Texas Law Review

Articles

Essay

Notes

December 2017 Vol. 96 No. 2 Pages 205 to 424

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: (800) 828-7571.

Single issues in the current volume may be purchased from the *Texas Law Review* Publications Office for $15.00 per copy shipping included. 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 one side only, with footnotes rather than endnotes. Citations should conform with *The Greenbook: Texas Rules of Form* (13th ed. 2015) and *The Bluebook: A Uniform System of Citation* (20th ed. 2015). Except when content suggests otherwise, the *Texas Law Review* follows the guidelines set forth in the *Texas Law Review Manual on Usage & Style* (14th ed. 2017), *The Chicago Manual of Style* (17th ed. 2017), and Bryan A. Garner, *Black's Law Dictionary* (10th ed. 2014).

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

THE UNIVERSITY OF TEXAS SCHOOL OF LAW

ADMINISTRATIVE OFFICERS

WARD FARNSWORTH, B.A., J.D.; *Dean, John Jeffers Research Chair in Law.*
JOHN B. BECKWORTH, B.A., J.D.; *Associate Dean for Administration and Strategic Planning, Lecturer.*
ROBERT M. CHESNEY, B.S., J.D.; *Associate Dean for Academic Affairs, Charles I. Francis Professor in Law.*
WILLIAM E. FORBATH, A.B., B.A., Ph.D., J.D.; *Associate Dean for Research, Lloyd M. Bentsen Chair in Law.*
EDEN E. HARRINGTON, B.A., J.D.; *Associate Dean for Experiential Education, Director of William Wayne Justice Center for Public Interest Law, Clinical Professor.*
ELIZABETH T. BANGS, A.B., J.D.; *Assistant Dean for Student Affairs.*
LAUREN FIELDER, B.A., J.D., LL.M.; *Assistant Dean for Graduate and International Programs, Senior Lecturer.*
MICHAEL G. HARVEY, B.A., B.S.; *Assistant Dean for Technology.*
REBECCA E. MELTON, B.A., J.D.; *Assistant Dean for Alumni Relations and Development.*
DAVID A. MONTOYA, B.A., J.D.; *Assistant Dean for Career Services.*
GREGORY J. SMITH, B.A., J.D.; *Assistant Dean for Continuing Legal Education.*

FACULTY EMERITI

RICHARD V. BARNDT, B.S.L., LL.B.; *Professor Emeritus.*
FRANK B. CROSS, B.A., J.D.; *Professor Emeritus.*
JULIUS G. GETMAN, B.A., LL.B., LL.M.; *Earl E. Sheffield Regents Chair 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. LEON LEBOWITZ, A.B., J.D., LL.M.; *Joseph C. Hutcheson Professor Emeritus.*
BASIL S. MARKESINIS, LL.B., Ph.D., D.C.L., LL.D.; *Jamail Regents Chair Emeritus in Law.*
JOHN T. RATLIFF, JR., B.A., LL.B.; *Ben Gardner Sewell Professor Emeritus in Civil Trial Advocacy.*
JOHN J. SAMPSON, B.B.A., LL.B.; *William Benjamin Wynne Professor Emeritus in Law.*
JAMES M. TREECE, B.S., J.D., M.A.; *Charles I. Francis Professor Emeritus in Law.*
ZIPPORAH B. WISEMAN, B.A., M.A., LL.B.; *Thos. H. Law Centennial Professor Emerita in Law.*

PROFESSORS

JEFFREY B. ABRAMSON, B.A., J.D., Ph.D.; *Professor of Government and Law.*
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.*
MARILYN ARMOUR, B.A., M.S.W., Ph.D.; *Associate Professor of Social Work.*
MARK L. ASCHER, B.A., M.A., J.D., LL.M.; *Hayden W. Head Regents Chair for Faculty Excellence.*
RONEN AVRAHAM, M.B.A., LL.B., LL.M., S.J.D.; *Senior Lecturer.*
LYNN A. BAKER, B.A., B.A., J.D.; *Frederick M. Baron Chair in Law, Co-Director of Center on Lawyers, Civil Justice, and the Media.*
BARBARA A. BINTLIFF, M.A., J.D.; *Joseph C. Hutcheson Professor in Law, Director of Tarlton Law Library and the Jamail Center for Legal Research.*
LYNN E. BLAIS, A.B., J.D.; *Leroy G. Denman Jr. Regents Professor in Real Property Law.*
ROBERT G. BONE, B.A., J.D.; *G. Rollie White Teaching Excellence Chair in Law.*
OREN BRACHA, LL.B., S.J.D.; *Howrey LLP and Arnold, White & Durkee Centennial Professor.*
DANIEL M. BRINKS, A.B., J.D., Ph.D.; *Associate Professor, Co-Director of Bernard and Audre Rapoport Center for Human Rights and Justice.*
J. BUDZISZEWSKI, B.A., M.A., Ph.D.; *Professor of Government and Philosophy.*
NORMA V. CANTU, B.A., J.D.; *Professor of Education and Law.*
MICHAEL J. CHURGIN, A.B., J.D.; *Raybourne Thompson Centennial Professor.*
JANE M. COHEN, B.A., J.D.; *Edward Clark Centennial Professor.*
WILLIAM H. CUNNINGHAM, B.A., M.B.A., Ph.D.; *Professor of Marketing Administration.*
JENS C. DAMMANN, J.D., LL.M., Dr. Jur., J.S.D.; *William Stamps Farish Professor in Law.*
JOHN DEIGH, B.A., M.A., Ph.D.; *Professor of Philosophy and Law.*
MECHELE DICKERSON, B.A., J.D.; *Arthur L. Moller Chair in Bankruptcy Law and Practice, University Distinguished Teaching Professor.*
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.*
DAVID J. EATON, B.A., M.Sc., M.A., Ph.D.; *Professor of Public Affairs.*
ZACHARY S. ELKINS, B.A., M.A., Ph.D.; *Associate Professor of Government.*
KAREN L. ENGLE, B.A., J.D.; *Minerva House Drysdale Regents Chair in Law, Founder and Co-Director of Bernard and Audre Rapoport Center for Human Rights and Justice.*
KENNETH FLAMM, A.B., Ph.D.; *Professor of Public Affairs.*
JOSEPH R. FISHKIN, B.A., M.Phil., D.Phil., J.D.; *Professor of Law.*
CARY C. FRANKLIN, B.A., M.S.T., D.Phil., J.D.; *Professor of Law.*
MIRA GANOR, B.A., M.B.A., LL.B., LL.M., J.S.D.; *Professor of Law.*
CHARLES E. GHOLZ, B.S., B.S., Ph.D.; *Associate Professor of Public Affairs.*
JOHN M. GOLDEN, A.B., J.D., Ph.D.; *Loomer Family Professor in Law.*
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. W. Walker Centennial Chair in Law.*
BENJAMIN G. GREGG, B.A., M.A., Ph.D.; *Associate Professor of Government.*
PATRICIA I. HANSEN, A.B., M.P.A., J.D.; *J. Waddy Bullion Professor.*
HENRY T. C. HU, B.S., M.A., J.D.; *Allan Shivers Chair in the Law of Banking and Finance.*

BOBBY R. INMAN, B.A.; *Professor of Public Affairs.*
GARY J. JACOBSOHN, B.A., M.A., Ph.D.; *Professor of Government and Law.*
DEREK P. JINKS, B.A., M.A., J.D.; *The Marrs McLean Professor in Law.*
STANLEY M. JOHANSON, B.S., LL.B., LL.M.; *James A. Elkins Centennial Chair in Law, University Distinguished Teaching Professor.*
CALVIN H. JOHNSON, B.A., J.D.; *John T. Kipp Chair in Corporate and Business Law.*
SUSAN R. KLEIN, B.A., J.D.; *Alice McKean Young Regents Chair in Law.*
ALAN J. KUPERMAN, B.A., M.A., Ph.D.; *Associate Professor of Public Affairs.*
JENNIFER E. LAURIN, B.A., J.D.; *Professor of 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.*
ANGELA K. LITTWIN, B.A., J.D.; *Professor of Law.*
VIJAY MAHAJAN, M.S.Ch.E., Ph.D.; *Professor of Marketing Administration.*
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.*
STEVEN A. MOORE, B.A., Ph.D.; *Professor of Architecture.*
SUSAN C. MORSE, A.B., J.D.; *Professor.*
LINDA S. MULLENIX, B.A., M.Phil., J.D., Ph.D.; *Morris and Rita Atlas Chair in Advocacy.*
STEVEN P. NICHOLS, B.S.M.E., M.S.M.E., J.D., Ph.D.; *Professor of Engineering.*
ROBERT J. PERONI, B.S.C., J.D., LL.M.; *The Fondren Foundation Centennial Chair for Faculty Excellence.*
H. W. PERRY JR., B.A., M.A., Ph.D.; *Associate Professor of Government and Law.*
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.; *Joseph D. Jamail Centennial Chair in Law, 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.; *Mark G. and Judy G. Yudof Chair in Law.*
DAVID W. ROBERTSON, B.A., LL.B., LL.M., J.S.D.; *William Powers Jr. and Kim L. Heilbrun Chair in Tort Law, University Distinguished Teaching Professor.*
MARY ROSE, A.B., M.A., Ph.D.; *Associate Professor of Sociology.*
WILLIAM M. SAGE, A.B., M.D., J.D.; *James R. Dougherty Chair for Faculty Excellence, Professor of Surgery and Perioperative Care.*
LAWRENCE SAGER, B.A., LL.B.; *Alice Jane Drysdale Sheffield Regents Chair.*
CHARLES M. SILVER, B.A., M.A., J.D.; *Roy W. and Eugenia C. McDonald Endowed Chair in Civil Procedure, Professor of Government, 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.*
TARA A. SMITH, B.A., PH.D.; *Professor of Philosophy.*
DAVID B. SPENCE, B.A., J.D., M.A., Ph.D.; *Professor of Business, Government and Society, and Law.*
JAMES C. SPINDLER, B.A., M.A., J.D., Ph.D.; *The Sylvan Lang Professor of Law, Professor of Business.*
JORDAN M. STEIKER, B.A., J.D.; *Judge Robert M. Parker Endowed Chair in Law, Director of Capital Punishment Center.*
MICHAEL F. STURLEY, B.A., J.D.; *Fannie Coplin Regents Chair.*
JEREMI SURI, A.B., M.A., Ph.D.; *Professor of Public Affairs.*
JEFFREY K. TULIS, B.A., M.A., Ph.D.; *Associate Professor of Government.*
SRIRAM VISHWANATH, B.S., M.S., Ph.D.; *Associate Professor of Electrical and Computer Engineering.*
STEPHEN I. VLADECK, B.A., J.D., *Professor.*
WENDY E. WAGNER, B.A., M.E.S., J.D.; *Joe A. Worsham Centennial Professor.*
MELISSA F. WASSERMAN, B.S., Ph.D., J.D.; *Professor of Law.*
LOUISE WEINBERG, A.B., LL.M., J.D.; *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., J.D., Ph.D.; *Bernard J. Ward Centennial Professor in Law.*
SEAN H. WILLIAMS, B.A., J.D.; *Professor of Law.*
PATRICK WOOLLEY, A.B., J.D.; *Beck, Redden & Secrest Professor in Law.*

ASSISTANT PROFESSORS

JAMES W. MCCLELLAND, B.S., Ph.D.

TIMOTHY D. WERNER, B.A., M.A., Ph.D.

SENIOR LECTURERS, WRITING LECTURERS, AND CLINICAL PROFESSORS

WILLIAM H. BEARDALL JR., B.A., J.D.; *Clinical Professor, Director of Transnational Worker Rights Clinic.*
NATALIA V. BLINKOVA, B.A., M.A., J.D.; *Lecturer.*
PHILIP C. BOBBITT, A.B., J.D., Ph.D.; *Distinguished Senior Lecturer.*
HUGH L. BRADY, B.A., J.D.; *Clinical Professor, Director of Legislative Lawyering Clinic.*
KAMELA S. BRIDGES, B.A., B.J., J.D.; *Lecturer, Director of The David J. Beck Center for Legal Research, Writing and Appellate Advocacy.*
JOHN C. BUTLER, B.B.A., Ph.D.; *Clinical Associate Professor.*
MARY R. CROUTER, A.B., J.D.; *Clinical Professor, Assistant Director of William Wayne Justice Center for Public Interest Law.*
MICHELE Y. DEITCH, B.A., M.S., J.D.; *Senior Lecturer.*
TIFFANY J. DOWLING, B.A., J.D.; *Clinical Instructor, Director of Actual Innocence Clinic.*
LORI K. DUKE, B.A., J.D.; *Clinical Professor.*
ARIEL E. DULITZKY, J.D., LL.M.; *Clinical Professor, Director of Human Rights Clinic.*

LISA R. ESKOW, A.B., J.D.; *Lecturer.*
LYNDA E. FROST, B.A., M.Ed., J.D., Ph.D.; *Clinical Associate Professor.*
DENISE L. GILMAN, B.A., J.D.; *Clinical Professor, Director of Immigration Clinic.*
KELLY L. HARAGAN, B.A., J.D.; *Clinical Professor, Director of Environmental Law Clinic.*
HARRISON KELLER, B.A., M.A., Ph.D.; *Senior Lecturer, Vice Provost for Higher Education Policy at the University of Texas at Austin.*
ANDREW KULL, B.A., B.A., M.A., J.D.; *Distinguished Senior Lecturer.*
BRIAN R. LENDECKY, B.B.A., M.P.A.; *Senior Lecturer.*
JEANA A. LUNGWITZ, B.A., J.D.; *Clinical Professor, Director of Domestic Violence Clinic.*
JIM MARCUS, B.A., J.D.; *Clinical Professor, Co-Director of Capital Punishment Clinic.*
FRANCES L. MARTINEZ, B.A., J.D.; *Clinical Professor.*
TRACY W. MCCORMACK, B.A., J.D.; *Senior Lecturer, Director of Advocacy Programs.*
F. SCOTT MCCOWN, B.S., J.D.; *Clinical Professor, Director of Children's Rights Clinic.*
RANJANA NATARAJAN, B.A., J.D.; *Clinical Professor, Director of Civil Rights Clinic.*
SEAN J. PETRIE, B.A., J.D.; *Lecturer.*
ELIZA T. PLATTS-MILLS, B.A., J.D.; *Clinical Professor.*
RACHAEL RAWLINS, B.A., M.R.P., J.D.; *Senior Lecturer.*
CHRIS ROBERTS, B.S., J.D.; *Clinical Professor, Director of Criminal Defense Clinic.*
AMANDA M. SCHAEFFER, B.A., J.D.; *Lecturer.*
WAYNE SCHIESS, B.A., J.D.; *Senior Lecturer.*
RAOUL D. SCHONEMANN, B.A., J.D. LL.M.; *Clinical Professor.*
PAMELA J. SIGMAN, B.A., J.D.; *Clinical Professor, Director of Juvenile Justice Clinic.*
DAVID S. SOKOLOW, B.A., M.A., J.D., M.B.A.; *Distinguished Senior Lecturer.*
ELISSA C. STEGLICH, B.A., J.D.; *Clinical Professor.*
STEPHEN SLICK, B.A., M.P.P., J.D.; *Clinical Professor, The Robert S. Strauss Center and The Clements Center for National Security.*
LESLIE L. STRAUCH, B.A., J.D.; *Clinical Professor.*
GRAHAM B. STRONG, B.A., J.D., LL.M.; *Distinguished Senior Lecturer.*
MELINDA E. TAYLOR, B.A., J.D.; *Senior Lecturer, Executive Director of Kay Bailey Hutchison Center for Energy, Law and Business.*
HEATHER K. WAY, B.A., B.J., J.D.; *Clinical Professor, Director of Entrepreneurship and Community Development Clinic.*
LUCILLE D. WOOD, B.A., J.D.; *Clinical Professor.*
ELIZABETH M. YOUNGDALE, B.A., M.L.I.S., J.D.; *Lecturer.*

ADJUNCT PROFESSORS AND OTHER LECTURERS

ROBERT J. ADAMS JR., B.S., M.B.A., Ph.D.
JAMES B. ADKINS JR., B.A., J.D.
ELIZABETH AEBERSOLD, B.A., M.S.
RICKY ALBERS, B.B.A., M.B.A., J.D.
WILLIAM R. ALLENSWORTH, B.A., J.D.
OWEN L. ANDERSON, B.A., J.D.
ANDREW W. AUSTIN, B.A., M.Phil., J.D.
SAMY AYOUB, B.A., M.Sc. Ph.D.
JACK BALAGIA, B.A., J.D.
CRAIG D. BALL, B.A., J.D.
SHARON C. BAXTER, B.S., J.D.
JERRY A. BELL, B.A., J.D.
MICHAEL L. BENEDICT
ALLISON H. BENESCH, B.A., M.S.W., J.D.
CRAIG R. BENNETT, B.S., J.D.
NADIA BETTAC, B.A., J.D.
MURFF F. BLEDSOE, B.A., J.D.
SUSAN L. BLOUNT, B.A., J.D.
ANNA C. BOCCHINI, B.A., J.D.
DIANA K. BORDEN, B.A., J.D.
WILLIAM P. BOWERS, B.B.A., J.D., LL.M.
STACY L. BRAININ, B.A., J.D.
ANTHONY W. BROWN, B.A., J.D.
JAMES E. BROWN, B.S., LL.B., J.D.
TOMMY L. BROYLES, B.A., J.D.
PAUL J. BURKA, B.A., LL.B.
ERIN G. BUSBY, B.A., J.D.
DAVID J. CAMPBELL, B.A., J.D.
AGNES E. CASAS, B.A., J.D.
RUBEN V. CASTANEDA, B.A., J.D.
EDWARD A. CAVAZOS, B.A., J.D.
LINDA BRAY CHANOW, B.A., J.D.
JEFF CIVINS, A.B., M.S., J.D.
REED CLAY JR., B.A., J.D.
ELIZABETH COHEN, B.A., M.S.W., J.D.
KEVIN D. COLLINS, B.A., J.D.
JULIO C. COLON, A.A., B.A., J.D.
JOSEPH E. COSGROVE JR., A.A., B.B.A., J.D.
STEPHEN E. COURTER, B.S., M.S.B.A.
KASIA SOLON CRISTOBAL, B.A., M.S., J.D.
KEITH B. DAVIS, B.S., J.D.
SCOTT D. DEATHERAGE, B.A., J.D.
DICK DEGUERIN, B.A., LL.B.
MELONIE M. DEROSE, B.A., J.D.
RICHARD D. DEUTSCH, B.A., B.A., J.D.
REBECCA H. DIFFEN, B.A., J.D.
PHILIP DURST, B.A., M.A., J.D.
ELANA S. EINHORN, B.A., J.D.
RACHEL A. EKERY, A.B., J.D.
LUKE J. ELLIS, B.A., J.D.
JAY D. ELLWANGER, B.A., J.D.
RANDALL H. ERBEN, B.A., J.D.
EDWARD Z. FAIR, B.A., M.S.W., J.D.
KAY FIRTH-BUTTERFIELD, B.A., M.B.A., LL.M.
ANDREW R, FLORANCE, B.A., M.S.I., B.A.
JAMES G. FOWLER, B.A., M.A., J.D.
KYLE K. FOX, B.A., J.D.
DAVID C. FREDERICK, B.A., Ph.D., J.D.
GREGORY D. FREED, B.A., J.D.
JENNIFER S. FREEL, B.J., J.D.
FRED J. FUCHS, B.A., J.D.
HELEN A. GAEBLER, B.A., J.D.
MICHELLE M. GALAVIZ, B.A., J.D.
RYAN M. GARCIA, B.G.S., J.D.
BRYAN A. GARNER, B.A., J.D.
MICHAEL S. GOLDBERG, B.A., J.D.
DAVID M. GONZALEZ, B.A., J.D.
JOHN F. GREENMAN, B.A., M.F.A., J.D.
DAVID HALPERN, B.A., J.D.
ELIZABETH HALUSKA-RAUSCH, B.A., M.A., M.S., Ph.D.
CLINT A. HARBOUR, B.A., B.A., J.D., LL.M.
ROBERT L. HARGETT, B.B.A., J.D.
MARY L. HARRELL, B.S., J.D.
WILLIAM M. HART, B.A., J.D.
KEVIN V. HAYNES, B.A., J.D.
JOHN R. HAYS, JR., B.A., J.D.
ELIZABETH E. HILKIN, B.A., M.S., J.D.
BARBARA HINES, B.A., J.D.
KENNETH E. HOUP, JR., B.A., J.D.
RANDY R. HOWRY, B.J., J.D.
BART W. HUFFMAN, B.S.E., J.D.
MONTY G. HUMBLE, B.A., J.D.
JENNIFER D. JASPER, B.S., M.A., J.D.
WALLACE B. JEFFERSON, B.A., J.D.

CHRISTOPHER S. JOHNS, B.A., L.L.M., J.D.
AARON M. JOHNSON, B.A., J.D.
DIRK M. JORDAN, B.A., J.D.
JEFFREY R. JURY, B.A., J.D.
PATRICK O. KEEL, B.A., J.D.
DOUGLAS L. KEENE, B.A., M.Ed., Ph.D.
SCOTT A. KELLER, B.A., J.D.
CHARI L. KELLY, B.A., J.D.
JEAN A. KELLY, B.A., J.D.
ROBERT N. KEPPLE, B.A., J.D.
PAUL S. KIMBOL, B.A., J.D.
MICHAEL R. KRAWZSENEK, B.S., J.D.
LARRY LAUDAN, B.A., M.A., Ph.D.
JODI R. LAZAR, B.A., J.D.
KEVIN L. LEAHY, B.A., J.D.
CYNTHIA C. LEE, B.S., B.A., M.A., J.D.
DAVID P. LEIN, B.A., M.P.A., J.D.
KEVIN R. LEISKE, B.S., L.L.M., J.D.
ANDRES J. LINETZKY, B.A., LL.M.
JAMES LLOYD LOFTIS, B.B.A., J.D.
ANDREW F. MACRAE, B.J., J.D.
MARK F. MAI, B.B.A., J.D.
ANDREA M. MARSH, B.A., J.D.
HARRY S. MARTIN, A.B., M.L.S., J.D.
ERIN D. MARTINSON, B.A., J.D.
MIMI MARZIANI, B.A., J.D.
LORI R. MASON, B.A., J.D.
PETER C. MCCABE, B.A., J.D.
BARRY F. MCNEIL, B.A., J.D.
MARGARET M. MENICUCCI, B.A., J.D.
JO ANN MERICA, B.A., J.D.
RANELLE M. MERONEY, B.A., J.D.
EDWIN G. MORRIS, B.S., J.D.
JAMES C. MORRISS III, B.S., J.D.
SARAH J. MUNSON, B.A., J.D.
GEORGE D. MURPHY JR., B.B.A., J.D.
JOHN A. NEAL, B.A., J.D.
CLARK M. NEILY, B.A. J.D.
JUSTIN A. NELSON, B.A., J.D.
MANUEL H. NEWBURGER, B.A., J.D.
MARTHA G. NEWTON, B.A., J.D.
HOWARD D. NIRKEN, B.A., M.P.Aff, J.D.
CHRISTINE S. NISHIMURA, B.A., J.D.
DAVID G. NIX, B.S.E., LL.M., J.D.
JOSEPH W. NOEL, B.S.E., J.D., M.S.L.S.
JANE A. O'CONNELL, B.A., M.S., J.D.
PATRICK L. O'DANIEL, B.B.A., J.D.
LISA M. PALIN, B.A., M.Ed., M.F.A., J.D.
MARK L. PERLMUTTER, B.S., J.D.
EDSON PETERS, LL.B., LL.M., Ph.D.
JONATHAN PRATTER, B.A., M.S.L.I.S., J.D.
MARGARET K. REIN, B.A., J.D.
DAWN REVELEY, B.A., J.D.
CLARK W. RICHARDS, B.A., LL.M., J.D.
BRIAN C. RIDER, B.A., J.D.
ROBERT M. ROACH JR., B.A., J.D.
BETTY E. RODRIGUEZ, B.S.W., J.D.
MICHELLE L. ROSENBLATT, B.A., J.D.
JAMES D. ROWE, B.A., J.D.
MATTHEW C. RYAN, B.A., J.D.
MARK A. SANTOS, B.A., J.D.
MICHAEL J. SCHLESS, B.A., J.D.
AMY J. SCHUMACHER, B.A., J.D.
SUZANNE SCHWARTZ, B.J., J.D.
RICHARD J. SEGURA Jr., B.A., J.D.
STACEY ROGERS SHARP, B.S., J.D.
DAVID A. SHEPPARD, B.A., J.D.
HON. ERIC M. SHEPPERD, B.A., J.D.
TRAVIS J. SIEBENEICHER, B.B.A, J.D.
RONALD J. SIEVERT, B.A., J.D.
AMBROSIO A. SILVA, B.S., J.D.
STUART R. SINGER, A.B., J.D.
STEPHEN T. SMITH, B.A., M.S.
BARRY T. SMITHERMAN, B.B.A., M.P.A., J.D.
LYDIA N. SOLIZ, B.B.A., J.D.
JAMES M. SPELLINGS Jr., B.S., J.D.
MATTHEW R. STEINKE, B.A., M.L.I.S., J.D.
WILLIAM F. STUTTS, B.A., J.D.
MATTHEW J. SULLIVAN, B.S., J.D.
JEREMY S. SYLESTINE, B.A., J.D.
KEITH P. SYSKA, B.S., J.D.
BRADLEY P. TEMPLE, B.A., J.D.
ROSA THEOFANIS, B.A., J.D.
SHERINE E. THOMAS, B.A., J.D.
CARLY M. TOEPKE, B.A., J.D.
MICHAEL J. TOMSU, B.A., M.B.A., J.D.
TERRY O. TOTTENHAM, B.S., LL.M., J.D.
CARLOS R. TREVINO, M.B.A., LL.M., J.D.
TIMOTHY J. TYLER, B.A., J.D.
SUSAN S. VANCE, B.B.A., J.D.
LANA K. VARNEY, B.A., J.D.
ELEANOR K. VERNON, B.A., J.D.
KEEGAN D. WARREN-CLEM, B.A., J.D., LL.M.
CHRISTOPHER M. WEIMER, B.A., J.D.
WARE V. WENDELL, A.B., J.D.
RODERICK E. WETSEL, B.A., J.D.
BENJAMIN B. WHITTENBURG, B.B.A., M.P.A., J.D.
RANDALL B. WILHITE, B.B.A., J.D.
DAVID G. WILLE, B.S.E.E., M.S.E.E., J.D.
ANDREW M. WILLIAMS, B.A., J.D.
JAMES C. WINTERS, B.A., J.D.
TRAVIS M. WOHLERS, B.S., Ph.D., J.D.
DENNEY L. WRIGHT, B.B.A., J.D., LL.M.
DANIEL J. YOUNG, B.A., J.D.
EVAN A. YOUNG, A.B., B.A., J.D.

VISITING PROFESSORS

VICTOR FERRERES, J.D., LL.M., J.S.D.

* * *

Texas Law Review

Volume 96 2017–2018

* * *

Texas Law Review

Volume 96, Number 2, December 2017

ARTICLES

ESSAY

NOTES

* * *

CEO Pay Redux

K.J. Martijn Cremers,* Saura Masconale,** and Simone M. Sepe***

Managerial power theory holds that structural flaws in corporate governance, such as board defenses, enable opportunistic managers to extract excessive pay. While this theory has proven highly influential, this Article argues that it fails to answer important questions. For example, how does managerial power theory relate to the prevailing economic paradigm of CEO pay as reflecting competition for scarce managerial talent? Further, how can one reconcile the theory's negative account of board protection with recent empirical studies showing that such protection is value increasing?

In investigating these and other questions—both theoretically and empirically—this Article makes four contributions. First, it shows that adopting defensive measures (such as the staggered board) is not associated with significant changes in CEO pay. Second, it documents that greater competition for managerial talent is positively associated with CEO pay. Third, it shows that higher CEO pay is associated with higher firm value, especially in firms with a staggered board. Fourth, it provides plausible causal evidence that the decline in stock options (i.e., high-powered incentives) that followed the 2005 mandate to expense options is associated with increased firm value.

These results suggest that high executive pay serves to attract talented managers, rather than reflecting managerial opportunism. They also suggest

---

*Professor of Finance, Mendoza College of Business, University of Notre Dame.

**Visiting Professor, University of Notre Dame Law School and Research Fellow at the Institute for Advanced Study in Toulouse—Fondation Jean-Jacques Laffont—Toulouse School of Economics.

***Professor of Law and Finance, James E. Rogers College of Law, University of Arizona; and Institute for Advanced Study in Toulouse—Fondation Jean-Jacques Laffont—Toulouse School of Economics. Email address: sms234@email.arizona.edu. Corresponding author. We are grateful for the criticisms and suggestions made by Scott Altman, Andrea Attar, Michelle Anderson, Shyam Balganesh, Jonathan Barnett, Matt Barrett, Bruno Biais, Scott Bice, Kenworthey Bilz, William Bratton, Peg Brinig, Rebecca Brown, Alex Capron, Howard Chang, Jacques Crémer, David Cruz, Ronald Garet, Sam Erman, Andrew Guzman, Gillian Hadfield, Dave Hoffman, Gregory Keating, Ed Klenbard, Dan Klerman, Dan Kelly, Michael Knoll, Martin Levine, Daniel Markovits, Edward McCaffery, Robert Mundheim, Kevin J. Murphy, Peter Nascenzi, Sebastian Pouget, Robert Rasmussen, Donald Scotten, Michael Simkovic, Dan Simon, Nomi Stolzenberg, Avishalom Tor, Michael Wachter, Mark Weinstein, Paul Weithman, and Abby Wood.

that board protection might be beneficial to prevent market pressure from introducing value-reducing distortions in executive pay, such as an excessive use of high-powered incentives emphasizing short- over long-term performance. The Article concludes by discussing the policy implications of the analysis.

Introduction

During the highly conflictual 2016 presidential race, the candidates agreed on at least one issue: the need to curb excessive executive pay. Then-candidate Donald Trump described CEO pay as "disgraceful . . . [and] a total and complete joke."[1] Along similar lines, Hillary Clinton lamented that it

1. Krista Hughes & Toni Clarke, *Trump Says High Pay for CEOs is a Joke and 'Disgraceful'*, REUTERS (Sept. 13, 2015), http://www.reuters.com/article/us-usa-election-trump-idUSKCN0RD0PA20150913 [https://perma.cc/8JCM-YTVQ].

"just doesn't make sense" that today's CEOs make 300 times more than the typical worker.[2]

The current political hostility toward CEO pay is unsurprising when considering the rising public outrage over executive compensation, especially after the 2007–2009 financial crisis. General discontent with executive pay is also not a recent trend; rather, "scrutinizing, criticizing, and regulating high levels of executive pay has been an American pastime for nearly a century."[3] Since the early 2000s, however, this trend has found a systematic theoretical framework in the "managerial power theory" of executive compensation, espoused most prominently by Harvard law professors Lucian A. Bebchuk and Jesse M. Fried.[4]

Managerial power theory views the typical CEO pay package as a reflection of managerial moral hazard. Managerial moral hazard is the risk that managers may fail to exert sufficient effort and abuse their corporate power for personal gains.[5] In response to this risk, corporate law grants ultimate control over corporate affairs to the board of directors, as the institution charged with monitoring management decisions in the interest of shareholders.[6] Managerial power scholars, however, argue that structural flaws in corporate governance, such as board insulation from shareholder discipline through the use of defensive measures, make boards largely beholden to managers.[7] As a result, among other inefficiencies, managers can extract excessive pay or pay that is not tied to performance—in economic terms, "inefficient rents."[8]

Managerial power theory has now become the dominant view in the law and economics literature.[9] It has also led to major regulatory changes

---

2. Dan Merica & Eric Bradner, *Clinton Focuses on Economy Amid Email Controversy*, CNN (July 24, 2015), http://www.cnn.com/2015/07/24/politics/hillary-clinton-economic-speech-2016-new-york [https://perma.cc/Y64Z-94EZ].

3. Kevin J. Murphy, *Executive Compensation: Where We Are, and How We Got There*, *in* 2A HANDBOOK OF THE ECONOMICS OF FINANCE 211, 213 (George M. Constantinides et al. eds., 2013).

4. *See* sources cited *infra* note 48.

5. *See* JEAN-JACQUES LAFFONT & DAVID MARTIMORT, THE THEORY OF INCENTIVES: THE PRINCIPAL-AGENT MODEL 145 (2002) ("By the mere fact of delegation, the principal often loses any ability to control those actions [of the agent] that are no longer observable. . . . Those actions cannot be contracted upon because no one can verify their value. In such cases we will say that there is moral hazard.").

6. *See* Lucian A. Bebchuk, *The Myth of the Shareholder Franchise*, 93 VA. L. REV. 675, 679–80 (2007) (describing the classic view that the board is charged with addressing the vertical agency problem between shareholders and managers).

7. LUCIAN BEBCHUK & JESSE FRIED, PAY WITHOUT PERFORMANCE: THE UNFULFILLED PROMISE OF EXECUTIVE COMPENSATION 23–44, 55–56 (2004).

8. *See id.* at 62 (using the term "rents" to refer "to the additional value that managers obtain beyond what they would get in arm's-length bargaining with a board that had both the inclination to maximize shareholder value and the necessary time and information to perform that task properly").

9. *See, e.g.*, William W. Bratton, *The Academic Tournament over Executive Compensation*, 93 CALIF. L. REV. 1557, 1560 (2005) (reviewing BEBCHUK & FRIED, *supra* note 7) (explaining that

promoting shareholder empowerment, which managerial power scholars defend as the most effective remedy to address excessive executive pay.[10] Most notably, in 2010, the Dodd-Frank Wall Street Reform and Consumer Protection Act yielded a series of new executive-compensation rules that, among other requirements, mandated that all U.S. public companies introduce nonbinding shareholder votes on CEO pay.[11]

Notwithstanding the far-reaching success of managerial power theory, this Article argues that this theory fails to convincingly answer crucial questions about executive compensation. In the first place, managerial power scholars fail to explain how one should reconcile their view with what this Article refers to as the "managerial talent theory" of executive compensation. The latter theory is the prevailing *economic* paradigm of executive compensation, pursuant to which CEO pay reflects compensation for scarce managerial talent in competitive markets.[12] Managerial talent theory thus challenges the view that the executive-compensation process is isolated from competitive pressure, as managerial power scholars seem to assume.

Further, managerial power theory is premised on a static, one-period model of executive compensation, where the manager initially makes decisions, and then investments are liquidated, gains or losses are realized, and the manager gets paid within the end of the single period.[13] In the real corporate world, however, the relationships between managers, boards, and shareholders tend to be "dynamic," as investments typically play out along multi-period horizons and top executives hold their positions for several years.[14] Consequently, we raise the question of how moving from a static to a dynamic setting affects the positive or normative conclusions of managerial power theory.

Still, recent empirical work on the value impact of defensive measures documents that temporary board protection from shareholder interference,

---

Bebchuk and Fried's managerial power theory "change[d] the terms of discourse" in the field); Randall S. Thomas & Harwell Wells, *Executive Compensation in the Courts: Board Capture, Optimal Contracting, and Officers' Fiduciary Duties*, 95 MINN. L. REV. 846, 847–48, 852 (2011) (observing that "Board Capture" theory, or managerial power theory, is "[f]ar more popular" than "the belief that the American executive compensation system works well").

10. *See infra* notes 86–90 and accompanying text.

11. *See* Dodd-Frank Wall Street Reform and Consumer Protection Act, Pub. L. No. 111-203, § 951(a)(1), 124 Stat. 1376, 1899 (2010) (codified as amended at 15 U.S.C. § 78n-1).

12. *See infra* note 70-72 and accompanying text.

13. *See* Bengt Holmström, Pay Without Performance *and the Managerial Power Hypothesis: A Comment*, 30 J. CORP. L. 703, 708 (2005) (generally observing that managerial power theory is premised on "the traditional moral hazard model," which is a static model); Lucian A. Bebchuk & Holger Spamann, *Regulating Bankers' Pay*, 98 GEO. L.J. 247, 249 (2010) (explicitly employing a one-period model of executive compensation).

14. *See, e.g.*, Alex Edmans & Xavier Gabaix, *Executive Compensation: A Modern Primer*, 54 J. ECON. LITERATURE 1232, 1260 (2016) (analyzing "dynamic models of moral hazard").

such as the protection granted by a staggered board,[15] is associated with increased firm value.[16] But if board protection can serve a positive governance function, how can it also be the main culprit for inefficient executive pay, as claimed by managerial power scholars?

In investigating these and other questions—both theoretically and empirically—this Article makes four basic contributions. First, it shows that the data do not support the managerial power claim that defensive measures are a source of distortions in CEO pay, as the adoption of such measures is not associated with significant changes in CEO pay levels or structure. Second, the Article documents that greater competition in the market for managerial talent is associated with statistically and economically significant increases in CEO pay, consistent with the predictions of managerial talent theory. Third, it documents that higher CEO pay is associated with higher firm value, especially in firms with a staggered board. Fourth, it provides plausible causal evidence that firms were overusing option-based pay (that is, high-powered incentives) in the early 2000s.

Overall, our analysis challenges the view that high executive wages generally constitute inefficient rents, indicating that such wages are instead often necessary to attract talented executives. It also suggests that protecting boards from short-term market and shareholder interference may promote a more positive relationship between CEO pay and firm value. This is because board protection makes it less likely that market forces introduce distortions in incentive schemes, such as an excessive use of high-powered incentives that inefficiently emphasizes short- over long-term performance.

The Article proceeds in five parts. In Part I, we provide the background necessary to understanding the context and importance of the executive-compensation debate, discussing the legal and economic foundations of the main theories of executive compensation.

