. Anna Driver & Chris Baltimore, *Stanford Vastly Overstated Assets: U.S. Receiver*, REUTERS, Apr. 23, 2009, http://www.reuters.com/articlePrint?articleId=USTRE53M5PJ20090423.

66. *Id.*

local auditor on a small island ought to have been aware of local land values, particularly given the magnitude of the discrepancy.

b. Stanford's Relationship with the Local Community Disincentivizes Investigation.—A disconcerting undertone begins to emerge from the wreckage of Stanford's activities in Antigua—regardless of ability, the nation did not appear willing to question the conman's success. This unwillingness was linked to Stanford's prominent role in the Antiguan community. While Stanford "was just another wealthy financier" in Texas, in Antigua "he was lord of an influential financial fief, decorated with a knighthood, courted by government officials and basking in the spotlight of sports and charity events on which he generously showered his fortune."[67] Thus, where Stanford was successful, so was the island.

Stanford's prominence in Antigua is evident in his holdings, which included "a newspaper, two banks, two restaurants, a spectacular cricket stadium that bears his name, a cricket tournament that turns players into millionaires, and some of the best real estate in the nation," in addition to his "theme-park-like" headquarters located "yards from V.C. Bird International Airport."[68] Not only were the people of Antigua enamored with Stanford for bringing thousands of jobs[69]—and international cricket—to the island, but the government itself had suspicious entanglements with the financier. During the late 1990s, amidst an international crackdown on offshore banking, Stanford came to Antigua's aid by advocating the creation of a domestic regulatory board to improve the nation's image.[70] Stanford, owner of the largest bank regulated by the board, influenced the initiative and financed the entire operation,[71] epitomizing the proverbial conundrum: "*Quis custodiet ipsos custodes*?"[72]—colloquialized as "Who shall watch the watchmen?"[73] As a result of his efforts, however, Antigua was removed from the financial watch list in 2001, enabling it to avoid public shaming and higher scrutiny under the PCAOB's sliding-scale analysis.[74]

Given Antigua's financial stake in Stanford's operation at a national level and Stanford's connection to the personal interests of leading Antiguans, the nation appears incapable of second-guessing the suspicious

67. Creswell et al., *supra* note 34.

68. Jacqueline Charles, *Billionaire Stanford's Troubles Cause a Headache on Antigua*, MCCLATCHY, Feb. 22, 2009, http://www.mcclatchydc.com/world/v-print/story/62615.html.

69. *See id.* (noting fears that Stanford's downfall may cost Antigua thousands of jobs).

70. *See* Creswell et al., *supra* note 34 (attributing the creation of the board, in part, to Stanford's role as an adviser to Antigua's prime minister).

71. *Id.*

72. JUVENAL, DECII JUNII JUVENALIS ET A. PERSII FLACCI SATIRAE 139 (Arthur John Macleane ed., Whittaker & Co. 2d ed. 1867) (n.d.).

73. ALAN MOORE, WATCHMEN, epigraph (2005).

74. *See* Creswell et al., *supra* note 34 (describing how Antigua's removal from the watch list highlighted its apparent reform efforts).

business practices of its favorite son. Even in the face of public upheaval in the United States, Antigua's former Prime Minister came to the financier's defense.[75] The robust nexus between Stanford and Antigua's financial interests is perhaps nowhere more evident than in the nation's unwillingness to assist SEC investigators looking into Stanford's dealings before the fraud became public, as is discussed below. By currying favor among the local community with the proceeds of his fraudulent conduct, Stanford created a systemic regulatory advantage for himself, placing himself beyond reproach from the only regulators capable of reaching him abroad.

c. Fraud Flourishes Beyond the SEC's Jurisdiction.—Finally, Stanford was able to play on the SEC's lack of independent jurisdiction in Antigua to execute his fraud and hide assets after the fraud became public.[76] In 2005, following a series of whistle-blowers drawing attention to Stanford's suspicious returns,[77] the SEC initiated an enforcement investigation.[78] However, because of the lack of reliability of Antiguan regulators, the SEC did not receive sufficient information to bring charges for four years.[79]

During this period, the fraud ballooned from $3.8 billion in 2005 to $8.4 billion in 2009.[80] Moreover, if a fraudster were to learn from a local regulator that the SEC was investigating her company, she might be more inclined to make risky or speculative investments to cover the asset deficiency on corporate balance sheets. Alternatively, realizing that the proceeds from fraud are reaching a terminal point, the fraudster may elect to ship her assets to another OFC in an effort to bury the funds even farther away from SEC jurisdiction. The latter problem currently manifests itself in the Stanford

75. *See* Charles, *supra* note 68 ("'Stanford has done a reasonable job for the people of this country, and you don't have nothing [sic] to be ashamed of,' former Prime Minister Lester Bird said . . . . 'If he's done something wrong, let the law take its course. But when we brought him here, he helped to develop this nation.'").

76. *See* Michael Sallah & Rob Barry, *As Feds Closed In, Allen Stanford Scrambled to Keep Fraud Secret, Money Flowing*, MIAMIHERALD.COM, Dec. 6, 2009, http://www.miamiherald.com/business/v-fullstory/story/1368018.html (detailing Stanford's creative attempts to destroy financial records and disguise his company's assets).

77. *See* Hennessey, *supra* note 46 (detailing a series of fraud allegations dating back to 2002).

78. *Cf. id.* (quoting SEC spokesman John Nester, who stated, "The SEC was unable to obtain detailed information about the offshore CDs in part because of complicating factors involving jurisdiction, including whether CDs issued by a foreign bank are securities covered by the federal securities laws . . . and the application of Antiguan laws governing confidentiality . . . .").

79. *See* Hennessey, *supra* note 46 (quoting an SEC official who acknowledged "significant regulatory obstacles" that delayed the filing of a fraud report until 2009, though the investigation began in 2005); *cf.* Clifford Knauss, *Antigua Dismisses Regulator Charged in Stanford Case*, N.Y. TIMES, June 24, 2009, at B2 (discussing the events surrounding a top Antiguan regulator being investigated for helping to cover up the Stanford fraud).

80. Stanford Complaint, *supra* note 35, at 8.

drama, as the proceeds from the fraud appear to have been dispersed from SIB in the final months before the scandal became public knowledge.[81]

3. Red Flags.—From the still-settling dust of Stanford's fraud, a series of red flags emerge, providing a guide for what triggers might warrant SEC investigation under the model advocated by this Note.

a. The Business Model.—Stanford's business model and compensation scheme were premised on the ability to produce twice-market returns.[82] However, upon viewing the low-risk, liquid instruments allegedly utilized by Stanford, such returns would appear "unsustainable."[83] Moreover, the professed returns themselves were admittedly "improbable," particularly where SIB produced identical returns of 15.71% in 1995 and 1996.[84] However, the SEC's inability to verify local data and the false representations by the local auditor[85] made a proper investigation of these claims more difficult.

b. Suspicious Oversight.—SIB relied on a small, foreign auditor in a nation where the financier was known to be prominently involved in the community.[86] Ultimately, Stanford would use the auditor to feign the veracity of his returns and deceive investors.[87] Additionally, SIB's seven-member board lacked sufficient independence.[88] For example, the board included both Stanford's father and a childhood friend, who, following a stroke in 2000, "[could]n't string seven words together."[89] Disclosure of the purported auditor—a small, residential operation monitoring a multibillion-dollar investment conglomerate[90]—and a board beholden to its CEO ought to have independently, and certainly collectively, drawn substantial regulatory attention.

---

81. Driver & Baltimore, *supra* note 65.

82. *See* Matthew Goldstein & David Polek, *Are These CD Rates Too Good to Be True?*, BUS. WK., Feb. 23, 2009, at 22, 22 (questioning the "supposedly super-safe" CDs offered by Stanford that returned twice the market average).

83. Andy Meek, *Stanford Receiver Details Findings*, MEMPHIS DAILY NEWS, Apr. 24, 2009, http://www.memphisdailynews.com/editorial/Article.aspx?id=42138.

84. Stanford Complaint, *supra* note 35, at 2.

85. *See supra* notes 52–66 and accompanying text.

86. *See supra* notes 61–69 and accompanying text.

87. *See supra* notes 61–66 and accompanying text.

88. Marie Colvin & Dominic Rushe, *$12.6bn Has Disappeared from Texan Mogul's Offshore Bank*, THE AUSTRALIAN, Feb. 23, 2009, http://www.theaustralian.com.au/news/bn-disappears-from-moguls-bank/story-e6frg6tf-1111118931577.

89. Henry Blodget, *Stanford Bank's Very Independent Board of Directors*, BUS. INSIDER, Feb. 20, 2009, http://www.businessinsider.com/stanford-financials-independent-board-of-directors-2009-2.

90. *See supra* note 64 and accompanying text.

c. Where There's Smoke: Whistle-blowers and Other Investigations.—In 2002, an accountant in Mexico alerted the SEC to potential disparities in Stanford's business model.[91] The accountant alleged that his mother, a Stanford investor, had received little information regarding the actual performance of SIB CDs.[92] In particular, the accountant expressed concerns regarding the twice-market returns alleged by SIB.[93] In 2003, an unknown insider claimed that Stanford was perpetuating a "massive Ponzi scheme," noting that the CDs were marketed as safe despite being linked to speculative investments.[94] Additionally, the insider claimed that SIB had not received a "legitimate audit" in seventeen years.[95] A second whistle-blower came forward two years later to allege that "the CDs were 'simply a hedge fund.'"[96] Finally, in 2007, FINRA, the broker–dealer regulator, fined the Stanford Group $20,000 "for failing to adequately state the risks involved in the CD investments or to disclose that an affiliation between the broker-dealer and the bank could pose a conflict of interest."[97] Although these individual claims may be little more than efforts by disgruntled employees to undermine SIB's reputation separately, they collectively form a reasonable level of suspicion as to the veracity of Stanford's business model. Moreover, the weight of the allegations takes even greater force in light of the dubious consistency and high returns Stanford claimed to achieve.

d. Fraudster's Relationship with OFC.—Aside from exercising dual citizenship—a documented indicator of financial crime[98]—with Antigua,[99] Stanford had a variety of suspicious links to the people and government of Antigua.[100] Most notably, Stanford exercised control over the creation of Antigua's regulatory framework,[101] providing him with the opportunity to condition a favorable regulatory environment. As a result, the local government had inadequate incentives to regulate Stanford on its own

91. Hennessey, *supra* note 46.

92. *Id.*

93. *Id.*

94. *Id.*

95. *Id.*

96. *Id.*

97. David Scheer & Alison Fitzgerald, *Allen Stanford Accused of "Massive, Ongoing" Fraud*, BLOOMBERG.COM, Feb. 17, 2009, http://www.bloomberg.com/apps/news?pid=20670001&sid=aELwAubcAvR8.

98. *Cf.* Stephen Gray, *Second Citizenships: What's on the Market*, *in* LEGAL ISSUES IN OFFSHORE FINANCIAL SERVICES 41, 41 (Rose-Marie B. Antoine ed., 2004) ("Individuals[] who are concerned with divorce, bankruptcy, government expropriation, violent personal creditors, unwarranted government investigation, and repressive local government[] want the security of a second passport . . . .").

99. Scheer & Fitzgerald, *supra* note 97.

100. *See supra* notes 67–74 and accompanying text.

101. *See supra* notes 70–74 and accompanying text.

initiative and to assist the SEC in its investigations, rendering local regulation ineffective.

B. Samuel Israel III and Daniel E. Marino

> For the past seven years, I have committed a fraud of a great magnitude.[102]
>
> If there is a hell I will be there for eternity.[103]

1. The Con.—Unlike many cons, Bayou Management's CEO Sam Israel and CFO Dan Marino founded their hedge fund without any apparent intent to defraud investors.[104] After two initial years of investment losses, however, the duo initiated the fraud that would cost investors more than $450 million.[105]

The fraudulent conduct centered on issuing false statements regarding the firm's performance to cover up mounting losses and lure investors. According to a complaint filed by the Commodities Futures Trading Commission (CFTC), the nation's predominant commodities regulator, Israel and Marino "misappropriated customer funds, acquired funds through false pretenses, engaged in unauthorized trading, and misrepresented material facts to actual and prospective investors."[106] These misrepresentations extended to the "rates of return the hedge funds earned, the value of assets under management, and the existence and identity of the accounting firms that had purportedly audited the hedge funds."[107] Israel and Marino began misreporting financial data in 1998, two years after the fund's inception.[108] Realizing they could not withstand a genuine audit, the hedge fund fired its existing auditor and replaced it with Richmond-Fairfield Associates, a sham auditor constructed by the two conmen.[109]

---

102. Katherine Burton & Rob Urban, *Bayou Fraud Exposes Tale of Lies, Drugs, Violence*, BLOOMBERG.COM, Oct. 27, 2005, http://www.bloomberg.com/apps/news?pid=71000001&refer=&sid=aq3TjcUbSyX0.

103. Gretchen Morgenson, *What Really Happened at Bayou*, N.Y. TIMES, Sept. 17, 2005, at C1.

104. *See* Press Release, U.S. Attorney's Office S.D.N.Y., Chief Executive Officer of Bayou Funds Sentenced to 20 Years in Federal Prison for Massive Investor Fraud (Apr. 14, 2008), *available at* http://www.usdoj.gov/usao/nys/pressreleases/April08/israelsamuelsentencepr.pdf (reporting that Israel admitted that he and Marino hatched a scheme to lie to investors only after two years of sustained losses).

105. *Id.*

106. Press Release, Commodity Futures Trading Comm'n, Hedge Fund Operator Bayou Management, Its Employees Samuel Israel III and Daniel E. Marino, and Accounting Firm Richmond Fairfield Associates, Are Charged with Misappropriation and Fraud in an Action Brought by U.S. Commodity Futures Trading Commission (Sept. 29, 2005), *available at* http://cftc.gov/opa/enf05/opa5121-05.htm.

107. *Id.*

108. Kathleen E. Lange, *The New Antifraud Rule: Is SEC Enforcement the Most Effective Way to Protect Investors from Hedge Fund Fraud?*, 77 FORDHAM L. REV. 851, 851 (2008).

109. *Id.* at 852.

Over seven years, the firm would take in nearly $450 million of investments, relying on repeated misstatements regarding the firm's performance that were rubber-stamped by the phony auditor.[110] Despite the fictitious returns, Israel and Marino took fees on the trades, expanding their personal wealth at the expense of their investors.[111] The fraud would continue until the Arizona Attorney General detected a high volume of asset transfers between the United States, Europe, and Asia and froze the assets while they were briefly under U.S. jurisdiction, ultimately tracing the funds to a partnership established by the two conmen.[112]

2. The OFC Element.—While Bayou Management began as a legitimate hedge fund, it quickly turned fraudulent and made use of OFCs in two significant ways. First, the fraudsters created a series of offshore entities in the Cayman Islands and the Isle of Man to engage in high-risk, speculative investments that would not be disclosed to American investors in the hopes of raising real assets to replace the fictitious ones.[113] Second, recognizing their inability to cover their mounting losses and fearing exposure, the duo began hiding assets in obscure offshore partnerships, presumably to avoid seizure and repatriation.[114] It was during this latter stage that the Arizona Attorney General detected a high volume of international transfers and exposed the fraud.

a. Using OFC Secrecy to Replace Fictitious Assets with Real Ones.—Israel and Marino established seven offshore accounts in the Cayman Islands to solicit funds from foreign investors.[115] The money in these accounts was used to offset mounting losses in the United States.[116] The duo "routinely" transferred money raised in offshore accounts into the Bayou

110. *See* Jenny Anderson, *A Modest Proposal to Prevent Hedge Fund Fraud*, N.Y. TIMES, Oct. 7, 2005, at C6 ("[Bayou's] auditor, Richmond-Fairfield Associates, was a fake accounting firm created to produce false audits of Bayou. Bayou Securities was the broker-dealer through which trades were made to create real commissions for Bayou's principals, who used them as compensation on top of 20 percent incentive fee they made on their fraudulent returns."); David Elman, *Bayou Sinks into Chapter 11*, DAILY DEAL, June 1, 2006, LEXIS, News Library, DADEAL File ("Marino formed accounting firm Richmond-Fairfield Associates CPA PLLC to approve the false financial statements, which enabled Bayou to attract new investors.").

111. *See* Burton & Urban, *supra* note 102 (noting that Israel and Marino collected $23.3 million dollars in fictitious trade fees, while the former rented Donald Trump's house and the latter purchased several new luxury cars).

112. *Id.*

113. *See* Complaint at 7–8, CFTC v. Bayou Mgmt., No. 05 Civ. 8374 (S.D.N.Y. Sept. 29, 2005) [hereinafter Bayou Complaint] (relating the creation of four "successor" entities to Bayou Management that facilitated the fraudulent scheme); Burton & Urban, *supra* note 102 (relating the high-frequency trades that ultimately destroyed the value of the assets under the fund's supervision).

114. *See infra* notes 126–30 and accompanying text.

115. Complaint at 4–5, SEC v. Samuel Israel III, No. 05 Civ. 8376 (S.D.N.Y. Sept. 29, 2005) [hereinafter Israel Complaint].

116. *Id.* at 5.

funds.[117] The same fund was liquidated later that year, and the investment proceeds were redistributed into four separate successor funds.[118] Functionally, this series of transfers and liquidations enabled Israel and Marino to raise money while generating sufficient obscurity to move the funds as they deemed necessary without leaving the proverbial paper trail.

In the meantime, Israel and Marino established IM Partners under their exclusive ownership and financed their new endeavor with the proceeds from the successor funds.[119] Through IM Partners, the two conmen entered a series of high-risk private placements in the United States and abroad, which they never disclosed to investors, in the hopes of generating real assets to replace the fictitious ones on Bayou's books.[120] The duo relied on the strong secrecy laws of the Isle of Man and the Cayman Islands to protect the anonymity of their investments as IM Partners.[121] For example, the partnership invested in two companies based out of the Isle of Man, Kycos Ltd. and Debit Direct Ltd., both of which liquidated within a year of receiving IM Partners' investment, without having to provide any disclosure of the failures to investors.[122]

Additionally, Israel and Marino invested in a prime bank investment (PBI) scheme,[123] where investors are told that the fund is able to access a "supposedly secret market for the world's prime banks" that guarantees huge returns.[124] IM Partners invested Bayou's last $150 million into a PBI scheme based out of the Isle of Man.[125] Had investors been aware of these speculative offshore investments, they would likely have withdrawn their investments and brought suits against the two conmen for failing to disclose

117. *Id.* at 4–5.

118. *Id.* at 5, 7.

119. *See* Burton & Urban, *supra* note 102 (speculating that as much as $40 million from the Bayou funds ended up in IM Partners).

120. *See* Rob Urban & Katherine Burton, *Bayou Group Investors Seek to Recover More than $100 Million*, BLOOMBERG.COM, Sept. 30, 2005, http://www.bloomberg.com/apps/news?pid=10000103&sid=aHAoQBN1TF9s (reporting that IM Partners invested in a range of questionable projects, including an offshore financial-services firm and film productions).

121. *See generally* LEWIS D. SOLOMON & LEWIS J. SARET, ASSET PROTECTION STRATEGIES § 9.04, at 553–60 (2009 ed. 2008) (summarizing the bank-secrecy laws of the Cayman Islands); ORG. FOR ECON. CO-OPERATION & DEV., BEHIND THE CORPORATE VEIL: USING CORPORATE ENTITIES FOR ILLICIT PURPOSES 68 (2001), *available at* http://www.oecd.org/dataoecd/0/3/43703185.pdf (noting that Isle of Man tax authorities are not empowered to collect information to assist foreign investigations).

122. Burton & Urban, *supra* note 102.

123. Israel Complaint, *supra* note 115, at 9.

124. Judy Nichols, *Fast Work in Arizona Halts Fraud, Freezes $100 Million*, ARIZ. REPUBLIC, Dec. 11, 2005, http://www.azcentral.com/arizonarepublic/business/articles/1211bayou11.html. In most PBIs, the fund wires the investments through a series of entities before claiming to have lost the funds at the hands of the financial institutions handling the transfers. *Id.* Investors, stripped of their cash, generally are advised then to pursue settlement claims with the highly lucrative financial institutions, rather than litigate with the allegedly insolvent fund. *Id.*

125. Israel Complaint, *supra* note 115, at 9.

the redistribution of their finances and the initiation of new, high-risk investments.

b. Hiding Assets in Undisclosed Offshore Partnerships.—According to Arizona Assistant Attorney General Cameron Holmes, Israel had an alternative objective for engaging in high-risk offshore investments: "'He was creating a disappearing act with the money.'"[126] IM Partners began transferring the fund's remaining assets through a series of bank accounts in foreign cities, including London, Hamburg, and ultimately Hong Kong.[127] During this attempt to "layer"[128] the assets and place them beyond the reach of U.S. investors, the Arizona Attorney General, who had been investigating a separate fraud case, received notice that a bank in Pennsylvania had rejected a suspicious transfer from a Hong Kong bank.[129] Ultimately, these funds were linked to Israel's "last big trade"—his effort to hide his remaining assets.[130]

3. Red Flags.—Unlike Stanford, who had a deep connection with the OFC he exploited as part of his fraud, Israel and Marino employed their OFC connections with less sophistication. Namely, the duo used the easy licensing and low oversight standards in the Cayman Islands and the Isle of Man to establish an undisclosed partnership beyond the SEC's jurisdiction and engage in speculative trading.

a. Offshore Partnership.—Low licensing standards form one of the principal advantages of an OFC.[131] Fraudsters see these low standards, accompanied by little direct government supervision, as an ideal means to establish a vehicle to carry their ill-gotten proceeds.[132] As a result, Israel and Marino were able to establish IM Partners without any questions regarding how the organization was financed.[133]

b. Undisclosed, Speculative Investing.—Due to the high levels of secrecy in most OFCs, these nations do not require notice to investors when a

126. Burton & Urban, *supra* note 102 (quoting Arizona Assistant Attorney General Cameron Holmes).

127. *Id.*

128. *See* Alison S. Bachus, *From Drugs to Terrorism: The Focus Shifts in the International Fight Against Money Laundering After September 11, 2001*, 21 ARIZ. J. INT'L & COMP. L. 835, 844 (2004) (describing the role of layering, where dirty money is cycled through a series of financial transactions to clean it, as the second of three steps in money laundering).

129. Burton & Urban, *supra* note 102.

130. *Id.*

131. *See* ORG. FOR ECON. CO-OPERATION & DEV., *supra* note 121, at 28–29 (noting that many jurisdictions have light registration requirements for partnerships, thereby opening the way for persons to anonymously hold assets and to frustrate creditors' claims).

132. *Id.*

133. *See supra* notes 115–22 and accompanying text.

company is engaged in high-risk trading, despite representations to the contrary.[134] Accordingly, nations that have reputations as breeding grounds for speculative investments, like the Cayman Islands,[135] should naturally be regarded with greater suspicion. Thus, the nation itself, in addition to the nature of the investment, functions as a red flag to indicate to regulators that repeated investments in the same jurisdiction may involve unusual, undisclosed risk-taking.

c. Suspicious Auditor.—Finally, the use of a small, local, unknown auditor[136] to oversee an international operation with the reputation of Bayou Management, in which some of the world's premier hedge funds had invested,[137] ought to have raised suspicions.

C. Bernard Madoff

It's all just one big lie.[138]

1. The Con.—Of the conmen discussed in this Note, Bernard Madoff without argument is the most famous—or, more properly, infamous. His alleged $65 billion[139] Ponzi scheme[140] transcended geographic borders, from local pension funds in Fairfield to the world's largest financial institutions,

---

134. *See* Lynnley Browning, *A Hamptons for Hedge Funds*, N.Y. TIMES, July 1, 2007, at C1 (writing on the competition among OFCs to attract secretive hedge funds and other investors).

135. *See id.* (reporting that the Cayman Islands has become the preferred home to hedge funds due to the jurisdiction's lax tax and regulatory regime and stiff secrecy laws).

136. *See* Burton & Urban, *supra* note 102 (noting that one of the conmen had invented a phony auditor whose name—Richmond-Fairfield Associates—reflected his daily commute between Richmond and Fairfield counties).

137. Tremont Capital Management, "a well-respected funds manager," invested in Bayou Management from 2000 to 2002. Morgenson, *supra* note 103. Tremont Group Holdings would go on to invest $3.3 billion in Bernie Madoff's Ponzi scheme. Lynnley Browning, *Madoff Spotlight Turns to Role of Offshore Funds*, N.Y. TIMES, Dec. 31, 2008, at B1.

138. Jonathan Lopez, *Bernard Madoff, A Con Man for Our Times*, U.S. NEWS & WORLD REP., Mar. 16, 2009, http://www.usnews.com/articles/opinion/2009/03/16/bernard-maddoff-a-con-man-for-our-times.html.

139. The widely cited $65 billion figure is inaccurate, yet informative. This value reflects the "stated value of all the account statements of every Madoff account holder" as of Nov. 30, 2008. James Bandler & Nicholas Varchaver, *How Bernie Did It*, FORTUNE, May 11, 2009, at 50, 70. The actual amount invested in Madoff's fraudulent scheme is "closer to $20 billion." *Id.* at 71. Nevertheless, this amount constitutes one of the largest frauds in history.

140. A Ponzi scheme's success relies on its ability to attract multiple levels of investors. The proceeds derived from subsequent investments are used to satisfy prior investments, creating a false impression of high, consistent returns and inducing further investment. Over the course of several rounds, the player orchestrating the Ponzi scheme—Madoff in this case—can take a commission on each alleged "trade," which is little more than siphoning off a piece of a subsequent victim's investment. For a historical explanation of the scheme's origins, as well as various examples, see MITCHELL ZUCKOFF, PONZI'S SCHEME (2005).

such as Banco Santander and HSBC.[141] In many ways, the scale of Madoff's fraud epitomizes the ambition and arrogance of the modern con artist. Unlike Bayou Management's Israel and Marino, Madoff ran a fraudulent operation from start to a well-publicized end.[142] Even the legitimate wing of Madoff's investment empire was repeatedly funded through his illegitimate proceeds.[143] Additionally, the fraud lasted decades, while Madoff himself expanded his empire by trading on his professed success as an investor[144] and his reputation as a business leader. He was a chairman of NASDAQ, a leading voice on SEC-sponsored industry panels, and an expert witness for Congress.[145]

Eventually however, declining market conditions during the financial crisis of 2008 prompted many investors to seek withdrawals of their supposedly secure Madoff investments to offset losses elsewhere in the market.[146] Realizing he could not return the alleged assets to his investors, as they had been fraudulently recycled to perpetuate his Ponzi scheme, Madoff decided to confess and face the consequences in December 2008.[147]

2. The OFC Element.—OFCs factored into Madoff's schemes in a variety of ways. First, Madoff capitalized on the high secrecy laws of OFCs to establish offshore feeder funds to recruit foreign investors, who formed the majority of his stakeholders.[148] Second, Madoff may have used OFCs as tax havens to avoid paying taxes on the commissions he earned from fictitious trading.[149] In conjunction with this activity, Madoff may have tried to hide assets abroad to avoid detection by American regulators.

a. Offshore Feeder Funds.—The use of offshore funds hinges on the ability to generate unverifiable amounts of capital, which can either be distributed to new investors to encourage them to invest larger amounts or be

---

141. *See Madoff's Victim List*, WSJ.COM, Mar. 6, 2009, http://s.wsj.net/public/resources/documents/st_madoff_victims_20081215.html (providing a list of Madoff's defrauded investors around the world).

142. *See* Lopez, *supra* note 138 (quoting Madoff's admission that his entire operation was "all just one big lie").

143. *See* Bandler & Varchaver, *supra* note 139, at 64 ("Madoff's illegal investment business was indeed subsidizing his legal trading operation.").

144. On paper, Madoff's investments blossomed from nearly $7 billion in 2000 "to as much as $50 billion by the end of 2005." *Id.* at 66.

145. James Bandler & Nicholas Varchaver, *How Bernie Did It*, FORTUNE, Apr. 30, 2009, http://money.cnn.com/2009/04/24/news/newsmakers/madoff.fortune/index.htm?postversion=2009042405.

146. Bandler & Varchaver, *supra* note 139, at 69.

147. *Id.*

148. *See* Taina Rosa, *Scandal Prompts Crackdown*, LATIN FIN., May 1, 2009, http://www.latinfinance.com/ArticlePrint.aspx?ArticleID=2188873 (recognizing significant investments in Madoff's feeder funds by Argentine and Mexican investors).

149. James Doran, *Madoff Probe Focuses on Tax Havens*, GUARDIAN (U.K.), Dec. 28, 2008, at 1.

used as purported collateral to acquire entirely new investors. The strong secrecy laws of OFCs play a critical role in providing the necessary obscurity.[150] Many Latin American investors are drawn to secretive investments, as it enables them to avoid disclosing the true depth of their wealth.[151] As a result, many investors choose to invest in a fraudster's offshore fund through their own offshore entities. In the Madoff case, for example, the Fairfield Greenwich Group fund of funds invested $7.4 billion in Madoff's fund—much through offshore entities.[152] Fraudsters, such as Madoff, often play on the ability of offshore funds to avoid the normal rules, regulations, and tax consequences associated with economic activity in higher regulation regions to attract these secretive investors.[153]

Additionally, the true extent of funds raised offshore is rarely known in these highly secretive regions.[154] The advantage is apparent in a Madoff-type scenario, when the fraudster does not want investors to have knowledge of the true value or source of real assets invested in the firm at any time. Rather, the fraudster can manipulate investors by convincing them that the numbers he alleges are accurate.

b. Tax Shelters and Hidden Assets.—The role of OFCs as tax shelters is well documented[155] and beyond the scope of this Note, which focuses on the use of OFCs in cases of fraud. However, avoiding tax payments may have contributed to Madoff's selection of particular nations in which he could establish his feeder funds. In particular, selling foreign investors on low-tax regimes may have induced them to participate in his unique fund.[156]

Moreover, Chief Deputy Frank DiPascali—Madoff's right-hand man in orchestrating his fraud—confessed to years of editing the financial returns investors gained from the Madoff fund to offset any gains or losses those investors suffered in other investments to assist them in avoiding higher

150. *See supra* subpart II(A).