Next, in Part II, we attempt to understand what difference a dynamic approach to executive compensation makes for the theoretical validity of managerial power theory.[17] As a positive matter, we argue that the more

---

15. In a staggered board, directors are grouped into different classes (usually three), and each class of directors stands for reelection in successive years, which requires challengers to win at least two election cycles to gain a board majority. *See* Richard H. Koppes et al., *Corporate Governance Out of Focus: The Debate over Classified Boards*, 54 BUS. LAW. 1023, 1029 n.19 (1999) (providing a list of relevant state law provisions allowing staggered boards).

16. *See* K.J. Martijn Cremers & Simone M. Sepe, *The Shareholder Value of Empowered Boards*, 68 STAN. L. REV. 67, 73–74, 114–15 (2016) (documenting evidence that staggered boards positively affect firm value); K.J. Martijn Cremers, Lubomir Litov & Simone M. Sepe, *Staggered Boards and Long-Term Firm Value, Revisited*, 126 J. FIN. ECON. 422 (2017) (providing robust evidence that staggered boards *are not* negatively related to firm value).

17. One of us explored the positive and normative implications of a dynamic approach to executive compensation in earlier work. *See* Simone M. Sepe, *Making Sense of Executive Compensation*, 36 DEL. J. CORP. L. 189, 212–13 (2011). As a positive matter, the analysis developed

realistic assumption of a multi-period relationship between managers, boards, and shareholders produces several complexities for the analysis of managerial incentives, which managerial power theory fails to incorporate. In particular, in a dynamic setting, managers may develop incentives for engaging in inefficient intertemporal tradeoffs. The result is that shareholders not only are exposed to the risk of suboptimal managerial effort, but to the additional agency problem of short-termism (or managerial myopia).[18] This problem arises because managers—especially those compensated with high-powered equity incentives, such as option grants—may prefer investments that boost short-term returns (and a manager's current pay) at the expense of losses only occurring in the future, which the manager discounts.

At the same time, a multi-period horizon enhances the opportunities for efficient incentive design, also challenging the normative conclusions of managerial power theory. When the board–manager relationship develops along multiple periods (as is commonly the case), the board can spread rewards for good managerial performance over time, while periodically reviewing the manager's performance. This dynamic context allows the board to exploit a manager's "continuation value," that is, a manager's expected payoffs from future employment periods.[19] Indeed, these expected payoffs provide a powerful bonding mechanism to ensure that exerting long-term effort is in the manager's interest, as the manager anticipates that exerting suboptimal effort or engaging in myopic strategies increases the risk of being terminated and, therefore, losing the continuation value.

This conclusion offers a plausible explanation for the positive governance function of board protection documented in recent empirical studies, while warning against the potential costs of increased market pressure, whether coming from enhanced shareholder power or intense

---

in this Article expands this earlier analysis, incorporating a competitive-market framework. As a normative matter, it revisits that analysis, based on the novel empirical evidence produced herein.

18. This Article uses the terms "short-termism" and "managerial myopia" interchangeably to refer to an excessive managerial focus on short-term results at the expense of long-term firm value. The standard reference for myopia studies is the pioneering work of Jeremy Stein. *See* Jeremy C. Stein, *Efficient Capital Markets, Inefficient Firms: A Model of Myopic Corporate Behavior*, 104 Q.J. ECON. 655, 667 (1989) (explaining that "myopia can arise from one of three 'imperfections': (1) invisibility of some managerial action, (2) ex ante superior information on the part of managers, or (3) inefficiencies in stock prices"); Jeremy C. Stein, *Takeover Threats and Managerial Myopia*, 96 J. POL. ECON. 61, 63–66 (1988) (modeling a scenario in which managers engage in "wasteful signaling" in order to boost current earnings).

19. The assumption here is that the manager is not terminated in earlier periods for poor performance. Put differently, for the continuation value mechanism to be viable, some level of tolerance for what may appear as "early failure" is required. *See* Gustavo Manso, *Motivating Innovation*, 66 J. FIN. 1823, 1823–24 (2011) (arguing that pay-for-performance schemes that reward or penalize managers based on near-term outcomes may have adverse consequences if the goal is to induce managers to explore new, untested investments).

product competition.[20] When managers are subject to such market forces, they will rationally anticipate a greater risk of being removed in the near future and, therefore, substantially discount their expected continuation value. By preventing market forces from interfering with a manager's continuation value, board protection might accordingly facilitate the design of pay schemes that promote long-term shareholder wealth.

Against this analytical background, Part III moves to the empirical investigation of the relationship existing between corporate governance and market forces, on the one hand, and CEO pay levels and structure, on the other. We begin by examining the relationship between the use of defensive measures and CEO pay. If the predictions of managerial power theory were accurate, we would expect to find that the adoption of such measures increases the likelihood that managers obtain higher pay or pay that is not sufficiently tied to performance. For example, pursuant to the claims made by managerial power scholars, we should find that CEO pay arrangements include larger portions of non-equity pay (e.g., cash, salary, and the like), which fail to directly align manager and shareholder interests.[21] Similarly, we would expect to find a greater use of restricted stock grants, which according to managerial power theory inefficiently provides managers with lower-powered incentives than the use of option grants.[22] In contrast with these predictions, we find no evidence that the adoption of defensive measures results in higher levels of executive compensation or changes in pay structure (neither before nor after the Dodd-Frank Act's introduction of new compensation rules).

Next, we examine the effect on CEO pay of various forms of competition. These forms include labor-market competition for managerial talent, product-market competition, and competition through merger and acquisition (M&A) activity, which we interpret as a proxy for an industry's shareholder pressure (as operating through the takeover channel).[23] Consistent with managerial talent theory, we document that competition for managerial talent has a substantial effect on both CEO pay levels and structure, as greater talent competition is positively associated with higher CEO pay levels and a larger proportion of restricted stock grants. This evidence suggests that the increase in CEO pay due to managerial talent competition largely comes from a greater use of restricted stock. On the contrary, greater product-market competition and M&A competition are associated with an increase in the option component of CEO pay and a

20. Similar to when they are exposed to enhanced shareholder power, managers are subject to intense market-driven discipline in environments with intense product competition due to relative performance evaluations under which the imperative for managers is beating competitors. *See* JEAN TIROLE, THE THEORY OF CORPORATE FINANCE 20, 28–29 (2006).

21. *See infra* text accompanying notes 63–64.

22. *See infra* text accompanying notes 65–69.

23. *See infra* subpart III(A).

decrease in the restricted stock component, indicating that boards tend to respond to stronger market discipline with more high-powered incentives.

Our analysis of the relationship between corporate governance, market forces, and CEO pay thus indicates that managerial talent competition is a critical source of increased CEO pay. Yet, the ultimate question, which we address in Part IV, is how current executive-pay levels and structure affect firms' financial performance. Contrary to the predictions of managerial power theory, we find that higher CEO pay is associated with higher firm value.[24] Combined with our result that greater managerial talent competition is associated with higher CEO pay, this finding seems to suggests that, in general, high executive pay is not "excessive" in the sense described by managerial power scholars (that is, in the sense that it dissipates shareholder value). Rather, high executive pay seems to ensure that the most talented CEOs are allocated to the most valuable firms.

Further, we document that CEO pay is more strongly positively associated with firm value in firms that adopt defensive measures and, especially, a staggered board. Conversely, greater product-market competition and higher M&A competition interacted with CEO pay are associated with lower firm value. Taken together, these results support our theoretical predictions about the positive role of board protection in the executive-compensation process and the corresponding costs of market and shareholder pressure, indicating that boards that are temporarily protected from such market forces are better positioned to design value-increasing CEO pay arrangements.

Nevertheless, it could be argued that the results of our value analysis are subject to endogeneity concerns—the ever-present risk that correlation might be mistaken for causation.[25] For example, it could be that enhanced CEO effectiveness or expectations of future positive performance result in higher CEO pay, rather than higher pay causing better performance. To mitigate such concerns, Part IV also considers an event study that focuses on the 2005 introduction of Financial Accounting Standard (FAS) regulation 123(R).[26] FAS123(R) mandated that all public firms expense stock options, eliminating the prior privileged accounting treatment of options relative to restricted stock. As the product of regulatory intervention, this event can be regarded as independent from firm-specific circumstances and, therefore, as plausibly

---

24. As it has become standard in the empirical literature, we employ Tobin's Q as a proxy for firm value. *See, e.g.*, Lucian A. Bebchuk & Alma Cohen, *The Costs of Entrenched Boards*, 78 J. FIN. ECON. 409, 419 (2005) (listing other studies that employ Tobin's Q). Tobin's Q is the ratio of the market value of assets to the book value of assets. Eugene F. Fama & Kenneth R. French, *Testing Trade-Off and Pecking Order Predictions About Dividends and Debt*, 15 REV. FIN. STUD. 1, 8 (2002).

25. *See* JEFFREY M. WOOLDRIDGE, ECONOMETRIC ANALYSIS OF CROSS SECTION AND PANEL DATA 50–51 (2002).

26. *See infra* notes 144–47 and accompanying text.

inducing exogenous changes in both the levels and structure of CEO pay (that is, changes that are outside a firm's direct control). It follows that examining the subsequent performance of firms that were and were not affected by the introduction of FAS123(R) can plausibly provide *causal* evidence about the impact of modifications in CEO pay levels and structure on firm value. In particular, as FAS123(R) leveled the playing field between the use of options and restricted stock from an accounting perspective, we expect such an event study to provide us with additional insights into the relative efficiency of more versus less powered incentives.

We show that FAS123(R) led to a significant reduction in the option component of executive pay for affected firms that had outstanding options. More importantly, we also show that the more firms reduced option grants in favor of restricted stock grants in the two years after the rule change, the greater the increase in firm value. Combined with our finding that greater market and shareholder pressures are associated with a larger use of option grants, these results confirm the above interpretation of our value analysis that market forces might introduce distortions in optimal incentive design, such as an excessive use of high-powered incentives that overemphasize short-term performance at the expense of long-term firm value.

From a normative perspective, as discussed in Part V, our analysis provides important insights for the social welfare implications of the executive-compensation debate. Our central results are that high executive pay is generally efficient in attracting talented managers, while enhanced market and shareholder pressure may produce value-reducing distortions in executive-pay schemes. Based on these results, we defend the traditional deference paid by Delaware courts to board decision-making in executive-compensation matters [27] as normatively desirable. We likewise defend temporary board protection as a means to promote, rather than jeopardize, the efficiency of executive-compensation plans—including by mitigating the risk that boards might overuse high-powered incentives in response to excessive market pressure. More broadly, we argue that policymakers would do well to reconsider the case for enhanced shareholder power in the executive-pay process, as this case emerges as both theoretically and empirically wanting.

I. Theories of Executive Compensation

This Part assesses the major theories of executive compensation, as developed in both legal and economic models. The agency problem between shareholders and managers is the common departure point of these theories. Indeed, as first famously espoused by Berle and Means, [28] and later

27. *See infra* note 34 and accompanying text.

28. ADOLF A. BERLE, JR. & GARDINER C. MEANS, THE MODERN CORPORATION AND PRIVATE PROPERTY 7, 84–89 (William S. Hein & Co. reprt. ed. 1982) (1933).

formalized by Jensen and Meckling,[29] the separation of ownership and control in the public corporation creates the risk that managers may engage in moral hazard, taking "hidden actions" in their own self-interest and at the expense of shareholders. The question, however, is whether executive compensation provides a means to address this quintessential agency problem or, rather, is part of it.

As explained in the discussion of what is commonly referred to as "optimal contracting theory" in subpart A below, executive-compensation research initially developed on the agency-theoretic premise that compensation plans provide an efficient contractual mechanism to mitigate managerial moral hazard. However, as discussed in subpart B, managerial power theory, which first emerged in the early 2000s, has challenged this premise, arguing that compensation plans reflect managerial moral hazard rather than mitigating it. Partially in response to this criticism, more recent "managerial talent theory" studies, which we discuss in subpart C, have proposed an expansion of the traditional agency model of executive compensation, moving beyond the bilateral contracting framework and emphasizing additional dimensions, such as competitive forces in the labor market for managerial talent.

A. Optimal Contracting Theory

1. Law.—The classic theory of executive compensation—referred to as "optimal (or efficient) contracting theory"—posits that compensation arrangements are designed by the board of directors to provide executives with incentives to refrain from moral hazard and maximize shareholder value.[30]

From a corporate law perspective, optimal contracting theory captures the central role served by the board of directors as the organizational body charged with mitigating the shareholder–manager agency problem.[31] Indeed, corporate law separates "decision management" and "decision control," delegating business decision-making to managers, while vesting the board with the authority to ratify and monitor management actions on behalf of shareholders.[32] To this end, directors are granted a vast array of powers,

---

29. Michael C. Jensen & William H. Meckling, *Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure*, 3 J. FIN. ECON. 305, 305–08 (1976).

30. For surveys of the vast literature on optimal contracting theory, see, for example, Kevin J. Murphy, *Executive Compensation*, *in* 3B HANDBOOK OF LABOR ECONOMICS 2485, 2519–22 (Orley C. Ashenfelter & David Card eds., 1999) (discussing classical contributions) and Alex Edmans & Xavier Gabaix, *Is CEO Pay Really Inefficient? A Survey of New Optimal Contracting Theories*, 15 EUR. FIN. MGMT. 486, 488–93 (2009) (discussing more recent contributions).

31. John E. Core et al., *Is U.S. CEO Compensation Inefficient Pay Without Performance?*, 103 MICH. L. REV. 1142, 1145–46 (2005) (reviewing BEBCHUK & FRIED, *supra* note 7).

32. Eugene F. Fama & Michael C. Jensen, *Separation of Ownership and Control*, 26 J.L. & ECON. 301, 308–10 (1983).

including the power of selecting and removing the CEO and other top executives, as well as the power of setting executive-compensation arrangements.

Reflecting the assumption that executive compensation is a matter efficiently delegated to board discretion, corporate law also provides for minimal judicial review of executive-compensation decisions, extending to such decisions the protection of the business judgment rule.[33]

2. Economics.—From an economic perspective, optimal contracting theory frames executive compensation as a remedy to reduce the moral hazard costs faced by shareholders.[34] Such costs arise because the principal (the shareholders) faces the problem of inducing effort by the agent (the managers)—or, put differently, avoiding the agent's moral hazard, for example in the form of private-benefits extraction[35]—despite being unable to observe all of the actions the agent takes.[36] The fundamental insight from the principal–agent model is that incentivizing effort requires giving the agent a monetary payoff, an *agency rent*, such that the exercise of effort is in the agent's own interest.[37]

Therefore, the board's essential problem is designing a contract that can maximize managerial effort, while minimizing the costs of providing incentives (the agency rent awarded to the manager). More specifically, the optimal contract maximizes shareholder value subject to two constraints. The first is the manager's *participation constraint*, which requires shareholders to pay the manager her reservation utility, that is, the value of her next-best employment opportunity available (also referred to as "outside option").[38] The second is the manager's *incentive constraint*, which requires shareholders to compensate the manager for choosing actions that maximize

33. *See* Brehm v. Eisner, 746 A.2d 244, 263 (Del. 2000) ("[A] board's decision on executive compensation is entitled to great deference. It is the essence of business judgment for a board to determine if a particular individual warrant[s] large amounts of money . . . ."). The sole exception is given by the doctrine of waste, under which judicial intervention in matters of executive compensation is warranted if "no person of ordinary, sound business judgment would deem it worth what the corporation has paid." Saxe v. Brady, 184 A.2d 602, 610 (Del. Ch. 1962). Since the 1990s, however, applications of the waste doctrine have become increasingly rare, so much that Leo E. Strine Jr., now chief justice of the Delaware Supreme Court, described waste as a "vestige" of discarded doctrines. Harbor Fin. Partners v. Huizenga, 751 A.2d 879, 897 (Del. Ch. 1999).

34. Michael C. Jensen & Kevin J. Murphy, *Performance Pay and Top-Management Incentives*, 98 J. POL. ECON. 225, 226 (1990).

35. *See* JOHN ROBERTS, THE MODERN FIRM 127 (2004) ("For simplicity we call the action being taken [by the agent] 'effort provision,' but numerous other interpretations are possible.").

36. *See* Jensen & Meckling, *supra* note 29, at 308 (describing the costs associated with ensuring that the agent's actions align with the principal's viewpoint and interests).

37. *See* LAFFONT & MARTIMORT, *supra* note 5, at 29 (referring to agency rent as "information rent").

38. *See* BERNARD SALANIÉ, THE ECONOMICS OF CONTRACTS: A PRIMER 122 (2d ed. 2005).

shareholder value rather than pursuing opportunities that the manager privately prefers but result in lower shareholder value.

The contracts that solve this optimization problem involve compensation schemes that tie CEO pay to shareholder value,[39] thus providing the theoretical justification for the current prevalence of equity-based compensation (such as stock options and restricted stock) over fixed compensation in CEO pay packages.[40] The intuitive case for equity-based compensation is straightforward: when financial rewards are guaranteed, managers are assumed to have no reason to avoid self-serving behavior. In contrast, when pay is anchored to increased shareholder value, managerial incentives are aligned with the interests of shareholders, mitigating moral hazard.

A simple example helps illustrate the economics of optimal contracting theory. We assume that a Board needs to set the compensation plan of a Manager who has an outside option valued at $5.[41] The Manager can take a project that generates revenues of $1,000 with probability 80% or revenues of $0 with the remaining 20% probability. To stylize the possibility of moral hazard, we assume that the Manager has the possibility of extracting a private benefit of $4, in which case the probability of the project's success drops to 60%.[42] This means that in order to maximize the project's chances of success, the Manager has to bear a cost of $4, which corresponds to the "disutility" of effort[43] of giving up private-benefit extraction.

---

39. The standard reference is to Michael C. Jensen & Kevin J. Murphy, *CEO Incentives—It's Not How Much You Pay, but How*, HARV. BUS. REV., May–June 1990, at 138, 143.

40. In the absence of more immediate proxies for evaluating corporate results, stock market value is described as the most reliable indicator of the "value of the entire future stream of expected cash flows." MICHAEL C. JENSEN, A THEORY OF THE FIRM: GOVERNANCE, RESIDUAL CLAIMS, AND ORGANIZATIONAL FORMS 146 (2000). Total compensation is generally comprised of six basic components: (1) base salaries; (2) discretionary bonuses; (3) non-equity incentives (based on both annual and multi-year performance measures); (4) stock options; (5) stock awards; and (6) other pay. Murphy, *supra* note 3, at 221. The "other pay" component usually includes items such as severance/change in control benefits, perquisites, pensions, post-retirement and consulting contracts.

41. By posing that the Manager's outside option is $5, we are assuming here an environment with virtually no competition. The modification of our example in section (I)(C)(2) below relaxes this assumption.

42. *See* Bengt Holmström, *Moral Hazard and Observability*, 10 BELL J. ECON. 74, 75–76 (1979) (modeling the principal's monetary payoff as a function of both the agent's unobservable actions, i.e., effort, and a random state of nature, with the expected realization of the principal's monetary payoff increasing in the agent's effort level). Of course, here we are assuming that corporate risk is unaffected by systemic changes; that is, we are only considering the idiosyncratic risk that cannot be diversified away.

43. LAFFONT & MARTIMORT, *supra* note 5, at 145–46.

Under these assumptions, the problem for the Board[44] is twofold. On the one hand, the Board needs to prevent the Manager from extracting the private benefit of $4, as this is detrimental to shareholder value.[45] On the other hand, under standard asymmetric information assumptions, the Board cannot observe whether the Manager engages in private-benefit extraction and therefore cannot make the Manager's payoff schedule contingent on her actions. Otherwise, assuming symmetric information, the Board could simply write a state-contingent contract under which it would pay the Manager $4 if she does not engage in private-benefit extraction.

In order to solve these problems and make the Manager's contract incentive compatible, the Board will need to provide the Manager an agency rent that awards the Manager part of the project's returns. Specifically, the efficient agency rent to be left to the Manager is the minimum amount of the project's returns that can satisfy both the Manager's participation constraint (*PC*) and incentive constraint (*IC*). Under the Manager's *PC*, the Manager needs to receive at least the value of her outside option ($5). Under the Manager's *IC*, the Manager must be at least as well off when she exerts effort as when she does not (that is, when she extracts the private benefit of $4).

In our setting, the Board can satisfy both these constraints by promising a percentage (%) of the project's returns to the Manager, so that the Manager's *IC* can be written as: $(0.8) \times [\% \times (\$1{,}000)] \geq (0.6) \times [\% \times (\$1{,}000)] + \$4$.[46] This condition implies that the Manager's *IC* is satisfied for $\% \geq 0.02$ (or a percentage of at least 2%); that is, by promising the Manager an agency rent of $20 in case of the project's success (i.e., $0.02 \times \$1{,}000$). This compensation schedule also satisfies the Manager's *PC* as $(0.8) \times (\$20) = \$16 \geq \$5$, therefore providing the Manager with incentives to exert optimal effort.[47]

B. Managerial Power Theory

1. Law.—Managerial power theory conceives of typical executive-compensation packages as part of, rather than a remedy to, the moral hazard problem arising in the public corporation. Conceptually, this theory—expounded in a series of studies by Lucian A. Bebchuk and Jesse M.

---

44. For simplicity, we assume here that no agency costs arise between the Board and the shareholders. As we explain in section (I)(B)(1) below, admitting or rejecting this assumption is at the center of the controversy between optimal contracting scholars and managerial power scholars.

45. Private-benefit extraction also reduces aggregate welfare, as $(0.8) \times (\$1{,}000) > (0.6) \times (\$1{,}000) + \$4$.

46. Under the assumption that the Manager's reservation utility is $5 as above, this compensation plan also automatically satisfies the Manager's participation constraint.

47. Note that while $20 is the actual agency rent the Manager will receive in case of success, $16 is the expected cost of the agency rent to the corporation (that is, the shareholders).

Fried[48]—relies on negating a central assumption of optimal contracting theory. The negated assumption is that the board is able to act as a faithful guardian of shareholder interests and, hence, to take an adversarial position against management in negotiating efficient compensation arrangements.[49] Instead, managerial power theorists argue, the delegation of "decision control" from shareholders to directors results in its own set of agency costs, which largely mirror those arising between shareholders and managers.[50]

Further, neither competitive forces outside the firm nor corporate law rules could mitigate directorial moral hazard costs. First, managerial power scholars argue, boards are largely immune from the forces arising from product- or labor-market competition as well as insufficiently interested in any financial benefits from M&A deals.[51] Therefore, market forces can impose at best light constraints on executive compensation.[52] Second, fiduciary rules provide only weak deterrence against directorial moral hazard especially in the executive-compensation context, due to both the traditional reluctance of courts to intervene in compensation matters and excessively high enforcement costs.[53] Third, and most importantly, the shareholders' power of removing directors (and their appointed officers, the managers) is "largely a myth"[54] as a result of managerial entrenchment and board capture.

---

48. *See generally* BEBCHUK & FRIED, *supra* note 7; Lucian Arye Bebchuk & Jesse M. Fried, *Executive Compensation as an Agency Problem*, J. ECON. PERSP., Summer 2003, at 71 [hereinafter Bebchuk & Fried, *Executive Compensation*]; Lucian Arye Bebchuk, Jesse M. Fried & David I. Walker, *Managerial Power and Rent Extraction in the Design of Executive Compensation*, 69 U. CHI. L. REV. 751 (2002) [hereinafter Bebchuk et al., *Managerial Power*]. For an anticipation of the managerial power view in legal literature, see Charles M. Elson, *Director Compensation and the Management-Captured Board—The History of a Symptom and a Cure*, 50 SMU L. REV. 127 (1996). Similarly, for earlier economic work on managerial power theory, see generally Marianne Bertrand & Sendhil Mullainathan, *Are CEOs Rewarded for Luck? The Ones Without Principals Are*, 116 Q.J. ECON. 901 (2001).

49. *See* BEBCHUK & FRIED, *supra* note 7, at 4, 5, 23 (pointing to time and information constraints as well as financial, psychological, and social factors as undermining directors' ability to negotiate compensation arrangements with managers that are in the shareholders' interest).

50. Bebchuk & Fried, *Executive Compensation*, *supra* note 48, at 72–73.

51. *See* BEBCHUK & FRIED, *supra* note 7, at 53–58 (discussing the degree to which market forces bear upon CEO compensation arrangements).

52. *Id.* at 58.

53. *Id.* at 45–48.

54. BEBCHUK & FRIED, *supra* note 7, at 207. Bebchuk and Fried hint at these problems in the policy part of their book. *See id.* at 201–16 ("[T]he safety valve of potential ouster via the ballot box—on which our corporate governance system is supposed to rely—has been all but shut off."). However, in the past decade, it has been Lucian Bebchuk that has emerged as the leading voice among so-called shareholder advocates. These advocates defend a model with stronger shareholder rights as the necessary response to the managerial entrenchment problem. *See* Lucian Arye Bebchuk, *The Case for Increasing Shareholder Power*, 118 HARV. L. REV. 833, 851–75 (2005) (advocating for the expansion of shareholder governance rights); Bebchuk, *supra* note 6, at 694–711 (advocating for a reform of corporate elections to make directors more accountable to shareholders).

Under the managerial power view, entrenchment arises when directors gain protection from the threat of removal through the adoption of defensive measures that make it difficult for shareholders to replace incumbents.[55] Classic examples of such measures include poison pills and staggered boards.

A poison pill (also called a shareholders' rights plan) is a defensive measure that, when implemented, so dilutes a bidder's economic rights that the only way to complete a hostile takeover is to first appoint a new majority of directors who can remove the pill.[56] When a board is staggered, however, a bidder's ability to do so is reduced because directors are grouped into different classes—usually three—serving staggered terms where each class stands for reelection in successive years. This requires a prospective hostile bidder to endure the costly delay of waiting through two election cycles before being able to replace a majority of the board. As a result, the adoption of a staggered board, in addition to a poison pill, provides directors with a very potent defense against hostile takeovers.[57]

Likewise, for managerial power scholars, entrenchment is favored by incumbents' exclusive access to the corporation's proxy machinery, which raises prohibitive procedural costs for prospective challengers.[58]

By weakening the disciplining effect of the threat of removal, entrenchment would allow top executives—who control the flow of information from lower corporate layers to the board and, more importantly, the board-appointment process—to capture directors, making them subservient to management.[59] As a result, managerial power scholars argue

---

55. *See* Lucian Bebchuk et al., *What Matters in Corporate Governance?*, 22 REV. FIN. STUD. 783, 788 (2009). In this article, the authors developed an entrenchment index (or "E index") based on six measures of board protection, including staggered boards and poison pills, and claim that the empirical evidence supports the view that such measures are value reducing. *See id.* at 786. *But see* K.J. Martijn Cremers, Saura Masconale & Simone M. Sepe, *Commitment and Entrenchment in Corporate Governance*, 110 NW. U. L. REV. 727, 732, 761–71 (2016) (revisiting the evidence obtained on the E-Index and showing that defensive measures benefit shareholders as long as such measures provide for shareholder approval).

56. A poison pill consists of stock purchase rights that are granted to existing shareholders in the event a corporate raider accumulates more than a certain threshold of outstanding stock and that entitle the existing shareholders (but not the raider) to acquire newly issued stock at a substantial discount from the market price. *See* Memorandum from Wachtell, Lipton, Rosen & Katz, Share Purchase Rights Plan (Mar. 1994), *in* ROBERT B. THOMPSON, MERGERS AND ACQUISITIONS: LAW AND FINANCE 204 (2010) (setting forth the terms of a standard poison pill).

57. Lucian Arye Bebchuk et al., *The Powerful Antitakeover Force of Staggered Boards: Theory, Evidence, and Policy*, 54 STAN. L. REV. 887, 893, 899, 902 (2002). Bebchuk et al. employed an event study to examine the wealth effects of staggered boards, finding that staggered boards have a negative effect on shareholder returns. *See id.* at 891. *But see* Cremers & Sepe, *supra* note 16, at 90–91 (arguing that the results of Bebchuk et al. are endogenous).

58. *See* Bebchuk, *supra* note 6, at 688–94 (detailing these procedural costs).

59. *See* BEBCHUK & FRIED, *supra* note 7, at 27–39, 80–86 (describing the various sources of executives' power to influence directors). A variation on the account of directors as captured by managers is the one considering directors' "structural bias." As put by Claire Hill and Brett McDonnell, "[t]he term 'structural bias' is used in corporate law cases to refer to excessive

that directors are unable to engage in arm's-length compensation negotiations on the shareholders' behalf, while managers can exploit the power they hold over boards to extract excessive remuneration.[60]

For these scholars, the cost of excessive CEO pay also is substantial and reaps a significant portion of shareholder wealth.[61] Further, managers' ability to influence the executive-pay process results in weak or even perverse incentives. That is, managers are able to obtain compensation plans that are decoupled from firm performance or even promote results misreporting, the suppression of bad news, and the undertaking of projects that are not transparent.[62]

2. Economics.—In economic terms, the central claim underpinning managerial power theory is straightforward: board capture enables executives to extract "entrenchment rents," that is, returns above the executives' agency rents.[63] This account of executive compensation would explain why a substantial part of executive compensation is provided in the form of non-equity components, which are not, or only poorly, tied to performance—including cash salary, bonus plans, signing bonuses, split-dollar life insurance policies, and severance payments.[64]

Accordingly, a managerial power theorist would claim that the Board in our example above would not just grant the Manager an agency rent equal to 2% of the project's expected returns. Rather, the Manager would be able to influence the Board's decision-making process to obtain additional pay—for example, in the form of a fixed bonus of $10 to be paid on top of the 2% agency rent. This bonus would correspond to an entrenchment rent, providing the Manager with value beyond what is needed to preserve the Manager's incentives to exert effort and hence decreasing shareholder value.

Of course, this is just one trivial representation of the many ways through which managers can extract entrenchment rents according to

---

deference by the directors to the management because of interlocking relationships or because they travel in the same social and economic circles." Claire Hill & Brett McDonnell, *Executive Compensation and the Optimal Penumbra of Delaware Law*, 4 VA. BUS. L. REV. 333, 335 n.2 (2009). This variation, however, does not affect this Article's analysis of managerial power theory.

60. *See* BEBCHUK & FRIED, *supra* note 7, at 201–16.

61. Bebchuk & Fried, *Executive Compensation*, *supra* note 48, at 88.

62. BEBCHUK & FRIED, *supra* note 7, at 183–85.

63. *See id.* at 62. As we explain below, under the realistic assumption of a competitive setting, an executive's agency (or total) rents include both what we refer to as "pure" agency rents and market rents. *See infra* section I(C)(2).

64. *Id.* at 121–36. This is not to say that any non-equity component of a manager's pay package should be regarded as an entrenchment rent. In general, optimal pay packages need to include both some fixed components and some equity components. *See* Sepe, *supra* note 17, at 217–19 (showing that equity components are desirable to promote effort, while fixed components help constrain excessive risk-taking). For managerial power scholars, however, fixed compensation is inherently less likely to promote effort and, hence, more likely to pay the manager entrenchment rents that are not tied to performance. BEBCHUK & FRIED, *supra* note 7, at 63.

managerial power scholars. For example, another way managers could extract rents is by bargaining for equity-based compensation in the form of restricted stock rather than stock options.[65] Both restricted stock and stock options tie managerial pay to stock-price performance and shareholder value. Restricted stock, however, does so in a linear way, moving dollar for dollar with the firm's share price. Therefore, the gains or losses of a manager that is compensated with restricted stock are the same as those of shareholders. Conversely, stock options (which are technically call options) deliver managers asymmetric payoffs. A call option is only valuable to managers if the share price at the option's exercise date is higher than the option's strike price. This makes managers highly sensitive to even small changes in share value above the option's strike price, while they are relatively insensitive to changes in share value below the strike price. It follows that option-based pay produces more high-powered incentives per dollar of compensation expenses than stock-based pay and, therefore, is seen as a less costly way of paying out necessary agency rents.[66]

A further modification of our example is useful to clarify this point. Recall that the Board's essential problem in designing the Manager's contract is that it cannot observe whether the Manager exerts effort or extracts a private benefit of $4. In order to stylize the use of restricted stock versus options, we assume that at the time the Board sets the Manager's compensation plan, each company share is worth $1. We further assume that if the Manager exerts effort, within a year the share value will increase to $1.2 with probability 80%, while it will drop to $0.8 with probability 20%. Under these assumptions, the expected value of the company's share price equals $1.12. Conversely, if the Manager does not exert effort, within a year the share value will increase to $1.2 with probability 30%, while it will drop to $0.8 with probability 70%. With the lack of managerial effort, the expected value of each share thus decreases to $0.92.

Let's now assume that the Board decides to pay the Manager's agency rent using restricted stock, granting a number of shares (# *Shares*) that can satisfy the Manager's incentive constraint (again, *IC*). Assuming, for simplicity, that the Manager is risk-neutral and does not discount future income (such that she is indifferent between receiving $4 now versus later), the Manager's *IC* can be written as: $[(\#\ Shares) \times (\$1.12)] - \$4 \geq [(\#\ Shares) \times (\$0.92)]$. Therefore, the incentives to be left to the Manager under a

65. *See* BEBCHUK & FRIED, *supra* note 7, at 170–73 ("The increased use of restricted stock is generally viewed as a response to shareholder concern about conventional options . . . . [But] firms replacing conventional options with restricted stock are ending up with an equity-based incentive plan that contains an even larger windfall element.").

66. *See id.* at 190 (arguing that with opportune "filtering, the same amount of incentives [that is provided, for example, through restricted stock] can be provided at a lower cost, or more incentives can be provided at the same cost").

restricted stock plan are equal to 20 ($4/0.2) shares, which translates into an expected cost for the shareholders equal to 20 × $1.12 = $22.4.[67]

On the contrary, should the Board decide to pay the Manager's agency rent through call options with a $1 strike price, the option would give the Manager the following payoff: Max[*Stock Value at Exercise* – $1, 0]. This means that the Manager will only exercise her option when *Stock Value at Exercise* exceeds $1, which materializes with probability 80% when the Manager exerts effort and with probability 30% when she does not exert it. Therefore, the option's expected value to the Manager equals 80% × [$1.2 – $1] = $0.16 when the Manager exerts effort, and 30% × [$1.2 – $1] = $0.06 when she does not. The Manager's *IC* under the option plan can thus be written as: (*# Options worth* $0.16) ≥ (*# Options worth* $0.06) + $4.[68] Consequently, the Manager needs to receive at least 40 options to exert effort ($4/0.1 = 40), with an expected cost to the company of ($0.16) × (40) = $6.4.

For managerial power theorists, then, the difference of $16 between restricted stock costing $22.4 and options costing $6.4 would be the entrenchment rent captured by the Manager.[69]

C. *Managerial Talent Theory*

1. Law.—What we refer to as the "managerial talent" theory of executive compensation[70] broadens the optimal contracting approach to account for additional dimensions, including market forces operating in a competitive equilibrium.[71] Thus, while for expositional convenience we treat these theories separately, managerial talent studies could also be regarded as part of optimal contracting theory.[72] Indeed, from a legal perspective, the assumptions underpinning optimal contracting theory continue to remain valid for managerial talent theory, which also conceives of compensation as a matter efficiently delegated to the central authority of the board. As explained below, however, managerial talent theory considers optimal board contracting in the executive-compensation process as not just resulting from

---

67. Recall that $1.12 is the expected share value when the Manager exerts effort.

68. In agency models, when the incentive constraint is "binding" (that is, the values on the two sides of the equation are the same), it is standard to assume that an agent will behave. *See* TIROLE, *supra* note 20, at 116–17.

69. *See supra* note 63 (dissecting the meaning of "entrenchment rents").

70. *Cf.* Edmans & Gabaix, *supra* note 14, at 1233 (using the term "shareholder value" theory).

71. *See* Xavier Gabaix & Augustin Landier, *Why Has CEO Pay Increased So Much?*, 123 Q.J. ECON. 49, 50 (2008) (building an equilibrium model in which the marginal product of managerial ability increases with firm size); Kevin J. Murphy & Ján Zábojník, *CEO Pay and Appointments: A Market-Based Explanation for Recent Trends*, AM. ECON. REV., May 2004, at 192, 192–93 (developing a general equilibrium model that explains the increase in executive pay as a result of the increased prevalence of general, transferrable executive skills).

72. *See* Murphy, *supra* note 3, at 214 (using the term "efficient contracting" theory to define a view of executive compensation that blends optimal contracting with managerial talent theory).

bargaining between boards and shareholders, but as influenced by exogenous market forces as well.

2. Economics.—Economically, managerial talent studies employ competitive equilibrium models, maintaining that the models of bilateral contracting employed in optimal contracting theory cannot capture the complexities of actual CEO-employment relationships.[73] These complexities arise from the fact that the incentives to be provided to executives are not determined "in a vacuum," as if boards and CEOs were insulated from market competition. Rather, such incentives reflect a competitive equilibrium in the market for scarce managerial talent, where "it may be optimal to pay high wages to attract talented CEOs, and implement high effort from them even though doing so requires paying a premium."[74] Under this view, high executive pay would thus reflect the high agency rents to be paid to preserve managerial incentives in a competitive labor market with a limited supply of managerial talent rather than entrenchment rents. For expositional clarity, we refer to such high but efficient executive rents as "market rents" to distinguish compensation premiums driven by competitive market forces from "pure" agency rents.[75]

A further modification of the basic example introduced above can serve to clarify the concept of market rent. As a stylized representation of a competitive market context, we add an additional corporation (Corporation II), referring to the board of our initial corporation (Corporation I) as Board I and the board of this additional corporation as Board II. As is standard in competitive labor-market models, we assume that there is scarcity of managerial talent and that our Manager has rare talent, such that both Board I and Board II are interested in hiring the Manager. Specifically, we assume that Board II is willing to pay the Manager a compensation package of $40.

Under these circumstances, the original expected compensation of $16 no longer suffices to satisfy the Manager's participation constraint,[76] which has now increased to $40 because of Board II's employment offer. Therefore, in order for Board I to retain the Manager, Board I will need to pay the Manager not just the pure agency rent of $(0.8) \times (\$20) = \$16$, but also a market rent of $24 to match the offer of Board II.[77] Importantly, under the

---

73. *See* Edmans & Gabaix, *supra* note 14, at 1233 (detailing the flaws in the models of bilateral contracting as applied to CEO-employment relationships).