151. *See Latin Investors Could Be Big Madoff Losers*, *supra* note 53 (citing the "private nature of Latin American fortunes" as a reason that many of Madoff's Latin American investors are reluctant to come forward).

152. Browning, *supra* note 137.

153. Jonathan Kent, *Fraudsters Go Offshore for Business, Warns Asset Recovery Specialist*, ROYAL GAZETTE, Apr. 30, 2009, http://www.royalgazette.com/siftology.royalgazette/Article/article.jsp?articleId=7d94f2b3003000b&sectionId=65; *see also* Browning, *supra* note 137 ("Nearly all hedge funds, including funds of funds, operate affiliates and partnerships offshore. Such havens offer low tax rates and light regulation. Offshore havens also help fund managers to defer or avoid American taxes on their personal profits by channeling the earnings through offshore affiliates.").

154. *See, e.g.*, Bandler & Varchaver, *supra* note 139, at 70 (describing investigators' difficulty in locating Madoff's assets in offshore locations).

155. *See, e.g.*, BRITTAIN-CATLIN, *supra* note 24, at 35 (discussing the role of tax shelters in a variety of high-profile cases).

156. *See* Patti S. Spencer, *Beware of the Dirty Dozen*, INTELLIGENCER J., Apr. 27, 2009, at B8 (discussing the allure of "hiding income offshore" for tax reasons).

taxes.[157] In this way, DiPascali used the secretive offshore Madoff funds to cover up fraudulent activity in other investments. Additionally, Madoff himself may have "sent large sums of money to offshore accounts in the Caribbean and Europe" from his New York Mellon Bank account over time to hide assets from the IRS and SEC.[158]

3. Red Flags.—Three red flags emerge from the debris of Madoff's fraud. First, he relied heavily on offshore vehicles to sustain his investments. Second, he had been the subject of repeated SEC inquiries. Finally, like Stanford, Israel, and Marino, Madoff employed a suspicious auditor.

a. Reliance on Offshore Investments.—Madoff's Ponzi scheme relied on a constant stream of new investors. To recruit new investors—particularly those who wished to keep the nature and extent of their investments secret—Madoff relied heavily on his offshore feeder funds to attract Latin American investors.[159] Moreover, "the use of multiple jurisdictions to carry out trades" creates jurisdictional problems for suspicious regulators, enabling a fraudster "'to transfer enormous sums of money and perhaps do it under the radar.'"[160] Thus, not only did the offshore funds create an aura of secrecy, which Madoff manipulated to attract foreign investors, but they also enabled him to transfer large sums of money without revealing the source or destination to regulators.

b. Prior Investigations and Allegations.—Madoff was known to the SEC not only as a successful investor but also as the subject of numerous prior SEC investigations. In 1992, the SEC interrogated Madoff as part of its investigation into Avellino & Bienes, an investment firm that had long worked with Madoff.[161] Ironically, Madoff was able to assure the SEC that Avellino & Bienes had not participated in a Ponzi scheme, as alleged, and despite the fact that all of the fund's money was in Madoff's hands, no further investigation was pursued.[162]

In 2001, two articles "raised serious questions about Madoff's investment operation."[163] In particular, one article noted that Madoff's alleged returns made his fund one of the two largest hedge funds in the world at the time, yet few knew of its existence.[164] That article further questioned

157. Bandler & Varchaver, *supra* note 139, at 71.

158. Doran, *supra* note 149.

159. *See supra* notes 147, 150–53 and accompanying text.

160. Browning, *supra* note 137.

161. Bandler & Varchaver, *supra* note 139, at 64.

162. *Id.*

163. *Id.*

164. *Id.* The article was Michael Ocrant, *Madoff Tops Charts; Skeptics Ask How*, MAR/HEDGE, May 2001, *available at* http://www.realclearmarkets.com/articles/MarHedge.pdf.

the "improbability of Madoff's smooth and steady 15% annual returns."[165] The second article echoed the first, arguing that Madoff used his market-making operations to "subsidize[] and smooth[] his hedge-fund returns."[166] This last claim, though denied by Madoff, turned out to be true.[167] Finally, in 2006, whistle-blower Harry Markopolos outright announced that Madoff had been running a Ponzi scheme.[168] In conjunction with this investigation, the SEC took a brief look at a number of feeder funds linked to Madoff investments but lacked the jurisdiction to go farther, ultimately electing not to pursue the matter beyond requiring Madoff's firm to register as an investment adviser.[169]

c. Suspicious Auditor.—Like Bayou Management and Stanford Financial, Bernie Madoff employed a little-known auditor to monitor his vast financial empire. In fact, Friehling & Horowitz, the alleged auditor, had not performed an audit since 1993 and had only one certified public accountant.[170] To compound matters, their inactivity was not a secret—the auditor had informed the American Institute of Certified Public Accountants of those facts fifteen years before the Madoff fraud became public.[171]

IV. Expanding the SEC's Authority to Overcome Jurisdictional Limits and Deter Conmen

> [O]ur commitment [is] to extending the reach of the securities laws to violators no matter where they hide.[172]

As demonstrated in the previous Part, modern cons are capable of transcending geographic borders and growing radically in only a few years. However, these cons often trigger red flags through their use of OFCs, making them detectable before they grow out of hand. Unfortunately, the SEC's current model of foreign reliance in combating fraud is outdated, permitting conmen to take advantage of the SEC's limited reach and cooperation delays by hiding activity in low regulation OFCs. To overcome

---

165. *Id.*

166. Evin E. Arvedlund, *Don't Ask, Don't Tell: Bernie Madoff Is So Secretive, He Even Asks Investors to Keep Mum*, BARRON'S, May 7, 2001, *available at* http://online.barrons.com/article/SB989019667829349012.html.

167. Bandler & Varchaver, *supra* note 139, at 64.

168. *Id.* at 65.

169. *Id.* Incidentally, Madoff's registration as an investment adviser coincided with a widespread initiative among hedge funds to register under the Investment Advisors Act of 1940 as a result of a change in the SEC's regulations (this change was later struck down in the *Goldstein* case). *See* Goldstein v. SEC, 451 F.3d 873, 874 (D.C. Cir. 2006) (holding that the SEC had failed to justify the change in its interpretation under the relevant statute).

170. Alyssa Abkowitz, *Madoff's Auditor . . . Doesn't Audit?*, CNNMONEY.COM, Dec. 19, 2008, http://money.cnn.com/2008/12/17/news/companies/madoff.auditor.fortune/index.htm.

171. *Id.*

172. Christopher Cox, Chairman, SEC, Remarks at the American Securitization Forum (June 7, 2006) (transcript available at http://www.sec.gov/news/speech/2006/spch060706cc.htm).

the hurdle that OFCs pose to timely investigation—in the modest hope of making the international fraudster's con at least slightly more difficult to execute—this Note advocates expanding the SEC's authority to conduct investigations in foreign territories under particular circumstances.

A. Expanding SEC Authority

1. Source of Authority and Enforcement.—As a federal agency, the SEC's authority is subject to its legislative mandate from Congress.[173] Accordingly, "the nature and scope of the SEC's authorized activities are determined by statute and are subject to subsequent congressional override and judicial review."[174] In order to expand the SEC's jurisdiction, the agency will require an endorsement to that effect by Congress.

In particular, Congress should grant the SEC authority to prohibit financial services linked to any nation that does not grant the SEC the right to initiate investigations into local threats that have the potential for adversely and substantially impacting the security of U.S. financial markets. A failure to grant the SEC investigative privileges would result in a "blacklisted" designation, which would then prevent any company linked to that region from engaging[175] the U.S. financial markets. Recognizing the extensive links that companies currently have to OFCs, the investigative requirement would be instituted through a series of graduated benchmarks, ultimately resulting in the investigative right described forthwith.

2. The Logistics—When, Where, and for How Long?—Certain qualifications must be placed on the investigative right. These limitations pertain to the scope of the investigation permitted, including access and duration, the triggers that warrant investigation, and the right of particular nations to be exempted from the requirement.

a. Scope.—In order to institute the investigative right in a manner that balances the SEC's enforcement interest and local sovereignty, the SEC will have to agree to access and duration limitations. SEC access will have to be limited to the particular matter under investigation. Accordingly, the SEC would have to notify local regulators of the specific nature of the SEC's pending investigation before initiating the inquiry. Additionally, the SEC

173. *See* Joan MacLeod Heminway, *Rock, Paper, Scissors: Choosing the Right Vehicle for Federal Corporate Governance Initiatives*, 10 FORDHAM J. CORP. & FIN. L. 225, 253 (2005) ("As a general matter, agency rulemaking is substantively and procedurally subsidiary to legislative rulemaking. In a real sense, it represents a delegation of lawmaking authority by the legislature to the regulatory authority, with the process resulting in regulations having the force of law.").

174. *Id.* at 254–55.

175. "Engaging" the market would involve activities designed to solicit funds in the United States, including sales, advertising, and trading. Moreover, listing on any U.S. exchange would be prohibited for any company connected to a blacklisted nation.

may exempt certain industries on a case-by-case basis to avoid encroaching on local security interests.

With respect to duration, the SEC ought to agree to a maximum period of occupation with each nation before instituting the program based on the nature of the matter under investigation for two reasons. First, the global community is unlikely to tolerate any measure that resembles a foreign occupation, even if confined to specific markets. A reasonably well-defined time limit would enable the SEC to mitigate claims that it is capturing the sovereignty of foreign nations. Additionally, durational limits would benefit the SEC, in that the extent of their diligence would be protected by time limitations—only so many rocks can be overturned in a specified period.

The objective of the investigative right is to permit the SEC to research particular concerns in the hopes of combating international fraud before those crimes threaten the integrity of global markets. Limiting the scope of this right balances the value of threatening potential fraudsters and diligently executing the SEC's responsibilities against the costs of undermining local sovereignty.

b. Triggers.—By agreeing to predetermined triggers, the SEC would limit the scope of its investigative authority to an internationally acceptable level, as well as put local regulators and potential fraudsters on notice regarding particular activity that would be targeted. The types of events that would trigger an investigation mirror the red flags discussed earlier in the case studies.

These factors include a business model articulating unusual[176] returns, suspicious auditing services, registration in known tax shelters,[177] credible whistle-blower allegations, repeated SEC investigations, a reliance on offshore feeder funds to induce foreign investment, and a significant nexus between a principal at the company and an offshore government. Although an individual red flag may be insufficient to trigger an investigation, an accumulation of such flags could activate scrutiny. Thus, the focus would turn to patterns of conduct, rather than individual events. For example, where a fraudster reports unusual returns, a high volume of foreign investment, and significant contacts with an offshore government, an investigation should be triggered.

176. The term "unusual" would be determined by context—a task that exceeds the scope of this Note. However, the case studies provide some guidance into a preliminary understanding of what might constitute unusual returns. Stanford's consistent twice-market returns and Madoff's high returns in the face of a crumbling financial landscape, for example, seem unusual in that they exceed any reasonably anticipated returns.

177. The OECD recently provided an analysis of the world's most suspicious tax standards. *See* ORG. FOR ECON. CO-OPERATION & DEV., *supra* note 21. As a result, a number of nations have committed themselves to improving their tax standards. *See supra* note 25.

c. Exempted Nations.—The SEC ought to employ a sliding-scale analysis to determine which nations must comply with its investigative-right requirement. This sliding-scale assessment resembles the PCAOB approach discussed earlier;[178] however, it would be based on the SEC's determination of local-regulator independence and would account for a nation's willingness to grant the SEC investigative authority. As discussed in the juxtaposition of Switzerland and the Seychelles in subpart II(A),[179] OFCs vary widely in their regulatory schemes and purported uses. Nations with a track record of fraudulent exploitation, like the Cayman Islands, ought to grant the SEC greatest deference, while intermediate deference may be more appropriate for a nation like the Isle of Man, which has demonstrated a willingness to comply with higher regulation.[180] Finally, nations like India and China, which offer less regulation than the United States but still present sophisticated regulatory regimes, ought to be granted complete exemption.

Naturally, a nation's political capital in global affairs will play into its exempted status. In the case of Switzerland, for example, the nation may be exempted based on its prominence in world markets, under the justification that the Swiss maintain high regulations despite their robust secrecy laws.[181] This justification distinguishes a prominent global player like the Caymans, which may have a strong global presence, but also has a weak regulatory structure with respect to foreign capital.[182]

B. Current Model of Foreign Reliance Ineffective

Given the robust incentives against self-regulation, offshore regulators generally make poor partners in combating fraud.[183] To that end, the principal advantage of widening the SEC's net is the ability to overcome the lack of self-regulation offshore. Unfortunately, the SEC's current landscape of international cooperation in securities enforcement relies exclusively on assistance from local regulators.

The SEC widely touts empirical evidence that international cooperation in securities enforcement is working.[184] Undoubtedly, the volume of

178. *See supra* notes 56–60 and accompanying text.

179. *See supra* notes 21–22 and accompanying text.

180. *See* Mark D. Ferbrache, *Offshore Financial Centers*, FBI L. ENFORCEMENT BULL., Feb. 2001, *available at* http://findarticles.com/p/articles/mi_m2194/is_2_70/ai_72299785 ("The consensus of these reviews [of the Channel Islands and the Isle of Man] found that the Islands have a well-regulated financial industry, money laundering legislation in place, and a demonstrated willingness to cooperate with and provide assistance to foreign authorities.").

181. *See supra* note 21 and accompanying text.

182. *See* BRITTAIN-CATLIN, *supra* note 24, at 21–22 ("Cayman companies are creatures endowed with quite incredible rights that make them equivalent to, but distinct from, real, living people. . . . [O]nly in the radical companies of Cayman does capitalism find true freedom and power.").

183. *See supra* subpart II(B).

184. *See, e.g.*, Office of Int'l Affairs, SEC, International Enforcement Assistance, http://sec.gov/about/offices/oia/oia_crossborder.shtml (last modified Jan. 20, 2010) ("In fiscal year

international assistance has increased, but substantial evidence indicates that it is not "working." Currently, the SEC's main method of expanding its jurisdiction to foreign soil involves the use of Memoranda of Understanding (MOUs).[185] Pursuant to the specific terms of an MOU, the SEC may have the right to assist a foreign regulator in investigating a fraud that has occurred.[186] These agreements are insufficient mechanisms to address the problem of international fraud because they still require the local regulator to grant the SEC access after the fact. Where a local regulator does not want the SEC to be involved, the SEC has no choice but to stand down. Even if the SEC receives assistance, a vitriolic fraud may have expanded and infected markets around the globe. Moreover, the nonbinding nature of MOUs underlines their fundamental inefficiency. In particular, nonbinding agreements, like the IOSCO MMOU, offer no incentive comparable to the economic benefits of aiding fraudulent actors.[187] In addition, the MOUs embody the SEC's "ad hoc and incremental approach" to upgrading its "national-based regulatory approach" to the modern, global marketplace.[188] As a result, the SEC adapts too slowly to changing market conditions. Notably, it took the SEC nearly twenty years to negotiate an MOU whereby

---

2008, the SEC made 594 requests to foreign authorities for enforcement assistance and responded to 414 requests from foreign authorities."); *see also* Emily Chasan, *SEC's Cox Urges Global Regulators to Cooperate More*, REUTERS, Nov. 18, 2008, http://www.reuters.com/article/idUSN1827001920081118 ("In the past year the SEC has made 556 requests of foreign regulators for assistance with SEC enforcement investigations, and in turn foreign regulators have made 454 requests from the SEC.").

185. The SEC is currently party to over thirty separate memoranda of understanding. *See* Office of Int'l Affairs, SEC, Cooperative Arrangements with Foreign Regulators, http://sec.gov/about/offices/oia/oia_cooparrangements.shtml (last modified Jan. 20, 2010) (listing the various MOUs to which the United States is a party). Most of these agreements are bilateral; however, the SEC is party to the International Organization of Securities Commissions (IOSCO) MMOU, which has been signed by nearly fifty nations. *See* Int'l Org. of Sec. Comm'ns, List of Signatories to the Multilateral Memorandum of Understanding Concerning Consultation and Cooperation and the Exchange of Information (May 2002), http://www.iosco.org/library/index.cfm?section=mou_siglist (listing the IOSCO MMOU signatories).

186. There are five types of MOUs. In increasing order of technical cooperation, they are enforcement cooperation, regulatory cooperation, technical assistance, terms of reference for bilateral dialogues, and mutual recognition. *See* Office of Int'l Affairs, *supra* note 184 (describing each MOU by type). Only the last provides a mechanism for asset freezing and repatriation, which is a critical element of international cooperation in securities enforcement. To date, only one mutual-recognition MOU exists—between the SEC and the Australian Securities and Investments Commission. Memorandum of Understanding Concerning Consultation, Cooperation and the Exchange of Information Related to the Enforcement of Securities Laws, SEC-Austl. Sec. & Invs. Comm'n, Aug. 25, 2008, *available at* http://www.sec.gov/about/offices/oia/oia_mututal_recognition/australia/enhanced_enforcement_mou.pdf [hereinafter U.S.–Austl. MOU]; Christopher Cox, Chairman, SEC, Speech at the SEC International Enforcement Institute: The Importance of International Enforcement Cooperation in Today's Markets (Nov. 7, 2008) (transcript available at http://www.sec.gov/news/speech/2008/spch110708cc.htm).

187. *See* Verdier, *supra* note 26, at 117 ("Other initiatives by IOSCO, such as its failed effort to establish global capital standards for securities firms, point to the limits of informal cooperation when domestic interests clash.").

188. Edward F. Greene & Omer S. Oztan, *The Attack on National Regulation: Why We Need a Global Framework for Domestic Regulation*, 4 CAP. MARKETS L.J. 6, 9, 23 (2009).

the SEC could secure an agreement featuring an asset-freeze-and-repatriation clause.[189] This failing is particularly relevant in the fraudster scenario, where conmen use OFCs to create jurisdictional limits that slow the SEC's response. Furthermore, when a fraudulent actor ingratiates himself with a foreign government, as in the Stanford case, the SEC is unlikely to receive the regulatory invitation it desires.

Moreover, international coordination in securities enforcement among high-regulation nations does not appear promising. The G-20, composed of the world's most economically developed nations, has attempted to engage in a meaningful dialogue to impede the proliferation of low-regulation offshore financial industries.[190] However, the G-20 appears fixated on the role of OFCs as tax havens[191] rather than vehicles of securities fraud. While the G-20's recognition of prominent OFCs as a threat to global market integrity is a promising sign, it falls short of recognizing the effect of OFC-based securities fraud on global capital markets. As a result, the United States cannot rely on unsatisfying international efforts to increase offshore tax regulations in combating domestic securities fraud.

Nevertheless, in conjunction with an investigative mechanism like the kind advocated in this Note, MOUs may become increasingly relevant because they can describe the type of documentation and procedures that a foreign regulator must follow to make the SEC's investigation more efficient.[192] Most importantly, reinforcing the SEC's MOU initiative with a firm investigative right would overcome the slow, market-trailing philosophy

---

189. *See* U.S.–Austl. MOU, *supra* note 186 (providing for the freezing of assets and repatriation on August 25, 2008); Michelle Grattan, *Bilateral Deal Will Simplify Trading of Australian Shares in the US*, AGE (Austl.), Mar. 31, 2008, at 1 (indicating that this MOU was under negotiation for "some years"); Press Release No. 2008-52, SEC, SEC Chairman Cox, Prime Minister Rudd Meet Amid U.S.–Australia Mutual Recognition Talks (Mar. 29, 2008), *available at* http://www.sec.gov/news/press/2008/2008-52.htm (explaining that the partnership between the Australia Securities and Investment Commission and the SEC has been developed over "two decades").

190. *See* Jonathan Wiseman, *G-20 to Set New Rules for Tax Havens Under Regulatory Shake-Up*, WSJ.COM, Mar. 30, 2009, http://online.wsj.com/article_email/SB123825462934465293-lMyQjAxMDI5MzI4MzIyNTMOWj.html ("The 20 largest economic nations in the world are expected to produce a new set of rules for oversight, transparency and conduct for offshore tax havens . . . as part of a broader effort to overhaul the regulatory structure of the world economy . . . .").

191. *Cf., e.g.*, *OECD Names and Shames Tax Havens*, BBC NEWS, Apr. 3, 2009, http://news.bbc.co.uk/2/hi/7980848.stm ("G20 leaders agreed to take sanctions against tax havens using the OECD list as its basis.").

192. *See, e.g.*, Memorandum of Understanding Regarding Cooperation, Consultation and the Provision of Technical Assistance, SEC-Sec. & Exch. Bd. of India, Mar. 6, 1998, *available at* http://www.sec.gov/about/offices/oia/oia_bilateral/india.pdf (establishing some initial procedural mechanisms by which the SEC and SEBI can assist one another in investigating potential securities fraud). For an example of this MOU in practice, see *The Fraud at Satyam Is an Accident*, REDIFF.COM, Jan. 10, 2009, http://in.rediff.com/money/2009/jan/10satyam-the-fraud-at-satyam-is-an-accident.htm, discussing the cooperative efforts between the SEC and SEBI following the $1.4 billion Satyam Computer Services, Ltd. fraud case.

of SEC enforcement abroad and create a binding requirement to embrace SEC authorities.

C. *Existing Precedent*

Although the investigative right advocated in this analysis would be the greatest expansion of SEC authority to date, it is a proposal with precedent. Specifically, the development of the PATRIOT Act as a tool to combat international money laundering and the grant of expanded authority to the U.S. Treasury's Financial Crimes Enforcement Network (FinCEN) establish a foundation for the SEC to request enhanced authority to fulfill its mandate from American investors.

1. The SEC Mandate.—In a globalized world of high-speed technology,[193] the SEC must upgrade its enforcement efforts to handle the pace of modern transactions—and modern frauds—to honor its congressional mandate.[194] This mandate is embodied in the Securities Exchange Act of 1934,[195] which charges the SEC with the authority to maintain "fair and honest markets . . . vital to the national public interest of the United States."[196] Moreover, the SEC believes that its purpose is to "protect investors; maintain fair, orderly, and efficient markets; and facilitate capital formation."[197] To fulfill its purpose, the SEC must be able to protect investors by providing transparency to the fullest extent permitted by Congress. Moreover, markets wrought by fraud threaten the nation's security and public interest. In particular, the recent financial crisis has demonstrated the sensitivity of global markets and the role that fraud can play in causing market volatility.[198] Deterring fraudulent conduct by conmen is not only a necessity to maintain market integrity and fiscal security, but it is a requirement of the SEC's mandate from Congress. Congress should recognize that expanding the SEC's preemptive investigatory powers abroad will go a long way towards vitiating the Commission's "weaker" overseas presence.[199]

---

193. Ethiopis Tafara & Robert J. Peterson, *A Blueprint for Cross-Border Access to U.S. Investors: A New International Framework*, 48 HARV. INT'L L.J. 31, 36 (2007).

194. *Id.* at 42.

195. Ch. 404, 48 Stat. 881 (codified as amended in scattered sections of 15 U.S.C. (2006)).

196. Tafara & Peterson, *supra* note 193, at 43.

197. SEC, 2004–2009 STRATEGIC PLAN 4 (2009), *available at* http://sec.gov/about/secstratplan0409.pdf.

198. *See* Amitai Aviram, *Counter-cyclical Enforcement of Corporate Law*, 25 YALE J. ON REG. 1, 5 (2008) ("Downward trends are perceived very frequently to involve fraud and to be caused by fraud."); *see also* CHARLES P. KINDLEBERGER, MANIAS, PANICS, AND CRASHES: A HISTORY OF FINANCIAL CRISES 165–202 (2005) (discussing the role of fraud in economic panics and crashes).

199. *See* Tafara & Peterson, *supra* note 193, at 48–49 (claiming that the SEC is hesitant to permit overseas firms to access U.S. financial markets due to its supposedly weaker ability to supervise foreign activities).

This argument becomes even more significant in light of Congress's recent criticism of the SEC's handling of the Madoff and Stanford crises.[200] Given Congress's ongoing exploration of whether the SEC needs to upgrade its funding and policing practices,[201] the political climate is ripe for an evaluation of the SEC mandate in light of a more integrated, faster paced global economy.

2. House Committee Suggestions from 1990.—The current crisis would not be the first time Congress has radically recommended that the SEC upgrade its international cooperative model. In 1988, the House Committee on Government Operations widely criticized the SEC's handling of financial-crime cases tied to suspicious foreign activity.[202] The report recommended that the SEC, instead of entering nonbinding MOUs with suspicious nations, should "obtain congressional authority to prohibit securities trading from states that will not enter into an MOU."[203] Additionally, the report advocated granting "expanded long arm jurisdiction to the Commission . . . [, which] would authorize the Commission's service of subpoenas outside the United States while respecting the sovereignty of foreign states."[204] At the time, Congress rejected endorsing a "waiver-by-conduct" model[205] under concerns that it could lead to an "extraterritorial application of United States laws, which would eventually hinder international cooperation."[206] The motivation behind this rejection was a flight-of-capital concern based around a suspicion that the model would motivate foreign investors to withdraw from the U.S. markets and invest elsewhere, crippling America's financial industry.[207]

In many ways this concern seems outdated. First, markets have integrated beyond their 1987 levels.[208] Fraud in the United States

200. *See* Rachelle Younglai, *Congress to Examine SEC Enforcement Unit: Source*, REUTERS, Apr. 16, 2009, http://www.reuters.com/article/idUSTRE53F6HM20090416 (discussing a scheduled subcommittee hearing on the SEC's handling of the Madoff case and "whether the agency has enough resources to properly police the markets").

201. *Id.*; *see also infra* notes 228–29 and accompanying text.

202. In fact, the report was titled "Problems with the SEC's Enforcement of U.S. Securities Laws in Cases Involving Suspicious Trades Originating from Abroad," and it blasted the SEC's policy of relying on MOUs to monitor international securities violations. *See* John T. Thomas, Note, *Icarus and His Waxen Wings: Congress Attempts to Address the Challenges of Insider Trading in a Globalized Securities Market*, 23 VAND. J. TRANSNAT'L L. 99, 129–30 (1990) (summarizing the House Committee's suggestions). For the actual report, see H.R. REP. NO. 100-1065 (1988).

203. Thomas, *supra* note 202, at 129.

204. *Id.*

205. The "waiver-by-conduct" model treats investment in U.S. securities markets as tacit consent to disclose trade-related information, regardless of the protection of foreign law. John M. Fedders et al., *Waiver by Conduct—A Possible Response to the Internationalization of the Securities Markets*, 6 J. COMP. BUS. & CAP. MARKET L. 1, 25 (1984).

206. Thomas, *supra* note 202, at 130.

207. *Id.*

208. *See* Kuntara Pukthuanthong & Richard Roll, Global Market Integration: An Alternative Measure and Its Application 1 (Jan. 10, 2009) (unpublished manuscript, on file at http://papers.ssrn.

substantially affects foreign markets[209]—increasing the likelihood that greater regulation in the United States would be appreciated in other sophisticated financial markets. Moreover, as foreign markets have become more robust, they have become more susceptible to fraud as domestic investors seek arbitrage opportunities that transcend national boundaries. Second, the anticipated flight of capital, though empirically unclear, may not be as severe as one may have perceived in the past for two reasons. First, investors may feel greater confidence in markets with higher regulation, particularly after the well-publicized frauds of recent months. Second, the United States may have to sacrifice a portion of its collective wealth to protect the integrity of its markets—embodying the no-gain-without-pain philosophy.

3. The PATRIOT Act.—In response to the terrorist attacks of 9/11, Congress passed the PATRIOT Act, which included a section dedicated to expanding the Treasury Secretary's "long-arm jurisdiction" in pursuing money launderers.[210] Specifically, the International Money Laundering Abatement and Financial Anti-terrorism Act of 2001[211] (IMLA or Title III) recognized the propensity for "criminals . . . to avoid using traditional [U.S.] financial institutions . . . [and to] move large quantities of currency . . . [which] can be smuggled out of the United States."[212] Functionally, the Treasury Secretary has the right to subpoena "any foreign bank that maintains a correspondent account in the United States and request records related to such correspondent account, including records maintained outside of the United States relating to the deposit of funds into the foreign bank."[213] As a result, foreign "financial institutions will be required to submit to the jurisdiction which can exert the most pressure on the financial institution to comply with their laws," granting U.S. courts the authority to try potential money launderers.[214] Additionally, foreign banks are now required to maintain "identifying information," including names and addresses of depositors.[215] IMLA also grants legal immunity to financial

---

com/sol3/papers.cfm?abstract_id=1108487) (discussing the "marked increase in measured integration" despite flawed measurement techniques in global markets).

209. *See, e.g.*, *Latin America Takes Action over Local Stanford Companies*, ASSOCIATED FOREIGN PRESS, Feb. 19, 2009, http://www.google.com/hostednews/afp/article/ALeqM5jojgnWFEWfP4cS3JskYJ7P-tPY2Q (discussing the repercussions of the Stanford debacle on Latin America, with a particular focus on bank runs in response to the fraud).

210. USA PATRIOT Act, Pub. L. No. 107-56, tit. 3, 115 Stat. 272, 296–342 (2001) (codified as amended in scattered titles of U.S.C.).

211. *Id.* tit. 3, sec. 301 (codified as amended at 31 U.S.C. § 5301 note (2006)).

212. 31 U.S.C. § 5332 (2006).

213. *Id.* § 5318.

214. Evan Metaxatos, *Thunder in Paradise: The Interplay of Broadening United States Anti-Money Laundering Legislation and Jurisprudence with the Caribbean Law Governing Offshore Asset Preservation Trusts*, 40 U. MIAMI INTER-AM. L. REV. 169, 183 (2008).

215. 31 U.S.C. § 5318.

institutions that voluntarily disclose suspicious transactions;[216] however, this type of disclosure is unlikely in a highly secretive region like the Cayman Islands that places a premium on confidentiality. As a result, the provision fails where it is most needed. Nevertheless, IMLA represents a significant expansion of the Treasury Secretary's power to enforce anti-money-laundering laws.[217]

Although IMLA's coverage of money laundering differs substantively from securities fraud, the former is often a part of the layering efforts to hide proceeds from the latter. Moreover, Congress's willingness to act in the face of a recognized threat to American security demonstrates its capacity to authorize an agency, in the face of new developments in the international arena, to take on a greater role in preserving the national public interest.