74. *Id.*

75. It is worth emphasizing that agency rents are driven by the Manager's incentive constraint, while market rents are driven by competitive forces that impact the Manager's participation constraint.

76. *See supra* note 47.

77. Note that the form this market rent takes (e.g., salary, bonus, or contingent compensation) is indifferent as long as the Manager expects to receive additional compensation for $24.

assumption of scarce managerial talent, it is reasonable to pose that should Board I not pay the Manager this market rent, the Manager would join Corporation II, while Board I would be forced to hire a less talented manager with lower productivity, resulting in a loss in shareholder wealth.[78]

II. Towards a Dynamic Approach

A. Unanswered Questions

Part I has outlined the essential elements of the main theories of executive compensation. It has not, however, surveyed the empirical literature that examined whether the different predictions generated by each theory are confirmed by the data because space constraints would force us to omit many important contributions. We thus refer the reader to the excellent surveys by John Core, Wayne Guay, and David Larcker,[79] and Carola Frydman and Dirk Jenter.[80]

As documented by these surveys, the empirical literature on executive compensation has generally yielded mixed results, failing to identify a consistent relationship between one set of theoretical predictions and the data.[81] For example, while board capture constitutes an essential element of managerial power theory, the available empirical evidence seems not to support the existence of systematically passive and captured boards.[82] On the contrary, indicators of board independence have steadily improved since the

78. For example, assume that with a less talented manager, Manager II, Corporation I's project would generate revenues of $1,000 with probability 70% and zero revenues with probability 30%. Like Manager I, Manager II can also extract a private benefit of $4, in which case the probability of the project's success drops to 50%. Under the assumption that Board I will offer Manager II the equilibrium agency rent of $20, Manager II will then avoid private-benefit extraction, but deliver a lower outcome than Manager I as she is less talented: (0.7) × ($1,000) instead of (0.8) × ($1,000). Consequently, if Board I refused to pay Manager I the market rent of $24, this would result in an expected revenue loss of $100 for the shareholders, which makes the payment of the market rent the only rational choice for Board I.

79. *See* John E. Core, Wayne R. Guay & David F. Larcker, *Executive Equity Compensation and Incentives: A Survey*, FED. RES. BANK N.Y. ECON. POL'Y REV., Apr. 2003, at 27, 27–28 (focusing, in particular, on the literature on equity-based compensation).

80. *See generally* Carola Frydman & Dirk Jenter, *CEO Compensation*, 2 ANN. REV. FIN. ECON. 75 (2010).

81. *See* Core, Guay & Larcker, *supra* note 79, at 34–35, 44 (discussing the mixed evidence on executive compensation and performance); Frydman & Jenter, *supra* note 80, at 76 (noting the inconsistencies between executive-compensation theories and the existing empirical evidence).

82. *See* Benjamin E. Hermalin, *Trends in Corporate Governance*, 60 J. FIN. 2351, 2351, 2376 (2005) (analyzing the potential consequences of a possible trend towards "more diligent boards of directors"); Bengt Holmström & Steven N. Kaplan, *The State of U.S. Corporate Governance: What's Right and What's Wrong?*, J. APPLIED CORP. FIN., Spring 2003, at 8, 15 (commenting favorably on certain improvements made to board governance while cautioning that further improvements remain to be made).

1980s,[83] and CEO turnover rates have also increased,[84] thus suggesting that boards do not refrain from disciplining poorly performing executives. Conversely, the data generally confirm that the CEO labor market has become increasingly more competitive[85]—consistent with the assumptions of managerial talent theory. Yet, it remains empirically uncertain whether growing competitiveness in the managerial labor market can fully explain the increase in executive-pay levels and the changes in the structure of executive pay that have occurred over the past decades.

As a matter of fact, however, managerial power theory has largely gained the upper hand in recent years, proving very influential in both academic and political circles and leading to major regulatory changes, especially in the United States. In 2006, for example, the Securities and Exchange Commission (SEC) mandated increased disclosure of compensation policies, requiring, among other things, that companies specifically address the role of executives in the pay-setting process.[86] Further, and explicitly taking into account the concern of managerial power scholars, the SEC also required a more detailed articulation of compensation items that might potentially constitute a perquisite.[87] Most notably, in 2010, the Dodd-Frank Wall Street Reform and Consumer Protection Act introduced a series of new executive-compensation rules affecting all public companies, including rules designed to bolster the independence of compensation committees and mandating that companies introduce nonbinding shareholder votes on executive compensation.[88]

These reforms share one common feature: they all embrace the call for shareholder empowerment that managerial power scholars have long

---

83. *See* Jeffrey N. Gordon, *The Rise of Independent Directors in the United States, 1950-2005: Of Shareholder Value and Stock Market Prices*, 59 STAN. L. REV. 1465, 1473–76 (2007) (discussing the changes in board compositions, specifically the percentage of inside directors, over the past several decades).

84. Mark R. Huson et al., *Internal Monitoring Mechanisms and CEO Turnover: A Long-Term Perspective*, 56 J. FIN. 2265, 2266 (2001).

85. *See* K.J. Martijn Cremers & Yaniv Grinstein, *Does the Market for CEO Talent Explain Controversial CEO Pay Practices?*, 18 REV. FIN. 921, 923–24 (2014) (examining the role of labor-market competition in compensation patterns for CEOs); Kevin J. Murphy & Ján Zábojník, Managerial Capital and the Market for CEOs 31 (Apr. 2007) (unpublished manuscript), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=984376 [https://perma.cc/BNW8-L6XH]:

> In our theory, the level of CEO pay is determined by competition among firms for top performing managers, and depends upon the portion of the CEOs' skills that is transferable across firms and industries. We suggest that the increase in executive compensation can be explained by an increase in the importance of general managerial skills, as opposed to firm-specific knowledge, in managing the modern corporation.

86. Executive Compensation and Related Person Disclosure, Exchange Act Release Nos. 33–8732A, 34–54302A, IC–27444A, 71 Fed. Reg. 53,158, 53,165 (Sept. 8, 2006), https://www.gpo.gov/fdsys/pkg/FR-2006-09-08/pdf/06-6968.pdf [https://perma.cc/7Z3B-PJEJ].

87. *See id.* at 53,176.

88. Dodd-Frank Wall Street Reform and Consumer Protection Act, Pub. L. No. 111-203, § 951(a)(1), 124 Stat. 1376, 1899 (2010) (codified as amended at 15 U.S.C. § 78n-1).

defended as the necessary remedy to what they see as the widespread problem of excessive executive pay.[89] Underpinning that call is these scholars' belief that only by strengthening the governance rights of shareholders can directors be trusted to faithfully serve shareholder interests, rather than be captured by opportunistic managers. To this end, in addition to favoring an enhanced role of shareholders in the pay-setting process, managerial power scholars defend the need to remove the barriers that insulate directors from shareholder discipline, such as defensive measures.[90]

This claim notwithstanding, recent work on the value impact of defensive measures, including our own research, does not support the view that stronger shareholder rights are an all-purpose remedy in corporate governance.[91] In particular, two of us have showed elsewhere that the adoption of a staggered board is associated, on average, with increased firm value, while not being associated with a lower CEO turnover.[92] This evidence poses a challenge to managerial power theory. Under this theory, one would expect to find, on the one hand, that the adoption of takeover defenses decreases firm value. On the other hand, such measures should be associated with a lower CEO turnover, as managerial power scholars argue that insulated boards are more likely to be captured by executives and hence less likely to fire them in case of poor firm performance.[93] In contrast, both these theoretical predictions are rejected by the data.

More generally, this evidence raises several questions about the explanatory power of managerial power theory. First, if boards that are protected from short-term shareholder interference seem better placed to promote long-term shareholder wealth, how can board protection be the main source of allegedly inefficient executive pay, as argued by managerial power scholars? If, on the other hand, board protection does not hamper the executive-pay process, perhaps it can serve a positive function within this process? Further, is there any link between the results we obtain on board protection and managerial talent theory? The attempt to answer these questions provides the motivation for the analysis we develop in this Article. Our interest lies in these questions' theoretical implications, which we discuss in this Part, as well as their empirical implications, which we explore in Parts III and IV below.

Theoretically, these questions point to a complex relationship between corporate governance, market forces, and executive compensation. In

89. *See supra* text accompanying notes 50–51.

90. *See supra* text accompanying notes 55–58.

91. Cremers, Masconale & Sepe, *supra* note 55, at 749–51.

92. Cremers, Litov & Sepe, *supra* note 16, at 423–24, 433.

93. *See* BEBCHUK & FRIED, *supra* note 7, at 88–89 (asserting that managers' influence over the board contributes to a disinclination to terminate managers and to "the practice of gratuitous goodbye payments" in cases where a manager is fired).

analyzing this complexity, we draw on some critical analysis that Bengt Holmström, 2016 Nobel laureate in economics, offered on incentive design about a decade ago.[94] First, Holmström observed that the right economic framework for studying executive compensation is that of "[d]ynamic models, where commitment problems and implicit incentives arise out of incomplete contracting and renegotiation."[95] This remark emphasizes that the nature of the relationship between managers, shareholders, and boards is not that of a static, one-shot transaction as envisioned within managerial power theory. Rather, such a relationship tends to develop along a multi-period horizon, with managers typically holding their positions for several investment periods.[96] As we argue in subpart B below, taking into account the dynamic nature of executive-compensation contracts sheds new light on the relationship existing between agency rents, entrenchment rents, and market rents.

Second, and relatedly, Holmström argued that the claim "that shareholders know what is best for them breaks down in a world where commitment to an ex post inefficient course of action is valuable ex ante."[97] As explained in subpart C, this analytical framework suggests that attempts at reducing moral hazard costs *ex post* (that is, *after* a manager's hire) by strengthening shareholder governance rights might impair a board's ability to design efficient incentive schemes *ex ante* (that is, *upon* a manager's hire), in turn calling for a novel evaluation of the respective costs and benefits of shareholder empowerment and board protection.

B. Dynamic Compensation Contracts

Convenience and tractability issues explain why scholars often prefer static settings to dynamic ones. The study of executive compensation is no exception. After all, so the common argument goes, as long as using a more tractable single-period setting does not change the general conclusions of the analysis, there is no need to overly complicate executive-compensation discussions.[98] However, the more recent economic literature on executive compensation emphasizes that taking into account the complexities of dynamic, strategic, and repeated interactions between executives, firms, and markets may lead to significantly different conclusions about whether executive-compensation theories can accurately predict observed

94. *See generally* Holmström, *supra* note 13.

95. *Id.* at 708.

96. The average tenure of U.S. CEOs is around seven years. *See* Steven N. Kaplan & Bernadette A. Minton, *How Has CEO Turnover Changed?*, 12 INT'L REV. FIN. 57, 58 (2012) (studying the period from 1992 to 2007).

97. Holmström, *supra* note 13, at 713.

98. *See* Bebchuk & Spamann, *supra* note 13, at 256 (assuming only present and future periods for the sake of simplicity because "[w]ith multiple periods, the analysis would become more complex, but our general conclusions would not change").

compensation practices.[99] In other words, the risk is that static models may be not just simple, but ultimately simplistic.

Drawing on this recent literature, we argue that once the relationship between managers, boards, and shareholders is more realistically represented as taking place in a dynamic and competitive setting, several important implications follow—regarding both managerial incentives and the opportunities of boards and shareholders to design efficient compensation schemes.

1. Dynamic Incentives.—On the incentives side, a dynamic setting can incorporate essential features of executive-compensation contracts as they are negotiated and implemented in the actual corporate world. First, real-world settings show that the relationship between incentives, rents, and competition can be more complex than described in Part I. We can again use the example we introduced above to better illustrate this point. For simplicity, in that example we considered only a single period, where at the beginning of the period, in "the present," the board implements the compensation schedule and the manager makes investment decisions, and at the end of the period, in "the future," gains or losses are realized and the manager gets paid. In this static setting, competitive market forces only operate in "the present," causing an increase in the manager's reservation utility and, accordingly, requiring that the manager be paid market rents, in addition to agency rents.[100] However, once one allows for multiple investment periods and hence the availability of interim information based on the realization of interim payoffs, increased competition for managerial talent produces additional complexities.

On the one hand, shareholders will be interested in both the manager's current contribution and those arising from her continued employment. On the other, initially optimal incentives may lose power over time. Indeed, if interim payoffs are positive, the manager's reservation utility will likely increase over subsequent periods, as positive interim payoffs can be expected to make the manager more appealing to competitors.[101] Accordingly, competition may produce a negative externality, rendering the initial compensation contract no longer apt to satisfy the manager's participation constraint in future periods.

---

99. *See* Edmans & Gabaix, *supra* note 14, at 1234, 1260 (explaining that a dynamic multi-period model adds certain challenges but also provides advantages unavailable in a single-period model).

100. *See supra* section I(C)(2).

101. The realization of positive interim payoffs provides "hard" information that is easily incorporated into the stock price and that, therefore, investors can easily verify. *See* TIROLE, *supra* note 20, at 249–50 (observing that "hard" information, which "can be verified by the investors once disclosed by the issuer," reduces asymmetry of information between issuers and investors).

In the context of our example, this means that paying the Manager a market rent of $24 in addition to an agency rent of $16 might not be sufficient to provide the Manager with long-term incentives to exert optimal effort in Corporation I.[102] This is because a successful Manager might have additional employment options available after the positive realization of interim payoffs, with such offers likely exceeding the current Manager's total rent of $40.

Further, a dynamic setting also accounts for a manager's ability to engage in inefficient intertemporal tradeoffs, which potentially exposes shareholders to an additional agency problem: short-termism (or managerial myopia).[103] In a static setting, the main concern for shareholders is that a manager may fail to exert optimal effort. The provision of equity incentives, which tie CEO pay directly to shareholder wealth, is the standard solution to this problem.[104] In a dynamic setting, however, a manager compensated through equity-based pay may develop incentives for investments that boost short-term returns (and thus a manager's pay), at the expense of lower gains or even losses occurring in the longer-term and which, therefore, the manager discounts.[105]

Excessive risk-taking provides a classic example of short-termism, as it became evident during the 2007–2009 financial crisis, when the huge underlying risks associated with U.S. banks' investments in the highly remunerative subprime market finally materialized.[106] Another classic example includes cutting specific investments—for example, investments in R&D or employee training—that would pay off later on. Importantly, increased competition may exacerbate short-termist incentives. Indeed, a manager who can depart to a new employer before any long-term loss materializes will be naturally more inclined to engage in short-termist strategies. In this environment, improving short-term performance at the expense of long-term value will not only boost a manager's current pay, but also her chances at upward mobility, at once increasing the value of the manager's outside option and the likelihood of sidestepping losses arising in the future.[107]

---

102. *See supra* section I(C)(2).

103. *See supra* note 18.

104. *See supra* notes 39–40 and accompanying text.

105. A dynamic setting necessarily modifies the assumption made above about no managerial discounting. *See supra* section I(B)(2).

106. *See* Simone M. Sepe, *Regulating Risk and Governance in Banks: A Contractarian Perspective*, 62 EMORY L.J. 327, 338–46 (2012) (discussing risk-taking incentives in the banking sector).

107. Simone M. Sepe & Charles K. Whitehead, *Paying for Risk: Bankers, Compensation, and Competition*, 100 CORNELL L. REV. 655, 659–60, 668–74 (2015).

2. Dynamic Pay and Committed Managers.—In addition to the above challenges, a dynamic setting simultaneously offers additional avenues for efficient incentive design, as it allows boards to spread the rewards for good managerial performance over time while periodically reviewing the manager's performance. In other words, in a dynamic setting, the board can "exploit" a manager's continuation value—that is, a manager's expected payoffs from future employment periods (assuming that the manager is not fired for poor performance)—as a powerful bonding mechanism to commit the manager to the creation of long-term firm value.

This dynamic approach to executive-pay design sheds new light on the function served by allegedly inefficient entrenchment rents—that is, compensation premiums that exceed a manager's total rent, as given by the sum of agency rents and market rents. Contrary to the assumptions of managerial power scholars, this approach shows that such rents can help to commit executives to the exercise of long-term effort.

Consider, for example, the use of fixed compensation, which managerial power scholars describe as the quintessential form of "pay-without-performance" and, hence, as primary evidence of inefficient rent extraction.[108] In contrast with this view, in a multi-period setting, fixed compensation can provide efficient incentives, as long as the manager's expected gains from future employment periods offset her disutility cost of effort. A rational manager anticipates that low effort increases the likelihood of poor interim performance[109] and, therefore, that she might be removed in the near future. Because managerial removal triggers the loss of continuation payoffs, the manager will then have incentives to exert effort.[110]

Fixed compensation may likewise provide a mechanism to mitigate the negative externalities that competition may introduce in incentive schemes. Consider again our example. Recall that in a competitive, dynamic setting, paying the Manager the total rent of $40 (that is, the $16 agency rent plus the $24 market rent) is no longer sufficient to satisfy the participation constraint of a talented Manager over time. In this context, paying the Manager more than the $40 total rent—for example, by granting the Manager an additional fixed bonus of $10 to be paid out each period the Manager continues to be employed—should thus be regarded as an efficient, rather than inefficient, rent. Under this compensation arrangement, the Manager will trade off any potential future increase in her reservation utility against her expected continuation payoffs, including expected future bonuses. It follows that providing the Manager with extra compensation may help promote the

108. *See supra* section I(B)(2).

109. *See supra* note 19.

110. *See* Sepe, *supra* note 17, at 219–23 (providing a formal model that explores the incentive function of fixed-compensation contracts in a dynamic setting).

Manager's long-term commitment to Corporation I and, ultimately, long-term shareholder wealth.

Similar considerations apply to the use of restricted stock instead of stock options to compensate executives. As discussed above, managerial power theorists are critical of restricted stock, arguing that providing incentives through this compensation component is more costly to shareholders than doing so through options.[111] When considered from a dynamic perspective, however, the use of restricted stock can serve a positive function, as the higher expected value that restricted stock provides to managers—relative to option grants—might help secure the long-term commitment of talented managers. Thus in our example above when the Manager is compensated through restricted stock she expects to receive a value of $22.4, which is considerably higher than the expected value of $6.4 the Manager receives under an option plan.[112] What managerial power scholars generally do not consider, however, is that in a dynamic, competitive setting, the higher value accruing to the Manager increases the likelihood that a successful Manager will stick to Corporation I over time.

Designing executive pay dynamically can similarly help mitigate myopic managerial incentives, as long as a manager's pay package is designed to ensure that the manager's expected payoffs from future employment periods offset the gains she might obtain out of a short-termist strategy today. Our example is again helpful to clarify this point. As a stylized representation of a dynamic setting, we assume that after "the present," there are two periods, where in each period $1,000 can be generated with some probability. For simplicity, we pose that the Manager cannot engage in private-benefit extraction in this modified setting, so that, under the assumption that the Manager will exert effort, the project she undertakes has an 80% probability of success in each period.

Here, however, the Manager can make a myopic choice, meaning that she can boost the payoff from the first investment period at the expense of the payoff from the second period. In particular, assume that the Manager can choose an investment strategy (e.g., cutting long-term investments in R&D) that increases the probability of a successful first-period realization from 80% to 85%, but simultaneously reduces the likelihood of a successful second-period realization from 80% to 60%. In our example, this choice is clearly inefficient, as it decreases the shareholders' expected returns. Nevertheless, for the Manager, such behavior is the most profitable choice when she does not expect to receive a continuation payoff. Assuming market rents away—for simplicity, but with no loss of generality—when the Manager only expects to receive her agency rent, she can increase her expected first-period

111. *See supra* text accompanying notes 65–69.

112. *See supra* text accompanying note 69.

compensation from (0.8) × ($20) = $16[113] to (0.85) × ($20) = $17 by behaving myopically. Further, a successful first-period performance will also enhance the Manager's chances for improved outside mobility and, therefore, for sidestepping the longer-term losses arising from a short-termist choice.

Conversely, if the Manager expects to receive a continuation value in the form of an additional rent of $10 per period, the Manager will trade off the short-term increment from the myopic strategy against the higher likelihood of losing her continuation value. In the case of myopic managerial behavior, the Manager reduces the likelihood of receiving the second-period rent of $10 from 80% to 60%. By undertaking a myopic strategy, the Manager can thus expect to lose $2 in the second period, which is more than what she gets out of the myopic strategy in the first period. Accordingly, a rational Manager will not engage in myopic behavior.

Lastly, it is important to emphasize that any form of compensation above a manager's total rent can be used to provide managers with efficient incentives to mitigate the risk of managerial myopia. In particular, results analogous to those described above for the use of fixed compensation can be achieved by using restricted stock instead of stock options. As is extensively discussed in the executive-compensation literature, especially after the recent financial crisis, the asymmetric payoff of options may give managers greater incentives to opportunistically exploit intertemporal tradeoffs—particularly in the form of excessive risk-taking.[114] This is because the sensitivity of option-compensated managers to even small changes in the share price around a call option's strike price tends to exacerbate a manager's incentives to undertake strategies that boost short-term stock prices.[115] Conversely, the linear payoff of restricted stock, combined with the instrument's vesting-period restrictions, anchor a manager's payoff to long-term shareholder value, therefore mitigating managers' incentives for short-termism.

C. Commitment, Competition, and Shareholder Power

The above discussion has shown that under the realistic assumption of a dynamic setting with competition for scarce managerial talent, boards can design a manager's expected rents from future employment periods to

---

113. *See supra* text accompanying note 47.

114. For an economic account of the distorted risk-incentives that managers compensated with equity-based pay may develop, see, for example, Patrick Bolton et al., *Executive Compensation and Short-Termist Behaviour in Speculative Markets*, 73 REV. ECON. STUD. 577, 577–78 (2006) and Ing-Haw Cheng et al., *Yesterday's Heroes: Compensation and Creative Risk-Taking* 2 (Nat'l Bureau of Econ. Research, Working Paper No. 16176, 2010), http://www.nber.org/papers/w16176.pdf [https://perma.cc/Y4YN-37J5]. For a legal account, see, for example, Bebchuk & Spamann, *supra* note 13, at 257–58 and Sanjai Bhagat & Roberta Romano, Essay, *Reforming Executive Compensation: Focusing and Committing to the Long-Term*, 26 YALE J. REG. 359, 363 (2009).

115. *See, e.g.*, TIROLE, *supra* note 20, at 23–24.

commit talented managers to the long-term exercise of effort, as well as to mitigate the risk of managerial myopia. Nonetheless, a managerial power theorist could object that this conclusion rests on the wrong conceptualization of the shareholders' power of removal. Indeed, the design of dynamic compensation contracts rests on a board's ability to both reward successful managers and remove poorly performing managers. According to managerial power scholars, however, in the real corporate world, the use of defensive measures and other barriers undermines the shareholders' own power of removing directors, which would produce entrenched boards beholden to top executives.[116] Consequently, the assumption that boards would fire an underperforming manager is misplaced. Thus, even under a dynamic approach one should conclude that efficient executive pay requires shareholder empowerment, as only in this way could directors and managers be held accountable for poor firm performance.

The above argument, however, neglects to consider that the creation of long-term firm value presupposes a bilateral commitment, from *both* managers *and* shareholders. This requires, on the one hand, providing managers with rents that make it in a manager's self-interest to exert long-term effort. On the other hand, it requires overcoming what we have referred to in prior work as the shareholders' "limited commitment problem."[117]

This problem arises because, upon the realization of a disappointing firm outcome (that is, low short-term earnings), shareholders always have the option to challenge the board and its appointed officials, the managers. They can do so in several ways—for example, supporting activist hedge funds, submitting shareholder proposals,[118] voting against management, or selling their shares in a hostile takeover attempt. However, under the conditions of informational asymmetry existing in the real corporate world, a disappointing firm outcome might be only temporary and reflect an investment whose value will materialize only later on. Put differently, the assumption that an opportunistic manager is generally more likely to be associated with disappointing firm outcomes in the short term[119] breaks down once one considers that directors and managers have private information that cannot easily be shared with outside shareholders. Therefore, as market prices may

116. *See supra* notes 55–60 and accompanying text.

117. Cremers & Sepe, *supra* note 16, at 73 & n.30, 114–15.

118. Under Rule 14a-8 of the Securities Exchange Act of 1934, 15 U.S.C. §§ 78a–78pp (2012), a public company is required to include a shareholder proposal (and related supporting statements) in its proxy statement and allow shareholders to vote on the proposal unless either the shareholders have not complied with eligibility or procedural requirements or some other named exception applies. *See* 17 C.F.R. § 240.14a-8 (2016).

119. The theoretical underpinning here is the semi-strong form of the Efficient Capital Market Hypothesis (ECMH), under which market prices accurately reflect all available public information on firm outcomes. *See* Eugene F. Fama, *Efficient Capital Markets: A Review of Theory and Empirical Work*, 25 J. FIN. 383, 383 (1970).

fail to capture the implications of private information until those implications begin to show up in cash flows over time, shareholders would benefit from supporting, rather than challenging, board actions in the short term.[120]

This analytical framework suggests that the benefits of shareholder empowerment—and, more generally, increased market discipline—need to be evaluated alongside their costs, while also indicating that temporary board protection might serve a positive governance function within the executive-pay process. When shareholders are empowered to intervene in corporate affairs "at all times"—as advocated by managerial power scholars—they are unable to provide a credible *ex ante* commitment *not* to challenge the board and managers in the short term. This problem is likely to be exacerbated in firms operating in markets with intense product competition. This is because in a context where managerial performance is evaluated in relative terms over fairly short periods of time and under the imperative of "beating competitors," shareholders might become even more impatient. In turn, they will be even less willing to support current board and managerial actions when faced with disappointing short-term firm performance, regardless of whether this outcome might be only temporary.

In an environment with increased shareholder power or intense product-market competition, a rational manager will accordingly anticipate a much higher risk of being removed in the near future. Under this risk, the manager will substantially discount her expected continuation payoff. The manager will also be more likely to develop short-termist incentives, as positive short-term results can be expected to reduce the risk of future managerial removal. As a result, a manager's expected continuation value might no longer be sufficient to offset her current cost of effort, or her future reservation utility increases, or, still, near-term gains from myopic investments—all circumstances that jeopardize a board's ability to design effective pay schemes under the dynamic approach described above.

By limiting the impact of short-term shareholder and market pressure on board decision-making, temporary board protection may help preserve a manager's continuation value and thus the ability of boards to design efficient pay schemes. When a board is protected from short-term shareholder interference and, more broadly, market pressure, managers will rationally

120. Market prices are especially likely to be uninformative in the case of firm-specific investments—such as investments in innovation or other intangible assets—as information about the fundamental value of these investments tends to be "soft," that is, nonverifiable by outsiders even *if* insiders share their views with them. TIROLE, *supra* note 20, at 249–50. At the same time, channeling resources to such investments tends to require large capital expenditures up front and, hence, to decrease earnings in the short term. This decrease in present earnings is a type of "hard" information, so that decreased earnings will tend to lead to lower short-term stock prices. *E.g.*, Alex Edmans et al., *The Real Costs of Financial Efficiency when Some Information Is Soft*, 20 REV. FIN. 2151, 2152–53, 2158 (2016). As a result, shareholders may take the fall in short-term stock prices following the undertaking of a profitable long-term project to signal managerial underperformance and rationally, although mistakenly, decide to remove the manager.

anticipate a lower risk of being removed in the short term. It follows that it will be easier for boards to efficiently design a manager's continuation payoff to promote long-term managerial effort.[121]

This does not mean that board protection should be perpetual. On the one hand, market prices can be expected to better incorporate uncertainty about current projects and a board's strategy over time, as the implications of directorial decisions materialize into actual cash flows. On the other hand, perpetual board protection would make the threat of removing underperforming managers no longer credible, which would likewise sabotage incentive schemes. However, in contrast to what is argued by managerial power scholars, most defensive measures are not perpetual. A staggered board, for example, does not permanently limit market discipline, but rather provides a longer time frame for shareholder evaluation of board and managerial performance. This is consistent with our conclusion that temporary board protection might increase the efficiency of compensation plans by ensuring that shareholder and market discipline take place periodically, rather than "at all times."

III. Organizations, Markets, and CEO Pay

In this Part, we seek to incorporate the above theoretical insights in the empirical investigation of the relationship between corporate governance features—and in particular board defenses—and competitive market forces, on the one hand, and CEO pay levels and structure, on the other. Our primary interests are the two issues that any explanation of executive compensation is called to address: the increase in executive-pay levels over time and the changes in the structure of pay over time.[122]

In developing our empirical investigation of CEO pay, we proceed through the following steps. In subpart A, we present our dataset. In subpart B, we provide a brief description of the time trends in average CEO pay over the past twenty-three years. In subpart C, in order to better understand the implications of board protection for incentive design, we focus on the relationship between defensive measures and executive pay. Lastly, in subpart D, we turn to examine the relationship between different forms of competition—namely, competition in the product and labor markets, as well as competition arising out of M&A activity—and executive pay.

---

121. As explained by one of us in a prior article, a board that is protected from short-term market pressure will be more likely to exhibit a bias toward tolerating the occurrence of disappointing firm outcomes and, hence, retaining an apparently underperforming manager. *See* Simone M. Sepe, *Board and Shareholder Power, Revisited*, 101 MINN. L. REV. 1377, 1420 (2017).

122. Murphy, *supra* note 3, at 330.

A. Data Description

Our universe of firms includes all firms in the ExecuComp database for the years 1993–2015,[123] excluding firms with dual-class stock and firms in regulated industries. Our final dataset consists of 18,511 firm-year observations for a total of 1,929 firms.

In examining executive compensation, we focus on four measures: *CEO Pay*, *Equity Portion*, *Option Portion*, and *Stock Portion*.[124] *CEO Pay* is measured as the natural logarithm of total CEO pay as reported in ExecuComp and is a proxy for the overall compensation payments received by the firm's top executives—whether in the form of salary, bonuses, other annual compensation components, restricted stock grants, long-term incentive plans, option grants, or any other form of compensation. *Equity Portion* is the percentage of the CEO's total compensation that is equity-based, including both restricted stock and option grants. *Option Portion* is the percentage of executive pay that is provided in the form of option grants, and *Stock Portion* is the percentage of executive pay that is provided in the form of restricted stock grants, where the sum of *Option Portion* and *Stock Portion* equals *Equity Portion*.

Further, we use two proxies for the equity-based incentives of the CEO's current stock and option holdings: *Pay-Performance Sensitivity* (*PPS*) and *Pay-Performance Volatility* (*PVS*).[125] *PPS* measures the sensitivity of the CEO's total equity holdings,[126] including both stock holdings and all options currently owned, to stock price.[127] That is, *PPS* measures how much the CEO's wealth increases upon a 1% increase in stock value (or, alternatively, how much it decreases upon a 1% decrease in stock value). Accordingly, this measure can be interpreted as capturing the degree of alignment between shareholder and manager interests, so that a higher *PPS* indicates a closer alignment of interests between shareholders and managers.

PVS measures the sensitivity of the CEO's current total option holdings to stock-return volatility. That is, *PVS* measures how much the value of the CEO's total portfolio of stock options increases with a 1% increase in the

123. The ExecuComp database provides information on the five highest paid top executives for firms included in the S&P 1500 composite index. The first year ExecuComp included a complete set of data was 1993.

124. As specified above, CEO-pay packages include many more components. *See supra* note 40. Over time, however, ExecuComp has changed the way in which it accounts for many of these components. Because of this limitation, we only consider the pay components that exhibit consistent reporting in ExecuComp across the years.

125. Both *PPS* and *PVS* are aggregated across a CEO's most recent and all previous years' pay packages, but only considering those stocks and options that the CEO currently still owns.

126. By definition, *PPS* does not capture the variability of fixed compensation, as fixed-pay components are not tied to changes in shareholder wealth.

127. In the literature, this is a rather standard, although by no means exclusive, way of measuring pay-for-performance sensitivity. *See* Murphy, *supra* note 3, at 234–37.

annualized stock-return volatility (or, alternatively, how much the CEO's wealth decreases with a 1% decrease in the firm's annualized stock-return volatility). Accordingly, *PVS* can be interpreted as capturing the "power" of executive incentives or, more practically, the level of "optionality" of incentives.[128]

For the data on corporate governance features, we use several sources. Data on defensive measures—*Staggered Board* and *Poison Pill*—for 1993–2015 come from RiskMetrics, SharkRepellent.net, and hand collection. *Staggered Board* is an indicator variable equal to one if the firm has a board that is staggered (zero otherwise). Similarly, *Poison Pill* is an indicator variable equal to one if the firm has a poison pill in place (zero otherwise).

In addition to investigating defensive measures, we also examine *Institutional Ownership*, which measures the percentage of outstanding shares held by institutional investors. Indeed, institutional investors could potentially play an important role in monitoring the executive-pay process.[129] We obtain data for *Institutional Ownership* from Thomson Reuters (for institutional ownership data) and from the Center for Research in Security Prices (CRSP) database (for the number of outstanding shares).

For investigating the effects of competition, we use several proxies and several sources. As a proxy for the level of labor-market competition for managerial talent, we use *Talent Competition.* This measure is based on the research of Xavier Gabaix and Augustin Landier, who show that general matching models (explaining which CEOs end up at which firms) imply that executive compensation is increasing with both the size of the CEO's firm and the general size of other firms in the economy.[130] On the one hand, CEO talent becomes more valuable to the firm as the firm becomes larger; on the other hand, as other firms become larger, stronger competition for scarce managerial talent bids up compensation for the most talented managers. Accordingly, *Talent Competition* is measured as the natural logarithm of the market capitalization of the outstanding shares of the 250th ranked firm in terms of market capitalization,[131] with data obtained from CRSP. Increases

128. While the ultimate source of the data for *PPS* and *PVS* is ExecuComp, we use the *PPS* and *PVS* data as used in previously published work by Jeffrey Coles et al., which they made publicly available with a sample extended to 2014. *See generally* Jeffrey L. Coles et al., *Co-Opted Boards*, 27 REV. FIN. STUD. 1751 (2014).

129. *See* BEBCHUK & FRIED, *supra* note 7, at 80, 82–83 (arguing that CEOs can be expected to have more power, and hence an easier time capturing boards, in companies with fewer institutional investors).

130. *See* Gabaix & Landier, *supra* note 71, at 50 ("The sixfold increase in CEO pay between 1980 and 2003 can be attributed to the sixfold increase in market capitalization of large U.S. companies during that period.").

131. Although our focus is on the S&P 1500, we still consider the S&P 500 for consistency with the work of Gabaix and Landier, which has come to constitute a reference in the empirical literature. However, while Gabaix and Landier use the asset market cap of the reference firm, we

in the market capitalization of this reference firm indicate increased competition for managerial talent.[132]

To capture product-market competition, *Product Competition*, we use the Herfindahl–Hirschman concentration index, with data obtained from the Compustat annual data file. The Herfindahl–Hirschman index is a measure of industry concentration based on sales (using all publicly traded firms in the industry).[133] For convenience, we present *Product Competition* as the negative of the Herfindahl–Hirschman concentration index, so that a lower Herfindahl–Hirschman index indicates a lower industry concentration and hence higher *Product Competition*.[134]

Further, we also employ *M&A Competition*, which is a proxy for the level of M&A activity per industry.[135] We retrieve data for *M&A Competition* from the SDC Platinum database offered by Thomson Reuters, which provides comprehensive information on all M&A activity in the United States. The rationale behind the use of this measure is that the level of M&A activity captures the level of an industry's shareholder pressure as operating through the takeover channel.[136]

Lastly, we add a set of standard controls, including leverage (*Leverage*), the ratio of capital expenditures over the book value of total assets (*CAPX/Assets*), the ratio of research and development expenditures over sales (*R&D/Sales*), an indicator variable for whether the firm is incorporated in Delaware (*Delaware Incorporation*), and the firm's profitability (*Profitability*). Appendix Tables A–B provide a brief definition of all our variables,[137] while descriptive statistics are presented in Appendix Table C.

---

use the equity market cap, consistent with the prior work of one of us on managerial talent competition. *See* Cremers & Grinstein, *supra* note 85, at 932.

132. *See* Gabaix & Landier, *supra* note 71, at 49–51.

133. *See* Cremers & Grinstein, *supra* note 85, at 932, 937. Since *Product Competition* is at industry level, it arguably has the advantage of not being (fully) under the firm's control, which mitigates endogeneity concerns.

134. We acknowledge that our measure of product competition is imperfect as it is actually a measure of market concentration. Indeed, there could be concentrated industries that are also competitive. In general, however, more concentrated industries will be less likely to be competitive (and vice versa), which makes our measure a plausible proxy for competition.

135. We adapt this measure from the prior work of one of us with Allen Ferrell. *See* Martijn Cremers & Allen Ferrell, *Thirty Years of Shareholder Rights and Firm Value*, 69 J. FIN. 1167, 1172 (2014) (introducing a new dataset on shareholder governance rights, spanning from 1978 to 2006). Like *Product Competition*, *Market Competition* is at industry level, and, therefore, we can assume plausible exogeneity at the firm level.