Furthermore, despite the fact that the investigative right articulated in this Note goes farther than IMLA in encroaching on foreign sovereignty, the justification remains the same—the SEC's mandate to protect the national public interest extends to providing financial security through open and fair markets. A failure to provide these protections results not only in potential market inefficiencies but also places the financial security of the nation and its citizens at risk.

V. Applying the Investigative Right to the Techniques of the Three Conmen

The investigative privilege advocated by this analysis is intended simultaneously to upgrade the SEC's enforcement power into a new era of technology and transactional speed while attempting to dissuade conmen from instigating frauds by making one of their tools less useful: the obscurity and regulatory inaccessibility of OFCs. Although hindsight can only take a regulator so far, applying the investigative right to the case studies discussed above demonstrates that, while the cons may not have been avoidable, they certainly would have been more difficult to execute in the manner by which they occurred.

A. Sir Robert Allen Stanford

The Stanford fraud would have been mitigated greatly through the SEC's enhanced investigative authority. First, the SEC had knowledge of whistle-blower claims tied to Stanford's unsustainable investment model and dubious auditor, prior investigations of suspicious activity, and an awareness

---

216. Metaxatos, *supra* note 214, at 184.

217. The role of FinCEN also has expanded the Treasury's ability to alert investors to potentially fraudulent moneymaking schemes. *See generally* Fin. Crimes Enforcement Network, U.S. Dep't of the Treasury, What We Do, http://www.fincen.gov/about_fincen/wwd (detailing FinCEN's authority to collect and analyze information from a large number of financial institutions).

of Stanford's pervasive relationship with Antiguan regulators.[218] Moreover, the Commission had this information in 2005, when the fraud was valued at $3.8 billion.[219] Had the SEC entered Antigua directly, rather than relying on the self-interested regulator to produce information, the Commission likely would have detected the fraudulent auditing scheme—it was grossly apparent to anyone who simply visited the auditor's office[220]—thereby disarming a critical element used to falsify investor returns. Instead, the SEC was forced to rely on local regulators, costing investors another $4.6 billion.[221]

B. Samuel Israel III and Daniel E. Marino

Of the three studies analyzed herein, the Bayou fraud would have been the least susceptible to the investigative right; however, it still would have been materially affected. First, investigative authority alone may not overcome the high secrecy and involvement of offshore centers in committing fraud. For instance, even if the SEC suspected fraud based on Bayou's suspicious returns and radical decision to terminate operations in 2004,[222] it would have had a difficult time initially determining which OFC to investigate. Moreover, once an OFC had been identified, the SEC would have had to articulate a specific scope to their investigation with limited knowledge of the fraud. Nevertheless, through the signals identified by credible whistle-blowers,[223] the SEC would have had some direction in determining which sectors to investigate. Ultimately, though the Bayou fraud may not have been avoided, the SEC could have instituted investigations that might have illuminated suspicious wire transfers between Bayou's bank accounts and IM Partners' accounts in obscure OFCs. Moreover, a threat of deeper investigation might have motivated the duo to utilize a legitimate auditor. This threat alone might have mitigated the desperate con employed by Sam Israel.

C. Bernard Madoff

The investigative right would have had an intermediate effect on Madoff's expansive fraud by significantly handicapping the fraudster's ability to employ offshore funds to raise and hide capital. Relying on whistle-blower allegations of a Ponzi scheme and an awareness that the Madoff business model relied heavily on foreign investment into its offshore feeder funds, the SEC could have initiated investigations in nations housing

---

218. *See supra* notes 91–101 and accompanying text.

219. *See supra* notes 77–80 and accompanying text.

220. *See* Szep, *supra* note 61 (highlighting the unreliability of Stanford's Antiguan auditors).

221. Stanford Complaint, *supra* note 35, at 8.

222. Burton & Urban, *supra* note 102.

223. *See id.* (discussing the activity of two potential whistle-blowers, Paul Westervelt and Greg Lopak, who noted potential securities-law violations at Bayou).

these funds.[224] Moreover, the presence of regulators in Barbados and other OFCs employed by Madoff as part of his alleged asset-hiding scheme might have dissuaded the conman from attempting to bury funds rather than cycling them back to investors. In addition, investors concerned with hiding their own personal wealth might have elected not to invest in a fund that would be subject to such potential scrutiny. Ultimately, a single change cannot overcome a fraud in excess of $20 billion; however, the threat of expanded jurisdiction might decrease the scale of a considered con or aid in the repatriation efforts following a con.

VI. Conclusion—The Obama Administration and Protecting the National Public Interest

> If financial institutions won't cooperate with us, we will assume that they are sheltering money . . . and act accordingly.[225]

Over a half century since Bonnie and Clyde terrorized the American South,[226] a new breed of con artist has emerged. Rather than blowing up bank vaults for a few thousand dollars, these individuals perpetrate frauds that cost investors billions. The tools of the trade have changed as well—from Tommy guns and plastic explosives to fake auditors and obscure offshore regulations.

The danger, however, remains the same. These con artists pose a threat to the national public interest, undermining the notion of fair and open markets with obscure avenues to hide clandestine activity. The sheriff in this scenario, charged with the mandate to promote fair disclosure, is the SEC. However, it currently lacks the ability to follow these conmen out of town, preserving the fraudster's ability to work around the SEC through nations that lack the impetus to prosecute financial fraud. In fact, these fringe players often appear complicit in the crimes, demonstrating the infeasibility of offshore self-regulation. As a result, the sheriff needs the authority to pursue these con artists in nearby towns, particularly those that appear to invite fraudulent conduct.

Given the Obama Administration's dedication to combating offshore tax havens,[227] the SEC may have its best opportunity to push for greater offshore jurisdiction. Not only does the President seem open to tackling offshore tax havens, but Congress also recently investigated the SEC's ability to pursue

---

224. These nations include Barbados and several Latin American states. Rosa, *supra* note 148.

225. President Barack Obama, Remarks on International Tax Policy Reform (May 4, 2009) (transcript available at http://www.whitehouse.gov/the-press-office/remarks-president-international-tax-policy-reform).

226. *See* Fed. Bureau of Investigations, Famous Cases: Bonnie and Clyde, http://www.fbi.gov/libref/historic/famcases/clyde/clyde.htm (summarizing the infamous duo's criminal exploits in the 1930s).

227. *See, e.g.*, Jackie Calmes & Edmund L. Andrews, *Obama Asks Curb on Use of Havens to Reduce Taxes*, N.Y. TIMES, May 5, 2009, at A1 (reporting on President Obama's speech on May 4, 2009).

fraudsters.[228] Moreover, the SEC's own purpose and productivity is under scrutiny,[229] leading to the creation of an entirely novel division focused for the first time on prospective regulation—the Division of Risk, Strategy, and Financial Innovation.[230] Given the publicity, public outrage, and loss of confidence stemming from the acts of recent conmen, the SEC has its best chance to advocate a greater role for itself in fulfilling its mandate to preserve the national public interest and to reestablish its credibility among critics.

Eight years ago, Congress recognized financial crime as a potential threat to the nation's security. Now Congress has the ability to take another step forward in preserving the integrity of American markets, and by extension, the security of American investors. Congress now has both the motive and the opportunity, so it should empower the sheriff to take down the conman.

—Nick S. Dhesi

228. *See S.E.C. Vows to Reorganize Unit to Head Off Fraud*, N.Y. TIMES, Sept. 11, 2009, at B5 (describing the inquiry into the SEC's failures).

229. *See* John C. Coffee, Jr. & Hillary A. Sale, *Redesigning the SEC: Does the Treasury Have a Better Idea?*, 95 VA. L. REV. 707, 714–17 (2009) (evaluating the merits of both the SEC and the Treasury and ultimately advocating regulatory consolidation); *see also* Jill E. Fisch, *Top Cop or Regulatory Flop? The SEC at 75*, 95 VA. L. REV. 785, 796, 823 (2009) (responding to Coffee and Sale and concluding that whatever regulatory changes occur should "emphasize transparency and enforcement").

230. Press Release, SEC, SEC Announces New Division of Risk, Strategy, and Financial Innovation (Sept. 16, 2009), *available at* http://www.sec.gov/news/press/2009/2009-199.htm.

A Comparative "Hard Look" at *Chevron*: What the United Kingdom and Australia Reveal About American Administrative Law*

Students of American administrative law often notice a striking anomaly in judicial deference to agency decisions: Judges are experts in interpreting statutes yet give great deference to agencies' legal interpretations, whereas judges closely review agencies' policy determinations despite the agencies' superior expertise and accountability.[1] Scholars and judges have long noted this anomaly, observing that one might expect precisely the opposite—that courts would review questions of law de novo but generally defer to agencies' policy decisions.[2] After all, why should courts give substantial deference to the statutory interpretations of agencies when statutory interpretation is a core judicial function?[3] Conversely, why should unelected judges exercise a strict standard of review over policy decisions made by agencies when agencies are subject to political pressures from the two elected branches of American government?[4]

This Note offers one explanation of the anomaly through a comparative analysis of American, British, and Australian political institutions. The anomaly is conspicuously absent in the United Kingdom and Australia.[5] Judges in those countries give no deference to administrative interpretations

---

* I am extremely grateful to Professor Peter Cane of The Australian National University College of Law for his invaluable guidance in researching and revising this Note. Likewise, I thank Professor Matthew Spitzer for his very helpful comments and suggestions. I also greatly appreciate the hard work and thoughtful recommendations of the members and editors of the Texas Law Review, particularly Daniel Aguilar, Daniel Cohen, Katherine Hacker, James Hughes, Jessica Miller, and Brice Wilkinson—thank you all. Finally, I would like to thank my family for their continued support.

1. *See, e.g.*, Stephen Breyer, *Judicial Review of Questions of Law and Policy*, 38 ADMIN. L. REV. 363, 364–65 (1986) ("[C]urrent doctrine is anomalous [because i]t urges courts to defer to administrative interpretations of regulatory statutes, while also urging them to review agency decisions of regulatory policy strictly.").

2. *See, e.g.*, STEPHEN G. BREYER ET AL., ADMINISTRATIVE LAW AND REGULATORY POLICY: PROBLEMS, TEXTS, AND CASES 403 (6th ed. 2006) ("Might a foreigner . . . exclaim: 'How odd. The American courts defer to agencies on questions of law, where courts are expert, but they conduct "in-depth" reviews of policy, where agencies are expert. They seem to have it backwards.'").

3. *Id.*

4. *See infra* notes 89–90 and accompanying text. This strict standard of review over policy is especially puzzling, given that a principal justification for *Chevron* deference rests on the superior democratic accountability of agencies over courts. *See* Chevron U.S.A. Inc. v. Natural Res. Def. Council, Inc., 467 U.S. 837, 865 (1984) ("While agencies are not directly accountable to the people, the Chief Executive is, and it is appropriate for this political branch of the Government to make such policy choices . . . ."). Why should this accountability argument not apply *a fortiori* with respect to agencies' policy decisions?

5. *See infra* Part II.

of law but grant substantial deference on policy decisions.[6] One important factor underlying this dichotomy is that the emergence of the administrative state has generated more significant institutional tensions in the United States than in the United Kingdom and Australia. In the British and Australian parliamentary systems, the executive and the legislature are functionally fused: the Prime Minister controls both government as well as Parliament.[7] As a result, there are no interbranch power struggles over control of the administrative state. In contrast, American administrative agencies are often caught in a tug-of-war between the President and Congress.[8] Furthermore, unlike the unitary executives of both Australia and the United Kingdom, the American Executive is bifurcated between the President and the administrative agencies.[9] Therefore, the American system is also more prone to intra-Executive discord than are the British and Australian systems, where executive power is more consolidated.

Thus, the American anomaly has its origins in the U.S. Constitution's separation of powers and the difficulties of applying our three-branch system of checks and balances to a fourth branch—the administrative state. The framers of the U.S. Constitution devised an ingenious system of checks and balances between the three branches of government. However, the framers did not envision the massive growth of the administrative state. The British and Australian legal systems were able to assimilate the administrative state with less complexity, in that administrators are situated squarely under the control of the executive.[10] As a result, the United Kingdom and Australia have largely avoided the contentious issues of accountability and deference that plague American administrative law.[11]

The principal difference between the American system on the one hand and the two parliamentary systems on the other hand is the issue of electoral

---

6. *See infra* Part II.

7. *See* AREND LIJPHART, PATTERNS OF DEMOCRACY: GOVERNMENT FORMS AND PERFORMANCE IN THIRTY-SIX COUNTRIES 11–12 (1999) (indicating that although the House of Commons theoretically can vote a cabinet out of office, "in reality, the relationship is reversed" so that "[t]he cabinet is clearly dominant vis-à-vis Parliament"); Parliament of Austl., Parliament: An Overview, http://www.aph.gov.au/PARL.HTM ("The Prime Minister is appointed by the Governor-General, who by convention under the Constitution, must appoint the parliamentary leader of the party, or coalition of parties, which has a majority of seats in the House of Representatives.").

8. Scholars disagree about whether or not Congress controls agencies. *See, e.g.*, Barry R. Weingast & Mark J. Moran, *Bureaucratic Discretion or Congressional Control? Regulatory Policymaking by the Federal Trade Commission*, 91 J. POL. ECON. 765, 766–67 (1983) (dividing scholarly views of agency policy making into two approaches: the traditional approach, which considers agencies to be independent of Congress, and the congressional-dominance approach, which argues that congressional committees exert control over agency decisions).

9. *See infra* subpart III(B) (contrasting the American bifurcated Executive with the British and Australian systems).

10. *See supra* note 7 and accompanying text.

11. *See infra* Part III; *see also* Michael Asimow, *Delegated Legislation: United States and United Kingdom*, 3 OXFORD J. LEGAL STUD. 253, 253 (1983) (noting that in Britain "nearly everyone seems satisfied with . . . procedural and substantive aspects of delegated legislation").

accountability. Voters in the United Kingdom and Australia know whom to hold accountable for administrative action: the party in power. American voters, in contrast, cannot attribute the blame or credit for administrative action quite as clearly: is Congress primarily responsible—since it passed the relevant legislation—or is the President responsible, since the Executive implements the legislation?

The seemingly inconsistent American deference doctrines for questions of law and policy actually represent rational judicial responses to imbalances created by the unanticipated growth of a fourth branch in a three-branch system of checks and balances. The resulting institutional tensions inherent in American administrative law cannot be resolved by current judicial doctrines alone. Thus, these tensions will likely continue to manifest themselves through—at times—seemingly inconsistent applications of judicial deference to administrative action, as judges attempt to strike a balance between two often conflicting values: the desire to defer to the superior expertise and political accountability of administrators on the one hand, and the desire to ensure that administrative action accurately tracks congressional mandates on the other.[12]

This Note will proceed in six Parts. Part I reviews the nature of the American anomaly. Part II briefly describes the deference regimes in Australia and the United Kingdom. In Part III, an institutional comparison reveals structural differences that affect the deference regimes in each country. Next, Part IV attempts to help explain the American anomaly in light of this institutional comparison. Part V examines what the comparative analysis suggests about the future of American deference doctrines. Finally, Part VI concludes.

I. The American Anomaly

The American anomaly is that courts give substantial deference to agencies' interpretations of law but grant far less deference to agencies' policy decisions.[13] Yet statutory interpretation traditionally falls within the province of the courts,[14] and judges typically prefer to avoid second-guessing the policy choices of the political branches of government.[15] Why, then, do American courts turn this natural order on its head when reviewing administrative action? In considering that question, this Part discusses the

12. An analysis of the motivations of judges reviewing administrative action is far beyond the scope of this Note. Rather, I intend to show merely that institutional tensions inherent in the U.S. constitutional structure have generated these conflicting pressures on the Judiciary with respect to administrative action.

13. *See supra* note 1 and accompanying text.

14. Elizabeth V. Foote, *Statutory Interpretation or Public Administration: How* Chevron *Misconceives the Function of Agencies and Why It Matters*, 59 ADMIN. L. REV. 673, 674 (2007).

15. *See, e.g.*, Antonin Scalia, *Judicial Deference to Agency Interpretations of Law*, 1989 DUKE L.J. 511, 515 ("Under our democratic system, policy judgments are not for the courts but for the political branches . . . .").

influence of two seminal cases that embody the traditional doctrines of deference to agency determinations of law and policy: *Chevron U.S.A. Inc. v. Natural Resources Defense Council, Inc.*[16] (questions of law) and *Motor Vehicle Manufacturers Ass'n v. State Farm Mutual Automobile Insurance Co.*[17] (questions of policy). This Part thus provides a basic—and admittedly oversimplified[18]—outline of the American anomaly.

A. Questions of Law: Chevron *Deference*

In modern literature on American administrative law, any discussion of judicial review of agencies' interpretations of statutes almost always includes *Chevron*. Upon its decision in 1984, *Chevron* was immediately recognized as a particularly significant administrative law opinion.[19] Two decades later, Cass Sunstein described *Chevron* as "unquestionably and by far the most important case about legal interpretation in the last thirty years."[20] It has been cited more often than *Marbury v. Madison*,[21] *Brown v. Board of Education*,[22] and *Roe v. Wade*.[23] If it does not already hold the distinction, *Chevron* may become the most cited case in American public law.[24]

Although the Administrative Procedure Act (APA) states that "[t]he reviewing court shall decide all relevant questions of law" and "interpret constitutional and statutory provisions,"[25] courts—even before *Chevron*—

16. 467 U.S. 837 (1984).

17. 463 U.S. 29 (1983).

18. The deference actually afforded to agencies by courts is much more complex in practice than this simplified dichotomy suggests. *See infra* Part V; *see also, e.g.*, William N. Eskridge, Jr. & Lauren E. Baer, *The Continuum of Deference: Supreme Court Treatment of Agency Statutory Interpretations from* Chevron *to* Hamdan, 96 GEO. L.J. 1083, 1090 (2008) (finding that "*Chevron* was applied in only 8.3% of Supreme Court cases evaluating agency statutory interpretations" and demonstrating that the Supreme Court employs a "*continuum of deference regimes*"). Eskridge and Baer further claim that:

> This continuum is more complicated than the literature or even the Court's own opinions suggest, and it is a continuum in which *Chevron* plays a modest role. Indeed, our most striking finding is that in the majority of cases—53.6% of them—the Court does not apply any deference regime at all. Instead, it relies on ad hoc judicial reasoning of the sort that typifies the Court's methodology in regular statutory interpretation cases.

Id. The deference regimes are simplified in this discussion for purposes of highlighting the American anomaly. In Part V, *infra*, I argue that such complexity is to be expected, given the inherent institutional tensions in American administrative law.

19. RUTH ANN WATRY, ADMINISTRATIVE STATUTORY INTERPRETATION: THE AFTERMATH OF *CHEVRON V. NATURAL RESOURCES DEFENSE COUNCIL* 6 (Eric Rise ed., 2002).

20. CASS R. SUNSTEIN, ONE CASE AT A TIME: JUDICIAL MINIMALISM ON THE SUPREME COURT 227 (1999).

21. 5 U.S. (1 Cranch) 137 (1803).

22. 347 U.S. 483 (1954).

23. 410 U.S. 113 (1973). By December 2005, federal courts had cited *Chevron* almost 8,000 times, far eclipsing the number of cites for *Brown v. Board of Education* (1,829 cites), *Roe v. Wade* (1,801 cites), and *Marbury v. Madison* (1,559 cites). BREYER ET AL., *supra* note 2, at 247.

24. SUNSTEIN, *supra* note 20, at 227.

25. 5 U.S.C. § 706 (2006).

have traditionally held that the agencies shall decide at least *some* questions of law.[26] Courts reasoned that in these cases, Congress has statutorily granted the agency the authority to make such a legal determination or interpretation.[27] But when such congressional intent is not clear, the courts remain the final arbiters of questions of law.[28]

Nonetheless, while remaining the ultimate arbiters of law, courts have given varying degrees of deference to agency determinations of questions of law. In the seminal pre-*Chevron* case, the Supreme Court in *Skidmore v. Swift & Co.*[29] noted that while the "rulings, interpretations and opinions" of an administrator are not controlling upon the courts, such determinations *do* "constitute a body of experience and informed judgment" to which courts can "properly resort for guidance."[30] In short, courts can rely on the agency determination as persuasive authority—a weak form of deference. The persuasive weight of an agency determination in a given case "will depend upon the thoroughness evident in its consideration, the validity of its reasoning, its consistency with earlier and later pronouncements, and all those factors which give it power to persuade, if lacking power to control."[31] These are the famous *Skidmore* factors, which courts would examine to determine how much deference to afford an agency's interpretation of law.[32]

However, the application of these factors was by no means an exact science. Forty years after *Skidmore*, the Supreme Court in *Chevron* seemingly adopted a simplified approach.[33] In *Chevron*, the Court famously set out a two-step test for reviewing an agency's interpretation of a statute.[34] In "Step One," the reviewing court must determine "whether Congress has directly spoken to the precise question at issue."[35] If Congress has directly spoken, then the court must apply Congress's clear meaning.[36] If not, then the court proceeds to "Step Two."

---

26. BREYER ET AL., *supra* note 2, at 232.

27. *Id.* (citing Henry P. Monahan, Marbury *and the Administrative State*, 83 COLUM. L. REV. 1, 25 (1983)).

28. *See, e.g.*, FEC v. Democratic Senatorial Campaign Comm., 454 U.S. 27, 32 (1981) ("[C]ourts are the final authorities on issues of statutory construction [and] must reject administrative constructions of the statute . . . that are inconsistent with the statutory mandate or that frustrate the policy that Congress sought to implement.").

29. 323 U.S. 134 (1944).

30. *Id.* at 140.

31. *Id.*

32. *See, e.g.*, United States v. Mead Corp., 533 U.S. 218, 228 (2001); River St. Donuts, L.L.C. v. Napolitano, 558 F.3d 111, 116 (1st Cir. 2009); Choin v. Mukasey, 537 F.3d 1116, 1120 (9th Cir. 2009) (all quoting this language from *Skidmore*).

33. BREYER ET AL., *supra* note 2, at 241.

34. Chevron U.S.A. Inc. v. Natural Res. Def. Council, Inc., 467 U.S. 837, 842 (1984).

35. *Id.*

36. *Id.* at 842–43.

The Step Two inquiry is whether the agency's interpretation "is based on a permissible construction of the statute."[37] The Court further explained that "a court may not substitute its own construction of a statutory provision for a reasonable interpretation made by the administrator of an agency."[38] Therefore, "the *Chevron* two-step" instructs reviewing courts to determine: (1) whether Congress's intent is clear or ambiguous and (2) if ambiguous, whether the agency's interpretation is "reasonable."[39] If the agency's interpretation is reasonable, the court will uphold the agency's decision.

Chevron thus established an extraordinarily deferential standard of review for agency statutory interpretations. As Professor Elizabeth Foote has observed, "[c]ourts rarely invoke unreasonableness as a ground for setting aside agency action."[40] Indeed, Foote concluded that "it is hard to find a single case in which the Court deemed the agency action unreasonable at the second step of the [*Chevron*] test."[41] Accordingly, *Chevron* created a judicial review doctrine that holds that the interpretation of ambiguous regulatory and administrative statutory provisions is primarily the province of the agencies and not the courts.[42]

Chevron does not tell the whole story, however. Recent Supreme Court decisions have provided means of narrowing *Chevron*'s scope, including the application of another inquiry that is sometimes referred to as *Chevron* Step Zero: the gateway question of whether or not the *Chevron* framework applies at all.[43] Furthermore, the Supreme Court does not always apply *Chevron* deference when reviewing agency statutory interpretations.[44] Nonetheless, *Chevron* is still good law, and as the most famous and most cited case in

37. *Id.* at 843.

38. *Id.* at 844.

39. *Id.* at 842–44. Matthew Stephenson and Adrian Vermeule have argued that, logically, the *Chevron* two-step really only has one step. Matthew C. Stephenson & Adrian Vermeule, Chevron *Has Only One Step*, 95 VA. L. REV. 587, 599–600 (2009). If an agency's interpretation is contrary to clear congressional intent (Step One), then it is also necessarily not a permissible construction (Step Two). *Id.* Therefore, all cases that could be addressed by Step One are merely a subset of cases that could be decided under Step Two, so the *Chevron* test could be reduced to just the Step Two inquiry. *Id.*

40. Foote, *supra* note 14, at 708.

41. *Id.* at 709; *see also* BREYER ET AL., *supra* note 2, at 247 ("As of this writing, no Supreme Court decision has invalidated an agency decision under step 2, though several courts of appeals decisions do so.").

42. *See* BREYER ET AL., *supra* note 2, at 247 ("If *Chevron*'s basic holding is that ambiguous statutory terms should be interpreted by agencies rather than courts, *Chevron* can be seen as a kind of counter-*Marbury* for the administrative state.").

43. Cass R. Sunstein, Chevron *Step Zero*, 92 VA. L. REV. 187, 191 (2006); *see also, e.g.*, United States v. Mead Corp., 533 U.S. 218, 226–27 (2001) ("[A]dministrative implementation of a particular statutory provision qualifies for *Chevron* deference when it appears that Congress delegated authority to the agency generally to make rules carrying the force of law, and that the agency interpretation claiming deference was promulgated in the exercise of that authority.").

44. *See* Eskridge & Baer, *supra* note 18, at 1090 (finding that the Supreme Court does not consistently apply *Chevron* but rather applies a continuum of deference doctrines to agency statutory interpretations).

American administrative law,[45] it represents the doctrine that courts should give substantial deference to agencies' legal interpretations.

B. *Questions of Policy:* State Farm *Deference*

A year before *Chevron*, the Supreme Court handed down another of the most important cases in American administrative law: *State Farm*. Whereas *Chevron* is the best known case on review of agencies' determinations of law, *State Farm* is perhaps the best known case[46] concerning judicial review of agencies' policy choices.[47] And as the *Chevron* doctrine grants deference to agencies' legal interpretations despite the APA's language that courts "shall decide all relevant questions of law [and statutory interpretation],"[48] the *State Farm* doctrine provides for more searching judicial review than is provided by the APA's "arbitrary and capricious" standard.

Under the APA,[49] reviewing courts shall set aside agency actions or conclusions that are "arbitrary, capricious, an abuse of discretion, or otherwise not in accordance with law."[50] By the 1960s, however, this "arbitrary and capricious" standard increasingly seemed inadequate to address concerns about "agency capture"—which posits that agency decision making is often hijacked by regulated industries and interest groups.[51] In response, federal appellate courts—especially the D.C. Circuit—developed the "hard look doctrine."[52] Courts began demanding that agencies show that they had taken a "hard look" at the substantive issues before deciding on a particular policy.[53]

---

45. *See* BREYER ET AL., *supra* note 2, at 247 (noting that *Chevron* had been cited in federal courts nearly 8,000 times by the end of 2005 and "may well qualify as the most influential case in the history of American public law" by "sheer number of citations").

46. *Id.* at 378.

47. *Cf.* M. Elizabeth Magill, *Step Two of* Chevron v. Natural Resources Defense Council, *in* A GUIDE TO JUDICIAL AND POLITICAL REVIEW OF FEDERAL AGENCIES 85, 98 (John F. Duffy & Michael Herz eds., 2005) ("The two steps of *Chevron* apply to questions of statutory interpretation, while *State Farm* applies to exercises of policy judgment.").

48. 5 U.S.C. § 706 (2006).

49. Pub. L. No. 79-404, 60 Stat. 238 (1946) (codified at 5 U.S.C. §§ 551–559, 701–706).

50. 5 U.S.C. § 706(2)(A).

51. BREYER ET AL., *supra* note 2, at 348.

52. Thomas J. Miles & Cass R. Sunstein, *The Real World of Arbitrariness Review*, 75 U. CHI. L. REV. 761, 761 (2008); *see also* Reuel E. Schiller, *Enlarging the Administrative Polity: Administrative Law and the Changing Definition of Pluralism, 1945–1970*, 53 VAND. L. REV. 1389, 1421 (2000) ("The decline of the group pluralist model of policy-making brought to administrative law and theory a profound fear that agencies had stopped serving the public interest and were simply representing the interests of the industries they regulated.").

53. Miles & Sunstein, *supra* note 52, at 761 (citing Ethyl Corp. v. EPA, 541 F.2d 1, 35 (D.C. Cir. 1976) (en banc)).

As set out in the landmark *State Farm* case,[54] the doctrine calls for courts to take a hard look at whether the agency has sufficiently considered all the available evidence and alternatives before making (and explaining) a well-reasoned policy choice.[55] In contrast to the *Chevron* doctrine, which involves an outcome-focused inquiry, the hard look doctrine endorsed by the Court in *State Farm* is a procedural test: it evaluates the agency's decision-making process.[56] Under *State Farm* and its progeny, the "failure of an agency to consider obvious [significant] alternatives has led uniformly to reversal."[57] Often, deciding whether an agency has sufficiently considered the alternatives as a procedural matter necessarily requires the court to examine the substantive issues in order to evaluate to what degree the agency should have investigated its options.[58] Thus, the hard look doctrine can be considered a fairly intrusive review of agency decision making.

C. *Summary of the American Anomaly*

The coexistence of two different approaches to judicial deference to agency action—the broad deference toward agency interpretations of law epitomized by *Chevron*, contrasted with the lower level of deference afforded to agency policy choices under *State Farm*—creates an anomaly in American administrative law. Following these doctrines, judges presumably give more deference to agencies' decisions of law than they do to agencies' policy choices.

The Judiciary is the least democratic branch of American government: federal judges are neither elected nor removed by the people.[59] Yet under the hard look doctrine, unelected judges review the policy decisions made by administrative agencies—often under the authority of the President—that implement statutes passed by Congress. A reviewing court may void an agency decision that did not amount to a well-reasoned (and explained)

---

54. *See* Motor Vehicle Mfrs. Ass'n v. State Farm Mut. Auto. Ins. Co., 463 U.S. 29, 46, 48–49 (1983) (agreeing with the D.C. Circuit that the agency acted arbitrarily and capriciously and requiring that the agency "cogently explain why it . . . exercised its discretion in a given manner"); BREYER ET AL., *supra* note 2, at 383 ("*State Farm* is regarded as having endorsed a relatively intensive version of 'hard look' review, of the type pioneered by the D.C. Circuit.").