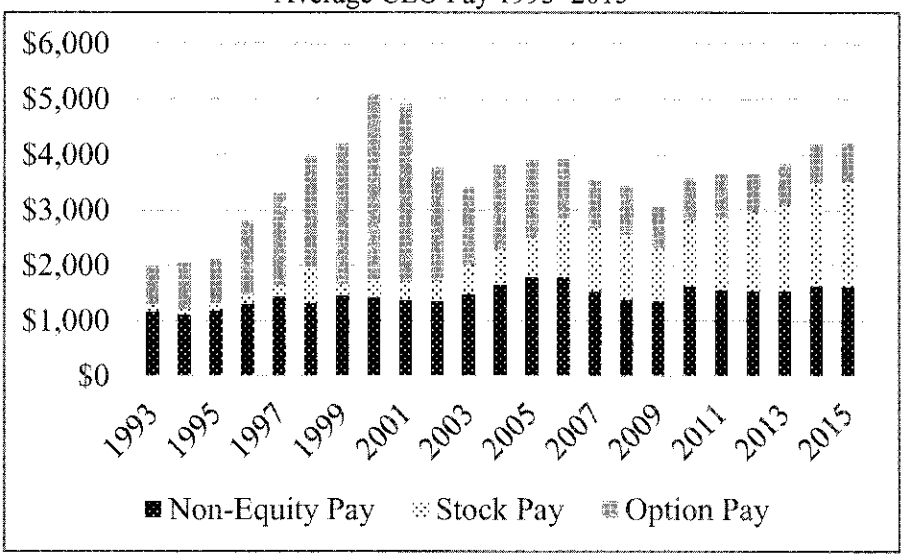
136. *See id.* at 1190–92 (suggesting that "the association between shareholder rights and firm value arises primarily through a 'takeover channel,' that is, through affecting the probability of an offer being received and accepted").

137. We emphasize that we reduce the importance of outliers in the data by consistently winsorizing each continuous variable at the 1% level on both sides of the distribution.

B. Time Trends in CEO Pay Level, Structure, and Incentives

We begin our empirical investigation by providing a description of the time trends in average CEO pay over the past twenty-three years. To this end, Figure 1 below shows the level and composition of average CEO pay in S&P 1500 firms for the period 1993–2015.[138] In particular, for each year, the figure shows the average level of three different components of executive compensation. The first is *Non-Equity Pay*, which includes salary, bonuses, and other forms of deferred compensation—essentially capturing the fixed component of executive compensation. We then have *Stock Pay*, which reflects the value of restricted stock grants, and lastly, *Option Pay*, which provides the average value of annual option grants.

Figure 1
Average CEO Pay 1993–2015



As shown in Figure 1, most of the dramatic increase in the level of CEO pay from the early 1990s to the early 2000s can be attributed to an escalation in *Option Pay*. The average *Option Pay* grew from $0.69 million in 1993 to $3.22 million in 2001, amounting to an increase of $2.53 million. In contrast, *Non-Equity Pay* only registers a growth of $0.21 million over the same nine-year period.

Notably, what caused the growth in the use of stock options is a challenging, and multifaceted, question. Commentators, however, tend to

138. In order to be able to compare compensation levels across years, the monetary amount for each year is converted to thousands of 2015-constant U.S. dollars using the Consumer Price Index deflator.

agree that the modification by the Clinton Administration, in 1993, of the tax code (which ironically was meant to reduce the level of CEO pay) offers at least a partial explanation, as the introduction of a $1 million deductibility cap for executive compensation with the exception of stock options[139] helped to provide a favorable tax treatment for options.[140]

Beginning in 2003, and even more since 2004, we observe a substantial shift away from *Option Pay* to *Stock Pay*, which progressively increases in the years afterward, with average *Option Pay* falling to $0.70 million and average *Stock Pay* swelling to $1.9 million by 2015.

This trend away from *Option Pay* to *Stock Pay* can partly be attributed to the introduction of a new, and less favorable, accounting treatment of options.[141] Indeed, the corporate scandals of the early 2000s renewed pressures for changing old accounting rules that permitted avoiding option expensing.[142] Under this pressure, many companies began voluntarily to transition to a regime of option expensing.[143] In response, in December 2004, the Financial Accounting Standards Board (FASB)—the private regulator responsible for standards of financial accounting and reporting in the United States—introduced a new accounting standard, FAS123(R) (revising rule FAS123),[144] under which the expensing rule for stock options became similar

---

139. *See* I.R.C. § 162(m) (2012) (excluding "performance-based compensation" from the $1 million deductibility cap).

140. Murphy, *supra* note 3, at 277–79.

141. *Id.* at 298; *see also* David I. Walker, *Evolving Executive Equity Compensation and the Limits of Optimal Contracting*, 64 VAND. L. REV. 611, 632–39 (2011) (discussing several possible causes of the decline in option compensation and the increase in stock compensation in recent times, including "a return to normalcy" following the dot-com crash of 2000–2002).

142. Murphy, *supra* note 3, at 297.

143. David Aboody et al., *Firms' Voluntary Recognition of Stock-Based Compensation Expense*, 42 J. ACCT. RES. 123, 124 (2004).

144. The changes in the accounting treatment of options have had a long story. In 1993, FASB proposed rules that would have required the value of stock options to be treated as compensation expenses on the income statement. In particular, the original FAS123 encouraged corporations to treat stock options as an expense based on their fair value on the date that the company granted the options. Corporations, however, could continue to use the previous intrinsic value method as long as they supplied additional disclosures about the options' fair value in the notes to the financial statements, which most of them continued to do until the corporate scandals of the 2000s. It was only in December 2004 that FASB issued FAS123(R), which mandated expensing. The Exposure Draft proposed application to new awards or modified awards in fiscal years beginning after December 15, 2004, but the final pronouncement postponed the effective date for large public entities after the start of the first interim or annual reporting period that began after June 15, 2005 (meaning July 1, 2005, for calendar-year corporations). *See* Tor-Erik Bakke et al., *The Causal Effect of Option Pay on Corporate Risk Management*, 120 J. FIN. ECON. 623, 625 (2016). In the midst of struggling with the demands of Sarbanes-Oxley, especially with respect to internal controls over financial reporting, issuers sought a further delay in the effective date, but FASB refused, even after the SEC joined in the request. The SEC then decided to act on its own, adopting a new rule that gave most large public companies a six-month reprieve. That rule allowed those companies to implement FAS123(R) for new or modified awards in the next fiscal year, that is, after January 1, 2006 (rather than the next reporting period that began after June 15, 2005). *See* Bakke et al., *supra*, at 625.

to that in place for restricted stock since 1972.[145] The new rule requires all U.S. public companies to recognize an accounting expense corresponding to the grant-date value of the shares amortized over the period when the options are not exercisable.[146] This change "significantly leveled the playing field between stock and options from an accounting perspective,"[147] plausibly inducing companies to increasingly substitute *Option Pay* with *Stock Pay*.

Next, we more closely examine the changing structure of equity-based compensation in Figure 2, which presents the proportions of stock-based executive pay and option-based executive pay (called, respectively, *Stock Portion* and *Option Portion*) in CEO compensation each year for all S&P 1500 companies in our sample for 1993–2015.

Figure 2
Stock Portion and *Option Portion*

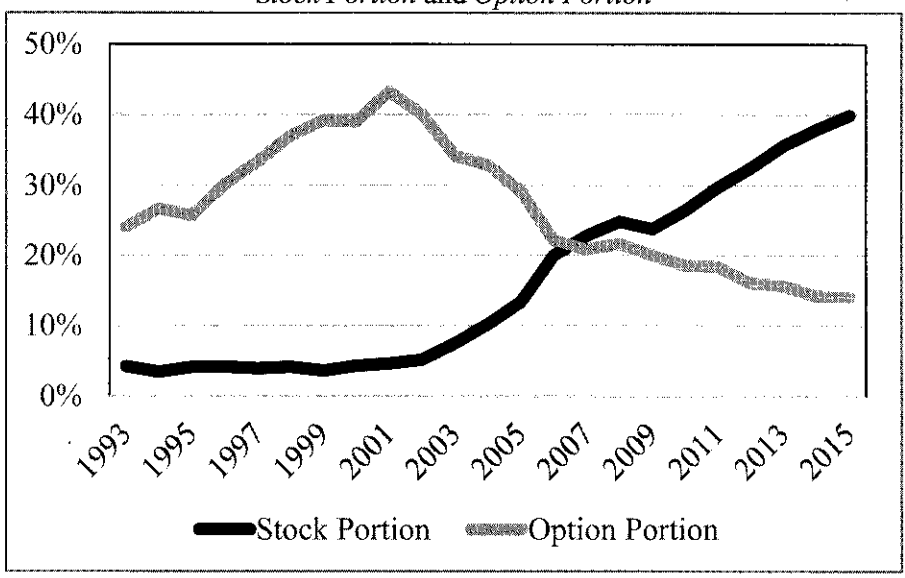


Figure 2 confirms that companies have been increasingly replacing options with stock grants in recent years. While *Option Portion* increased from about 24% in 1993 to 43% by 2001, it then fell to 22% in 2005 and then further to 14% in 2015. Conversely, the percentage of *Stock Portion* remained stable at around 5% from 1993 to 2002, then began to slowly increase in 2003 (when it went up to 8%) and more so in the years afterwards,

145. *See* SHARE-BASED PAYMENT, Statement of Fin. Accounting Standards No. 123 (Fin. Accounting Standards Bd. 2004), http://www.fasb.org/cs/BlobServer?blobkey=id&blobnocache=true&blobwhere=1175823287357&blobheader=application%2Fpdf&blobcol=urldata&blobtable=MungoBlobs [https://perma.cc/NU7H-GVMR].

146. *See id.*; Murphy, *supra* note 3, at 297–98.

147. Murphy, *supra* note 3, at 298.

reaching a percentage of 20% in 2006 and then surpassing the percentage of options in 2007 (i.e., 23% versus 21%). The gap in the use of *Stock Portion* relative to *Option Portion* further increased in the following years, with *Stock Portion* swelling to 40% and *Option Portion* declining to below 15% of total CEO pay by 2015.

C. *Corporate Governance*

Moving to the core of our empirical analysis, in this subpart we examine the relationship between corporate governance features—focusing on board defenses—and executive-compensation levels, structure, and incentives. As discussed above, under managerial power theory, such defenses provide the central means through which directors are insulated from the shareholders' power of removal and hence captured by managers.[148] Conversely, under our dynamic approach to executive compensation, temporary board protection from shareholder discipline serves a positive governance function. This function is mitigating the risk that the shareholders' limited commitment problem might strengthen a manager's incentives toward short-termism and impair a board's ability to design pay schemes that can efficiently induce managers to focus on long-term value creation.[149]

These opposite views of the relationship between board protection and executive compensation yield contrasting predictions. Under managerial power theory, one would expect to find that the adoption of defensive measures increases the likelihood that managers may extract inefficient entrenchment rents. This could happen, for example, through the overuse of fixed compensation or through a lower pay-for-performance sensitivity of the CEO pay package, which could take the form of a preference for the use of restricted stock grants over option grants.

On the contrary, under our dynamic approach to executive compensation, board protection is instrumental to efficient bargaining. As a result, our approach does not necessarily predict any change in the CEO pay packages of firms depending on their level of defensive measures. In other words, the dynamic approach suggests that whether or not we observe changes in these firms' CEO pay packages, such changes (or lack thereof) should be generally regarded as the result of optimal contracting and hence, on average, as value increasing.

Therefore, our empirical investigation in this subpart is more concerned with the explanatory power of managerial power theory, while we will empirically investigate our theoretical predictions about board protection in Part IV below, which focuses on financial value analysis. We also note that prior studies have challenged managerial power theory from an empirical

148. *See supra* text accompanying notes 55–60.

149. *See supra* section II (C).

perspective; for example, as discussed above, documenting an improvement in the indicators of board independence or an increase in CEO turnover rates.[150] However, very few studies have directly examined the relationship between the use of defensive measures and executive compensation.[151]

Empirically, we proceed, in Table 1 below, by investigating the time-series association between *Staggered Board* and *Poison Pill* and the proxies described above for, respectively, the level (*CEO Pay*) and the structure (*Equity Portion*, *Stock Portion*, and *Option Portion*) of annual executive compensation, as well as the cumulative incentives of the executive's current total holdings of stocks and options (*PPS* and *PVS*). As is well known in the literature, the advantage of using a time-series analysis is that of capturing intertemporal variation within the same firm—rather than across firms, as in cross-sectional analysis—with the result that time-series analysis is better placed to mitigate endogeneity concerns.[152]

In addition to our standard controls (which we do not show for the sake of brevity), in Table 1 we also control for *Institutional Ownership*. This control is meant to verify the prediction of managerial power theory that managers should be less able to extract inefficient entrenchment rents in firms with more institutional investors, as these investors can arguably count on more powerful governance levers to discipline directors and managers. We further include interactions between *Staggered Board*, *Poison Pill*, and *Institutional Ownership* with *Post 2010*, an indicator variable equal to one if the fiscal year included in the analysis falls after 2010 (zero otherwise). This additional test is meant to capture any variation in executive-compensation patterns that might have followed the 2010 enactment of the Dodd-Frank Act. Indeed, under managerial power theory, the Act's introduction of measures favoring an enhanced role for shareholders in the pay-setting process, such as say-on-pay shareholder votes, should arguably have reduced inefficient entrenchment rents.

---

150. *See supra* notes 82–84 and accompanying text.

151. *See, e.g.*, Rüdiger Fahlenbrach, *Shareholder Rights, Boards, and CEO Compensation*, 13 REV. FIN. 81 (2009) (similarly documenting empirical evidence inconsistent with managerial power theory). Fahlenbrach employs governance indices to measure the relative strength of shareholder and manager rights, including the E-Index. See *supra* note 55 for an explanation of the E-Index. However, he only performs a cross-sectional analysis, rather than a time-series analysis with firm fixed effects. See *infra* note 152 for a comparison of these two analyses. This is probably due to very limited time-series variation in the governance indices over the time period he considers, i.e., 1993–2004.

152. Unlike a cross-sectional analysis, a time-series analysis controls for any and all firm variables that do not change over time—that is, a firm's "fixed effects"—for each firm included in a panel dataset, adding a separate dummy variable for each unique firm. *See* Cremers, Masconale & Sepe, *supra* note 55, at 752–53 for further explanation.

Table 1
CEO Pay and Corporate Governance

In this table, we present pooled panel regressions of six proxies for the level, structure, and incentives of the CEO's compensation package (*CEO Pay*, *Equity Portion*, *Stock Portion*, *Option Portion*, *PPS*, *PVS*; see Appendix Table A for descriptions) on *Staggered Board* and *Poison Pill* plus a set of control variables (*Institutional Ownership*, *Assets*, *Stock-Return Volatility*, *Leverage*, *Delaware Incorporation*, *Profitability*, *CAPX/Assets*, *R&D/Sales*; see Appendix Table B for descriptions). Coefficients on the controls are not shown to save space, except for *Institutional Ownership*. We further include interactions between *Staggered Board*, *Poison Pill*, and *Institutional Ownership* with *Post 2010*, which is an indicator variable equal to one for years after 2010 (zero otherwise). All specifications include year and firm fixed effects (not shown). Statistical significance of the coefficients is indicated at the 1%, 5%, and 10% levels by ***, **, and *, respectively, based on robust standard errors clustered by firm.

Dep. Variable:	*CEO Pay*	*Equity Portion*	*Stock Portion*	*Option Portion*	*PPS*	*PVS*
Variables	(1)	(2)	(3)	(4)	(5)	(6)
Staggered Board	-0.0127	-0.0139	-0.0141	0.0001	-0.0559	-0.2150**
	(-0.37)	(-1.04)	(-1.15)	(0.01)	(-0.78)	(-2.09)
Poison Pill	0.0213	-0.0060	-0.0068	0.0008	-0.0976**	0.0461
	(1.03)	(-0.74)	(-0.97)	(0.09)	(-2.45)	(0.87)
Institutional Ownership	0.418***	0.134***	-0.0309	0.165***	0.400***	0.845***
	(5.59)	(4.84)	(-1.50)	(6.19)	(2.76)	(4.78)
Staggered Board	0.0171	-0.00243	-0.0179	0.0155	-0.00461	0.142
× *Post 2010*	(0.59)	(-0.21)	(-1.37)	(1.27)	(-0.07)	(1.35)
Poison Pill	-0.0174	0.0188	0.00718	0.0116	0.0956	0.214**
× *Post 2010*	(-0.52)	(1.41)	(0.43)	(0.72)	(1.15)	(1.99)
Institutional Ownership	0.164	-0.000758	0.156***	-0.157***	0.486*	0.285
× *Post 2010*	(1.62)	(-0.02)	(3.50)	(-3.91)	(1.80)	(0.71)
Fixed Effects	Firm,Year	Firm,Year	Firm,Year	Firm,Year	Firm,Year	Firm,Year
N	18,613	18,613	18,613	18,613	17,790	18,253
Adjusted R-Squared	0.752	0.458	0.566	0.475	0.744	0.715

As shown in Table 1, we find no significant changes in *CEO Pay* (shown in Column 1), *Equity Portion* (shown in Column 2), *Stock Portion* (shown in Column 3), *Option Portion* (shown in Column 4), and *PPS* (shown in Column 5) of firms with a staggered board relative to firms without a staggered board. The only compensation variable for which we find a significant change in firms with a staggered board is *PVS* (shown in Column 6), whose negative coefficient indicates that having a staggered board is associated with less powerful equity incentives.

Similarly, in the case of *Poison Pill*, we find no significant changes in any compensation proxy, except *PPS* (shown in Column 5). As *PPS* measures the sensitivity of the CEO's total equity holdings to stock price, this result could be taken to suggest that executive-compensation schemes in firms with a poison pill are less successful in aligning manager and shareholder interests. Yet, empirical difficulties and incongruences warn against this interpretation. First, since a board of directors can unilaterally adopt a pill at any time, it is difficult to gather any inference about the use of "visible" rather than "shadow" (or "off-the-rack") pills.[153] Second, even though we observe a lower sensitivity of pay to performance in firms with a poison pill, we cannot tell whether this change is inefficient, as we have not yet verified its effects on firm value. Third, we observe that *Poison Pill* is not associated with any significant change in *Equity Portion*, which makes it complicated to understand where the reduction in *PPS* comes from.

Taken as a whole, the results on *Staggered Board* and *Poison Pill* in Table 1 seem incompatible with managerial power theory, as there is no evidence that having these defenses in place results in higher levels of executive compensation or changes in pay structure.

On the contrary, these results seem compatible with our dynamic approach and the managerial talent hypothesis, as this approach does not necessarily predict any change in the CEO pay packages of firms depending on their level of defensive measures. Indeed, this approach recognizes that the level and structure of CEO pay, as well as the level of takeover defenses, investment policy, and the firm's overall strategy are simultaneously chosen, depending on the firm's specific circumstances (including both the firm's current performance as well as expectations of future performance) and the level of competition the firm faces in the product market and managerial talent market.

Our results on *Institutional Ownership* also seem incompatible with managerial power theory, as we find that the level of *CEO Pay* (shown in

153. *See* John C. Coates IV, *Takeover Defenses in the Shadow of the Pill: A Critique of the Scientific Evidence*, 79 TEXAS L. REV. 271, 288 (2000) (describing how a "shadow pill," a pill that *could* be adopted but has not yet been, may be as effective in impeding takeovers as a pill that has been adopted and commenting that "adoption of an actual pill by any given firm only brings this shadow pill into the light").

Column 1) is higher in firms with more institutional investors, contrary to the view that such investors would be better placed to rein in allegedly inefficient executive pay. We also find that *Institutional Ownership* is associated with a statistically significant increase in *Equity Portion* (shown in Column 2) and both *PPS* and *PVS* (respectively shown in Columns 5 and 6). In particular, the positive association between *Equity Portion* and *Institutional Ownership* seems to be driven by the use of options, as *Option Portion* has a large positive and statistically significant association with *Institutional Ownership* (shown in Column 4), while *Stock Portion* is negatively associated with *Institutional Ownership* (although the coefficient, shown in Column 3, is not statistically significant). These results suggest that firms with more institutional investors prefer the adoption of more-skewed incentive schemes.

Again, while we cannot yet tell whether more-skewed incentives generally serve shareholder interests, it is important to emphasize that managerial power theory predicts that such incentives should be associated with a reduction—rather than an increase—in overall CEO pay. This is because providing managers with more-skewed incentives (that is, options) should constitute a less costly way of paying a manager's agency rents.[154] Our data, however, seem to indicate otherwise, as *Institutional Ownership* has a strongly positive and statistically significant relationship with the level—and not just the structure—of executive compensation.

When we control for possible changes that might have occurred in executive-compensation trends after the 2010 introduction of the Dodd-Frank Act (*Post 2010*), we obtain results that are in line with our considerations above for each of the variables of interest. First, having a staggered board in place after 2010 is not associated with any statistically significant change in executive compensation. Second, in firms with a poison pill in place after 2010, the only statistically significant change is an increase in *PVS* (shown in Column 6). Yet, the above considerations on the difficulty of interpreting *Poison Pill* results also apply here. Third, concerning the post-2010 impact of *Institutional Ownership* on executive compensation, we document a statistically significant increase of *Stock Portion* (shown in Column 3), accompanied by a statistically significant decrease of *Option Portion* (shown in Column 4). This seems to suggest that the 2010 executive-compensation reform has possibly contributed to producing a shift from *Option Portion* to *Stock Portion* in firms with more institutional investors.

When coupled with the additional evidence that this shift has increased, rather than decreased, *PPS*, the results on the post-2010 impact of *Institutional Ownership* on executive compensation seem incompatible with the managerial power view that the payment to the executive of restricted stock, rather than options, should be regarded as a form of entrenchment rent

154. *See supra* text accompanying notes 65–69.

that is detrimental to shareholder interests. Otherwise, why would we observe that firms with more institutional investors have chosen to increasingly substitute *Stock Portion* for *Option Portion* after such investors were empowered with a greater say in the CEO pay process? Additionally, why would we find that this substitution improves the alignment of shareholder and manager interests?

To sum up, the results of Table 1 suggest that board decision-making concerning the CEO pay process is better explained by external governance factors, such as the presence of institutional investors, than by internal governance arrangements, such as the adoption of a staggered board or a poison pill. This conclusion is not just difficult to reconcile with managerial power theory, but it directly points to a role for market forces in the CEO pay process. We turn to the empirical investigation of that role next.

D. Product, Financial, and Labor Markets

In this subpart, we aim to explore what role market forces and competition play in shaping CEO pay levels, structure, and incentives. To this end, Table 2 presents pooled panel regressions (with firm fixed effects but no year fixed effects)[155] of three proxies of competition—*Talent Competition*, *Product Competition*, and *M&A Competition*—on CEO pay levels, structure, and incentives. As explained in our data description,[156] each of these competition proxies is designed to capture different market forces that may play a role in shaping executive compensation. Drawing on managerial talent theory, our primary interest is on labor-market competition for managerial talent (*Talent Competition*). Secondarily, we are interested in understanding whether competition in the product market (*Product Competition*) or competition in the M&A market (*M&A Competition*, which we interpret as a measure capturing the level of shareholder pressure in a given industry) have any role to play in the executive-pay process.

155. As the reference firm in *Talent Competition* is identified each year, we cannot use year fixed effects. Therefore, in order to control for inflation, we deflate *CEO Pay*, *Assets*, and *Market Cap* into 2015-constant dollars, using the GDP deflator from the Bureau of Economic Analysis.

156. *See supra* subpart III(A).

Table 2:
CEO Pay and Competition

In this table, we present pooled panel regressions of six proxies for the level, structure, and incentives of the CEO's compensation package (*CEO Pay*, *Equity Portion*, *Stock Portion*, *Option Portion*, *PPS*, and *PVS*; see Appendix Table A for descriptions) on three proxies for market competition (*Talent Competition*, *Product Competition*, and *M&A Competition*; see Appendix Table B for descriptions). We further include *Market Cap*, i.e., the natural logarithm of the market capitalization of the outstanding shares of the firm, plus a set of standard controls (*Assets*, *Stock-Return Volatility*, *Leverage*, *Delaware Incorporation*, *Profitability*, *CAPX/Assets*, and *R&D/Sales*; see Appendix Table B for descriptions).[157] Coefficients on the controls are not shown to save space (except for *Market Cap* and *Assets*). All specifications include firm fixed effects (also not shown). Statistical significance of the coefficients is indicated at the 1%, 5%, and 10% levels by ***, **, and *, respectively, based on robust standard errors clustered by firm. For *Talent Competition* and *M&A Competition*, we cluster by year, as this results in more conservative (i.e., smaller) *t*-statistics.

Dep. Variable:	*CEO Pay*	*Equity Portion*	*Stock Portion*	*Option Portion*	*PPS*	*PVS*
Variables	(1)	(2)	(3)	(4)	(5)	(6)
Talent Competition	0.341***	0.102***	0.0908***	0.0116	0.173	0.695***
	(7.04)	(5.69)	(2.84)	(0.34)	(1.21)	(3.39)
Product Competition	-0.421	-0.306**	-0.751***	0.445***	1.377**	0.835
	(-1.40)	-(2.49)	(-5.35)	(3.18)	(2.25)	(0.75)
M&A Competition	0.00619	0.0143	-0.117***	0.131***	0.420***	0.638***
	(0.12)	(0.48)	(-1.88)	(1.87)	(2.78)	(1.90)
Market Cap	0.0857***	0.0449***	-0.0371***	0.0820***	0.288***	0.121***
	(5.21)	(7.56)	(-6.88)	(12.36)	(10.85)	(3.15)
Assets	0.295***	0.0230***	0.120***	-0.0970***	0.0756*	0.273***
	(13.52)	(2.88)	(14.46)	(-10.69)	(1.94)	(5.00)
Fixed Effects	Firm	Firm	Firm	Firm	Firm	Firm
N	18,692	18,692	18,692	18,692	17,871	18,309
Adjusted R-Squared	0.729	0.446	0.479	0.430	0.721	0.672

Table 2 shows that *Talent Competition* has a substantial effect on executive compensation. First, greater *Talent Competition* is associated with a statistically significant increase in *CEO Pay* (shown in Column 1),

157. Controlling our results for the market capitalization of the firm's equity (*Market Cap*), as well as the book value of the firm's total assets (*Assets*), allows us to also control for the direct effects of firm size—that larger firms pay their CEOs more and that firms where market capitalization or assets have gone up increase their CEO pay.

supporting the prediction of managerial talent theory that increasingly higher CEO pay levels can be largely explained in terms of market rents. The magnitude of the association between the level of competition for managerial talent and the various aspects of CEO pay also is economically meaningful. For example, a standard deviation increase in *Talent Competition*[158] is associated with an increase in *CEO Pay* of 0.14[159] which constitutes about 14.1% of the standard deviation in *CEO Pay*. Overall, the results in Table 2 on the association between *Talent Competition* and *CEO Pay* thus seem to suggest that the explosion in CEO pay is more likely to be the result of competitive forces than a dramatic increase in the managerial moral hazard problem.[160]

Talent Competition is also positively associated with a statistically significant increase in *Equity Portion* (shown in Column 2), *Stock Portion* (shown in Column 3), and *PVS* (shown in Column 6). The coefficients on *Option Portion* (shown in Column 4) and *PPS* (shown in Column 5) are also positive but not statistically significant. Thus, the increase in CEO pay levels due to competition in the managerial labor market seems largely attributable to an increase in the stock component of CEO pay. In other words, restricted stock grants emerge as the most common way of paying out the market rents demanded by increased competition for managerial talent.

Conversely, greater *Product Competition* is associated with a decrease of *CEO Pay* (shown in Column 1, though statistically insignificant) and both *Equity Portion* (shown in Column 2 and statistically significant) and *Stock Portion* (shown in Column 3, and strongly statistically significant). Possibly, the negative association between competition in the product market and executive compensation might be explained by the fact that firms in more competitive industries tend to be less profitable, which would justify an average CEO pay that is lower. At the same time, greater *Product Competition* is associated with an increase in both *Option Portion* (shown in Column 4) and *PPS* (shown in Column 5). The coefficient on *PVS* (shown in Column 6) is also positive but not statistically significant. Consistent with our earlier assumptions,[161] this result could be explained by the fact that firms in more competitive industries are continuously required to outperform their

158. This is equal to 0.417. *See infra* Appendix Table C.

159. This is calculated by multiplying a standard deviation increase in *Talent Competition*, 0.417, by the coefficient on *Talent Competition* of 0.341.

160. When we control the specification of Table 2, Column 1 for *Staggered Board*, the stand-alone impact of *Staggered Board* on *CEO Pay* is negative and statistically significant. However, the impact of *Talent Competition* on *CEO Pay* conditional on the presence of a staggered board is positive and statistically significant. This means that while in general a staggered board helps control *CEO Pay*, when *Talent Competition* increases, a staggered board efficiently reacts to external pressure by readjusting *CEO Pay* in line with that of firms without a staggered board.

161. *See supra* text accompanying notes 117–21.

competitors, which could justify why such firms have a preference for more-skewed incentives.

The results for *M&A Competition* are quite similar to those for *Product Competition*, except that *M&A Competition* is not associated with any significant change in *CEO Pay* (shown in Column 1) and *Equity Portion* (shown in Column 2) and is positively and significantly associated with *PVS* (shown in Column 6). This finding also seems consistent with our assumption that firms that are subject to greater shareholder pressure might prefer skewed incentives to promote positive short-term performance and, therefore, reduce the likelihood of a future takeover and other short-term shareholder interferences.[162]

Combined, the results on *Product Competition* and *M&A Competition* indicate that greater market discipline (whether in the form of intense product-market competition or enhanced shareholder pressure) produces an increase in the sensitivity of pay to performance through the use of less pay in the form of restricted stock and more in the form of option grants. Indeed, the evidence that *Option Portion* increases, while *Stock Portion* decreases points to a substitution effect between *Option Portion* and *Stock Portion*. In contrast, greater *Talent Competition* has the opposite effect of increasing *Stock Portion*. However, we cannot tell yet which of these effects is efficient. It could be that decreasing stock and increasing options enhances shareholder value, as suggested by managerial power theorists,[163] or, on the contrary, that increasing stock can better serve shareholder interests by mitigating short-term preferences, which would be consistent with our dynamic approach. To answer this and additional questions, we move to the value analysis of executive compensation.

IV. Value Analysis

This Part examines the interacted impact on firm value of corporate governance features, competition, and executive compensation. Two main research interests motivate this additional analysis. First, our investigation of the empirical relationship between defensive measures and executive compensation seems incompatible with the managerial power view that strong defenses, such as the staggered board, are the main source of inefficient entrenchment rents, as we document that staggered boards are not associated with any of the executive-compensation patterns predicted by this view. Per se, however, this analysis does not give us information on how defensive measures interact with executive compensation to impact firm value. Our hypothesis is that once one considers the complexities arising from competition and the dynamic nature of executive-compensation

162. *See id.*

163. *See supra* text accompanying notes 65–69.

contracts, the use of defensive measures might be instrumental to promoting optimal bargaining between boards and managers.

Second, and relatedly, our theoretical and empirical analysis of the relationship between competition and executive-compensation points to a major impact of various forms of talent, product, and M&A competition on both CEO pay levels and structure. Therefore, for each of these forms of competition, we are interested in understanding whether its impact on executive compensation also produces an impact on firm performance and what kind of impact it may produce.

Of course, we are aware of the endogeneity concerns that such investigation might raise, which a time-series analysis can mitigate but not fully address. Indeed, firm performance and CEO pay are jointly determined, which makes it difficult to understand the direction of causality, if any, in the relationship between these variables. In response to such concerns, Part IV also considers an event study that focuses on the 2005 introduction of the FAS123(R) mandate to expense stock options.[164] As the product of regulatory intervention, this event can be regarded as independent from firm-specific circumstances and, hence, as plausibly inducing exogenous changes in both the levels and structure of CEO pay. Therefore, examining the subsequent performance of firms that were and were not affected by the introduction of FAS123(R) can arguably provide *causal* evidence about the impact of modifications in CEO pay levels and structure on firm value. Further, as FAS123(R) leveled the playing field between the use of options and restricted stock from an accounting perspective, we expect such an event study to provide us with additional insights into the relative efficiency of more versus less skewed compensation arrangements.

A. The Value of CEO Pay

Table 3 presents time-series results for our value analysis of executive compensation.[165] As a proxy for firm value, we use Tobin's *Q* (*Q*) (the ratio of a firm's market value of assets to its book value of assets) as it has become standard in the empirical literature.[166] Because firm valuations are generally strongly related to market-wide movements, in addition to firm fixed effects, we also include year fixed effects. This, in turn, prevents us from including *Talent Competition* in our value analysis, as this proxy corresponds to the market cap of the 250th firm in the S&P 500 and, thus, only varies by year.

---

164. *See supra* notes 144–47 and accompanying text.

165. All specifications include firm fixed effects such that we are effectively estimating how changes in firm value are associated with changes in the level of competition and the use of takeover defenses in the time series. Since our corporate governance focus is on *Staggered Board* and *Poison Pill*, and given space constraints, in Table 3 we do not run regression for *Institutional Ownership*. If we include *Institutional Ownership* as a control, however, our results remain unchanged.

166. *See supra* note 24.

We note, however, that as *Talent Competition* seems to constitute a primary source of the increase in CEO pay levels, testing the value association of changes in the level of *CEO Pay* with changes in firm value provides indirect information on the efficiency implications of labor-market competition for managerial talent.

Table 3
CEO Pay and Firm Value

In this table, we present pooled panel *Q* regressions on one-year lagged values of proxies for the level, structure, and incentives of CEO pay (*CEO Pay*, *Equity Portion*, *PPS*, and *PVS*; see Appendix Table A for descriptions), corporate governance (*Staggered Board* and *Poison Pill*), competition (*Product Competition* and *M&A Competition*), and a set of standard controls (*Assets*, *Stock-Return Volatility*, *Leverage*, *Delaware Incorporation*, *Profitability*, *CAPX/Assets*, and *R&D/Sales*). See Appendix Table B for all non-compensation variable descriptions. In Columns 5 and 7, we add the interactions of *Staggered Board* and *Poison Pill* with *CEO Pay*, and in Columns 6 and 7, we further add the interactions of *Product Competition* and *M&A Competition* with *CEO Pay*. Coefficients on the standard controls are not shown to save space. All specifications include firm and year fixed effects (also not shown). Statistical significance of the coefficients is indicated at the 1%, 5%, and 10% levels by ***, **, and *, respectively, based on robust standard errors clustered by firm.

Dep. Variable: Firm Value (*Q*)							
Variables (Lagged)	(1)	(2)	(3)	(4)	(5)	(6)	(7)
CEO Pay	0.100***	0.134***		0.0946***	0.0524	0.0871**	0.0344
	(4.35)	(4.80)		(3.81)	(1.49)	(2.55)	(0.78)
Equity Portion		-0.132**		-0.0329			
		(-2.56)		(-0.66)			
PPS			0.193***	0.186***			
			(13.30)	(12.76)			
PVS			-0.0675***	-0.0749***			
			(-5.65)	(-6.42)			
Staggered Board					0.0565*		0.0577*
× *CEO Pay*					(1.74)		(1.78)
Poison Pill					0.0321		0.0348
× *CEO Pay*					(1.20)		(1.30)
Product Competition						-0.288	-0.329
× *CEO Pay*						(-1.12)	(-1.28)
M&A Competition						-0.220**	-0.230**
× *CEO Pay*						(-2.02)	(-2.10)
Staggered Board	0.179***	0.177***	0.184***	0.185***	-0.305	0.181***	-0.312
	(2.93)	(2.92)	(3.32)	(3.38)	(-1.14)	(2.99)	(-1.18)
Poison Pill	-0.0285	-0.0295	0.00591	0.00478	-0.287	-0.0275	-0.308
	(-0.84)	(-0.87)	(0.18)	(0.15)	(-1.33)	(-0.82)	(-1.42)
Product Competition	0.566	0.568	0.369	0.383	0.513	1.774	2.166
	(1.23)	(1.24)	(0.84)	(0.88)	(1.12)	(0.88)	(1.07)
M&A Competition	-0.154	-0.152	-0.192*	-0.189*	-0.156	1.589*	1.658*
	(-1.56)	(-1.55)	(-1.94)	(-1.90)	(-1.58)	(1.84)	(1.92)
Fixed Effects	Firm,Year	Firm,Year	Firm,Year	Firm,Year	Firm,Year	Firm,Year	Firm,Year
N	17,905	17,905	17,112	17,071	17,905	17,905	17,905
Adjusted R-Squared	0.707	0.708	0.720	0.722	0.708	0.707	0.708

As shown in Table 3, Column 1, higher *CEO Pay* is associated with higher firm value. For example, multiplying the coefficient in Column 1 of 0.100 with the standard deviation of *CEO Pay* (equal to 0.998) suggests that one standard deviation increase in executive pay is associated with a future

and permanent increase in firm value of about 0.10, which amounts to an increase of about 6.5% of the median firm value in the sample.[167]

The positive association between CEO pay levels and firm value seems inconsistent with the notion that a large portion of CEOs receive substantially "excessive" levels of pay in the sense described by managerial power theory.[168] Indeed, under such a theory, we would have expected a negative association between increased CEO pay levels and firm value.

Of course, *CEO Pay* in Table 3 is determined endogenously along with firm value so that causation can run both ways rather than just from *CEO Pay* to firm value. In other words, enhanced CEO effectiveness could result in higher CEO pay rather than higher pay causing better performance. We address endogeneity concerns in the next section. Still, it is worth observing here that to the extent that causation runs from firm value to *CEO Pay*, this result would still be more consistent with managerial talent theory and our dynamic approach to executive compensation than with managerial power theory. Indeed, when considered in combination with the results of Table 2 on *Talent Competition*, high CEO pay levels seem to largely reflect market rents that are necessary to ensure that the most talented CEOs are allocated to the most valuable firms, consistent with the claims of managerial talent theory.[169] Further, the positive association between *CEO Pay* and higher long-term firm value also seems consistent with our dynamic approach, under which granting managers high rents may help commit talented managers to the exercise of long-term effort and, therefore, increase long-term shareholder wealth.[170]

Next, Column 2 shows that a higher *Equity Portion* is associated with a statistically significant reduction in firm value.[171] Combined with the results we obtain for *PPS* and *PVS* respectively (shown in Column 3), this suggests that the source of reduced firm value is not the provision of equity incentives per se, but rather the provision of equity incentives that are excessively skewed (that is, a higher *Option Portion*), as a higher *PVS* is associated with lower firm value, while a higher *PPS* is associated with higher firm value.

This is consistent with our theoretical prediction about the effects of intertemporal tradeoffs and the related risk of short-termism.[172] Given this risk, providing managers with more powerful incentives (more options and less restricted stock) may decrease, rather than increase, shareholder wealth.

---

167. The percentage effect is calculated as the coefficient estimate of *CEO Pay* (0.100), multiplied by the standard deviation (0.998, see *infra* Appendix Table C) of *CEO Pay*, and then dividing the product by the median Tobin's *Q* in the sample (1.524, see *infra* Appendix Table C).

168. *See supra* subpart I(B).

169. *See supra* subpart I(C).

170. *See supra* section II(B)(2).

171. Note that in Table 3, we do not disentangle the stock and option components of equity-based compensation, as this analysis is at the core of our event study in subpart IV(B) below.

172. *See supra* section II(B)(1).

Indeed, such incentives may induce managers to prefer short-termist investment projects at the expense of long-term firm value, possibly explaining why a higher *PVS* is associated with lower firm value in Table 3.

The results of the interacted impact of *CEO Pay* with *Staggered Board* and *Poison Pill* respectively (shown in Column 5) are also consistent with our theoretical predictions. Combining a staggered board with high CEO pay is associated with higher firm value, consistent with our theory that the adoption of a staggered board might serve a positive function in the executive-compensation context.[173] Indeed, while neither staggered boards nor poison pills are associated with statistically significant changes in either the level or structure of CEO pay,[174] the results of Table 3 are strongly suggestive that firms with a staggered board have more efficient CEO pay (that is, the positive association between higher *CEO Pay* and higher firm value seems to be primarily driven by firms with a staggered board). Even if this is driven by selection, then it seems to be "good selection."

Interpreted differently, adopting a staggered board seems to lead to—or at least be indicative of—a more positive relationship between CEO pay and firm value, suggesting that talented CEOs (who presumably receive higher pay) are more effective at firms with a staggered board. Our commitment theory provides a rational explanation for this result, suggesting that the staggered board can serve as a bilateral commitment device towards long-term value creation by both shareholders and managers, thereby mitigating the negative externalities arising from excessive market discipline or investor short-termism.[175] This, in turn, helps preserve a manager's continuation value and the ability of boards to exploit such value for efficient incentive design.[176]

Our results for *Poison Pill* are in line with those for *Staggered Board*, although the increase in firm value associated with the combination of a poison pill and high CEO pay is not statistically significant. Overall, our results on the interacted impact of defensive measures and CEO pay on firm value thus seem to suggest that such measures complement the use of pecuniary incentives to promote long-term value creation.

Next, in Column 6, we examine the interacted impact of both *Product Competition* and *M&A Competition* with *CEO Pay* on firm value. We find that higher *Product Competition* is associated with lower firm value (although the association is not statistically significant). The coefficient on *M&A Competition* in Column 6 is similar, where a higher level of M&A

173. The stand-alone coefficient on *Staggered Board* is also positive and statistically significant, in line with the results obtained in recent research on staggered boards. See Cremers & Sepe, *supra* note 16, at 100–05 and Cremers, Litov & Sepe, *supra* note 16, at 425-30 for more detailed discussions.

174. *See supra* Table 1.

175. *See supra* subpart II(C).

176. *See id.*

activity per industry interacted with *CEO Pay* is associated with a statistically significant lower firm value.

These results are, again, consistent with our theoretical predictions, supporting the view that market forces such as greater product-market competition and increased shareholder pressure might introduce value-reducing distortions in incentive design. On the other hand, we also document that the stand-alone coefficient on *M&A Competition* is positive and statistically significant. This is consistent with the assumption that the higher likelihood of a future takeover increases firm value, as shareholders have a greater probability of selling their shares at a premium.[177] This expectation, however, comes at the cost of disrupting compensation schemes, further confirming that the benefits of greater market discipline need to be evaluated along with its cost.[178]

B. Endogeneity Concerns

As hinted to above, the results of our value analysis need to be carefully interpreted due to possible endogeneity concerns. In particular, two of such concerns seem more severe. First, there could be important factors that affect *both* firm performance *and* executive compensation, which could result in a "specification problem." Empiricists refer to such a problem when changes in the dependent variable might be attributable to factors other than changes in the independent variable. [179] For example, actual CEO talent is unobservable and may be time-varying. As a result, our finding that firm value tends to go up if CEO pay increases may at least partly reflect an enhanced CEO effectiveness over time.

Second, executive compensation and firm performance are jointly determined. This means that executive compensation affects firm performance, but firm performance also affects executive compensation, with the added complexity that some aspects of firm performance are unobservable to outsiders, although they are presumably observable to the board. This, in turn, may involve what empirical scholars refer to as a "reverse causality" problem, which arises when the dependent variable causes changes in the independent variable, rather than the other way

---

177. *See, e.g.*, Alex Edmans et al., *The Real Effects of Financial Markets: The Impact of Prices on Takeovers*, 67 J. FIN. 933, 934 (2012) ("[A]n *anticipation effect* may lead to reverse causality from takeover activity to market valuations, with forward-looking prices inflated by the probability of a future takeover.").

178. Lastly, for robustness, Column 7 verifies the impact on firm value of each of our corporate governance, competition, and compensation proxies while simultaneously controlling for the other proxies, confirming the results described in subpart IV(A).

179. *See* WILLIAM H. GREENE, ECONOMETRIC ANALYSIS 56–58 (7th ed. 2012) (discussing the two most common specification errors: the omission of relevant variables and the inclusion of irrelevant variables).

around.[180] In our case, it could be that expectations of future outperformance might lead to increases in current compensation, based on the board's observation that the CEO has been more effective than perhaps recognized by the shareholders. Therefore, the predictability in firm value that we document (where increases in the level of pay predict better future performance) may be due to improved CEO effectiveness that is not yet recognized by the stock market, that is, be due to reverse causality. As with possible specification issues affecting our results, however, it is worth emphasizing that the possibility of reverse causality does not go in the direction of supporting managerial power theory, as this theory falls short of predicting that CEO pay is largely driven by the board's well-informed expectations about future performance.[181]

In response to these concerns, in this subpart we employ an event study, examining the effects on executive compensation and firm value of the exogenous changes produced by the 2005 introduction of the mandate to expense stock options, namely accounting standard FAS123(R).[182] As we document below, the results we obtain are consistent with the above interpretation of our value analysis as supporting the view that greater market and shareholder pressure might lead to value-reducing distortions in CEO pay, such as an overuse of option grants that place excessive emphasis on short-term performance at the expense of long-term firm value.

1. Identification Strategy.—As briefly discussed in subpart III(B), FAS123(R)'s mandate to expense stock options significantly reduced the attractiveness of stock options from an accounting perspective, leading most U.S. public companies to increasingly substitute option grants with restricted stock grants. It is important to emphasize, however, that FAS123(R) only impacted the accounting performance of firms, but did not directly affect corporate cash flows. Indeed, while firms granting larger option awards to their executives now bear higher accounting expenses (and, hence, lower net accounting profits), expensing stock options does not lead to any cash

180. *See* MITCHELL H. KATZ, STUDY DESIGN AND STATISTICAL ANALYSIS 25 (reprt. 2009) is report supposed to be abbreviated lice this? (defining "reverse causality" as when "the 'effect' causes the 'cause'").

181. One could argue that since we do not present a theory of pay determination under managerial power theory, we cannot exclude the possibility that managers extract higher rents when firm value is high. This is why we consider a quasi-natural experiment in this subpart—in order to get some plausibly exogenous variation in executive compensation that is unlikely to be driven by such reverse causality.

182. A prior study by Tor-Erik Bakke et al. employed FAS123(R) as a quasi-natural experiment to study how CEOs' option compensation affects the hedging behavior of energy firms. Bakke et al., *supra* note 144. The basic design of our empirical methodology closely follows the empirical design by Bakke et al. *See id.* at 626–29. However, to the best of our knowledge, we are the first to employ the option-expensing mandate as a quasi-natural experiment to study the association between firm performance and executive compensation.

outflow at the time of the option award. That is, the cash flows associated with option awards remain unaltered. Therefore, FAS123(R) should not be expected to have directly affected long-term firm value or long-term shareholder value, as shareholders should ultimately only care about a firm's cash flows rather than accounting performance.

Hence, our main identifying assumption is that FAS123(R) has arguably impacted firm value only indirectly, through the changes that it produced on the level and structure of CEO pay.[183] This means that any direct effect of FAS123(R) on firm value should be regarded as second-order, or minor, relative to the effect on firm value the rule produced through the changes it caused in the level and structure of executive compensation. Under these assumptions, the association between the changes in compensation that were caused by FAS123(R) and the changes in firm value that took place after the rule's introduction can accordingly be interpreted as providing plausible causal evidence for how changes in compensation affect changes in firm value.

2. High-Powered Incentives: A Negative Result.—Our empirical analysis proceeds in two steps. In the first step, we consider how the level and structure of CEO pay changed in the two years before (fiscal years 2003 and 2004) versus after (fiscal years 2005 and 2006) the introduction of FAS123(R). In the second step, we consider how the changes in firm value occurring before versus after the introduction of FAS123(R) relate to the plausibly exogenous changes that companies made to the level and structure of their CEOs' compensation in response to the new accounting rule.

In order to isolate compensation changes that were made in response to FAS123(R) from compensation changes made for other reasons, we always include firm fixed effects, year fixed effects, and the resulting impact on firm value. To the same end, we compare the changes in firm value and compensation for firms that were affected by the introduction of FAS123(R)—the "treated firms"—to a set of "control firms," which did not award any options to their CEOs in either 2003 or 2004 and were therefore arguably unaffected by the introduction of FAS123(R) over the time period that we study (2003–2006). Thus, all our results for how FAS123(R) affected either compensation or firm value compare how specific changes affected the "treated firms" differently from the "control firms."[184]

Table 4, Panel A presents the results for the first step of our analysis, essentially showing the effect of the mandate to expense stock options on CEO pay levels and structure.

---

183. While we recognize that shareholders may learn about future cash flows from current accounting performance, the direct effect of expensing stock options seems easily identifiable.

184. One could argue that some firms only adopted FAS123(R) after Jan. 1, 2006. *See supra* note 144. What matters for our analysis, however, is the comparison between firms that were

Table 4, Panel A
CEO Pay Around the Introduction of FAS123(R)

In this table, we present pooled panel regressions of four proxies for the level, structure, and incentives of the CEO's compensation package (*CEO Pay*, *Equity Portion*, *Stock Portion*, and *Option Portion*; see Appendix Table A for descriptions) in the four years (2003–2006) surrounding the introduction of the mandate to expense option grants starting in fiscal year 2005 (FAS123(R)) on *Treated* × *2005–2006*, plus proxies for competition, corporate governance, and the standard controls. *Treated* is an indicator variable equal to one if the firm awarded options to the CEO in 2003 or 2004, and *2005–2006* is an indicator variable that equals one in fiscal years 2005 and 2006. The included competition variables are *Product Competition*, *M&A Competition*, and *Talent Competition*; the included governance variables are *Staggered Board* and *Poison Pill*; and the set of controls includes *Market Cap*, *Institutional Ownership*, *Assets*, *Stock-Return Volatility*, *Leverage*, *Delaware Incorporation*, *Profitability*, *CAPX/Assets*, and *R&D/Sales*. See Appendix Table B for descriptions. We do not show the coefficients on some of the controls to save space. All specifications include year and firm fixed effects (not shown). Statistical significance of the coefficients is indicated at the 1%, 5%, and 10% levels by ***, **, and *, respectively, based on robust standard errors clustered by firm.

affected by FAS123(R) because they had outstanding options and firms that were not affected by the new rule because they did not have outstanding options. Therefore, even if our sample of treated firms may not include all the firms that were affected by FAS123(R) (because some firms only adhered to the mandate in 2006), this does not affect our ability to compare treated firms and control firms.

Dep. Variable:	*CEO Pay*	*Equity Portion*	*Stock Portion*	*Option Portion*
Variables	(1)	(2)	(3)	(4)
Treated × 2005–2006	-0.296***	-0.193***	0.0271	-0.220***
	(-4.44)	(-7.69)	(1.60)	(-10.66)
Staggered Board	-0.0105	-0.0559	-0.0463	-0.00958
	(-0.10)	(-1.41)	(-1.31)	(-0.25)
Poison Pill	-0.0449	-0.0126	-0.0190	0.00632
	(-0.72)	(-0.44)	(-0.77)	(0.20)
Institutional Ownership	0.00441**	0.000950	-0.000165	0.00111
	(2.04)	(1.08)	(-0.28)	(1.32)
Talent Competition	0.0718	-0.177	0.466***	-0.643***
	(0.27)	(-1.43)	(4.57)	(-5.96)
Product Competition	-1.266	-0.912	-0.367	-0.545
	(-0.81)	(-1.60)	(-0.88)	(-1.20)
M&A Competition	-0.0155	-0.0187	0.00241	-0.0211
	(-0.12)	(-0.33)	(0.08)	(-0.47)
Market Cap	0.0375	0.0590***	-0.0226	0.0817***
	(0.69)	(2.84)	(-1.46)	(3.86)
Fixed Effects	Firm,Year	Firm,Year	Firm,Year	Firm,Year
N	4,188	4,188	4,188	4,188
Adjusted R-Squared	0.845	0.630	0.576	0.676

As shown in Table 4, Panel A, Column 1, the introduction of FAS123(R) (that is, *Treated × 2005–2006*) is associated with a statistically significant reduction in *CEO Pay*. In particular, as shown in Column 2, the reduction in *CEO Pay* is attributable to a statistically significant reduction of *Equity Portion*, which, unsurprisingly, can be exclusively attributed to a statistically significant decline of *Option Portion* (Column 4).[185] Indeed, while *Stock Portion* increases (at an almost significant level), such an increase is not sufficient to offset the decline in *Option Portion*, explaining the overall reduction in *CEO Pay*.

Next, Table 4, Panel B shows the results for the second step of our analysis, considering how firm value changed from before and after the introduction of FAS123(R), depending on one-year lagged changes in CEO compensation.

185. *See supra* subpart III(B) (providing a description of the time trends in executive compensation over the period 1993–2015).

Table 4, Panel B

CEO Pay and Firm Value Around the Introduction of FAS123(R)

In this table, we present pooled panel *Q* regressions on one-year lagged values of proxies for the level and structure of the CEO's compensation package (*CEO Pay*, *Equity Portion*, *Stock Portion*, and *Option Portion*; see Appendix Table A for descriptions) and their interaction with *Treated* and *Treated* × *2005–2006*. The time period for the independent (lagged) variables is 2003–2006. *Treated* is an indicator variable equal to one if the firm awarded options to the CEO in 2003 or 2004. We include *Staggered Board*, *Poison Pill*, and the set of controls (*Institutional Ownership*, *Assets*, *Stock-Return Volatility*, *Leverage*, *Delaware Incorporation*, *Profitability*, *CAPX/Assets*, and *R&D/Sales*). See Appendix Table B for descriptions. We do not show the coefficients on the controls to save space. All specifications include year and firm fixed effects (not shown). Statistical significance of the coefficients is indicated at the 1%, 5%, and 10% levels by ***, **, and *, respectively, based on robust standard errors clustered by firm.

Dep. Variable: Firm Value (Tobin's *Q*)			
Variables (Lagged)	(1)	(2)	(3)
Treated × *2005–2006*	0.0350	-0.151	-0.0750
	(0.18)	(-0.78)	(-0.37)
CEO Pay	-0.0286	-0.0819	-0.0830
	(-0.34)	(-0.85)	(-0.85)
CEO Pay × *Treated*	0.114	0.151	0.155
	(1.32)	(1.49)	(1.52)
CEO Pay × *Treated* × *2005–2006*	-0.0192	0.0228	0.0125
	(-0.85)	(0.94)	(0.50)
Equity Portion	-0.105*	0.109	
	(-1.68)	(0.54)	
Equity Portion × *Treated*		-0.0977	
		(-0.45)	
Equity Portion × *Treated* × *2005–2006*		-0.302***	
		(-3.22)	
Stock Portion			0.0684
			(0.25)
Stock Portion × *Treated*			-0.212
			(-0.73)
Stock Portion × *Treated* × *2005–2006*			-0.0434
			(-0.36)
Option Portion			0.164
			(0.76)
Option Portion × *Treated*			-0.143
			(-0.63)
Option Portion × *Treated* × *2005–2006*			-0.357***
			(-3.43)
Fixed Effects	Firm,Year	Firm,Year	Firm,Year
N	4,188	4,188	4,188
Adjusted R-Squared	0.893	0.893	0.894

As shown in Table 4, Panel B—and consistent with our identifying assumption above—the introduction of FAS123(R) (that is, *Treated* × *2005–2006*) did not directly affect firm value (Columns 1 through 3). Similarly, *CEO Pay* per se did not affect firm value in our period (Columns 1 through

3). Further, as consistently shown by the interaction coefficient of *CEO Pay* × *Treated* × *2005–2006* (Columns 1 through 3), the changes in compensation levels attributable to the change in the accounting rule did not affect firm value. Consistent with the results we obtain in our value analysis in subpart IV(A) above, these results suggest that the overall level of compensation before the introduction of FAS123(R) was set "efficiently" on average, in the sense that compensation levels before the accounting rule modification do not seem to reflect "excessive rents" that fail to promote incentives for firm value maximization, as predicted by managerial power theory.[186]

Conversely, the triple interaction *Equity Portion* × *Treated* × *2005–2006* (in Column 2) has a large negative and statistically significant coefficient of –0.302. This means that the more firms affected by the introduction of FAS123(R) reduced the use of equity-based pay in the two years after the rule change, the greater the increase in firm value. Economically, the coefficient of *Equity Portion* × *Treated* × *2005–2006* indicates that a standard deviation decline in *Equity Portion* in affected firms is associated with an increase in firm value of 4.4%.[187] To the extent that these reductions in equity-based pay were caused by the introduction of FAS123(R) and were thus exogenous to firm-specific circumstances, this result points to an inefficiency in executive-compensation *structure*, indicating that firms had been paying their CEOs with excessive proportions of equity-based pay prior to the rule change.[188]

Most importantly, in Column 3, we find that the negative association between equity-based pay and firm value is driven by the negative association between option-based pay and firm value, rather than the association between stock-based pay and firm value. Specifically, the triple

186. One could argue that the changes induced by the new rule all flow from the reduced use of options, so that it would not be possible to disentangle the effects of reduced CEO pay and the reduced use of options because those reductions coincide in the data. Yet, if compensation levels (and not just *Option Portion*, which is a percentage measure) before the new rule were excessive in the sense defended by managerial power scholars, should we not observe a mechanical effect on firm value (i.e., an increase in value) after the new rule? Conversely, *CEO Pay* × *Treated* × *2005–2006* does not have a negative association with firm value.

187. This percentage is calculated by multiplying the coefficient on *Equity Portion* × *Treated* × *2005–2006*, –0.302, by the standard deviation of *Equity Portion* (2.777, see Appendix Table C) and then dividing the product by the average *Q* in the sample (1.89, see Appendix Table C).

188. The double interaction *Equity Portion* × *Treated* (shown in Column 2) has a small negative and insignificant coefficient of –0.0977 (*t*-statistic of 0.45). Therefore, the reductions in equity-based pay following the introduction of FAS123(R) were not associated with changes in performance in the two years *before* the rule change—when such reductions would have been more likely to be due to firm-specific circumstances rather than the upcoming rule change. Instead, we observe such association only in the two years *after* the rule change. Therefore, we document a discontinuity in the association between *Equity Portion* and firm value. This suggests that changes in the proportion of equity-based pay *before versus after* the rule change were caused by different factors, consistent with a causal interpretation of our event study.

interaction *Stock Portion* × *Treated* × *2005–2006* has a small and insignificant coefficient, while the triple interaction *Option Portion* × *Treated* × *2005–2006* has a large negative and statistically significant coefficient of –0.357 (*t*-statistic of 3.43). Economically, this coefficient indicates that a standard deviation decline in *Option Portion* in the firms affected by the accounting rule change is associated with an increase in firm value of 5.1%.[189]

Thus, only reductions in option-based pay were associated with increases in firm value for the subset of firms affected by the introduction of FAS123(R).[190] This is consistent with a plausible causal interpretation of our event study, as Panel A of Table 4 documents that treated firms affected by the accounting rule change experienced a reduction in the use of option-based pay relative to control firms unaffected by such change, while the use of stock-based pay remained basically unchanged.[191]

To sum up, the main takeaway from our event study is that option-based pay seems to have been overused in recent years, which raises the question of what might have determined this distortion. Our results that greater product-market competition and shareholder pressure are associated with an increase in option-based pay (see results for *M&A Competition* in Table 2),[192] as well as the negative association of such market forces with firm value (see Table 3), point to a plausible answer. Taken together, and consistent with our theoretical predictions in subpart II(C), these results seem to indicate that the

189. This percentage is calculated by multiplying the coefficient on *Option Portion* × *Treated* × *2005–2006*, –0.357, by the standard deviation of *Option Portion* (0.269, see Appendix Table C) divided by the average *Q* in the sample (1.89, see Appendix Table C).

190. A possible concern with our specification could be that firms in our control group may not be perfectly comparable to the treated firms affected by the accounting rule change. For example, our control firms could operate in different industries than our treated firms. More generally, as options are more likely to be used in certain industries than others, it is possible that industry-level shocks could partly explain our performance results. In order to mitigate this concern, we run the following robustness check. For each of the specifications in Panel B of Table 4, we add time-varying industry fixed effects by interacting two-digit SIC industry groups with year fixed effects. This adds about 200 more dummies to each specification. Our results are robust to the addition of such *higher-order fixed effects*. More specifically, once we control for higher-order fixed effects, the interacting variable *Equity Portion* × *Treated* × *2005–2006* in Column 2 of Panel B of Table 4 has a coefficient of –0.271 (with a *t*-statistic of 2.74), and the interacting variable *Option Portion* × *Treated* × *2005–2006* in Column 3 has a coefficient of –0.324 (with a *t*-statistic of 2.95). These results are very similar in terms of statistical and economic significance to those appearing in Panel B of Table 4.

191. Figure 2 documents that over 2003–2006, firms generally were increasing their proportions of stock-based pay. Consistently, the results in Panel A of Table 4 show that control firms were on average increasing their use of stock-based pay in 2005 and 2006 as much as the treated firms. Therefore, the overall trend in Figure 2 masks an important diversion in 2005 and 2006 between firms affected and unaffected by the introduction of FAS123(R): both groups of firms increased their use of stock-based pay, but only the firms affected by the accounting rule change substantially lowered their overall use of equity-based pay, on average.

192. It is worth emphasizing again that *M&A Competition* is a measure at industry level; therefore, it substantially reduces endogeneity concerns. *See supra* note 135.

overuse of option-based pay can be at least partly attributed to distortions arising from excessive market and shareholder pressure. Indeed, boards of firms that were more exposed to such market forces may have felt forced to substantially increase option-based compensation that overemphasizes short-term performance, regardless of the long-term effects of such incentive design. To this extent, the results we obtain in Table 4, Panels A and B also support the hypothesis that long-term value creation requires a strong commitment from *both* managers *and* shareholders towards long-term cooperation. The results do so by indicating that such cooperation is less likely to take place in firms that are subject to excessive market pressure, as such pressure challenges a board's ability to provide a credible commitment to stable managerial relationships, in turn jeopardizing efficient incentive design.

V. Going Forward

The above discussion has exposed the managerial power account of high executive pay as a reflection of inefficient rent extraction as both theoretically and empirically wanting, documenting empirical evidence that such pay is generally necessary to attract talented executives and increase firm value. It has further shown that protecting boards from short-term market and shareholder interference seems to lead to a more positive relationship between CEO pay and firm value. From a normative perspective, as we explain in this Part, these results carry crucial insights for the social welfare implications of the executive-compensation debate.

A. Rethinking "Entrenchment" Rents

Managerial power theory has been very influential in promoting the view that shareholder empowerment is the main solution to improve executive-compensation policies. In practice, this view has resulted in major executive-compensation reforms, including the introduction by the SEC of new rules on the disclosure of compensation policies, and most importantly, the provision of say-on-pay shareholder votes.[193] Yet, according to managerial power scholars, such reforms can at best help ameliorate the alleged inefficiencies of executive pay but not fully solve them. To this end, they argue, it is also necessary to make directors dependent on shareholders by changing the legal arrangements that insulate boards from market discipline.[194] Therefore, these scholars' more ambitious reform agenda focuses on introducing new rules to eliminate the staggered board and other

193. *See supra* notes 86–88 and accompanying text.

194. BEBCHUK & FRIED, *supra* note 7, at 11.

defensive measures, in addition to removing the procedural obstacles that favor incumbents' access to the corporation's proxy machinery.[195]

The theoretical and empirical analysis developed in this Article, however, implies that the above reform agenda is undesirable. Indeed, the managerial power scholars' critique of entrenchment rents is grounded on reductionist assumptions about the transactional environment in which the relationships between boards, managers, and shareholders take place.[196] When examined against more realistic assumptions about the constraints arising from competitive market forces in a dynamic transactional context, allegedly inefficient entrenchment rents may help commit managers to long-term value creation, serving to offset both the distortions arising from increased talent competition and managerial incentives for inefficient intertemporal tradeoffs.[197] Therefore, measures designed to eliminate such rents should be regarded as counterproductive.

More concretely, the results of our event study in subpart IV(B) directly contradict the central policy recommendation of managerial power scholars that option grants should be preferred to restricted stock grants, using the argument that options are less costly to shareholders. This account of options overlooks the drawbacks that this form of compensation involves for long-term shareholder wealth. Indeed, the asymmetric payoffs of options may give managers incentives for taking risks that increase the likelihood that the stock price at the option's exercise date will be higher than the option's strike price, even if taking such risks may be detrimental to long-term firm value.[198] Conversely, the linear payoff of restricted stock helps mitigate short-termist incentives. Most important for the purpose of this discussion, our event study supports the view that the short-termist costs of options exceed their benefits, showing that the decreased use of option grants that followed the mandate to expense such pay components was associated with increased firm value.

Therefore, the position of managerial power scholars that compensating managers through restricted stock, rather than options, reflects the extraction of inefficient rents is invalid. Our results indicate that less skewed incentive plans better serve shareholder interests, as such plans help mitigate the distortions that may arise from excessive market and shareholder pressure for short-term value. The introduction of FAS123(R)—and the new trend it promoted towards the use of less option-based and more stock-based pay—should accordingly be regarded as helpful to improve executive-compensation practices.

---

195. *See id.* at 207–13.

196. *See supra* subpart I(B).

197. *See supra* section II(B)(2).

198. *See supra* text accompanying notes 114–15.

B. Rethinking Board Authority vs. Shareholder Empowerment

More broadly, our analysis challenges the call of managerial power scholars for reforming corporate governance to eliminate a board's defensive measures and empower shareholders. Contrary to what is argued by such scholars, the above discussion has shown that temporary board protection from short-term market discipline makes it more likely that the necessary conditions for a bilateral commitment toward long-term value creation by both shareholders and managers will be in place, in turn promoting efficient incentive design. In other words, when examined in the context of the informational asymmetry existing between firm insiders and outsiders and the consequential limited commitment problem of shareholders, the traditional board-centric model of the corporation emerges as economically rational. On the contrary, making directors more dependent on shareholders is likely to jeopardize the ability of that model to deliver efficient outcomes,[199] including value-increasing pay schemes.

We accordingly defend the traditional approach of the Delaware courts to executive-compensation matters as normatively desirable. Delaware's approach has consistently accorded boards of directors a great deal of authority in setting executive pay by granting directors the protection of the business judgment rule.[200] This approach is thus consistent with the evidence provided in this Article that, in general, boards of directors efficiently design executive pay to attract the most talented executives, as well as the evidence that protecting boards from the interference of shareholders (including by means of shareholder litigation) helps promote a positive relationship between CEO pay and firm value.

Conversely, we challenge the current approach toward executive compensation of influential proxy advisory firms. These firms provide investors with voting recommendations on virtually any matter on which shareholders vote, including say-on-pay votes, playing today a major role in influencing corporate governance policies at many U.S. corporations.[201]

---

199. *See* Cremers & Sepe, *supra* note 16, at 142 (similarly defending the rationality of the traditional board-centric model on economically grounded reasons); Leo E. Strine, Jr., *Toward a True Corporate Republic: A Traditionalist Response to Bebchuk's Solution for Improving Corporate America*, 119 HARV. L. REV. 1759, 1777–82 (2006) (illustrating how a traditionalist would defend the "republican" board-centric model of the corporation against the proposals by shareholder advocates to move to a "direct democracy" model); William W. Bratton & Michael L. Wachter, *The Case Against Shareholder Empowerment*, 158 U. PA. L. REV. 653, 658–60 (2010) (defending the received board-centric model of the corporation based on the board's informational advantage).

200. *See, e.g.*, Freedman v. Adams, 58 A.3d 414, 417 (Del. 2013) (rejecting a claim for waste alleging that the board had failed to adopt a compensation plan under § 162(m) of the Internal Revenue Code and reaffirming that decisions concerning both the level and structure of executive compensation are protected by the business judgment rule).

201. David F. Larcker et al., *The Influence of Proxy Advisory Firm Voting Recommendations on Say-on-Pay Votes and Executive Compensation Decisions*, CONF. BOARD, Mar. 2012, at 1, 1–2,

While say-on-pay votes are nonbinding, low approval rates tend to bring about increased scrutiny of a firm's corporate governance practices by activist shareholders, the media, and proxy advisory firms.[202] Importantly, this attention risks not just embarrassing directors, but fueling activist hedge fund campaigns, which may result in major disruptions to a company's management.[203] Additionally, in a growing number of cases, negative shareholder votes on executive plans have triggered derivative shareholder lawsuits.[204]

In their recommendations, proxy advisory firms consider features that focus on the link between pay and performance and peer groups in benchmarking executive pay.[205] This is troubling for two reasons. First, the focus on shareholder returns may promote what has been called "earnings hysteria,"[206] inducing boards to adopt compensation that overemphasizes short-term performance. Consistent with this prediction, in a recent survey, companies reported making broad changes to executive-pay packages in response to advisory recommendations, with such changes affecting all areas of compensation programs and, in particular, compensation structures.[207] Second, peer-group benchmarking may also introduce distortions in executive pay, as under standard asymmetric information assumptions between firm insiders and outsiders,[208] the stock price might be temporarily off even in firms with talented management and good financial performance. Nevertheless, as put by one commentator, under the current influence exerted by proxy advisors on executive pay, temporary disappointing earnings may

---

https://www.gsb.stanford.edu/sites/gsb/files/publication-pdf/cgri-survey-2012-proxy-voting_0.pdf [https://perma.cc/PKH2-TDHQ]. This is especially true for Institutional Shareholder Services (ISS), which advises clients that manage over $25 trillion in assets. *See* Robert D. Hershey, Jr., *A Little Industry with a Lot of Sway on Proxy Votes*, N.Y. TIMES (June 18, 2006), http://www.nytimes.com/2006/06/18/business/yourmoney/18proxy.html?mcubz=3 [https://perma.cc/6MW8-893C] (describing ISS as the most prominent advisory firm). Further, ISS also offers consulting services on executive-compensation programs, which are designed to increase a firm's likelihood of receiving a favorable say-on-pay vote. Larcker et al., *supra*, at 201.

202. Larcker et al., *supra* note 201, at 1–2.

203. *See* Martijn Cremers, Saura Masconale & Simone M. Sepe, *Activist Hedge Funds and the Corporation*, 94 WASH. U. L. REV. 261, 270, 292–94 (2017) (describing the rise of activist hedge funds and providing real-world examples of hedge funds' governance activism).

204. *See* Larcker et al., *supra* note 201, at 2.

205. *Id.*

206. Anders Melin, *'Earnings Hysteria' Pits ISS Against Clinton and Fink on CEO Pay*, BLOOMBERG (Aug. 23, 2016), https://www.bloomberg.com/news/articles/2016-08-23/-earnings-hysteria-pits-iss-against-clinton-and-fink-on-ceo-pay [https://perma.cc/PUX3-9BLV] (reporting a statement by Larry Fink, CEO of BlackRock, the largest asset management firm in the world).

207. *See* Larcker et al., *supra* note 201, at 3, 5 (documenting these changes, which, among others, included enhanced disclosure, reduced severance benefits, and the introduction of performance-based equity awards).

208. *See supra* note 120.

mean that "you fall below the peer group's median return, fail ISS's quantitative test and all of a sudden 30% of shareholders go against you."[209]

It follows that policymakers and corporate actors would do well to reconsider the current case for enhanced shareholder power in corporate governance and the executive-pay process. Claiming that board protection from market discipline produces, among other drawbacks, inefficient CEO pay arrangements, managerial power scholars and other shareholder advocates have successfully advocated for a shift in corporate power from boards of directors to shareholders since the early 2000s. The introduction of say-on-pay votes is just one example of the several regulatory changes that have contributed to shareholder empowerment in the past fifteen years. Other significant changes have included, for example, modifications to proxy filing requirements that have facilitated the use of shareholder proposals[210] and amendments to the Delaware General Corporation Law that have granted shareholders greater access to the ballot box.[211] These regulatory reforms were also accompanied by changes in capital markets and corporate practices, including not just the rise of activist hedge funds[212] and proxy advisory firms,[213] but also the introduction of "universal" majority voting and accompanying withholding campaigns,[214] and the growing use and success of shareholder proposals.[215]

As a result of these developments, board power on corporate affairs, including the executive-pay process, has been significantly eroded. The decline in the use of staggered boards provides a vivid example of this erosion of board authority. While directors theoretically retain a veto power over destaggering decisions (as long as the board is established in the charter),[216]

---

209. *See* Melin, *supra* note 206 (reporting a statement by Ira Kay, managing partner at compensation consultant Pay Governance).

210. *See* Marcel Kahan & Edward Rock, *Embattled CEOs*, 88 TEXAS L. REV. 987, 1013–15, 1017–22 (2010) (providing a thorough discussion of the changes that have occurred in proxy rules in the past ten to twenty years); *see also supra* note 118 (summarizing Rule 14a-8, which governs shareholder proposals).

211. *See, e.g.*, DEL. CODE ANN. tit. 8, § 112 (2011) (providing that a company's bylaws may give shareholders the right to nominate dissident slates of directors); *id.* § 113(a) (permitting a company's bylaws to "provide for the reimbursement by the corporation of expenses incurred by a stockholder in soliciting proxies in connection with an election of directors").

212. *See supra* note 203.

213. Larcker et al., *supra* note 201, at 1.

214. Under majority voting, only nominees who receive a majority of the votes cast are elected to the board. As a result, vote-withholding campaigns—involving the withholding of votes on specific governance issues, including the election of directors—have become potent weapons in the arsenal of activist shareholders, since shareholders can effectively use this process to throw incumbents out of office without having to file a proxy statement with the SEC. *See* Leo E. Strine, Jr., *Toward Common Sense and Common Ground? Reflections on the Shared Interests of Managers and Labor in a More Rational System of Corporate Governance*, 33 J. CORP. L. 1, 11–12 (2007).

215. Kahan & Rock, *supra* note 210, at 1038.

216. Charter amendments can only be initiated by the board and require shareholder approval. *See, e.g.*, DEL. CODE ANN. tit. 8, § 242(b) (2011); MODEL BUS. CORP. ACT § 10.03 (AM. BAR ASS'N

in recent years boards have increasingly acquiesced to destaggering proposals[217] under the pressure exerted by both proxy advisors[218] and activist shareholders.[219] Accordingly, the ability of boards to effectively use defensive measures to gain temporary protection from shareholder and market interferences is significantly diminished in the current corporate environment, potentially jeopardizing boards' ability to design efficient pay schemes.

In response, we advocate measures that can restore the defensive value of staggered boards and other board protections, such as, for example, the requirement of supermajority requirements for dismantling such protections. As we explained in more detail in prior work, similar measures would help re-empower U.S. boards vis-à-vis shareholders and thereby help promote the stability that is necessary for long-term firm value creation.[220] Among other benefits, this institutional stability would help improve executive-compensation practices, which is in the long-term interest of shareholders and society as a whole. At the same time, to address concerns that board protection might become perpetual—and hence equally detrimental to efficient compensation design and shareholder interests—one of us has also explained elsewhere that board defenses should be designed to have a finite, predetermined life, with the option for shareholders to approve their extension for subsequent periods.[221]

While we are aware of the practical difficulties that the implementation of these proposals may encounter, it is noteworthy to observe that there have been recent signs that a demand for recalibrating corporate governance arrangements—including executive-pay schemes—toward the pursuing of long-term firm value seems to be emerging among the largest institutional

---

2010). When the staggered board is established in the bylaws, instead, shareholders can unilaterally dismiss it, as board initiative is not required for bylaw amendments. *See, e.g.*, DEL. CODE ANN. tit. 8, § 109(a); MODEL BUS. CORP. ACT § 10.20.

217. *See* Cremers & Sepe, *supra* note 16, at 98–99 (documenting that the decline in staggered boards over time is largely due to increased destaggering rather than a fall in staggering events).

218. The recommendation that companies should have a unitary board, or else shareholders should seek a destaggering proposal, figures among the most important voting guidelines that proxy advisors routinely provide to investors. *See, e.g.*, INSTITUTIONAL S'HOLDER SERVS., 2014 U.S. PROXY VOTING SUMMARY GUIDELINES 10, 17 (2013), https://www.issgovernance.com/file/files/2014ISSUSSummaryGuidelines.pdf [https://perma.cc/3E5U-LML3].

219. *See, e.g.*, Cremers & Sepe, *supra* note 16, at 92 (discussing the intense destaggering activity of the Harvard Shareholder Rights Project, a clinical program established at Harvard Law School to assist institutional investors in the submission of destaggering proposals).

220. *See* Cremers, Masconale & Sepe, *supra* note 55, at 771 (discussing supermajority requirements to amend a firm's charter or a firm's provisions for approving mergers); Sepe, *supra* note 121, at 1437–38 (suggesting "that a charter provision requiring the approval of a supermajority—two-thirds or more—of the shareholders for the dismissal of defensive tactics would be beneficial").