55. *See* Elena Kagan, *Presidential Administration*, 114 HARV. L. REV. 2245, 2380 (2001) ("The current version of [the hard look] doctrine subjects all agency decisionmaking, irrespective of provenance or pedigree, to a wide-ranging judicial review for errors of process: the courts take a hard look at whether the agencies themselves have taken a hard look at the range of evidence, arguments, and alternatives relevant to an issue, and have made and explained a reasoned policy choice based on these considerations.").

56. Magill, *supra* note 47, at 98.

57. City of Brookings Mun. Tel. Co. v. FCC, 822 F.2d 1153, 1169 & n.46 (D.C. Cir. 1987).

58. *See, e.g.*, *State Farm*, 463 U.S. at 53–54 & nn.16–19 (examining the empirical evidence in the record and venturing into sophisticated statistical analysis).

59. Of course, the Judiciary is not entirely free from political pressures. *See* John A. Ferejohn & Barry W. Weingast, *A Positive Theory of Statutory Interpretation*, 12 INT'L REV. L. & ECON. 263, 263 (1992) ("If a court's decision fails to reflect external political reality, it cannot stand for long.").

choice among carefully considered policy alternatives.[60] Thus, unelected judges review policy decisions that can be considered to be the product of the two elected branches.

Meanwhile, under *Chevron*, judges give great deference to agencies' interpretations of law. This deference might seem incongruous with the separation of powers enshrined in the Constitution and the intent of the framers that "[t]he interpretation of the laws is the proper and peculiar province of the courts."[61] Therefore, the American anomaly is that judges defer to agencies on questions of law—where judges have expertise and constitutional legitimacy—but take a hard look at policy decisions by agencies that have superior expertise and democratic legitimacy.[62] This Note attempts to help explain this anomaly by means of a comparative institutional analysis, so it will next turn to a discussion of the analogous deference regimes in Australia and the United Kingdom.

II. Judicial Deference in Australia and the United Kingdom

The American anomaly is brought into even sharper relief by a comparative look at the Australian and British systems, which invert the American model of deference. Courts in both Australia and the United Kingdom give *no deference* to administrative interpretations of statutes, while granting substantially more deference to administrative policy decisions.

A. Australia

The Australian deference regime is exactly the reverse of the American system: Australian courts retain the exclusive authority to decide questions of law but generally give near-complete deference to administrators' policy decisions.[63] Justice Sackville has called these the "twin pillars" of judicial review of Australian administrative action: that legal interpretation is the province of the courts and not the executive, and that courts may not infringe upon the power of the executive to make policy or to make administrative decisions on the "merits."[64] The Australian High Court has flatly rejected *Chevron*, holding that "there is very limited scope for the notion of 'judicial

60. Kagan, *supra* note 55, at 2380.

61. THE FEDERALIST NO. 78, at 467 (Alexander Hamilton) (Clinton Rossiter ed., 1961).

62. BREYER ET AL., *supra* note 2, at 403.

63. Ronald Sackville, *The Limits of Judicial Review of Executive Action—Some Comparisons Between Australia and the United States*, 28 FED. L. REV. 315, 329 (2000). Justice Sackville also observed that "[t]he situation in the United States is itself paradoxical; but the nature of the paradox is the mirror image of that in Australia." *Id.* at 323. Unlike this Note, which argues that the American anomaly is tied to the U.S. presidential system, Justice Sackville highlights the American example in order to question whether Australia might learn from the American approach. *Id.* at 323, 328–30.

64. *Id.* The twin pillars derive from the strict separation of the judicial power from the other powers of government. *Id.* at 319.

deference'" to administrative interpretations of law.[65] The High Court reasoned that "an essential characteristic of the judicature is that it declares and enforces the law which determines the limits of the power conferred by statute upon administrative decision-makers."[66] For a court to fail to make its own, independent legal determination would be "to abdicate judicial responsibility."[67] Therefore, Australian courts grant no deference to legal interpretations by administrators.

With respect to administrators' policy choices, review by Australian courts does not extend to the "merits" of the administrators' decisions.[68] Australian courts observe a strict distinction between the legality and the merits of administrative decisions.[69] Administrative actions are only reviewable for their legality and not for their merits.[70] Thus, the Australian deference regime presents a mirror image of the American anomaly. With its sharp divide between judicial review of administrators' legal interpretations and policy choices, the Australian system comports with our intuitions about judicial deference to administrative action: judges defer to administrative policy choices while retaining the exclusive authority to interpret the law.

B. United Kingdom

As in Australia, courts in the United Kingdom afford no deference to administrative interpretations of law.[71] In the *Anisminic* case,[72] the House of Lords ruled that all "errors of law" by administrative tribunals or authorities are reviewable de novo.[73]

With respect to judicial review of administrative policy decisions in the United Kingdom, the degree of deference differs based on the subject matter involved.[74] For example, English decisions after the 1998 passage of the Human Rights Act[75] have been trending in favor of higher scrutiny in cases implicating fundamental human rights.[76] Thus, depending on the subject

65. City of Enfield v. Dev. Assessment Comm'n (2000) 199 C.L.R. 135, 158.

66. *Id.* at 153.

67. *Id.* at 158.

68. Sackville, *supra* note 63, at 315.

69. *Id.* at 321.

70. *Id.* at 321–22.

71. *See* Michael C. Tolley, *Judicial Review of Agency Interpretation of Statutes: Deference Doctrines in Comparative Perspective*, 31 POL'Y STUD. J. 421, 428 (2003) ("In the 1960s, courts in Britain took assertive steps that made them the conclusive arbiters on all questions of law.").

72. Anisminic Ltd. v. Foreign Compensation Comm., [1969] 2 A.C. 147 (H.L. 1968) (appeal taken from Eng.).

73. *Id.* at 182. "Error of law" is the traditional basis for judicial review of administrative action in England, and error of law claims generally involve statutory interpretation. Tolley, *supra* note 71, at 428.

74. R. v. Sec'y of State for the Home Dep't, *ex parte* Daly [2001] UKHL 26, [2001] 2 A.C. 532, 549 (H.L. 2001) (appeal taken from Eng.).

75. 1998 c. 42 (U.K.).

76. Tolley, *supra* note 71, at 430.

matter, courts may review administrative policy decisions under standards ranging from reasonableness under *Associated Provincial Picture Houses Ltd. v. Wednesbury Corp.*,[77] to proportionality, to heightened scrutiny.[78] The default standard, however, is *Wednesbury* reasonableness—a low threshold that upholds administrative action unless the decision is "so unreasonable that no reasonable decision maker could ever have come to it."[79]

Therefore, British judges generally extend significant deference to most administrative action. This deference comports with the fundamental English doctrine of the legislative supremacy of Parliament.[80] Under this doctrine, Parliament enjoys unlimited and exclusive legislative authority.[81] Accordingly, the judiciary reviews executive action chiefly to ensure that government is acting within the powers granted by Parliament through legislation—the ultra vires principle.[82] Thus, the British deference regime is the reverse of the American anomaly: judges in the United Kingdom review errors of law de novo and give significant deference to administrative policy actions.

III. A Comparative Institutional and Political Analysis: Australia, the United Kingdom, and the United States

The rise of the administrative state has generated more significant institutional tensions in the United States than in the United Kingdom and Australia. In the British and Australian parliamentary systems, control over the executive and legislative branches is consolidated, with the Prime Minister controlling both government and Parliament. Furthermore, the executive is more unified in Britain and Australia than in the United States, where there are significant constitutional and statutory divisions between the President and administrative agencies.

In Britain and Australia, then, the executive controls both the legislature and administrators.[83] Accordingly, this parliamentary structure minimizes interbranch power struggles over control of the administrative state. In contrast, American administrative agencies are often the subject of struggles for influence between the President and Congress.[84] Especially during periods of divided government, the agencies are at the forefront of President–

77. [1948] 1 K.B. 223 (C.A. 1947).

78. Tolley, *supra* note 71, at 430.

79. ANDREW LE SUEUE ET AL., PRINCIPLES OF PUBLIC LAW 302 (2d ed. 1999).

80. *See* Eric Barendt, *Fundamental Principles*, *in* ENGLISH PUBLIC LAW 30, 30 (2004) (maintaining that English constitutional law is dominated by three principles—the legislative supremacy of Parliament, the rule of law, and the separation of powers—and that the first principle has traditionally been considered the most important).

81. *Id.*

82. *Id.* at 41.

83. *See supra* note 7 and accompanying text.

84. *See* CHARLES H. KOCH ET AL., ADMINISTRATIVE LAW: CASES AND MATERIALS 90 (5th ed. 2006) (noting interbranch conflicts on the appointment and removal of officials).

Congress conflicts.[85] The Judiciary, as the natural constitutional arbiter of these disputes, thus finds itself in a position to mediate between the other two branches.

A. Separation of Powers

The separation of legislative and executive powers is more pronounced in the American context than it is in the United Kingdom and Australia. One reason is that the separation of powers is more important to the American approach to protecting individual freedom. The British–Australian approach—the "rule of law" approach—seeks to protect citizens through robust legal protections for individual liberties.[86] The American approach—checks and balances—aims to protect individual freedoms indirectly, by limiting the power of government generally.[87] The three branches of government are played against each other in a struggle for power; attempts at tyranny by any one branch can be thwarted by the other two.[88] Thus, the power of government collectively is also restrained.[89]

The American constitutional system has three defining characteristics: (1) separation of powers, (2) checks and balances, and (3) federalism.[90] The importance of the concept of separation of powers is made apparent by the organization of the U.S. Constitution itself, which assigns the "legislative

85. *See* Kagan, *supra* note 55, at 2346–47 ("Presidential administration may alter congressional oversight of agency action, making this activity less effective yet more vehement, especially in periods of divided government.").

86. *See, e.g.*, John McMillan, *An Overview of the Australian Legal System*, *in* ASPECTS OF ADMINISTRATIVE LAW IN AUSTRALIA AND INDONESIA 36, 44 (Robin Creyke et al. eds., 1996) ("Under the banner of [the rule of law] principle the courts have always declared that they have a special duty to protect the interests of the individual.").

87. *See, e.g.*, Peter L. Strauss, *The Place of Agencies in Government: Separation of Powers and the Fourth Branch*, 84 COLUM. L. REV. 573, 578 (1984) ("Like separation of powers, [checks and balances] seeks to protect the citizens from the emergence of tyrannical government by establishing multiple heads of authority in government, which are then pitted one against another in a continuous struggle; the intent of that struggle is to deny to any one (or two) of them the capacity ever to consolidate all governmental authority in itself, while permitting the whole effectively to carry forward the work of government."); *see also, e.g.*, PATRICK M. GARRY, AN ENTRENCHED LEGACY: HOW THE NEW DEAL CONSTITUTIONAL REVOLUTION CONTINUES TO SHAPE THE ROLE OF THE SUPREME COURT 35 (2008) (recounting that the framers believed that the best way to safeguard individual liberties against government tyranny was to design weak institutions and diffuse government power).

88. Strauss, *supra* note 87, at 578; *see also* THE FEDERALIST NO. 51 (James Madison), *supra* note 61, at 321–22 ("But the great security against a gradual concentration of the several powers in the same department, consists in giving to those who administer each department the necessary constitutional means and personal motives to resist encroachments of the others.").

89. *See supra* note 87 and accompanying text.

90. *See* BREYER ET AL., *supra* note 2, at 31 (describing the American constitutional system as one of separation of powers and of checks and balances); *see also* GARRY, *supra* note 87, at 27 ("The most important governing structures created by the Constitution are federalism and separation of powers.").

Powers" to Congress in Article I, the "executive Power" to the President in Article II, and the "judicial Power" to the Judiciary in Article III.[91]

In contrast to the American strict separation-of-powers regime, the separation of powers is far less stringent in England[92] and Australia.[93] In these parliamentary systems, the formal separation of legislative and executive powers breaks down as a practical matter. Government is formed by the party who wins a majority in Parliament.[94] This institutional structure, together with strong party discipline, ensures that the governing party can pass and implement legislation as it sees fit.[95] As a result, Australia and the United Kingdom generally do not experience the significant Executive–Legislative conflicts that occur in the United States.[96]

B. The Executive and the Administrative State

A crucial difference between the American administrative state on the one hand and the Australian and British administrative states on the other is that America has a bifurcated Executive whereas Australia and the United Kingdom each have a unified executive.

The Constitution grants the President the right to appoint the major executive officials,[97] and the President currently appoints over 8,000 other executive officials.[98] This power to appoint gives the President considerable power over administrative agencies.[99] However, the President does not have

---

91. U.S. CONST. arts. I–III; *see also, e.g.*, GARRY, *supra* note 87, at 29 (observing that "a general principle of separation of power can be constructed from various provisions in the Constitution," which "outlines the powers of each branch and then creates a complex system of checks and balances on the respective branches").

92. *See* Terry M. Moe & Michael Caldwell, *The Institutional Foundations of Democratic Government: A Comparison of Presidential and Parliamentary Systems*, 150 J. INSTITUTIONAL & THEORETICAL ECON. 171, 177 (1994), *reprinted in* ECONOMICS OF ADMINISTRATIVE LAW 567, 573 (Susan Rose-Ackerman ed., 2007) (describing the Westminster parliamentary system, in which the same party controls the executive and legislature).

93. *See* McMillan, *supra* note 86, at 43 ("Separation of powers could never be practiced in any pure form in Australia. Responsible government is itself a fundamental breach, since the executive's Ministers also sit in the Parliament.").

94. Moe & Caldwell, *supra* note 92, at 573.

95. *Id.*

96. *See id.* at 574 ("[In parliamentary systems, t]he executive is the leadership of the majority party, and thus the leadership of the legislature itself. What the executive wants, it gets.").

97. *See* U.S. CONST. art. II, § 2, cl. 2 (granting the President the power to nominate ambassadors, ministers, judges, and "all other Officers of the United States").

98. David Fontana, *The Second American Revolution in the Separation of Powers*, 87 TEXAS L. REV. 1401, 1415 (2009) (citing Lois Romano, *In Any Guise, Podesta a Smooth Master of the Transition Game*, WASH. POST, Nov. 25, 2008, at C1).

99. *See, e.g.*, JEFFREY PFEFFER, MANAGING WITH POWER: POLITICS AND INFLUENCE IN ORGANIZATIONS 130 (1994) (identifying the "power to hire, fire, and reward" as central to the power associated with a formal leadership position within an organizational hierarchy). Nevertheless, the agencies still maintain a significant degree of independence from the President, and their power is limited by the organic and regulatory statutes passed by Congress. Strauss, *supra* note 87, at 586.

complete power over the administrative agencies. Instead, there is a separation between the President and the rest of the Executive. The American executive power is divided between the President and the agencies, in contrast to the unitary executives in Australia and the United Kingdom. Therefore, not only is American governmental power divided between the Executive and the Legislative Branches, but the executive power is further divided within the Executive between the President and the agencies. This is yet another reason why the administrative state generates greater institutional tensions in the United States than in Australia and the United Kingdom.