221. Sepe, *supra* note 121, at 1437–39.

investors.[222] Hopefully, the framework of analysis offered in this Article, and the conclusion it achieves, will prove useful to support this demand, while also providing policymakers with tangible reasons for reconsidering the current direction of corporate governance and executive-compensation policies.

Conclusion

"Managerial power theory"—the view that structural flaws in corporate governance, such as board defenses, enable managers to extract inefficient entrenchment rents—has risen to the forefront of the executive-compensation debate. The executive-compensation reforms that have taken place in recent years, such as the introduction of say-on-pay shareholder votes, provide perhaps the clearest evidence of this theory's enormous influence. The common denominator of such reforms is their attempt to increase shareholder power vis-à-vis directors, consistent with the central claim of managerial power scholars that only by empowering shareholders can directors be made truly accountable and executive-compensation practices be improved.

In spite of having gained the upper hand, managerial power theory surprisingly lacks robust empirical testing. Equally surprising is the failure of managerial power scholars to thoroughly confront the currently prevailing *economic* paradigm of executive compensation: "managerial talent theory," according to which high executive pay reflects compensation for scarce managerial talent in competitive markets.

This Article remedies these failures, reviewing the main arguments of managerial power theory both theoretically and empirically and finding them wanting. Supporting the managerial talent view of executive pay, we conclude that high executive pay is generally consistent with optimal board contracting towards the attraction and retention of scarce, talented managers, rather than with opportunistic rent extraction by managers.

Our study shows that once one incorporates competitive market forces and the dynamic nature of managerial employment relationships into the analysis, granting managers high rents emerges as beneficial, rather than detrimental, to shareholder interests. Indeed, paying managers these rents helps ensure that a manager's continuation payoff (that is, the expectation of future compensation) can be "exploited" as a bonding mechanism to commit the manager to the creation of long-term firm value. There are two reasons. First, the expectation of high future rents helps offset the potential distortions that a manager's chances at mobility may introduce in managerial incentive schemes within a competitive environment. Second, expected high rents mitigate the short-termist incentives that managers may develop in the context of dynamic employment relationships, as the prospect of losing these

222. *Id.* at 1440–41.

rents disincentivizes managers to undertake strategies that boost short-term performance at the expense of future losses.

Within this analytical framework, we show that board protection from market discipline, whether in the form of intense product-market competition or greater shareholder power, is likely to increase, rather than decrease, the efficiency of executive-compensation plans. Board protection does so by reducing the risk that market discipline will interfere with a manager's continuation value, thereby making shareholders' own commitment to long-term value creation credible and avoiding value-decreasing, short-termist distortions in executive-pay schemes. The evidence we provide on the overuse of option grants in the past decade offers a vivid illustration of such distortions, as this evidence suggests that boards that are more exposed to market forces are more likely to use option-based compensation that emphasizes short-term performance at the expense of long-term firm value.

Going forward, policymakers would do well to reconsider the case for enhanced shareholder power in corporate governance, which has driven recent executive-compensation reforms and which has boosted the influence of proxy advisory firms on the executive-pay process. As this Article has shown, this case is theoretically lacking, unsupported by the data, and seems detrimental to both the interests of shareholders and society as a whole.

Future research would also do well to reexamine proposals to reform executive pay in the attempt to address the rising income inequality in the United States.[223] Our analysis suggests that blaming boards of directors and managers for this rising inequality[224] reflects a fundamentally reductionist understanding of the dynamics that have led to the rise in CEO pay. Accordingly, proposals that penalize firms for high executive pay will

223. The clearest evidence is provided by the 2010 Dodd-Frank Act's provision directing the SEC to introduce a CEO-pay-ratio disclosure requirement. Dodd-Frank Wall Street Reform and Consumer Protection Act of 2010, Pub. L. No. 111-203, § 953(b), 124 Stat. 1376, 1899 (2010) (codified as amended at 15 U.S.C. § 78n-1). Under this requirement, all U.S. public companies will have to disclose the ratio of CEO pay to median employee pay starting in 2018 proxy statements (reporting on fiscal year 2017). *See* 17 C.F.R. § 229.402 (Executive Compensation). The theory behind this impeding mandate is that making the CEO pay ratio publicly available will make board capture by executives more difficult—or, at least, more evident to firm outsiders, including shareholders—ultimately leading to a reduction of both CEO pay and income inequality. *See* Ira Kay & Blaine Martin, *CEO Pay Ratio and Income Inequality: Perspectives for Compensation Committees*, HARV. L. SCH. F. ON CORP. GOVERNANCE & FIN. REG. (Oct. 25, 2016), https://corpgov.law.harvard.edu/2016/10/25/ceo-pay-ratio-and-income-inequality-perspectives-for compensation-committees/ [https://perma.cc/9LR7-46CU].

224. One of the most prominent advocates of this view has been French economist Thomas Piketty. In his *book du jour*, *Capital in the Twenty-First Century*, Piketty argues that "supermanagers, that is, top executives of large firms who have managed to obtain extremely high, historically unprecedented compensation packages for their labor," have come nowadays to make up most of the income hierarchy's top 0.1%. THOMAS PIKETTY, CAPITAL IN THE TWENTY-FIRST CENTURY 302 (2014). In searching for a rational basis for the explosion in managerial pay, Piketty suggests that one plausible explanation is that "social norms" have allowed senior managers to substantially set their own pay. *See id.* at 332.

perhaps placate the populist indignation over CEO compensation, but are unlikely to effectively address related inequality issues.

Rather than pursuing such proposals, it would be desirable to focus on exploring the many questions that still remain unanswered about the link between executive pay and labor income inequality. In particular, under this Article's analysis, the neoclassical question resurfaces of whether the increased demand for a super-skilled workforce might at least in part explain the widening labor income gap. Similarly, the question arises of whether a link exists between labor income inequality and sets of social norms other than those allegedly favoring managerial opportunism. Beginning to answer these questions will hopefully produce reform interventions that can effectively tackle inequality, rather than just feed populism.

Appendix Table A
Definitions of CEO Compensation Variables

Appendix Table A presents brief definitions of the CEO compensation variables that appear in the analysis. All variables are winsorized at one percent in both tails.

CEO Compensation Variables:

CEO Pay	Natural logarithm of total CEO compensation in that fiscal year, using the measure 'TDC1' in ExecuComp, which defines CEO Pay as the sum of salary, bonus, stock grants, the Black-Scholes value of new option grants, plus any other deferred compensation, including pension benefits.
Equity Portion	The percentage of total CEO compensation (measured by 'TDC1' in ExecuComp) that is awarded to the CEO in the fiscal year that comes in the form of stock grants ('rstkgrant' in ExecuComp until 2006, 'stock_awards_fv' in ExecuComp after 2006) or option grants ('option_awards_blk_value' in ExecuComp until 2006, 'option_awards_fv' in ExecuComp after 2006).
Stock Portion	The percentage of total CEO compensation (measured by 'TDC1' in ExecuComp) that is awarded to the CEO in the fiscal year that comes in the form of stock grants ('rstkgrant' in ExecuComp until 2006, 'stock_awards_fv' in ExecuComp after 2006).
Option Portion	The percentage of total CEO compensation (measured by 'TDC1' in ExecuComp) that is awarded to the CEO in the fiscal year that comes in the form of option grants ('option_awards_blk_value' in ExecuComp until 2006, 'option_awards_fv' in ExecuComp after 2006).
PPS (Pay Performance Sensitivity)	Natural logarithm of (1 + Delta), where Delta measures how much the dollar value (in thousands of US$) of the CEO's stock and options held changes if the firm's stock price increases by 1%. Data comes from the website of Lalitha Naveen and was derived from ExecuComp data, and is updated from the data as previously used in Coles, Daniel and Naveen (2006).
PVS (Pay Volatility Sensitivity)	Natural logarithm of (1 + Vega), where Vega measures how much the dollar value (in thousands of US$) of the CEO's options held changes if the annualized volatility of the firm's stock returns increases by 1%. Data comes from the website of Lalitha Naveen and was derived from ExecuComp data, and is updated from the data as previously used in Coles, Daniel and Naveen (2006).

Appendix Table B
Definitions of Other Variables

Appendix Table B presents brief definitions of the main variables that appear in the analysis other than the CEO compensation variables described in Appendix Table A. All continuous variables are winsorized at one percent in both tails.

Variables:	
Q	Tobin's *Q*, defined as the Market value of assets (i.e., Total Assets – Book Equity + Market Equity) divided by the book value of assets. Calculation follows Fama and French (1992). Source of data is Compustat annual data file.
Staggered Board	Indicator variable equal to one (zero otherwise) if the board is staggered. Data is obtained from RiskMetrics, SharkRepellent.net and hand collection for 1990–2015.
Poison Pill	Indicator variable equal to one (zero otherwise) if the firm has a poison pill. Data is obtained from from RiskMetrics, SharkRepellent.net and hand collection for 1990–2015.
Institutional Ownership	The percentage of outstanding shares held by institutional owners. Data is from CRSP (number of outstanding shares) and Thomson Reuters (institutional ownership from 13F filings).
Post 2010	Indicator variable equal to one if the fiscal year is after 2010, and equal to zero otherwise.
Assets	Logarithm of the book value of total assets from Compustat.
Stock Return Volatility	Standard deviation of the daily stock return, from CRSP.
Leverage	Ratio of the book value of total debt to the book value of total assets, from Computat.
Delaware Incorporation	Indicator variable if the company is incorporated in Delaware, from historical CRSP header files.
Profitability	Accounting profitability, measured as the ratio of the book value of earnings before interest, taxes, depreciation and amortization (EBITDA) over the book value of total assets, from Compustat.
CAPX/Assets	Ratio of the book value of capital expenditures over the book value of total assets, from Compustat.

R&D/Sales	Ratio of the book value of R&D over the book value of total sales, from Compustat.
Product Competition	Negative of the Herfindahl index, which is a measure of industry concentration of the sales in the industry across all publicly traded firms in the industry, such that more industry concentration means a higher Herfindahl index and thus a lower level of *Product Market Competition*. Source of data is Compustat annual data file.
M&A Competition	The ratio of mergers & acquisitions' dollar volume in SDC to the total market capitalization from CRSP for a calendar year, per industry. The CRSP annual industry market capitalization is for ordinary stocks only and excludes ADRs and REITs. If no M&A activity per given industry-year is reported in SDC, we assume it to be zero. We only include transactions in SDC where the buyer achieves control of the target.
Talent Competition	Natural logarithm of the market capitalization of the outstanding shares of the 250th ranked firm in terms of market capitalization, from CRSP.
Market Cap	Natural logarithm of the market capitalization of the outstanding shares of the firm, from CRSP.

Appendix Table C
Descriptive Statistics

Appendix Table C presents sample descriptive statistics for the main variables. The sample consists of all ExecuComp firms without dual class stock in non-regulated industries.

Variable	Mean	Median	St. Dev.	Min	Max	Obs.
CEO Pay	8.048	8.066	0.998	6.001	10.036	19,572
Equity Portion	0.442	0.482	0.277	0.000	1.000	19,572
Stock Portion	0.164	0.000	0.228	0.000	1.000	19,572
Option Portion	0.277	0.231	0.269	0.000	1.000	19,572
PPS (Pay Performance Sensitivity)	5.519	5.499	1.418	2.302	8.655	18,653
PVS (Pay Volatility Sensitivity)	3.862	4.092	1.740	0.000	6.796	19,119
Tobin's Q	1.890	1.524	1.186	0.482	11.714	19,572
Classified Board	0.565	1.000	0.496	0.000	1.000	19,572
Poison Pill	0.469	0.000	0.499	0.000	1.000	19,572
Institutional Ownership	0.717	0.745	0.178	0.001	0.999	17,984
Assets	7.612	7.444	1.561	3.862	12.444	19,572
Stock Return Volatility	0.026	0.023	0.012	0.009	0.063	19,349
Leverage	0.480	0.478	0.212	0.118	0.958	19,514
Delaware Incorporation	0.629	1.000	0.483	0.000	1.000	19,572
Profitability	0.142	0.138	0.092	-0.402	0.472	19,572
CAPX/Assets	0.053	0.038	0.051	0.000	0.344	19,572
R&D/Sales	0.044	0.001	0.114	0.000	3.338	19,572
Product Competition	0.075	0.059	0.055	0.018	0.310	19,572
M&A Competition	0.027	0.009	0.070	0.000	1.175	19,572
Talent Competition	9.653	9.788	0.417	8.378	10.289	19,572
Market Cap	7.636	7.518	1.594	0.980	13.348	19,572

International Law in the Post-Human Rights Era

Ingrid Wuerth*

Abstract

International law is in a period of transition. After World War II, but especially since the 1980s, human rights expanded to almost every corner of international law. In doing so, they changed core features of international law itself, including the definition of sovereignty and the sources of international legal rules. But what has been called the "age of human rights" is over, at least for now. Whether measured in terms of the increasing number of authoritarian governments, the decline in international human rights enforcement architecture such as the Responsibility to Protect and the Alien Tort Statute, the growing power of China and Russia over the content of international law, or the rising of nationalism and populism, international human rights law is in retreat.

The decline offers an opportunity to consider how human rights changed, or purported to change, international law and how international law as a whole can be made more effective in a post-human rights era. This Article is the first to argue that international human rights law as a whole—whatever its much disputed benefits for human rights themselves—appears to have expanded and changed international law in ways that have made it weaker, less likely to generate compliance, and more likely to produce interstate friction and conflict. The debate around international law and human rights should be reframed to consider these costs and to evaluate whether international law, including the work of the United Nations, should focus on a stronger, more limited core of international legal norms that protects international peace and security, not human rights. Human rights could be advanced through domestic and regional legal systems, through the the development of non-binding international norms, and through iterative processes of international reporting and monitoring—a model not unlike the Paris Climate Agreement.

---

*Professor of Law, Vanderbilt Law School. Feedback welcome: ingrid.wuerth@vanderbilt.edu. For helpful comments, thanks to Harlan Cohen, Jean d'Aspremont, Monica Hakimi, Larry Helfer, Neha Jain, Tim Meyer, Mike Newton, Eric Posner, Anthea Roberts, Brad Roth, Chris Serkin, Ganesh Sitaraman, Chris Slobogin, Kevin Stack, and Paul Stephan. Thanks also to participants in the 2017–2018 Cyber Colloquium and in faculty workshops at the Minnesota and Vanderbilt Law Schools. Jeremiah Cioffi, Chris Gracey, and Weike Guo provided excellent research assistance.

Introduction

During the past half-century, human rights have become central to international law. Substantive international human rights law expanded from a small core of initial protections after World War II to today's vast domain reaching from foreign direct investment to climate change. With this proliferation of norms came other, more fundamental changes. The meaning of sovereignty purportedly shifted to focus on the individual. The sources of international law changed so that international law now includes as "law" many norms, especially human rights norms, that are routinely violated. This Article argues that even as substantive human rights obligations continue to proliferate, the more fundamental transformations of international law through human rights have not fully taken hold and have proven costly to international law as a whole. Applications of a variety of interdisciplinary

tools, from empirical measures of the causes of war to domestic social psychology, suggest that human rights-related changes to both sovereignty and to the doctrine of sources tended to undermine interstate relations and the "territorial peace," and to make it harder to generate compliance with many norms of international law. These costs point to the conclusion that efforts to transform international law around human rights should be abandoned in favor of the development of a stronger core of international law dedicated to protecting international peace and security rather than human rights. At a minimum, the debate around human rights and international law should be reframed to consider the costs of human rights to international law as a whole.

The transformation of sovereignty and international law through human rights has stalled across several fronts, as described in Part I of this Article. The conceptual reorientation of sovereignty to focus on the individual or on humanity writ large provided—or was hoped to provide—the basis for doctrinal innovations in international law to allow for the coercive enforcement of international human rights law through foreign domestic courts, secession, and the use of force. Institutionally, the United Nations became increasingly focused on the enforcement of international human rights. The conceptual, doctrinal, and institutional aspects of the human rights enforcement architecture are all fading, however. Conceptually, traditional sovereignty is resurgent. Institutionally, the United Nations' human rights enforcement mechanisms appear to be largely ineffective despite decades of reform efforts, although here the trajectory may be more stagnation than decline. Doctrinally, innovations such as universal jurisdiction have encountered growing resistance and backlash. More broadly, human rights themselves appear to be in global decline.

The flagging efforts to transform international law and institutions around human rights raise an additional issue addressed in Part I: the costs of that transformation for international law. Methods of enforcing human rights through international law as a whole have the apparent effect of undermining the peaceful and friendly interaction of states, which is part (although not all) of what explains the failed efforts at transformation.[1] Key doctrinal innovations in international law, such as a lack of foreign state immunity in human rights cases, universal jurisdiction, a human right to democracy, remedial secession, and the responsibility to protect, all arose out of the human rights transformation of sovereignty. These enforcement doctrines were never widely adopted, were often applied politically, became associated with Western efforts to project power instead of universal values, and

1. This Article assumes that interstate cooperation is one objective of international law. *But cf.* Monica Hakimi, *The Work of International Law*, 58 HARV. INT'L L.J. 1, 5 (2017) (arguing that conflict is not necessarily a problem for international law to overcome). It also assumes that interstate cooperation, interstate peace, and compliance with international law provide benefits to human kind. Barriers to interstate cooperation and to compliance with international law, as well as war itself, thus impose what this Article terms "costs" on human kind.

undermined interstate cooperation. There may also have been a parallel institutional development: human rights enforcement may have contributed to polarization within the United Nations and made it less effective overall. Some of the doctrinal developments also undermine the territorial security provided by international law, which may threaten what political scientists term the "territorial peace."

Human rights did bring transformative changes to another core feature of international law: the doctrine of sources. As Part II of this Article describes, the effort to protect as many human rights as possible through law led to an expansion of the two primary sources of international law—treaties and custom. The fruits of this labor seem clear. From military intervention to gentrification and the Greek debt crisis, every international legal issue today is an international human rights issue.[2]

In practice, changes to the doctrine of sources have also meant, however, that human rights norms are deemed part of binding international law notwithstanding widespread violations and nonconforming behavior. Violations of international human rights law have become rampant, ranging from failures to comply with administrative reporting requirements to gross violations of human dignity. Again, there is a parallel development in the United Nations. The United Nations Security Council expanded its mandate to include redressing and preventing human rights violations, yet such violations continue to occur with alarming regularity. Others have questioned whether these developments are good for human rights, but their impact international law as a whole has gone unexamined.[3] Models of state and individual behavior, from rational choice theories, constructivism,

---

2. *See* Saki Knafo, *Is Gentrification a Human-Rights Violation?*, ATLANTIC (Sept. 2, 2015), https://www.theatlantic.com/business/archive/2015/09/gentrification-brooklyn-human-rights-violation/402460/ [https://perma.cc/T9QN-Y8LK]; Juan Pablo Bohoslavsky, *Greek Crisis: Human Rights Should Not Stop at Doors of International Institutions, Says UN Expert*, U.N. HUM. RTS. (June 2, 2015), http://www.ohchr.org/EN/NewsEvents/Pages/DisplayNews.aspx?NewsID=16032 [https://perma.cc/22GL-CQCB].

3. Other scholars have discussed the expansionist project of international human rights law, but generally without focusing on the potential effects on international law as a whole. *See, e.g.*, Jean d'Aspremont, *Expansionism and the Sources of International Human Rights Law*, 2016 ISR. Y.B. ON HUM. RTS. 223, 227–28 (describing the expansionist effect of international human rights law on the doctrine of sources and the paradoxes that this creates for international human rights law); Makau Mutua, *Is the Age of Human Rights Over?*, *in* THE ROUTLEDGE COMPANION TO LITERATURE AND HUMAN RIGHTS 450, 451 (Sophia A. McClennen & Alexandra S. Moore eds., 2015) ("Creeds and ideologies that overpromise—and inevitably underperform—are destined to suffer public fatigue. Human rights is one such ideology."); Jacob Mchangama & Guglielmo Verdirame, *The Danger of Human Rights Proliferation: When Defending Liberty, Less Is More*, FOREIGN AFF. (July 24, 2013), https://www.foreignaffairs.com/articles/europe/2013-07-24/danger-human-rights-proliferation [https://perma.cc/S3Y8-KHP7] (arguing that the expansion of international human rights law has diluted the value of human rights). *But cf.* Harlan Grant Cohen, *Finding International Law: Rethinking the Doctrine of Sources*, 92 IOWA L. REV. 65, 67 (2007) (noting that the expansion of human rights to include norms which are routinely violated provides "powerful fodder for those who believe that international law is meaningless").

organizational sociology, and social psychology, all show that widespread violations of some international legal rules likely make it more difficult to enforce others. We might call this a "broken windows" theory of international law.

The present decline of international human rights law—however broad and durable it may be—accordingly presents an opportunity to reframe the normative debate about human rights and international law, as discussed in Part III. That extensive and inconclusive debate has focused almost entirely on whether international law is an effective way to promote human rights. The broken windows problem, the effect of human rights enforcement on interstate relations, the empirical work on the "territorial peace," and problems with polarization and credibility all demonstrate, however, that human rights impose costs on the international legal *system*. Just as human rights may benefit from being part of an international legal system, this Article illustrates that the system itself may be weakened through the inclusion of human rights. Those costs suggest that international law should become focused on a core set of stronger, more effective norms that promote international peace and security, not human rights. They also suggest that we should explore with more creativity and vigor ways to advance human rights that do not rely upon the coercive enforcement of binding international law; proposals for doing so can draw on the important work of human rights scholars who have focused on domestic processes and insitutions and their relationship to international norms.

This Article can reframe the debate around international law and human rights, but it will not resolve it. The benefits of international human rights law for human rights are themselves contested, an issue that this Article discusses, but does not seek to resolve. A second important variable not fully explored in this Article is the extent to which human rights can today, thanks in part to the success of the international human rights movement, be enforced through domestic and regional law, and soft international commitments. This Article does not purport to resolve those issues; it endeavors instead to show that the debate about human rights should expand to consider the relationship between human rights and international law as a whole.

To begin, a few clarifications about terminology: "International human rights law" refers in this Article to international law governing the relationship between states and their own citizens; excluded are regional human rights instruments and most of international criminal law, although both fields see related developments,[4] and both are noted at various places in the Article.

---

4. Regional human rights systems and international criminal law are also undergoing some level of backlash and decline. *See, e.g.*, Mikael Rask Madsen, *The Challenging Authority of the European Court of Human Rights: From Cold War Legal Diplomacy to the Brighton Declaration and*

The terms "decline" and the "post-human rights era," do not mean we are at the end of human rights obligations themselves (indeed, they are multiplying), but instead that we have seen the end of an era: what Makau Mutua has termed the "age of human rights" or what Louis Henkin called the "Age of Rights" in international law.[5] Temporal baselines vary slightly from section to section, but for the most part the decline is measured from a baseline period in the late 1990s. In most of the doctrinal areas described here, decline began in the middle or end of the first decade of the twenty-first century.[6] In some sections stasis is a more accurate term than decline, and to some extent the transformations themselves were more aspirational than on-the-ground realities, so that perhaps the golden age was never so golden. Nor do the terms "decline" and "post-human rights era" preclude the possibility that international human rights enforcement is gaining ground in discrete areas, or that it will someday have a second golden age so that the process is really a dialectic, not heyday and crash. Nevertheless, contemporary political events and decades-long stasis and decline in enforcement architecture have come together to produce an especially difficult period for international human rights law. The Article focuses on the primary international legal enforcement mechanisms for human rights and on the United Nations because they are important to human rights, to international law, and to the claim that sovereignty has been transformed.

Finally, this Article is about the costs of human rights to international law as a whole. Its purpose is to advance our overall understanding of how best to strengthen international law while at the same time ensuring that people everywhere enjoy lives of dignity and well-being. It builds on the work of human rights scholars who have repeatedly questioned the

---

Backlash, 79 LAW & CONTEMP. PROBS., no. 1, 2016, at 167–68, 170; Karen Engle, *Anti-Impunity and the Turn to Criminal Law in Human Rights*, 100 CORNELL L. REV. 1069, 1119 (2015); Alexandra Huneeus, *Courts Resisting Courts: Lessons from the Inter-American Court's Struggle to Enforce Human Rights*, 44 CORNELL INT'L L.J. 493, 494 (2011); Somini Sengupta, *As 3 African Nations Vow to Exit, International Court Faces Its Own Trial*, N.Y. TIMES (Oct. 26, 2016), https://www.nytimes.com/2016/10/27/world/africa/africa-international-criminal-court.html [https://perma.cc/BBJ5-ZEBC]. *But cf.* Karen J. Alter et al., *A New International Human Rights Court for West Africa: The ECOWAS Community Court of Justice*, 107 AM. J. INT'L L. 737, 738 (2013). This Article discusses universal jurisdiction, which is one aspect of international criminal law, because there is no clear line between civil and criminal remedies in the universal-jurisdiction case law. *See* Donald Francis Donovan & Anthea Roberts, *The Emerging Recognition of Universal Civil Jurisdiction*, 100 AM. J. INT'L L. 142, 153 (2006).

5. Makau Mutua, *Is the Age of Human Rights Over?*, *in* ROUTLEDGE COMPANION TO LITERATURE AND HUMAN RIGHTS 450–58 (Sophia A. McClennen & Alexandra Schultheis Moore eds., 2016); *see generally* LOUIS HENKIN, THE AGE OF RIGHTS (1990).

6. Other scholars have asked whether there is a present "decline" in the "international rule of law," a term that describes changes in international law that accelerated during the 1990s and are based in part on a "liberal human rights vision." Heike Krieger & Georg Nolte, *The International Rule of Law—Rise or Decline?—Points of Departure* 8–9, 13 (KFG Working Paper Series, No. 1, 2016), https://papers.ssrn.com/sol3/Papers.cfm?abstract_id=2866940 [https://perma.cc/PAC7-YVXD].

effectiveness of various international enforcement mechanisms[7] and on scholarship analyzing how domestic politics and institutions, as well as iterative engagement with international institutions, contribute to the development and efficacy of human rights norms.[8] This Article provides additional reasons, beyond human rights themselves, to explore means other than binding international law through which human rights might be realized. Other authors have called for creativity and for the development of new frameworks to advance the cause of human rights[9]—this Article provides additional impetus for doing so.

I. International Human Rights: Costs to Peace and Friendly Relations

Human rights are often viewed as a foremost achievement of modern international law.[10] The protection and promotion of human rights are central to much of the work of the United Nations today, a striking expansion from both the interwar League of Nations and the immediate post-World War II period.[11] Human rights are even said to have transformed international law itself. International law was once understood as a discrete set of rules derived from consent of sovereign states and designed to facilitate their peaceful interaction.[12] Today, by contrast, sovereignty and international law purportedly derive their legitimacy from individuals.[13]

---

7. *See, e.g.*, *infra* text at notes 14–17, 185–90, 214–25.

8. *See, e.g.*, Cosette D. Creamer & Beth A. Simmons, *Do Self-Reporting Regimes Matter?: Evidence from the Convention Against Torture* 14–15 (Bos. Univ. Sch. of Law, Working Paper No. 15–55, 2016), http://scholar.harvard.edu/files/cosettecreamer/files/creamersimmons_catselfreporting_feb2016.pdf [https://perma.cc/4BXH-AX8M]; Gráinne de Búrca, *Human Rights Experimentalism*, 111 AM. J. INT'L L. 277, 299–304 (2017); Katerina Linos & Tom Pegram, *What Works in Human Rights Institutions?*, 112 AM. J. INT'L L. (forthcoming 2017).

9. *See, e.g.*, Samuel Moyn, *Beyond the Human Rights Measurement Controversy*, L. & CONTEMP. PROBS. (forthcoming 2018); Philip Alston, *Human Rights in the Populist Era*, JUSTSECURITY (Oct. 18, 2017), https://www.justsecurity.org/46049/human-rights-populist-era/ [https://perma.cc/XJ4U-K8QH].

10. *See* James Crawford & Marti Koskenniemi, *Introduction* to THE CAMBRIDGE COMPANION TO INTERNATIONAL LAW 1, 16 (James Crawford & Marti Koskenniemi eds., 2012); STEVE RATNER, THE THIN JUSTICE OF INTERNATIONAL LAW 177–78 (2015).

11. *See* STEVEN L.B. JENSEN, THE MAKING OF INTERNATIONAL HUMAN RIGHTS: THE 1960S, DECOLONIZATION AND THE RECONSTRUCTION OF GLOBAL VALUES 5–8, 37–47 (2016).

12. *See, e.g.*, Olivier Barsalou, *The Cold War and the Rise of an American Conception of Human Rights, 1945–8*, *in* REVISITING THE ORIGINS OF HUMAN RIGHTS 362, 364–65 (Pamela Slotte & Miia Halme-Tuomisaari eds., 2015); Krieger & Nolte, *supra* note 6, at 8; Prosper Weil, *Towards Relative Normativity in International Law?*, 77 AM. J. INT'L L. 413, 419 (1983).

13. *E.g.*, Harold Hongju Koh, *Humanitarian Intervention: Time for Better Law*, 111 AJIL UNBOUND 287, 288 (2017) (arguing in the context of humanitarian intervention that international law should "serve human purposes—including the protection of human rights, not just the territorial sovereignty of states"); RUTI TEITEL, HUMANITY'S LAW 8–11 (2011) (arguing that the "state-centered" vision of international law is being transformed by "a normative order that is grounded in the protection of humanity"); BRAD R. ROTH, SOVEREIGN EQUALITY AND MORAL DISAGREEMENT 91 (2011) (noting that "modern" sovereignty is often understood as circumscribed by human rights); Anne Peters, *Humanity as the A and Ω of Sovereignty*, 20 EUR. J. INT'L L. 513, 514 (2009) (arguing

Despite its ubiquity and its many successes, international human rights law is under growing attack. To begin, scholars question, with what appears to be increasing force and from a variety of perspectives, how effective international law actually is at safeguarding individual human rights;[14] new analyses have thrown into question the historical pedigree often claimed for international human rights law, including its "glorious, triumphalist" narrative;[15] and questions persist about the Western, imperial "civilizing mission" of the human rights movement.[16] A new study shows how the pursuit of human rights and other aspects of liberal internationalism undermined U.N. peacekeeping efforts.[17]

Second, according to many experts, human rights conditions are also in global decline. In the past decade, internationally protected civil and political rights have suffered a downturn as measured by a number of states experiencing a decline in rights protection.[18] The number of countries in

---

that "the normative status of sovereignty is derived from humanity," that "humanity is the A and Ω of sovereignty," and that "sovereignty remains foundational only in a historical or ontological sense"); W. Michael Reisman, Editorial Comments, *Sovereignty and Human Rights in Contemporary International Law*, 84 AM. J. INT'L L. 866, 876 (1990) (finding that it is no longer a "*per se*" violation of state sovereignty when some states make "externally motivated actions" to remove an "unpopular government," but requiring an inquiry into the actions' underlying motives); Kofi A. Annan, *Two Concepts of Sovereignty*, ECONOMIST (Sept. 16, 1999), http://www.economist.com/node/324795 [https://perma.cc/G8YA-L5Y5] ("States are now widely understood to be instruments at the service of their peoples, and not vice versa."); *see also* JEAN L. COHEN, GLOBALIZATION AND SOVEREIGNTY: RETHINKING LEGALITY, LEGITIMACY, AND CONSTITUTIONALISM 2–4 (2012) (summarizing the claim that the international legal order is no longer based upon the sovereign equality of states and state consent, but instead upon human dignity).

14. *E.g.*, SURYA SUBEDI, THE EFFECTIVENESS OF THE UN HUMAN RIGHTS SYSTEM: REFORM AND THE JUDICIALISATION OF HUMAN RIGHTS (2017); ROSA FREEDMAN, FAILING TO PROTECT: THE UN AND THE POLITICISATION OF HUMAN RIGHTS (2015) [hereinafter FREEDMAN, FAILING TO PROTECT]; STEPHEN HOPGOOD, THE ENDTIMES OF HUMAN RIGHTS ix (2013); ERIC POSNER, THE TWILIGHT OF HUMAN RIGHTS LAW (2014); *see also* de Búrca, *supra* note 8 at 279 (describing academic scholarship harshly critical of international human rights, some of it by scholars who are "'insiders' to the human rights system."). Earlier work in this vein includes Oona Hathaway, *Do Human Rights Treaties Make a Difference?*, 111 YALE L.J. 1935 (2002), and Laurence R. Helfer, *Overlegalizing Human Rights: International Relations Theory and the Commonwealth Caribbean Backlash Against Human Rights Regimes*, 102 COLUM. L. REV. 1832 (2002).

15. CHRISTOPHER N. J. ROBERTS, THE CONTENTIOUS HISTORY OF THE INTERNATIONAL BILL OF HUMAN RIGHTS 9 (2015); SAMUEL MOYN, THE LAST UTOPIA: HUMAN RIGHTS IN HISTORY 5–7 (2010).

16. Mutua, *supra* note 5, at 455; *see also* Makau Mutua, *Savages, Victims, and Saviors: The Metaphor of Human Rights*, 42 HARV. INT'L L.J. 201 (2000).

17. EMILY PADDON RHOADS, TAKING SIDES IN PEACEKEEPING: IMPARTIALITY AND THE FUTURE OF THE UNITED NATIONS 2–3 (2016).

18. *See, e.g.*, AMNESTY INT'L, AMNESTY INTERNATIONAL REPORT 2016/17: THE STATE OF THE WORLD'S HUMAN RIGHTS 12–13 (2017), https://www.amnesty.org/download/Documents/POL1048002017ENGLISH.PDF [https://perma.cc/B4X2-89LP] (describing 2016 as "a year of unrelenting misery and fear" and as witness to "the idea of human dignity and equality, the very notion of a human family, coming under vigorous and relentless assault"); FREEDOM HOUSE, 2017, at 1 (2017) [hereinafter FREEDOM IN THE WORLD 2017], https://freedomhouse

which human rights are at "extreme risk" has increased from twenty in 2008 to thirty-five in 2015.[19] Although harder to measure because most economic and social rights are framed in relative terms,[20] it appears that these kinds of rights have also suffered a setback.[21] Empirical measures of human rights are generally contested, in part because an apparent decline might be due to better reporting and monitoring.[22] It is clear, however, that a wide variety of observers report a recent and serious rise in human rights violations in many of the world's powerful and/or regionally important countries such as China, Russia, India, Turkey, Poland, Hungary, Venezuela, Egypt, and Philippines.[23] In these countries, power is increasingly centralized around a

---

.org/sites/default/files/FH_FIW_2017_Report_Final.pdf [https://perma.cc/X75N-FYS7] (describing 2016 as "the 11th consecutive year in which declines outnumbered improvements"); Kenneth Roth, *The Dangerous Rise of Populism: Global Attacks on Human Rights Values* ("[T]oday a new generation of populists is turning [human rights] protection on its head."), *in* HUMAN RIGHTS WATCH: WORLD REPORT 2016, at 1 (2017), https://www.hrw.org/sites/default/files/world_report_download/wr2017-web.pdf [https://perma.cc/9HVR-H3BX]; *see also* John Kerry, *Preface* to U.S. DEP'T OF STATE, COUNTRY REPORTS ON HUMAN RIGHTS PRACTICES (2015), https://www.state.gov/j/drl/rls/hrrpt/2015humanrightsreport/index.htm#wrapper [https://perma.cc/M49Y-CQPD] (emphasizing that the abuses of human rights originated from "governments crack[ing] down on the fundamental freedoms" and observing an "accelerating trend by both state and nonstate actors to close the space for civil society").

19. *Human Rights Risk Atlas 2015*, VERISK MAPLECROFT (Dec. 3, 2014), https://maplecroft.com/portfolio/new-analysis/2014/12/03/human-rights-deteriorating-most-ukraine-thailand-turkey-due-state-repression-civil-unrest-maplecroft-human-rights-risk-atlas/ [https://perma.cc/D6C9-38DX].

20. *See, e.g.*, International Covenant on Economic, Social and Cultural Rights, art. 2, Dec. 16, 1966, 993 U.N.T.S. 3 [hereinafter ICESCR] (requiring state parties to "take steps" to "progressively" achieve the realization of individuals' social and economic rights).

21. *See Rights in Crisis*, CTR. ECON. & SOC. RTS., http://www.cesr.org/rights-crisis [https://perma.cc/RM9Y-2BMM] (reporting on governments' failure to protect people's social and economic rights during serious economic crises).

22. *Compare* FREEDOM IN THE WORLD 2017, *supra* note 18 *with* MARK GIBNEY ET AL., THE POLITICAL TERROR SCALE 1976–2015 (2016), http://www.politicalterrorscale.org [https://perma.cc/EMV8-8TK5] (using different measurement systems). *See also* Christopher J. Fariss, *Respect for Human Rights Has Improved over Time: Modeling the Changing Standard of Accountability*, 108 AM. POL. SCI. REV. 297 (2014) (arguing that changes in how agencies interpret human rights reports may mask improvements in human rights outcomes); Christopher J. Fariss, *The Changing Standard of Accountability and the Positive Relationship Between Human Rights Treaty Ratification and Compliance*, BRIT. J. POL. SCI., July 2017, at 1, 28 (finding a robust positive relationship between improving levels of respect for human rights and ratification of human rights treaties when the changing standard of accountability is taken into consideration). *But cf.* Adam S. Chilton & Eric A. Posner, *The Influence of History on States' Compliance with Human Rights Obligations*, 56 Va. J. Int' L. 211 (2016) (arguing that historical institutions, events, and conditions, not the ratification of human rights treaties, may explain improvements in human rights practices).