C. Legislation

In Australia and the United Kingdom, legislation is drafted by the executive and then approved by Parliament.[100] Legislation is therefore an instrument of the executive. From a balance-of-powers perspective, legislation in the United States can be conceived of as a tool of the Legislature to control the Executive. Consistent with the American system of checks and balances, the Executive is then constitutionally tasked with implementing legislation.[101]

However, congressional delegations to executive agencies can create significant principal–agent problems. Congress passes statutes, and executive agencies are tasked with interpreting and executing those regulations. Yet the agencies are primarily accountable not to Congress but to the President.[102] Thus, there is strong potential for misalignment between the interests and objectives of Congress and those of the agency.[103] One logical result of this system is that Congress tends to delegate more authority to agencies when the Executive and Legislative Branches are politically aligned.[104]

In Australia and the United Kingdom, the executive and the lower legislative body are *always* politically aligned.[105] Therefore, Australian and

---

100. ANDY WILLIAMS, UK GOVERNMENT AND POLITICS 8 (2d ed. 1998).

101. *See* U.S. CONST. art. II, § 3 ("[The President] shall take Care that the Laws be faithfully executed . . . .").

102. *See* Lisa Schultz Bressman, *Beyond Accountability: Arbitrariness and Legitimacy in the Administrative State*, 78 N.Y.U. L. REV. 461, 485 (2003) (explaining how President Reagan and his successors have exerted both managerial and directorial control over administrative agencies).

103. *See id.* at 507–08 (illustrating "the ability of *any* president, Republican or Democrat, to use passive-aggressive strategies of controlling regulatory policy that are less likely to achieve their purported public purpose than to please private interests").

104. *See* Lisa Schultz Bressman, Chevron*'s Mistake*, 58 DUKE L.J. 549, 569 (2009) ("Congress is likely to delegate more authority to executive branch agencies when the president is of the same political party, that is, during periods of unified government. In such periods, Congress may reasonably assume that presidential preferences are less likely to diverge from legislative preferences.").

105. It is important to note that in Australia, while the lower house (the House of Representatives) and the Prime Minister are always aligned, the Senate can be controlled by a different political party. *See* Ben Sharples, *Rudd Seeks Election Trigger Over Health, Herald Says*, BUS. WK., Feb. 18, 2010, *available at* http://www.businessweek.com/news/2010-02-18/rudd-seeks-

British legislators do not have to doubt—at least to the same degree as their American counterparts—whether administrators will faithfully execute their statutes.[106] The U.S. model, on the other hand, provides ample incentive and opportunity for administrators to undermine congressional intent, particularly when different parties control the presidency and Congress.

IV. What the Australian and British Models Suggest About the American Anomaly

A. Why U.S. Courts Give More Deference to Administrators' Legal Interpretations

A significant reason why American courts defer to agencies' legal interpretations is that an agency's interpretation of a statute is often tied to its policy choices,[107] and—as expressed by the Supreme Court in *Chevron*—unelected judges should not substitute their own policy preferences for those of administrators accountable to the democratically elected president.[108] But why should this be more of a concern for American judges than for their Australian and British counterparts?

One response is that the American Judiciary is more politicized than the judiciaries of Australia and the United Kingdom.[109] This politicization of the

---

election-trigger-over-health-herald-says-update1-.html (stating that while the Labor government has a majority in the House of Representatives, "neither it nor the opposition Liberal–National party coalition control the Senate, with both sides needing the support of other parties to pass or block legislation"). Similarly, the English Prime Minister may not be able to exert as much control over the House of Lords as over the House of Commons. *Cf.* LIJPHART, *supra* note 7, at 18–19 (explaining that while the House of Commons is popularly elected, the House of Lords consists of the hereditary nobility and government-appointed life peers; while "almost all legislative power belongs to the House of Commons," the House of Lords can delay legislation). Furthermore, a notable difference between the two countries is that the Australian Senate has a legislative veto, while the British House of Lords does not. *See* George Tsebelis, *Veto Players and Institutional Analysis*, 13 GOVERNANCE 441, 459 (2000) (arguing that although the House of Lords lacks veto power, it is "able to abort legislation just by delaying passage until the election").

106. There is, of course, a temporal problem with the implementation of legislation under all three systems. With the passage of time and subsequent elections, control of the legislative and executive branches can change hands in all three countries. Yet even this problem is lessened in the parliamentary systems: since political accountability for administration action (or inaction) is so much clearer than in the United States, Australian and British judges can depend on the political process to resolve discrepancies with more confidence than can their American counterparts.

107. Michael Herz, *Deference Running Riot: Separating Interpretation and Lawmaking Under* Chevron, 6 ADMIN. L.J. AM. U. 187, 190 (1992) ("Separating interpretation from lawmaking is both easy and impossible. . . . Yet under almost any theory of statutory interpretation, the two overlap.").

108. Chevron U.S.A. Inc. v. Natural Res. Def. Council, Inc., 467 U.S. 837, 865 (1984).

109. *See* ERIC BARENDT, AN INTRODUCTION TO CONSTITUTIONAL LAW 134 (1998) (observing that with respect to the appointment of British judges, "[t]here is admittedly no evidence that party political considerations play any part in this process; Lord Mackay, the last Conservative Lord Chancellor, appointed some judges who were known to have liberal, left-wing sympathies"); Mita Bhattacharya & Russell Smyth, *The Determinants of Judicial Prestige and Influence: Some Empirical Evidence from the High Court of Australia*, 30 J. LEGAL STUD. 223, 230–31 (2001)

Judiciary may result—at least in part—from the American system of checks and balances, in which the Judiciary plays an important role in checking the power of the other two branches. With the Judiciary situated as a coordinate political branch to the Executive and Legislative Branches, American judges occupy a more overtly political role than do Australian and British judges. As such, the Supreme Court had cause to fear that judges' policy preferences could affect their review of agencies' statutory interpretation.[110]

It is also perfectly reasonable for generalist judges to defer to the legal interpretations of administrators who may have more expertise within their own specialized statutory regimes. For these reasons, judicial deference to agencies' legal interpretations makes sense in the American context. However, these reasons—the politicization of the Judiciary coupled with agencies' superior accountability and expertise—seem to support with even greater force the argument that judges should give significant deference to agencies' policy choices.

B. Why U.S. Courts Give Less Deference to Administrator's Policy Choices

Why then do American courts give comparatively less deference to policy choices? One explanation is that the U.S. Judiciary must fulfill a role in correcting political defects—defects that are largely absent in the Australian and British systems. The modern American "administrative state" has often been called the fourth branch of the U.S. government.[111] There are now over 100 federal regulatory agencies and subagencies, which promulgate approximately 4,500 new rules every year.[112] The existence of this expansive administrative state has upset the balance of power between the American political branches, leading to President–Congress conflicts over control of the agencies. For example, President Reagan centralized review of agency rules in the Office of Information and Regulatory Affairs, increasing presidential influence over agency rulemaking.[113] Later, through the passage

---

(contrasting the rarity of politically driven judicial appointments in Australia with the overtly partisan judicial-selection practices in the United States).

110. *Cf. Chevron*, 467 U.S. at 865 ("Courts, must in some cases, reconcile competing political interests, but not on the basis of the judges' personal policy preferences.").

111. Strauss, *supra* note 87, at 578.

112. John D. Graham, Adm'r, Office of Info. & Regulatory Affairs, Office of Mgmt. & Budget, Speech to Weidenbaum Center Forum: Executive Regulatory Review: Surveying the Record, Making It Work (Dec. 17, 2001) (transcript available at http://georgewbush-whitehouse.archives.gov/omb/inforeg/graham_speech121701.html).

113. *Oversight of the Congressional Review Act Before the H. Subcomm. on Commercial and Administrative Law*, 110th Cong. 3 (2007) (statement of Morton Rosenberg, Specialist in American Law, Congressional Research Service) [hereinafter *Hearings*] (describing the "effectiveness of President Reagan's executive orders centralizing review of agency rulemaking . . . in the [OMB's] Office of Information and Regulatory Affairs (OIRA) in the face of aggressive challenges of congressional committees").

of the Congressional Review Act,[114] many legislators hoped to shift the balance of power over administrative rulemaking back towards Congress.[115]

The growth of the administrative state in the United Kingdom and Australia has not created a corresponding imbalance in those countries, thanks in large part to two institutional factors: (1) that each country has a unitary executive (as opposed to the bifurcated American Executive, split between the presidency and the administrative agencies) and (2) control of the executive and legislative branches is largely consolidated in the hands of the governing party. Therefore, political accountability for administrative decisions in the United Kingdom and Australia is clear: responsibility for such decisions lies with the governing party.

In contrast, the American Congress has experienced great difficulty in ensuring that the administrative agencies faithfully execute the legislation they are charged with implementing.[116] The system of checks and balances established by the U.S. Constitution, which did not foresee the massive administrative state of the twentieth and twenty-first centuries, does not adequately address principal–agent concerns with respect to administrative action. Therefore, Congress relies on the courts to help police administrative action to ensure that the agencies serve as the "faithful agents" of Congress. U.S. courts help restrict the agencies' ability to circumvent legislative will. For these reasons, there is a greater need for judicial review of administrative policy choices in the United States than in Australia and the United Kingdom.

V. What This Comparative Analysis Predicts About Deference Regimes in American Administrative Law

The comparison with Australia and the United Kingdom demonstrates that the American anomaly is a result of tensions inherent in the U.S. political structure.[117] While the American anomaly is contrary to both our logical intuitions as well to the Australian and British models, it represents the natural outgrowth of competing judicial values: the desire to defer to agencies who may have greater expertise and political accountability than courts on the one hand, and the need to serve as a political check on abuses of discretion by agencies on the other.[118]

---

114. Pub. L. No. 104-121, sec. 251, 110 Stat. 847, 868 (1996) (codified as amended at 5 U.S.C. §§ 801–808 (2006)).

115. *See Hearings*, *supra* note 113, at 3 ("The expectation of many was that Congress, through the CRA, would again become a major player in influencing agency decisionmaking.").

116. *See, e.g.*, Bowsher v. Synar, 478 U.S. 714, 734 (1986) (holding that the Constitution does not allow congressional intrusion into executive functions and that retention of removal power constitutes such an intrusion); INS v. Chadha, 462 U.S. 919, 959 (1983) (striking down the use of the legislative veto).

117. *See supra* Parts III and IV.

118. *See* STEPHEN BREYER, ACTIVE LIBERTY 103, 108 (2005) (arguing that "[t]o reconcile democratically chosen ends with administrative expertise requires striking a balance—some

The American deference regime does not resolve the fundamental institutional tensions that led to its genesis: hence the proliferation of scholarly works continuing to debate fervently the proper degree of deference.[119] Therefore, we should expect to continue to see judicial flexibility and changes with respect to American deference doctrines. We might also expect a multiplicity of doctrines that seek to correct imbalances in the system of checks and balances.

Justice (then-Judge) Breyer advocated against a broad, simplistic reading of *Chevron* shortly after its decision.[120] Breyer presciently predicted that courts would instead follow "more varied approaches, sometimes deferring to agency interpretations, sometimes not, depending upon the statute, the question, the context, and what 'makes sense' in the particular litigation, in light of the basic statute and its purposes."[121] Indeed, this is exactly what has resulted. Eskridge and Baer have recently shown that even the Supreme Court does not religiously apply the *Chevron* doctrine, but rather it employs a continuum of deference doctrines depending on the particularities of each case.[122] In light of the institutional tensions in American administrative law, courts will likely continue to need the flexibility offered by such a continuum of doctrines in order to calibrate the borders of the President–Congress divide.

VI. Conclusion

It has been noted that in the United States, "the substance of regulations and the procedure by which they are made present issues which generate enormous controversy," whereas in the United Kingdom, "nearly everyone seems satisfied with . . . procedural and substantive aspects of delegated legislation."[123] This dichotomy can be traced back to the fundamental institutional differences between presidential and parliamentary systems. In the Australian and British parliamentary systems, political accountability for administrative action is clear, and principal–agent problems between the legislature and executive are minimized. In contrast, the American system involves significant principal–agent problems between Congress and executive agencies, as well as between the President and the agencies. As a result, political accountability is also unclear. Therefore, deference doctrines in American administrative law will continue to be flexible—and at times seemingly inconsistent—as they struggle to address the institutional tensions

---

delegation, but not too much" and that judges therefore should treat *Chevron* "not as an absolute rule, but as a rule of thumb").

119. *See supra* Part I.

120. *See* Breyer, *supra* note 1, at 373 ("To read *Chevron* as laying down a blanket rule, applicable to all agency interpretations of law . . . would be seriously overbroad, counterproductive and sometimes senseless.").

121. *Id.* at 381.

122. *See supra* note 18.

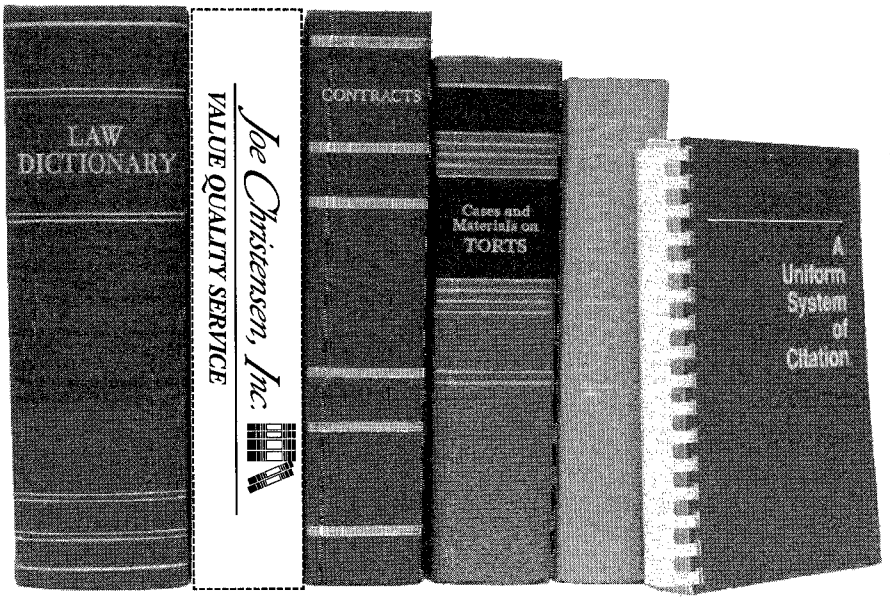
123. Asimow, *supra* note 11, at 253.

generated by the application of a three-branch system of checks and balances to a fourth branch—the administrative state.

—*Robert C. Dolehide*

* * *

* * *



We Complete the Picture.

In 1932, Joe Christensen founded a company based on Value, Quality and Service. With 77 years of printing experience, Joe Christensen, Inc. remains the most experienced Law Review printer in the country.

Our printing services bridge the gap between your editorial skills and the production of a high-quality publication. We ease the demands of your assignment by offering you the basis of our business—customer service.

Joe Christensen, Inc.

1540 Adams Street
Lincoln, Nebraska 68521-1819
Phone: 1-800-228-5030
FAX: 402-476-3094
email: sales@christensen.com

Value

Quality

Service

Your Service Specialists


TEXAS
TLR
LAW REVIEW