23. *See* HUMAN RIGHTS WATCH: WORLD REPORT 2016, at 466–77 (2016), https://www.hrw.org/sites/default/files/world_report_download/wr2016_web.pdf [https://perma.cc/9FG7-K8S3]; Steven A. Cook, *How Erdogan Made Turkey Authoritarian Again*, ATLANTIC (July 21, 2016), https://www.theatlantic.com/international/archive/2016/07/how-erdogan-made-turkey-authoritarian-again/492374/ [https://perma.cc/UT8U-KMQC]; Muhammad Mansour, *Why Sisi Fears Egypt's Liberals: Behind the Recent Crackdown on Civil Society*, FOREIGN AFF. (May 18, 2016), https://www.foreignaffairs.com/articles/egypt/2016-05-18/why-sisi-fears-egypts-

strong leader who tolerates little dissent.[24] Moreover, populist leaders recently elected in Britain and the United States are openly critical of international human rights norms.[25] Citizens of mature democracies are becoming less satisfied with their form of government.[26] Whether generally termed an "authoritarian resurgence," or "the coming illiberal order"—the trend appears clear.[27] As Philip Alston puts it: "[t]he world as we in the human rights movement have known it in the recent years is no longer."[28]

There is a third way of considering the general state of human rights and of international human rights law in particular: the decline of what one might term the "architecture" for the international legal enforcement of human rights. That enforcement architecture is built conceptually upon the redefinition of sovereignty as based upon a responsibility towards individuals and their universally acknowledged human rights. This conceptual foundation provided the basis for doctrinal innovations in international law designed to facilitate the coercive enforcement of a variety of human rights

---

liberals [https://perma.cc/3SWL-LYHJ]; Kati Marton, *Hungary's Authoritarian Descent*, N.Y. TIMES (Nov. 3, 2014), https://www.nytimes.com/2014/11/04/opinion/hungarys-authoritarian-descent.html?mcubz=1 [https://perma.cc/BP4H-3FD6]; Andrew J. Nathan, *Who is Xi?*, N.Y. REV. BOOKS (May 12, 2016), https://www.nybooks.com/articles/2016/05/12/who-is-xi/ [https://perma.cc/7HDV-24XS]; *Philippines President Rodrigo Duterte Says He 'Doesn't Give a S*** About Human Rights' as 3,500 Killed in War on Drugs*, INDEPENDENT (Oct. 17, 2016), https://www.independent.co.uk/news/world/asia/philippines-president-rodrigo-duterte-doesnt-give-a-s-about-human-rights-war-on-drugs-civilians-a7365156.html [https://perma.cc/KSL5-4XZY]; Orville Schell, *Crackdown in China: Worse and Worse*, N.Y. REV. BOOKS (Apr. 21, 2016), http://www.nybooks.com/articles/2016/04/21/crackdown-in-china-worse-and-worse/ [https://perma.cc/TC9D-YNGM]; Francisco Toro, *It's Official: Venezuela Is a Full-Blown Dictatorship*, WASH. POST (Oct. 21, 2016), https://www.washingtonpost.com/news/global-opinions/wp/2016/10/21/its-official-venezuela-is-a-dictatorship/?utm_term=.60fc6acef963 [https://perma.cc/WBN6-WMED].

24. *See, e.g.*, Javier Corrales, *Autocratic Legalism in Venezuela*, J. DEMOCRACY, Apr. 2015, at 37, 37–38; Abbas Milani, *Iran's Paradoxical Regime*, J. DEMOCRACY, Apr. 2015, at 52, 52; Lilia Shevtsova, *Forward to the Past in Russia*, J. DEMOCRACY, Apr. 2015, at 22, 23; Frederic Wehrey, *Saudi Arabia's Anxious Autocrats*, J. DEMOCRACY, Apr. 2015, at 71, 71–72; *see also* sources cited *supra* note 23.

25. *See* Anushka Asthana & Rowena Mason, *UK Must Leave European Convention on Human Rights, Says Theresa May*, GUARDIAN (Apr. 25, 2015), https://www.theguardian.com/politics/2016/apr/25/uk-must-leave-european-convention-on-human-rights-theresa-may-eu-referendum [https://perma.cc/FXK2-GV4X]; Jonathan Swan, *Trump Calls for "Hell of a Lot Worse than Waterboarding"*, HILL (Feb. 6, 2016), http://thehill.com/blogs/ballot-box/gop-primaries/268530-trump-calls-for-hell-of-a-lot-worse-than-waterboarding [https://perma.cc/NG7T-F78N]; Daniel White, *Read Donald Trump's Ohio Speech on Immigration and Terrorism*, TIME (Aug. 15, 2016), http://time.com/4453110/donald-trump-national-security-immigration-terrorism-speech/ [https://perma.cc/4XFV-SPXY].

26. Roberto Stefan Foa & Yascha Mounk, *The Democratic Disconnect*, J. DEMOCRACY, July 2016, at 5, 5.

27. Michael J. Boyle, *The Coming Illiberal Order*, SURVIVAL: GLOBAL POL. & STRATEGY, Apr.–May 2016, at 35, 36–39; *see also* Larry Diamond, *Facing Up to the Democratic Recession*, J. DEMOCRACY, Jan. 2015, at 141, 147 (describing the erosion of democracy that has occurred around the world with the rise of abusive rulers).

28. Alston, *supra* note 9.

norms through foreign domestic courts, secession, and the use of force. Institutionally, all branches of the United Nations became increasingly focused on the enforcement of human rights law as central to their mission, including the Security Council, which is nominally charged instead with the protection of international peace and security.

The doctrinal and institutional aspects of the human rights enforcement architecture have entered a period of setbacks and retrenchment, which is the focus of the rest of Part I. More important than the specifics of each doctrine, or the possibility that one or more may see a resurgence (or never had a heyday at all), is the broader pattern and the views of states as a whole. This Part shows that states lack a serious commitment to the enforcement of human rights as international law—such enforcement has often been perceived as selective and political and has imposed costs on the peaceful interaction of states.

A. Doctrinal Innovations and Their Costs

The human rights-driven transformation of international law and sovereignty has generated specific doctrinal changes—including limitations on state immunity, universal jurisdiction, the right to democracy, the right to remedial secession, and humanitarian intervention/Responsibility to Protect (R2P).[29] These innovations limit the protections afforded to sovereigns by doctrines such as immunity, jurisdiction, the sovereign equality of states, and prohibitions on the use of force. The changes are based on the conceptual redefinition of sovereignty (or of international law itself) as legitimate only to the extent it reflects popular choice and to the extent it protects and promotes individual human rights. The traditional understanding of sovereignty, by contrast, affords all sovereign states exclusive control over their territory and includes the "principles of sovereign immunity, domestic jurisdiction, and nonintervention."[30] The move from the traditional to the human rights-based understanding of sovereignty is described in different ways using different terminology, and it is often seen as part of a broader diminution in the significance of the nation-state itself in both international law and politics.[31] Whatever the terminology and however sweeping the

29. The human rights-related changes to public international law go beyond those explored here. Others include the development of *jus cogens* norms and *erga omnes* obligations, the growth of international criminal law, changes in the law of war, and an increase in non-consensual international legal norms. For a detailed discussion, see THEODORE MERON, THE HUMANIZATION OF INTERNATIONAL LAW 1–209, 247–301 (2006). These changes fall outside this Article's focus on international human rights enforcement mechanisms and on the doctrine of sources. *Jus cogens* norms are addressed briefly in the text at notes 43–45.

30. Jean L. Cohen, *Whose Sovereignty? Empire Versus International Law*, ETHICS & INT'L AFF., Dec. 2004, at 1, 7.

31. *See, e.g.*, Martti Koskenniemi, *What Use for Sovereignty Today*, 1 ASIAN J. INT'L L. 61, 62–65 (2011); Helen Stacy, *Relational Sovereignty*, 55 STAN. L. REV. 2029, 2040–42 (2003).

global changes, all acknowledge the shift in sovereignty and international law towards the individual and the protection of people.[32]

The five doctrinal innovations, which were advanced with most enthusiasm in a period from the early 1990s to the mid-2000s, are all experiencing decline or stasis. They all also appear to harm interstate relations—or at least states say that they do. A right to remedial secession, for example, would permit oppressed groups who are the victims of extreme human rights abuses to use violence and break apart from the state to which they belong. The cost to international peace is part of the doctrine itself. In other respects, however, the doctrines generate indirect costs for international relations, such as problems with selective enforcement or the diminution of the principle of sovereign equality of states. These problems increase the potential for international friction and thus increase the possibility of instability or even military conflict, as illustrated by the actual efforts to implement all five doctrinal innovations.

1. Immunity.—Immunity is a long-standing, classic doctrine of public international law based on "the sovereign equality of states" and described as "one of the clearest examples of the 'statist' nature of international law."[33] Beginning in the late 1990s, immunity appeared to be of declining application in cases alleging violations of international human rights law. A reversal of course began in early 2010, and today it seems well-established that immunity applies in human rights cases just as it does in other litigation against foreign states or their officials.

International law provides several kinds of immunity to states and the individuals who work on their behalf. The most important, at least in terms of economic impact, is the immunity states themselves enjoy from suit in foreign domestic courts, a form of immunity that also extends to some individual government officials.[34] State immunity has exceptions, including one for litigation related to a state's commercial activity.[35] The purpose of state immunity is to ensure the peaceful coexistence of states and to minimize

32. Louis Henkin, *That "S" Word: Sovereignty, and Globalization, and Human Rights, et Cetera*, 68 FORDHAM L. REV. 1, 8 (1999); Krieger & Nolte, *supra* note 6, at 13; *see also* MERON, *supra* note 29, at 5 (noting that, since the 19th century, the international law of war has increasingly embodied humanitarian constraints).

33. Menno Kamminga, *Final Report on the Impact of International Human Rights Law on General International Law*, *in* THE IMPACT OF HUMAN RIGHTS LAW ON GENERAL INTERNATIONAL LAW 1, 14 (Menno T. Kamminga & Martin Scheinin eds., 2009); *see also* TEITEL, *supra* note 13, at 38–39; Paul Christoph Bornkamm, *State Immunity Against Claims Arising from War Crimes: The Judgment of the International Court of Justice in* Jurisdictional Immunities of the State, 13 GER. L.J. 773, 779–82 (2012).

34. Chanaka Wickremasinghe, *Immunities Enjoyed by Officials of States and International Organizations*, *in* INTERNATIONAL LAW 388, 388–89 (Malcolm D. Evans ed., 2003).

35. RESTATEMENT (THIRD) OF THE FOREIGN RELATIONS LAW OF THE UNITED STATES § 453 (AM. LAW INST. 1986).

interstate friction. As a leading scholar and judge put the point: "[t]here is no more important way to avoid conflict than by providing clear norms as to which state can exercise authority over whom, and in what circumstances."[36]

Scholars and activists have long argued that state immunity should not apply in suits or prosecutions alleging human rights violations.[37] At the state-to-state level, efforts to develop such an exception never gained much traction.[38] But around the turn of the twenty-first century, developments in a handful of domestic and regional courts appeared to support the claim that an exception was in the making. Most prominently, the House of Lords denied immunity to Augusto Pinochet, the former dictator of Chile.[39] The decision cleared the way for the extradition of Pinochet to Spain for trial on criminal charges of torture committed during his rule, although the U.K. government ultimately refused to send him to Spain for health reasons.[40] The Law Lords offered a variety of rationales for their ruling, including some which would apply broadly to claims of human rights violations against states.[41] Shortly thereafter, the European Court of Human Rights held in a nine-to-eight decision that international law afforded Kuwait immunity from a suit in English courts alleging torture.[42] But the dissenters, and many scholars, reasoned that because torture violates *jus cogens* norms, a state which allegedly committed torture "cannot invoke hierarchically lower rules (in this case, those on State immunity) to avoid the consequences of the illegality of its actions."[43]

Cases from Italy and Greece pointed in the same direction, at least initially. The most significant, *Ferrini v. Germany*,[44] denied immunity to

---

36. ROSALYN HIGGINS, PROBLEMS AND PROCESSES: INTERNATIONAL LAW AND HOW WE USE IT 56 (1994); *see also* Wickremasinghe, *supra* note 34, at 389 (summarizing the underlying rationale for sovereign immunity as facilitating "processes of communication between States on which international relations and cooperation rely").

37. *See, e.g.*, Mathias Reimann, *A Human Rights Exception to Sovereign Immunity: Some Thoughts on* Princz v. Federal Republic of Germany, 16 MICH. J. INT'L L. 403, 406, 419–25 (1995) (arguing for an exception to foreign sovereign immunity in the United States in cases alleging "torture, genocide, or enslavement").

38. *See, e.g.*, Lorna McGregor, *State Immunity and* Jus Cogens, 55 INT'L & COMP. L.Q. 437, 441–43 (2006) (highlighting the reluctance of Canadian and British courts to allow exceptions to sovereign immunity even in cases of *jus cogens* violations).

39. Reg. v. Bow St. Metro. Stipendiary Magistrate, *Ex parte* Pinochet Ugarte (No. 3) [2000] 1 AC (HL) 147 (appeal taken from Eng.) [hereinafter *Ex parte Pinochet*]. For a detailed discussion of this case, see generally Ingrid Wuerth, Pinochet*'s Legacy Reassessed*, 106 AM. J. INT'L L. 731 (2012).

40. Wuerth, *supra* note 39, at 735–36.

41. *Ex parte Pinochet*, *supra* note 39, at 112 (joint dissenting opinion of Judges Rozakis and Caflisch joined by Judges Wildhaber, Costa, Cabral Barreto, and Vajić).

42. Al-Adsani v. United Kingdom, 2001-XI Eur. Ct. H.R. 79.

43. *Id.* at 30.

44. Ferrini v. Fed. Republic of Ger., Cass., sez. plen., 11 marzo 2004, n. 5044, 87 RIVISTA DI DIRITTO INTERNAZIONALE [RDI] 539 (2004) (It.) (reported by Andrea Bianchi at 99 AM. J. INT'L L. 242 (2005)).

Germany in a civil case alleging forced labor during World War II.[45] In response, Germany sued Italy before the International Court of Justice alleging that the Italian judgments violated customary international law.[46] The ICJ held for Germany. Its decision focused in part on the purpose and values served by state immunity, including the "sovereign equality of States, . . . one of the fundamental principles of the international legal order."[47] On the heels of the *Germany v. Italy* decision, the European Court of Human Rights held in *Jones v. United Kingdom*[48] that customary international law afforded immunity to an individual Saudi official who allegedly committed torture and who was sued in a civil suit in the U.K.[49]

As international law stands today, immunity applies in suits alleging human rights violations as it does in other cases. The reversal of momentum is due in part to interstate friction caused by domestic court cases against foreign individuals accused of human rights violations.[50] There continues to be a significant constituency seeking to limit immunity in human rights cases,[51] national court decisions in Italy have found the *Germany v. Italy* decision inconsistent with the Italian legal order,[52] and a South African court recently concluded that sitting heads of state could be prosecuted in South Africa for some crimes, although such prosecutions would violate customary international law.[53] The issue has certainly not gone away. But as a "fundamental principle of the international legal order," designed to facilitate

---

45. *Id.*; *see also* Prefecture of Victoria v. Fed. Republic of Ger., Areios Pagos [A.P.] [Supreme Court] 11/2000 (Greece) (reported by Maria Gavouneli & Ilias Bantekas at 95 AM. J. INT'L L. 1998 (2001)) (denying Germany's sovereign immunity defense in a human rights case).

46. *See generally* Jurisdictional Immunities of the State (Ger. v. It.), Judgment, 2012 I.C.J. Rep. 99 (Feb. 3).

47. *Id.* at 123, ¶ 57.

48. Jones v. United Kingdom, 2014-I Eur. Ct. H.R. 1.

49. *See id.* at 68.

50. *See* U.N. Secretariat, Immunity of State Officials from Foreign Criminal Jurisdiction, Memorandum by the Secretariat, ¶ 150 n.419, Int'l Law Comm'n, U.N. Doc. A/CN.4/596 (Mar. 31, 2008); *see also Belgium Scales Back Its War Crimes Law Under U.S. Pressure*, N.Y. TIMES (Aug. 2, 2003), http://www.nytimes.com/2003/08/02/world/belgium-scales-back-its-war-crimes-law-under-us-pressure.html?mcubz=1 [https://perma.cc/U33D-J2R7].

51. *See, e.g.*, Harold Hongju Koh, *Foreign Official Immunity After Samantar: A United States Government Perspective*, 44 VAND. J. TRANSNAT'L L. 1141, 1151 (2011) (". . . [O]fficial immunity law will need to take into account the human rights revolution.").

52. Simoncioni v. Repubblica Federale di Germania, Corte cost., 22 ottobre 2014, n.238, Gazzetta Ufficiale [G.U.] (ser. Spec.) n.45, 29 ottobre 2014, I, 1, http://www.cortecostituzionale.it/documenti/download/doc/recent_judgments/S238_2013_en.pdf [https://perma.cc/9SW4-LV9T] (reported by Riccardo Pavoni at 109 AM. J. INT'L L. 400 (2015)). For a detailed discussion of this case, see Note, *Constitutional Courts and International Law: Revisiting the Transatlantic Divide*, 129 HARV. L. REV. 1362, 1371–72 (2016).

53. *See Minister of Justice and Constitutional Dev. v. S. African Litig. Ctr.* 2016 (4) BCLR 487 (SCA) at 507 para. 49, 522 para. 84–85, 528–29 para. 103 (S. Afr.) (holding that the South African law that implemented the Rome Statute did not provide head-of-state immunity to those charged with international crimes under the statute).

the peaceful coexistence of states, the international law of immunity is of ongoing significance.[54]

2. Universal Jurisdiction.—Universal jurisdiction was another doctrinal facet of the changing definition of sovereignty.[55] It, too, began to decline about a decade ago and today appears to be in a period of stasis at best.[56] Universal jurisdiction allows any state to exercise prescriptive jurisdiction and apply its laws, even if there is no traditional basis (such as territory or nationality) for doing so.[57] Universal jurisdiction seemed like a promising way to ensure the prosecution of, and civil remedies against, many individuals who violated basic human rights norms such as torture and genocide.[58] In practice, however, the doctrine has—like other methods of enforcing international human rights—been applied selectively based on political calculations, leading to international friction.

Universal jurisdiction cases that go forward are generally those brought against defendants from weak, poor states or states that have been vilified by the world community as a whole—the former Yugoslavia and Rwanda serve as examples.[59] In many universal jurisdiction cases, the executive branch (including prosecutors) of the forum state plays an important role in case selection or has the power to veto cases altogether.[60] It often exercises its discretion to have cases dismissed when the defendant is from a powerful

54. Ger. v. It., 2012 I.C.J. at 123; *see also* ROTH, *supra* note 13, at 221–24 (arguing that nullification of immunity and the exercise of extraterritorial jurisdiction have the potential to erode the international legal order).

55. *See, e.g.*, TEITEL, *supra* note 13, at 58–61; *see also* Carly Nyst, *Solidarity in a Disaggregated World: Universal Jurisdiction and the Evolution of Sovereignty*, J. INT'L L. & INT'L REL., Fall 2012, at 36, 38–39, 47 (surveying the history of the universal jurisdiction doctrine and exploring its role within the larger debate on the changing conception of sovereignty).

56. *See, e.g.*, Mark Ellis, Exec. Dir., Int'l Bar Ass'n, The Decline of Universal Jurisdiction—Is It Reversible?, Remarks at the 10th Annual Ruth Steinkraus-Cohen International Law Lecture 5, 8 (Feb. 22, 2012), http://www.unawestminster.org.uk/pdf/grot12_mark_ellis_lecture.pdf [https://perma.cc/YT92-CNRE] ("As much of the international community promotes universal jurisdiction, state practice is limiting the scope and use of it.").

57. Kenneth C. Randall, *Universal Jurisdiction Under International Law*, 66 TEXAS L. REV. 785, 788 (1988).

58. *See id.* at 788, 828–29, 837 (tracing the rise of the universal jurisdiction doctrine as a tool "to combat egregious offenses that states universally have condemned").

59. *See* Máximo Langer, *The Diplomacy of Universal Jurisdiction: The Political Branches and the Transnational Prosecution of International Crimes*, 105 AM. J. INT'L L. 1, 3, 8–9 (2011) (collecting universal jurisdiction cases and finding that three-quarters of all defendants tried have been Rwandans, former Yugoslavs, or Nazis); *cf.* TRIAL INT'L ET AL., UNIVERSAL JURISDICTION ANNUAL REVIEW 2016: MAKE WAY FOR JUSTICE 49–51, 69 (2017) (reporting that seven out of twenty pending universal jurisdiction cases, including those under investigation, involved conduct in Iraq or Syria).

60. *See* Langer, *supra* note 59, at 11–12 (examining German legislation); *id.* at 15–16 (examining English and Welsh statutes); *cf. id.* at 19–20 (comparing discretion afforded to the investigating judges under French criminal law).

country or the case would otherwise impose foreign policy costs on the forum state.[61]

A survey of universal jurisdiction cases against Chinese nationals, for example, shows numerous potential cases arising out of severe human rights violations in Tibet and against the Falun Gong.[62] In case after case, however, prosecutors dismissed the claims or, failing that, legislatures passed new laws to ensure that Chinese defendants would not stand trial.[63] German litigation against Donald Rumsfeld on allegations of torture has similarly gone nowhere, again because of the intervention or control of German government officials.[64] Litigation in Spain against U.S. officials for conduct in Guantanamo has fared no better.[65]

The disproportionate focus on African defendants led to complaints from individual African countries and from the African Union.[66] The Assembly of the African Union noted that the "abuse" of universal jurisdiction could "endanger [i]nternational law, order and security" and declared that "[t]he political nature and abuse of the principle of universal jurisdiction by judges from some non-African States against African leaders, particularly Rwanda, is a clear violation of the sovereignty and territorial integrity of these States."[67] Putting aside whether these criticisms are valid, the point here is that, as with other doctrinal innovations, the selective application of universal jurisdiction has led to political friction and the doctrine itself is in apparent retreat.[68]

International human rights litigation in the United States followed a similar course. The Second Circuit's 1980 decision in *Filartiga v. Pena-Irala*[69] opened the door for a wide range of human rights cases to go forward

---

61. *Id.* at 5–6.

62. Eric L. Hwang, *China: The Growth of a New Superpower and the Extinction of Universal Jurisdiction*, WIS. INT'L L.J. 334, 344–46, 348 (2014).

63. *See, e.g.*, Stewart M. Patrick, *Spain's Welcome Retreat on Universal Jurisdiction*, COUNCIL ON FOREIGN REL. (Feb. 14, 2014), http://blogs.cfr.org/patrick/2014/02/14/spains-welcome-retreat-on-universal-jurisdiction/ [https://perma.cc/KDB3-PF82]; *cf.* Hwang, *supra* note 62, at 348–49, 353 (describing the hesitation to exercise universal jurisdiction over Chinese defendants).

64. *See* Katherine Gallagher, *Universal Jurisdiction in Practice: Efforts to Hold Donald Rumsfeld and Other High-Level United States Officials Accountable for Torture*, 7 J. INT'L CRIM. JUST. 1087, 1104–06 (2009).

65. Charles Chernor Jalloh, *Universal Jurisdiction, Universal Prescription? A Preliminary Assessment of the African Union Perspective on Universal Jurisdiction*, 21 CRIM. L.F. 1, 56 (2010).

66. Ernest A. Young, *Universal Jurisdiction, the Alien Tort Statute, and Transnational Public-Law Litigation After* Kiobel, 64 DUKE L.J. 1023, 1043 (2015).

67. Karinne Coombs, *Universal Jurisdiction: A Means to End Impunity or a Threat to Friendly International Relations?*, 43 GEO. WASH. INT'L L. REV. 419, 442 (2011) (quoting African Union, *Decision on the Report of the Commission on the Abuse of the Principle of Universal Jurisdiction*, ¶ 5 (i)–(ii), Assembly/AU/Dec.199 (XI) (July 2008)).

68. *See* Patrick, *supra* note 63 (noting that Spain's Parliament voted to curb its judges' authority to exercise universal jurisdiction). *See generally* ROTH, *supra* note 13, at 271.

69. Filartiga v. Pena-Irala, 630 F.2d 876 (2d Cir. 1980).

under the Alien Tort Statute (ATS). The ATS became the primary vehicle for expansive human rights litigation in the United States.[70] It also attracted worldwide attention and generated a cottage industry of litigators and scholars.[71] Jurisdictionally, the extraterritorial application of the statute was said to rest on universal civil jurisdiction.[72] But the Supreme Court substantially curtailed ATS litigation in *Kiobel v. Royal Dutch Petroleum Co.*, which applied the presumption against extraterritorial application of statutes to the ATS.[73] The Court justified its decision in part by noting that if the ATS provides a "cause of action for conduct occurring in the territory of another sovereign," it could result in "diplomatic strife."[74] In support, the Court cited to "[r]ecent experience" in the form of lower court ATS litigations, which drew objections from Canada, Germany, Indonesia, Papua New Guinea, South Africa, Switzerland, and the United Kingdom.[75]

The European countries that had been at the forefront of the universal jurisdiction movement have since amended their statutes to limit the doctrine's application, and it appears that the universal jurisdiction complaints have decreased so that the doctrine is on the decline.[76] On the other hand, universal jurisdiction may be reinvigorated through European cases against Syrians accused of torture and other crimes during that country's brutal civil war.[77] Even if successful, however, such prosecutions are likely to fuel, not quell, complaints that universal jurisdiction is applied selectively.

3. Right to Democracy.—A third doctrinal change brought by human rights was a purported right to democracy. This doctrine, too, was designed to make inroads in the protections afforded to sovereigns by international law. After a heyday in the immediate post-Cold War period, the purported

70. *See generally* Ingrid Wuerth, Kiobel v. Royal Dutch Petroleum Co.: *The Supreme Court and the Alien Tort Statute*, 107 AM. J. INT'L L. 601 (2013).

71. *Id.*

72. Donovan & Roberts, *supra* note 4, at 146–49, 153–54.

73. Kiobel v. Royal Dutch Petroleum Co., 569 U.S. 108, 108 (2013).

74. *Id.* at 124.

75. *Id.* (citing Doe v. Exxon Mobil Corp., 654 F.3d 11, 77–78 (D.C. Cir. 2011) (Kavanaugh, J., dissenting)); *see also* Young, *supra* note 66, at 1042–43.

76. *Cf.* Máximo Langer, *Universal Jurisdiction Is Not Disappearing: The Shift from 'Global Enforcer' to 'No Safe Haven' Universal Jurisdiction*, 13 J. INT'L CRIM. JUST. 245, 248–49 (2015) (arguing that the number of universal jurisdiction complaints has not substantially decreased in recent years).

77. *See* Christian Wenaweser & James Cockayne, *Justice for Syria?: The International, Impartial and Independent Mechanism and the Emergence of the UN General Assembly in the Realm of International Criminal Justice*, 15 J. INT'L CRIM. JUST. 211, 219 (2017); TRIAL INT'L ET AL., *supra* note 59.

right to democracy was largely abandoned by the first decade of the new century.

The breakup of the Soviet Union generated enthusiasm for democratic governance around the world.[78] Election-monitoring institutions and initiatives grew quickly in countries from Namibia to Nicaragua to Cambodia and Bulgaria.[79] The ouster of democratically elected Jean-Bertrand Aristide resulted in unprecedented condemnation by the Organization of American States and the U.N. General Assembly, both of which suggested that governmental authority is only legitimate to the extent it is a function of popular sovereignty in the form of democracy.[80] A series of other events unfolded over the decades that followed, including the restoration of Aristide to power through the use of force authorized by the U.N. Security Council,[81] which made clear a growing connection between political legitimacy and democratic governance. The support of "democratic institutions" and "democracy promotion"[82] became an important mission of the United Nations and other regional and international organizations.[83]

A wave of scholarly enthusiasm for the emerging right to democratic governance ensued, especially among U.S. scholars.[84] Based in part on international and state practice, the right also found support in regional treaty regimes, the International Covenant on Civil and Political Rights (which affords individuals a right to vote), and Article 21 of the Universal Declaration of Human Rights.[85] Article 21 protects the right of individuals to participate in "periodic and genuine elections," which express the will of the

78. *See* Thomas M. Franck, *The Emerging Right to Democratic Governance*, 86 AM. J. INT'L L. 46, 46–47 (1992) [hereinafter Franck, *The Emerging Right*]; *cf.* Reisman, *supra* note 13, at 860–61; *see also* Thomas M. Franck, *Legitimacy and the Democratic Entitlement*, *in* DEMOCRATIC GOVERNANCE AND INTERNATIONAL LAW 25, 36–40 (Gregory H. Fox & Brad R. Roth eds., 2000) [hereinafter DEMOCRATIC GOVERNANCE] (chronicling an increase in democratic entitlement through U.N. initiatives during the 1990s after the fall of communism). The events of the early 1990s augmented earlier efforts to ensure democratic governance. *Id.*

79. Franck, *The Emerging Right*, *supra* note 78, at 55, 70–74.

80. *Id.* at 47.

81. S.C. Res. 940, ¶¶ 5, 17 (July 31, 1994).

82. Jean d'Aspremont, *The Rise and Fall of Democracy Governance in International Law: A Reply to Susan Marks*, 22 EUR. J. INT'L L. 549, 554, 563 (2011) (citing, *inter alia*, Geoffrey Marston, *Guidelines on the Recognition of New States in Eastern Europe and in the Soviet Union*, 62 BRIT. Y.B. INT'L L. 559, 559–60 (1991)) (describing the period from 1989 to 2010).

83. *See* Susan Marks, *What Has Become of the Emerging Right to Democratic Governance?*, 22 EUR. J. INT'L L. 507, 516; Amichai Magen, *The Democratic Entitlement in an Era of Democratic Recession*, 4 CAMBRIDGE J. INT'L & COMP. L. 368, 375–76 (2015).

84. *E.g.*, Reisman, *supra* note 13; Magen, *supra* note 83, at 374–75; Fernando R. Tesón, *The Kantian Theory of International Law*, 92 COLUM. L. REV. 53, 60–74 (1992); *see also* B. BAUER, DER VÖLKERRECHTLICHE ANSPRUCH AUF DEMOKRATIE (1998).

85. International Covenant on Civil and Political Rights, Dec. 16, 1966, 999 U.T.S. 171 [hereinafter ICCPR]; G.A. Res. 217 (III) A, art. 21, Universal Declaration of Human Rights (Dec. 10, 1948) [hereinafter Universal Declaration]. For further discussion, see Franck, *The Emerging Right*, *supra* note 78.

people and further provides that "[t]he will of the people shall be the basis for the authority of government."[86] Thus, states' legitimate claim to the protections of "sovereignty" on the international level arguably became dependent upon an internal democratic legal order.[87] Understanding legitimate sovereign authority as necessarily based upon a democratic order led many to conclude that military and other intervention in foreign states is not prohibited by international law when used to restore or create democracy.[88] The costs to international peace and security are clear. To be sure, scholars offered prudential reasons for states to be cautious when using "intrusive political, economic, and military measures" to "implement democratization in a recalcitrant State," but such measures could "now be included on the menu of lawful options."[89]

Today, however, international practice—and even academic scholarship—has for some time retreated from the position that international law requires states to have a "democratic origin."[90] The retreat can be attributed to several factors. Like universal jurisdiction and human rights exceptions to immunity, a right to democracy caused interstate friction. African nations protested the emphasis on democracy in the European Union's trade and development policy.[91] China's (and to some extent Russia's) rise in power has also weakened the right to democracy, as China has rejected any purported requirement of international law that it or other states have a government that is democratic in origin.[92] More generally, democracy is in a worldwide decline.[93] Academics and other critics have increasingly questioned the purported benefits of democratic governance, especially the claim that it leads to peace, the protection of other human rights, and economic prosperity.[94] Finally, the promotion of and purported

86. Universal Declaration, *supra* note 85.

87. Reisman, *supra* note 13, at 869–70. *But see* Samantha Besson, *Sovereignty, International Law and Democracy*, 22 EUR. J. INT'L L. 373, 376–82 (2011).

88. Gregory H. Fox & Brad R. Roth, *The Spread of Liberal Democracy*, *in* DEMOCRATIC GOVERNANCE, *supra* note 78, at 10–13. Some writers viewed the right to democracy as part of a right to internal self-determination. *See* Jean d'Aspremont, *1989–2010: The Rise and Fall of Democratic Governance in International Law*, *in* 3 SELECT PROCEEDINGS OF THE EUROPEAN SOCIETY OF INTERNATIONAL LAW 61, 66 n.36 (James Crawford & Sarah Nouwen eds., 2012).

89. DEMOCRATIC GOVERNANCE, *supra* note 78, at 12. *But see generally* Michael Byers & Simon Chesterman, *"You, the People": Pro-Democratic Intervention in International Law*, *in* DEMOCRATIC GOVERNANCE, *supra* note 78 (arguing that unilateral armed intervention to support or restore democracy remains prohibited by international law).

90. *See* d'Aspremont, *supra* note 82, at 70; *see also* Thomas Carothers, *The Backlash Against Democracy Promotion*, 85 FOREIGN AFF., Mar.–Apr. 2006, at 55 (2006).

91. Päivi Leino, *European Universalism?—The EU and Human Rights Conditionality*, 24 Y.B. EUR. L. 329, 338–40 (2005).

92. *See* d'Aspremont, *supra* note 82, at 72.

93. *See* Magen, *supra* note 83, at 378–81.

94. *See, e.g.*, Marks, *supra* note 83, at 519, 523.

right to democracy has been used selectively to advance other foreign policy agendas of Western countries, especially the United States.[95]

4. Remedial Secession.—International human rights led to a fourth purported doctrinal change in international law: the emergence of a right to "remedial secession." Like the others, this doctrine experienced a surge in academic enthusiasm in the 1990s based on some indicia of state practice. Contemporary arguments about secession are closely linked to the "self-determination" of "peoples," a right protected by the U.N. Charter,[96] by the International Covenant on Civil and Political Rights,[97] by the International Covenant on Economic, Social and Cultural Rights,[98] and by a variety of other human rights instruments.[99] That there is a right to self-determination for peoples subjected to colonial rule is now well-settled.[100] The right to self-determination may also afford "peoples"—that is, groups united by some combination of race, ethnicity, territorial affiliation, language, and religion—a right to political representation within their state.[101] If the state does not afford peoples *internal* self-determination, the state is arguably not fully sovereign, giving rise to a right of unilateral external secession or "remedial" secession.[102] As Ruti Teitel puts it, the "values of stability of statehood" that had been "settled since the postwar period—entrenched in the UN Charter, and so on—are now in play with other values, such as those of the protection

95. *See id.* at 521–23.

96. U.N. Charter art. 43, ¶ 1.

97. ICCPR, *supra* note 85, art. 1.

98. ICESCR, *supra* note 20, art. 1.

99. U.N. Charter art. 1, ¶ 2, art. 55; International Covenant on Civil and Political Rights, art. 1, ¶ 1, *adopted* Dec. 19, 1966, 999 U.N.T.S. 14668; International Covenant on Economic, Social and Cultural Rights, art. 1, ¶ 1, *adopted* Dec. 16, 1966, 993 U.N.T.S. 14531; G.A. Res. 2625 (XXV), Declaration on Principles of International Law Concerning Friendly Relations and Co-Operation Among States in Accordance with the Charter of the United Nations (Oct. 24, 1970) [hereinafter Friendly Relations Declaration]; G.A. Res. 1514 (XV), Declaration on the Granting of Independence to Colonial Countries and Peoples (Dec. 14, 1960) [hereinafter Declaration on Granting of Independence].

100. *See* Accordance with International Law of the Unilateral Declaration of Independence in Respect of Kosovo, Advisory Opinion, 2010 I.C.J. Rep. 403, ¶ 79 (July 22). A series of ICJ decisions have addressed the right of self-determination in the colonial context. *See, e.g.*, Legal Consequences of the Construction of a Wall in the Occupied Palestinian Territory, Advisory Opinion, 2004 I.C.J. Rep. 136, ¶ 88; East Timor (Port. v. Austl.), Judgment, 1995 I.C.J. Rep. 90, ¶ 29; Western Sahara, Advisory Opinion, 1975 I.C.J. Rep. 13, ¶¶ 54–59; Legal Consequences for States of the Continued Presence of South Africa in Namibia (South West Africa) notwithstanding Security Council Resolution 276 (1970), Advisory Opinion, 1971 I.C.J. Rep. 16, ¶¶ 52–53.

101. *E.g.*, ICCPR, *supra* note 85, art. 1; *see also* DAVID RAIČ, STATEHOOD AND THE LAW OF SELF-DETERMINATION 237–43 (2002) (describing self-determination as including peoples' rights to "participate in the . . . decision-making process" of the state).

102. ALLEN BUCHANAN, JUSTICE, LEGITIMACY, AND SELF-DETERMINATION 357–400 (2004).

of persons and peoples, e.g., self-determination as a remedy for oppression."[103]

Teitel's argument is part of an important strand of philosophical work on international law and its relationship to human rights.[104] The right to remedial self-determination follows from the beginning premise that sovereignty and international law are only legitimate to the extent they represent and protect the individual; here, the individual's rights and well-being are reflected in a group-based or collective right to self-determination.

The difficulty with an international right to violent or "remedial" secession is that it threatens to increase armed conflicts and war, and to encourage the break-up of states into ever-smaller units. Descriptions of a right to remedial secession implicitly recognize these dangers by limiting it to situations in which peoples in question have suffered "extreme abuses" as in Teitel's formulation of the argument. In one sense, it is difficult to see the basis for such a limitation. To the extent that a state has inflicted "very bad" (but not "extreme") human rights abuses, the state is not acting as a legitimate sovereign, so that there is no sovereignty-based principle upon which to limit peoples' right to violent secession. Yet the articulation of the right appears to include a concern about the values that undergird a state-based system: stability and limits on the use of force.

Philosophers have taken various approaches to this question. Some, for example, defend a moral right to secession but limit it by what is feasible in the international legal order.[105] Others, like Allan Buchanan, explicitly argue that a legal right to secede is "inherently institutional" and can only be defined with reference to its harmony "with the other main elements of a morally defensible international legal system"—including those aspects of the international legal order which discourage armed conflict.[106] Again, this formulation of the right underscores the significance of the state-based international legal order, even when a state is violating the human rights of peoples within its territory.

---

103. TEITEL, *supra* note 13, at 194; *see also* Christian Walter & Antje von Ungern-Sternberg, *Introduction* to SELF-DETERMINATION AND SECESSION IN INTERNATIONAL LAW 1, 6 (Christian Walter et al. eds., 2014).

104. *See, e.g.*, BUCHANAN, *supra* note 102, at 353; Daniel Philpott, *In Defense of Self-Determination*, 105 ETHICS 352 (1995); *see also* Alan Patten, *Democratic Secession from a Multinational State*, 112 ETHICS 55, 563–64 (2002).

105. BUCHANAN, *supra* note 102, at 345–46 (noting the work of other scholars); *see also* Stefan Oeter, *The Role of Recognition and Non-Recognition with Regard to Secession*, *in* SELF-DETERMINATION AND SECESSION IN INTERNATIONAL LAW, 45, 60 ("[R]emedial secession is a tool for genuine cases of extreme emergency where the very survival of specific population groups is at stake.").

106. BUCHANAN, *supra* note 102, at 345–47.

Turning to contemporary international law, it is widely (but not universally) accepted that there is no "right" to remedial secession,[107] despite the push by many theorists and human rights activists, especially after the Cold War ended.[108] Secession is closely linked to recognition. Recognition is the formal acknowledgement that a particular entity possesses the qualifications for statehood or that a particular regime is the effective government of a state.[109] Recognition is generally followed by the establishment of diplomatic relations, the possibility of acceptance into international organizations, and other indicia of statehood.[110] Secessionist movements have recognition as their goal because it is an essential bridge to statehood.

International recognition practice shows there is a very strong norm against unilateral secession in general.[111] Indeed, states almost always refuse to recognize secessionist groups as new states if their home states do not recognize them, even if the secession enjoys long-term military success.[112] The norm against secession is so strong that even in the twenty-nine post-World War II cases—not limited to cases of remedial secession—in which a secessionist movement opposed by its home state was able to both gain control of territory and govern a population for at least two years, only three (Bangladesh, Eritrea, and South Sudan) were fully successful, meaning that their statehood was eventually recognized by their home state. An additional three (Kosovo, Taiwan, and Palestine) have been recognized by more than ten members of the United Nations, but not by the state from which they seek to secede; none are members of the United Nations. The remaining twenty-

107. *See, e.g.*, Walter & von Ungern-Sternberg, *supra* note 103, at 3; Christopher J. Borgen, *Law, Rhetoric, Strategy: Russia and Self-Determination Before and After Crimea*, 91 INT'L L. STUD. 216, 229–34 (2015); Jure Vidmar, *International Legal Responses to Kosovo's Declaration of Independence*, 42 VAND. J. TRANSNAT'L L. 779, 816 (2009); *cf.* MILENA STERIO, THE RIGHT TO SELF-DETERMINATION UNDER INTERNATIONAL LAW 134 (2013) ("[A]n argument can be advanced that international law tolerates a limited right of secession for peoples whose rights to internal self-determination have been egregiously disrespected.").

108. *See, e.g.*, Lawrence M. Frankel, *International Law of Secession: New Rules for a New Era*, 14 HOUS. J. INT'L L. 521, 526 (1992); *see also* BUCHANAN, *supra* note 102, at 397–98 (acknowledging that the right to secession he defended in 2004 was narrower than the one he defended in 1991; the later work reflects a greater appreciation of the limited institutional capacities of international law).

109. RESTATEMENT (THIRD) OF FOREIGN RELATIONS LAW OF THE UNITED STATES § 203 cmt. a (AM. LAW INST. 1986).

110. *See* MIKULAS FABRY, RECOGNIZING STATES: INTERNATIONAL SOCIETY AND THE ESTABLISHMENT OF NEW STATES SINCE 1776, at 7 (2010).

111. *Id.* at 165–66, 179; *see also* Tanisha M. Fazal & Ryan D. Griffiths, *Membership Has Its Privileges: The Changing Benefits of Statehood*, 16 INT'L STUD. REV. 76, 97 (2014); Borgen, *supra* note 107, at 229.

112. *See* JAMES CRAWFORD, THE CREATION OF STATES IN INTERNATIONAL LAW 415–18 (2d ed. 2006); FABRY, *supra* note 110, at 13.

three "states" have been recognized by fewer than ten other states, or were reabsorbed into their home states.[113]

Some of both the successful and the unsuccessful states involved claims of remedial secession, but there is widespread disagreement as to which situations present legitimate claims to remedial secession and which do not. Even in the cases that arguably provide the strongest support for remedial secession—Kosovo and South Sudan—the basis for secession is not clear and success in terms of recognition is best explained by power politics.[114] And although claims of human rights abuses directed at the group that seeks secession might improve their political arguments for recognition, international law does not require other states to accord recognition on this (or any other) basis.[115]

Despite the claim that human rights have displaced or transformed sovereignty as the basis of the international legal order and the arguments for a legal right to remedial secession which flows from that claim, international law and practice have not followed.

5. Humanitarian Intervention & Responsibility to Protect.—The claim that international human rights have transformed state sovereignty led to an explicit call for the use of force by the international community to prevent widespread human rights violations. This fifth doctrinal change goes under the heading of "humanitarian intervention" or (with somewhat different content) the "Responsibility to Protect" (R2P). Both are explicitly premised on a reorientation of sovereignty in favor of human rights couched as either the right of other states to intervene or as the responsibility of sovereigns to protect individual rights. Both have been diminished by the unsuccessful R2P

113. Kristy Buzard et al., *Unrecognized States: A Theory of Self-Determination and Foreign Influence*, 33 J.L. ECON. & ORG. 578, 582 (2017). Palestine and Taiwan arguably do not involve secession at all. *See* Winston P. Nagan & Aitza M. Haddad, *Recognition of Palestinian Statehood: A Clarification of the Interests of the Concerned Parties*, 40 GA. J. INT'L & COMP. L. 341, 377 (2012); CRAWFORD, *supra* note 112, at 206–21, 434.

114. *See* Milena Sterio, *Secession: A Proposal for a New Legal Framework*, GERM. Y.B. INT'L L. (forthcoming 2017) (describing uncertainty around remedial secession in the cases of Kosovo and Nagorno-Karabakh); Oeter, *supra* note 105, at 59–60 (arguing that Kosovo did not meet the criteria for remedial secession); STERIO, *supra* note 107, at 136 (arguing that Chechnya had a better claim to remedial secession than Kosovo); *id.* at 151–52 (arguing that claims by South Ossetia and Abkhazia to remedial secession were comparable to those of Kosovo); *id.* at 167 (arguing that South Sudan's claim for remedial secession is problematic and that the legally stronger argument would rest on a claim of delayed decolonization). With respect to Eritrea, see Gregory Fox, *Eritrea*, *in* SELF-DETERMINATION AND SECESSION, *supra* note 103, at 288–89, who argues that there are several legal bases for Eritrea's secession, including remedial secession, but the question of whether "the level of Ethiopian oppression was sufficiently egregious to trigger the right" to remedial secession is impossible to answer.

115. Jure Vidmar, *Remedial Secession in International Law: Theory and (Lack of) Practice*, 6 ST. ANTONY'S INT'L REV., May 2010, at 37, 51.

intervention in Libya and subsequent failure of the international community to respond effectively to the crisis in Syria.

The global failure to prevent war in Bosnia during the collapse of Yugoslavia and to prevent a horrific genocide in Rwanda during the 1990s led practitioners and scholars to look for new ways to improve the international response to such atrocities.[116] Then during the spring of 1999, human rights abuses escalated in the Serbian province of Kosovo.[117] Haunted by the events in Bosnia and in Rwanda, many state officials believed that it was morally unconscionable to watch the human rights situation deteriorate in Kosovo without responding.[118] The U.N. Security Council was unable to act because Russia, with its permanent member's veto, was a longtime ally of Serbia.[119] Russia argued that the unrest in Kosovo was a domestic issue, not one that justified international intervention.[120] These events led to an extensive debate about the wisdom and the legality of humanitarian intervention that lacked either the U.N. Security Council authorization or the consent of the territorial state.[121]

In the end, NATO launched a bombing campaign in Serbia, which NATO described partly in humanitarian terms. Some scholars have defended humanitarian intervention on the grounds that the protection and enforcement of human rights is "the ultimate justification of the existence of states."[122] States that violate human rights are accordingly not protected by sovereignty from external military intervention.[123]

---

116. Monica Hakimi, *Toward a Legal Theory on the Responsibility to Protect*, 39 YALE J. INT'L L. 247, 251–54 (2014).

117. *See* Mark Weisburd, *International Law and the Problem of Evil*, 34 VAND. J. TRANSNAT'L L. 225, 231–33 (2001) for a discussion of the factual background.

118. *See Law, Morality and the Use of Force*, NORTH ATLANTIC TREATY ORG., http://www.nato.int/cps/en/natohq/opinions_18418.htm?selectedLocale=en#top [https://perma.cc/NQ42-VYAQ]; Press Release, Security Council, NATO Action Against Serbian Military Targets Prompts Divergent Views as Security Council Holds Urgent Meeting on Situation in Kosovo, U.N. Press Release SC/6657 (Mar. 24, 1999).

119. *See* Abraham D. Sofaer, *International Law and Kosovo*, 36 STAN. J. INT'L L. 1, 2 n.7 (2000).

120. *Why Russia Opposes Intervention in Kosovo*, BBC NEWS (Oct. 13, 1998), http://news.bbc.co.uk/2/hi/europe/111585.stm [https://perma.cc/SMN3-96XJ]; *see also* Vladimir Putin, President of Russ., Address (Mar. 18, 2014), http://en.kremlin.ru/events/president/news/20603 [https://perma.cc/9R33-SXXC] (comparing the Crimea dispute to Kosovo and arguing that violations of domestic legislation are not equivalent to a violation of international law).

121. *See* Adam Roberts, *The So-Called 'Right' of Humanitarian Intervention*, 3 Y.B. INT'L HUMANITARIAN L. 3, 3 (2000); Simon Chesterman, *Legality Versus Legitimacy: Humanitarian Intervention, the Security Council, and the Rule of Law*, 33 SECURITY DIALOGUE 293, 294 (2002).

122. FERNANDO R. TESÓN, HUMANITARIAN INTERVENTION: AN INQUIRY INTO LAW AND MORALITY 16 (3d ed. 2005).

123. *See* Julie Mertus, *Reconsidering the Legality of Humanitarian Intervention: Lessons from Kosovo*, 5 WM. & MARY L. REV. 1743, 1764 (2000) ("[W]hen another state intervenes to protect human rights . . . it is not violating a principle of sovereignty. Rather, it is liberating a principle of sovereignty."); Koh, *supra* note 13, at 288 (arguing in favor of humanitarian intervention in part because international law should serve human purposes, including human rights); Kimberley N.

The Kosovo intervention had major ramifications. After the bombing campaign, the United Nations authorized a peacekeeping force in Kosovo. With the strong support of Western European countries and the United States, Kosovo declared its independence from Serbia in 2008.[124] Serbia and its allies, especially Russia, strongly condemned the declaration of independence and continue today to refuse to recognize Kosovo.[125] Russian officials in turn used the Kosovo precedent to support Russia's use of force in both Georgia and Ukraine.[126] Crimea, which was part of Ukraine, is today Russian. Russian justification for its actions against Georgia and Ukraine made repeated and clear references to precedent set by NATO in Kosovo.

The events in Kosovo were not just directly destabilizing in terms of international peace and security, but they also led to the formation of an independent commission (the International Commission on Intervention and State Sovereignty or ICISS), which published a report entitled "The Responsibility to Protect."[127] The United States and NATO had not explicitly defended the Kosovo intervention as consistent with international law, although the United Kingdom and Belgium did make an explicitly legal argument in favor of "humanitarian intervention."[128] The ICISS report made a comprehensive defense of R2P, built on (but slightly different from) humanitarian intervention.

The ICISS Report declared that states have a responsibility to protect not only their own citizens, but also those of foreign countries, from massive human rights violations.[129] It lists as one of the foundations of R2P the "specific legal obligations under human rights and human protection declarations, covenants and treaties, international humanitarian law and

---

Trapp, *Unauthorized Military Interventions for the Public Good: A Response to Harold Koh*, AJIL UNBOUND, Oct. 2017, at 292, 292–93 (arguing that state sovereignty should not "shield a state so as to provide it with the (legal and physical) space within which to violate individual rights to life and physical integrity").

124. *See* James Summers, *Kosovo*, *in* SELF-DETERMINATION AND SECESSION IN INTERNATIONAL LAW, *supra* note 103, at 236–44.

125. *Id.* at 235.

126. For Russia's position in the independence of South Ossetia and Abkhazia, see *id.* at 235–36 (citing *Statement by the Ministry of Foreign Affairs of the Russian Federation*, ORG. SECURITY CO-OPERATION EUR. (Aug. 28, 2008), http://www.osce.org/pc/33166?download=true [https://perma.cc/UG8D-W5XV]); for Russia's reference to Kosovo in supporting Ukraine's secession, see David Herszenhorn, *Crimea Votes to Secede from Ukraine as Russian Troops Keep Watch*, N.Y. TIMES (Mar. 16, 2014), https://www.nytimes.com/2014/03/17/world/europe/crimea-ukraine-secession-vote-referendum.html [https://perma.cc/8V38-ACBB].

127. GARETH J. EVANS ET AL., THE RESPONSIBILITY TO PROTECT: REPORT OF THE INTERNATIONAL COMMISSION ON INTERVENTION AND STATE SOVEREIGNTY (2001) [hereinafter ICISS REPORT].

128. Harold Hongju Koh, *The War Powers and Humanitarian Intervention*, 53 HOUS. L. REV. 971, 977–80 (2016).

129. ICISS REPORT, *supra* note 127, at 5, 81.

national law."[130] Unlike humanitarian intervention, R2P focuses on the responsibility of sovereigns rather than their rights.[131] Like humanitarian intervention, however, the ICISS report supported the unilateral use of force by states to prevent widespread human rights atrocities if the U.N. Security Council is unable to act.[132] The U.N. Secretary General subsequently endorsed R2P, but U.N. Member States did not accept that there was either a right or a duty to intervene militarily in humanitarian crises without the approval of the U.N. Security Council.[133]

The R2P principle initially appeared to receive a major boost from the 2011 Libya intervention. The U.N. Security Council authorized the use of force in Libya[134] in part because the Libyan government "forfeited" its "responsibility to protect Libyan citizens, implicitly inviting the United Nations to act"[135] for humanitarian purposes, namely to protect the population from grave human rights violations.[136] Libya was the first time that the Security Council approved the use of force as an application of the R2P doctrine.[137] By all assessments, the Libya intervention started out well. As President Obama put it: "So we actually executed this plan as well as I could have expected: We got a UN mandate, we built a coalition, it cost us $1 billion—which, when it comes to military operations, is very cheap. We averted large-scale civilian casualties, we prevented what almost surely would have been a prolonged and bloody civil conflict."[138]

But today, Libya, a "key test of [the R2P] principle,"[139] is widely recognized as a failure. First, Libya itself is now in effect a failing state, with no central government, warring factions, and mounting civilian casualties.[140] It is not at all clear that the Libyan intervention improved the lives or the

130. *Id.* at xi.

131. *Id.* at 16–17.

132. *Id.* at 54–55.

133. G.A. Res. 60/1, ¶ 30, 2005 World Summit Outcome (Sept. 16, 2005) (accepting state responsibility to "protect its populations from genocide, war crimes, ethnic cleansing and crimes against humanity," which entails the prevention of such crimes "through appropriate and necessary means").

134. *See* S.C. Res. 1973, ¶ 4 (Mar. 17, 2011) (authorizing member states to "take all necessary measures . . . to protect civilians and civilian populated areas" in Libya).

135. *See* Koh, *supra* note 128, at 982.

136. President Obama explained the U.S. action in Libya in terms of human rights. President Barack Obama, Remarks by the President in Address to the Nation on Libya (Mar. 28, 2011).

137. MICHAEL M. DOYLE, THE QUESTION OF INTERVENTION: JOHN STUART MILL AND THE RESPONSIBILITY TO PROTECT 127–28 (2015).

138. Jeffrey Goldberg, *The Obama Doctrine*, ATLANTIC (Apr. 2016), https://www.theatlantic.com/magazine/archive/2016/04/the-obama-doctrine/471525/ [https://perma.cc/5ZJB-52EG].

139. Alex J. Bellamy & Paul D. Williams, *Libya*, *in* THE UN SECURITY COUNCIL IN THE TWENTY-FIRST CENTURY 699, 700 (Sebastian von Einsiedel et al. eds., 2016).

140. Alan J. Kuperman, *Obama's Libya Debacle: How a Well-Meaning Intervention Ended in Failure*, 94 FOREIGN AFF., Jan.–Feb. 2015, at 66, 66–72.

human rights protection of Libyans in the medium-term.[141] Second, Libya has become a haven for terrorists and ISIS, a situation that has forced the United States to use airstrikes within Libya.[142] In that sense, the intervention has led to greater regional unrest, as the ISIS terrorists in Libya threaten other countries.[143] Third, the Libyan intervention launched a contentious debate about the selective use of force and about regime change as an aspect of R2P.[144] Past support of the Gadhafi regime by Western countries despite the regime's dismal record on human rights, along with other factors, led to questions about political and economic motives for the intervention.[145] Finally, the application of the R2P doctrine in Libya appears to be partly responsible for the decisions by Russia and China to block any U.N. Security Council resolution on the use of force for humanitarian purposes in Syria.[146] Syrian protestors and rebels waited for the U.N. Security Council to authorize the use of force—and to save them and their country—but that support never came.[147] The tragic humanitarian crisis and horrific violations of international law in Syria have undercut the R2P doctrine.[148]

More recently, however, in April, 2017, President Trump launched missile strikes in response to Syria's use of chemical weapons, which drew

---

141. President Obama calls Libya a "mess," but Jeffrey Goldberg writes that, "*Mess* is the president's diplomatic term; privately, he calls Libya a 'shit show.'" Goldberg, *supra* note 138; *see also* Seamus Milne, *If the Libyan War Was About Saving Lives, It Was a Catastrophic Failure*, GUARDIAN (Oct. 26, 2011), https://www.theguardian.com/commentisfree/2011/oct/26/libya-war-saving-lives-catastrophic-failure [https://perma.cc/F3BJ-TM3S] (discussing the increase in "mass abduction and detention, beating and routine torture, killings and atrocities," and civilian "massacres" that have occurred after the Libya intervention).

142. *See* Goldberg, *supra* note 138 ("[Libya has] subsequently become an ISIS haven—one that [the United States] has already targeted with air strikes.").

143. *Id.*; *see also* Kuperman, *supra* note 140, at 72 (arguing that "[a]nother unintended consequence of the Libya intervention has been to amplify the threat of terrorism from [Libya] and surrounding countries").

144. *See, e.g.*, Amy Baker Benjamin, *To Wreck a State: The New International Crime*, 19 NEW CRIM. L. REV. 208, 223 (2016) (proposing a new crime in international law that prevents states from using humanitarian concern as a pretext for invading or attacking another state).

145. *See id.* at 230–31 (arguing that potential motivations for Libyan regime change include gaining control of Libya's oil, protecting the status of the French Franc and French influence in Africa, and providing a boost to international banks by destroying Libya's civilian infrastructure).

146. *See* Salman Shaikh & Amanda Roberts, *Syria*, *in* THE UN SECURITY COUNCIL IN THE 21ST CENTURY, *supra* note 139, at 717, 719; Koh, *supra* note 128, at 998; *Russia and China Veto UN Resolution Against Syrian Regime*, GUARDIAN (Oct. 4, 2011), https://www.theguardian.com/world/2011/oct/05/russia-china-veto-syria-resolution [https://perma.cc/JQV8-8EDY].

147. Ian Black & Peter Walker, *Syrian Protestors Demand Action to Halt Killings by Bashar al-Assad Regime*, GUARDIAN (Aug. 1, 2011), https://www.theguardian.com/world/2011/aug/01/syria-demands-action-assad-killings [https://perma.cc/N32V-TNHR].

148. *See* Tom Esslemont, *As Syrian Deaths Mount, World's "Responsibility to Protect" Takes a Hit: Experts*, REUTERS (Oct. 24, 2016), https://www.reuters.com/article/us-mideast-crisis-syria-law/as-syrian-deaths-mount-worlds-responsibility-to-protect-takes-a-hit-experts-idUSKCN12O2S3 [https://perma.cc/NXA2-S3UC] (discussing the limited success of R2P, which is described as a "high moral aspiration" that has "floundered" on the complex realities of warfare today, especially in the Syrian conflict).

significant support from other states—in the form of both explicit approval and lack of condemnation.[149] Most important, perhaps, was China's decision not to condemn the strikes as a violation of international law.[150] These events provide some support for R2P, although the U.S. government has not defended its actions in those terms. Russia's condemnation of the strikes illustrate the difficulties of implementing R2P to the mutual satisfaction of countries with different strategic objectives.[151]

6. Conclusion.—Before leaving the five doctrinal innovations analyzed in this section, let us consider two overarching counterarguments. First, these five doctrines are not at the core of international human rights. In one sense this counterargument is correct. The core substantive human rights commitments that are made by and/or are binding upon states are set out in treaty instruments and in customary international law, not in the doctrines addressed in this section, which generally go to enforcement. Even if the five doctrines are of declining significance, they could be abandoned without changing substantive international human rights obligations. Enforcement is nevertheless, however, a core issue of international human rights law. In a sense, enforcement is the key challenge of international human rights law; if methods of enforcement are stripped away, then the promise of international human rights *law* has gone unfulfilled. As well, these doctrinal innovations are not "merely" about enforcement. Taken together, they are important components of the claim that sovereignty and international law have been reframed to have individual human rights and welfare at their core.[152] To abandon these doctrinal innovations is to gut much of the core of the purported transformation of sovereignty and international law, although other aspects of that transformation, including international criminal law, are not addressed here.

A second counterargument is that the international peace and security costs associated with these doctrinal innovations have not actually been so great, with the possible exceptions of humanitarian intervention, unilateral

149. *See* Madison Park, *Who's with the US on Syria Strike and Who Isn't*, CNN (Apr. 9, 2017), http://edition.cnn.com/2017/04/07/world/syria-us-strike-world-reaction/index.html [https://perma.cc/4ZYL-6NX8].

150. *See Foreign Ministry Spokesperson Hua Chunying's Regular Press Conference on April 7, 2017*, MINISTRY FOREIGN AFF. CHINA (Apr. 7, 2017), http://www.fmprc.gov.cn/mfa_eng/xwfw_665399/s2510_665401/2511_665403/t1452149.shtml [https://perma.cc/7VE3-HRZH].

151. *Briefing by Foreign Ministry Spokesperson Maria Zakharova, Moscow, April 12, 2017*, MINISTRY FOREIGN AFF. RUSS. (Apr. 12, 2017), http://www.mid.ru/en/web/guest/foreign_policy/news/-/asset_publisher/cKNonkJE02Bw/content/id/2725573#9 [https://perma.cc/LT7M-S5V6].

152. *See, e.g.*, TEITEL, *supra* note 13, at 8–11; William J. Aceves, *Liberalism and International Legal Scholarship: The* Pinochet *Case and the Move Toward a Universal System of Transnational Law Litigation*, 41 HARV. INT'L L.J. 129, 183 (2000); Peters, *supra* note 13, at 514.

secession in Kosovo, and the use of R2P in Libya. Moreover, even if states *say* that a particular doctrine undermines peace and security, that statement may be insincere and may instead mask a self-serving preference not to enforce human rights through international law.

Even assuming that these objections are correct, the descriptive point remains: the limited (and apparently dwindling) state support for these doctrines suggests the purported transformation of international law around human rights is on questionable footing. As well, to the extent that the costs of the doctrines have been low, that may be because these enforcement measures have been used so infrequently. Like the low levels of funding for the U.N. treaty bodies, discussed below, these doctrines have either been too costly or not nearly costly enough. If states were serious about a reorientation of international law around human rights, the foregoing doctrines would be employed far more often, reducing the problem of selective enforcement, but with potentially far higher costs.

B. Costs to the "Territorial Peace"

Political scientists provide another way to consider the costs of the human rights-related doctrinal innovations in international law. Empirical scholarship has linked peace with an absence of territorial conflict: a "territorial peace." Some of the doctrines described above undermine the ways in which international law tends to generate territorial security.

The occurrence of interstate war and overall mortality in war sharply decreased after 1815 and again after 1945.[153] Scholars do not fully agree on the reason for this decline, but the empirical evidence suggests that a decline in armed disputes may be caused by the "democratic peace" or the "territorial peace." The first, discussed in greater detail in Part III, posits that democracies are less likely to have conflicts with each other than they are with nondemocracies.[154] The second posits that war has declined because there has been a reduction in conflict over territory.[155] The democratic peace

153. Azar Gat, *Is War Declining—and Why?*, 50 J. PEACE RES. 149, 152 (2012); *see also* STEVEN PINKER, THE BETTER ANGELS OF OUR NATURE: WHY VIOLENCE HAS DECLINED 249–50 (2011) (describing the many categories of war which have not happened since 1945, including wars with nuclear weapons, battlefield combat between superpowers, and interstate wars between major developed countries).

154. DAVID SOBEK, THE CAUSES OF WAR 84–85 (2009); Johann Park & Patrick James, *Democracy, Territory, and Armed Conflict, 1919–1995*, 11 FOREIGN POL'Y ANALYSIS 85, 86 (2015). The democratic peace is sometimes attributed to the structural aspects of democracy such as elections, which allow voters to punish the leaders and the political parties that take a country into war. *See* David Lektzian & Mark Souva, *A Comparative Theory Test of Democratic Peace Arguments, 1946–2000*, 46 J. PEACE RES. 17, 32 (2009).

155. For a literature review of territorial peace theory, see John A. Vasquez & Emily E. Barrett, *Peace as the Absence of Militarized Conflict: Comparing the Democratic and Territorial Peace*, 2 J. TERRITORIAL & MAR. STUD., Winter/Spring 2015, at 5-8.

literature is long-standing and well-established; territory is the "rising star" among explanations of the interstate-conflict processes.[156]

Numerous empirical studies show that territorial disputes are a leading cause of war and that their absence is a significant predictor of peace. Since 1815, territorial disputes have had a higher probability of leading to war than any other kind of dispute.[157] Moreover, for wars between 1816–1997, territorial wars were the most prevalent form of war.[158] Another study of wars from 1648 through 2007 concluded that "[t]erritory has consistently constituted the issue over which states have most frequently gone to war, and this is true by a wide margin."[159] States able to resolve their territorial disputes[160] thus eliminate a significant cause of war. Territorial disputes in relationships between states dropped dramatically after 1945, providing one explanation for the decrease in interstate wars.[161] Additional studies have also concluded that the absence of territorial claims is associated with fewer militarized interstate disputes (MIDs), a measure that includes war, as well as military displays and threats of war.[162] The effect of territorial disputes on war and peace is felt in an additional way: eliminating territorial conflict apparently not only eliminates a war-generating issue, but it may also reduce hostility around *other* issues.[163] The association between the absence of territorial claims and peace is referred to as the "territorial peace."[164]

Why are territorial disputes apparently so detrimental to peace? Perhaps removing territorial disputes allows states to shrink their militaries, reducing the perception that they are threatening.[165] Another possibility: researchers posit that MIDs are more likely to result from territorial claims that are imbued with intangible values like ethnic worth than they are from territorial claims without intangible values.[166] Perhaps settling territorial claims with high intangible values has more "spill-over" effect than the settlement of

156. Park & James, *supra* note 154, at 86.

157. Vasquez & Barrett, *supra* note 155, at 6 (citing multiple studies, e.g., JOHN A. VASQUEZ, THE WAR PUZZLE (1993) and Paul R. Hensel, *Charting a Course to Conflict: Territorial Issues and Interstate Conflict, 1816–1992*, 15 CONFLICT MGMT. & PEACE SCI., at 43–73 (1996)).

158. *See* John A. Vasquez & Brandon Valeriano, *Classification of Interstate Wars*, 72 J. POL. 292, 300 (2010).

159. GARY GOERTZ ET AL., THE PUZZLE OF PEACE: THE EVOLUTION OF PEACE IN THE INTERNATIONAL SYSTEM 89 (2016).

160. For a discussion of what counts as a "territorial dispute," see *id.* at 79–80. Although studies vary some in how they define that term, it generally does not include peaceful boundary management or maritime disputes (unless they are accompanied by dispute over land). *Id.* at 80.

161. *Id.* at 93 tbl.4.4.

162. *See id.* at 91–92 (analyzing MIDs and showing that "the raw frequency of territorial disputes, as well as their percentage relative to other issues[,] . . . has declined in the post-World War II era").

163. Vasquez & Barrett, *supra* note 155, at 6.

164. GOERTZ ET AL., *supra* note 159, at 82.

165. *Id.* at 101 (citation omitted).

166. Vasquez & Barrett, *supra* note 155, at 9.

other kinds of claims because they tend to be especially inflammatory, making all other issues more intractable.

Whatever the reasons, the foregoing analysis has important implications for international law. Legal rules and institutions that reduce conflict over territory appear likely to reduce military conflicts. And they are likely to do so with greater positive effect than legal rules and institutions that reduce political or economic disagreements between states. International legal rules designed to reduce conflict over territory and borders include: Article 2(4) of the U.N. Charter and other international legal rules that prohibit the use of force to acquire territory; the doctrine of *uti possidetis*, which provides that newly formed states keep the territorial borders they had before decolonization or their previous internal borders; and the lack of a right to violent secession.

Not only are these legal rules partially designed to reduce territorial conflict, but we now also have a body of empirical evidence that suggests they have been successful. Territorial conquests, for example, declined during the twentieth century as the international rule limiting the use of force hardened.[167] The conclusion of the Kellogg–Briand Pact in 1928 arguably precipitated the decline in conquests;[168] in any event, prohibition on the use of force for territorial conquest was strengthened in the U.N. Charter and became a cornerstone of the post-World War II international legal order.[169] Second, the *uti possidetis* norm appears to have reduced territorial conflict.[170] Third, violent secessions have decreased during the twentieth century.[171]

The reorientation of the international legal system toward individual human rights and away from state sovereignty threatens to undermine all three of these doctrines. The focus on human rights may also threaten to undermine the ability of the U.N. Security Council to resolve territorial disputes peaceably. As described above, humanitarian intervention, Responsibility to Protect, a right to democracy, and a right to violent secession all weaken Article 2(4) of the U.N. Charter. The *uti possidetis* doctrine is challenged to the extent it violates the rights of people who are rendered a potentially oppressed minority in the new state.[172] The "territorial

167. *See* TANISHA M. FAZAL, STATE DEATH: THE POLITICS AND GEOGRAPHY OF CONQUEST, OCCUPATION, AND ANNEXATION 169–228 (2007); GOERTZ ET AL., *supra* note 159, at 112–17; Mark W. Zacher, *The Territorial Integrity Norm: International Boundaries and the Use of Force*, 55 INT'L ORG. 215 (2001).

168. OONA HATHAWAY & SCOTT SHAPIRO, THE INTERNATIONALISTS: HOW A RADICAL PLAN TO OUTLAW WAR REMADE THE WORLD 309–35 (2017).

169. U.N. Charter art. 2, ¶ 4.

170. David B. Carter & H. E. Goemans, *The Making of the Territorial Order: New Borders and the Emergence of Interstate Conflict*, 65 INT'L ORG. 275, 275–76 (2011).

171. GOERTZ ET AL., *supra* note 159, at 126–28.

172. *See, e.g.*, Steven R. Ratner, *Drawing a Better Line:* Uti Possidetis *and the Border of New States*, 90 AM. J. INT'L L. 590, 612 (1996). *Uti possidetis* is also challenged as potentially increasing the number of territorial disputes (rather than reducing them) because it may lead groups of people

peace" literature suggests that to the extent human rights-related doctrines render borders less secure and the ownership of territory less certain, there may be significant costs to international peace and security.

Using military force in the humanitarian intervention and R2P contexts, or to install democracy or effectuate violent secession, is not the same as using force for territorial acquisition. Arguably, using force in these human rights-promoting contexts does not destabilize borders, even if the use of force for territorial acquisition *does* lead to conflict about borders. But our actual experience with recent uses of force does not bear out this argument. Consider Kosovo—one of the key examples in the arguments that sovereignty and international law have been reconceptualized.[173] Although the intervention violated Article 2(4) of the U.N. Charter, the use of force by NATO was not intended as territorial acquisition. Its purpose was, instead, to prevent mass atrocities and human rights violations.[174] Nor was territorial acquisition the direct effect of the intervention: NATO countries did not claim the territory of Kosovo for themselves.

The intervention and subsequent secession by Kosovo from Serbia did nonetheless destabilize borders. The precedent set by the Western states with respect to Kosovo, which later seceded from Serbia and became a Western ally, became the legal basis for Russia's intervention in Georgia and in Crimea, Ukraine, ostensibly at the request of newly seceded regimes. The newly contested border between Ukraine and Russia has, in turn, generated broader tensions between Russia, the United States, and NATO.[175] The empirical data tells us that this is the kind of conflict—about borders and territory—most likely to escalate into a militarized dispute.

---

unhappy with the borders to make territorial claims, perhaps through the use of force. Empirical work does not appear to support this argument. *See* Carter & Goemans, *supra* note 170, at 304.

173. As another example, the U.S. invasion of Iraq, designed in part to remedy human rights violations and to promote democracy, has arguably failed to do either and has also destabilized the Iraqi border with Turkey and with the quasi- or pseudo-states of ISIS and Kurdistan. *See* Metin Gurcan, *Turkey Sticks Its Neck Out Again, This Time in Iraq*, AL-MONITOR (Dec. 7, 2015), https://www.al-monitor.com/pulse/originals/2015/12/turkey-iraq-becomes-third-largest-army.html [https://perma.cc/TAY3-W25L].

174. Press Release, North Atlantic Treaty Organization [NATO], The Situation in and Around Kosovo, Statement Issued at the Extraordinary Ministerial Meeting of the North Atlantic Council, NATO Press Communique M-NAC-1 (99)51, ¶ 2 (Apr. 12, 1999) ("[NATO] condemns these appalling violations of human rights and the indiscriminate use of force by the Yugoslav government. These extreme and criminally irresponsible polices . . . have made necessary and justify the military action by NATO.").

175. Bruce Blair, *Could U.S.–Russia Tensions Go Nuclear?*, POLITICO (Nov. 17, 2015), http://www.politico.eu/article/could-u-s-russia-tensions-go-nuclear/ [https://perma.cc/W6M4-R5EA]; Jeremy Diamond & Greg Botelho, *U.S.–Russia Military Tit for Tat Raises Fears of Greater Conflict*, CNN (June 19, 2015), http://www.cnn.com/2015/06/17/politics/russia-us-military-threats-rise-ukraine/ [https://perma.cc/M2SX-C5MG]; *Relations with Russia*, NORTH ATLANTIC TREATY ORG. (June 16, 2017), http://www.nato.int/cps/en/natolive/topics_50090.htm# [https://perma.cc/CH75-NJ3T].

Note that there are many potential reasons for the decrease in armed conflict in the post-World War II period and before, many of which are not closely tied to international law.[176] Nevertheless, the strength of the association between territorial disputes and war in the empirical literature suggests that if international law can reduce—or contribute to a reduction in—conflict over territory, it will contribute to a reduction in armed conflict, whatever the mixture of factors which may have contributed to peace historically or which may do so in the future.

C. *The United Nations and Polarization Costs*

The United Nations, too, has been changed by human rights in ways that appear to impose costs on international cooperation and international law. The United Nations and the treaty that created it, the U.N. Charter, are the cornerstones of the post-World War II international order.[177] Developments in the United Nations over the past several decades appear in some respects to parallel the developments in the doctrinal enforcement mechanisms discussed above: the enforcement of international human rights law through the United Nations is widely perceived as ineffective and selective, leading to polarization and a lack of credibility that may hamper the overall work of the United Nations. The trajectory is a bit different, however. The U.N. human rights bodies did not experience the same turn-of-the-century heyday, although there was a spurt of optimism around the reform of the Human Rights Council in 2006. Here, the story is mostly one of stasis. It is also important to note again that measuring the effectiveness of international human rights law and institutions is generally difficult and contested.

The purposes of the United Nations are "[t]o maintain international peace and security," and "[t]o achieve international cooperation in solving international problems of an economic, social, cultural, or humanitarian character, and in promoting and encouraging respect for human rights and for fundamental freedoms . . . ."[178] This choice of language subordinates the promotion and encouragement of respect for human rights to the achievement of international cooperation and to the maintenance of peace and security,[179] a choice underscored by the placement of the responsibility for human rights with the Economic and Social Council of the U.N. (ECOSOC), an "organ in

176. *See, e.g.*, Gat, *supra* note 153, at 2–8 (discussing many possibilities such as economic growth, commmercial interdependence, social attitude changes, and nuclear deterrence); Pinker, *supra* note 153, at 189–288 (discussing many possibilities including the value placed on human life, nuclear weapons, democracy, and international organizations).

177. ICISS REPORT, *supra* note 127, at 52.

178. U.N. Charter art. 1, ¶¶ 1, 3.

179. Martti Koskenniemi, *The Police in the Temple: Order, Justice and the UN: A Dialectical View*, 6 EUR. J. INT'L L. 325, 336–37 (1995).

decline almost from the start."[180] The language also does not necessarily mean human rights are or should be protected by or as *international law*. Indeed, contemporary observers understood that the Charter envisioned "voluntary cooperation" on human rights as coordinated and facilitated by the General Assembly and the ECOSOC.[181] Human rights are protected by domestic law and constitutions, regional organizations, and moral or religious beliefs and values. In all of these forms, human rights can be "promoted" and "encouraged" through international cooperation and international institutions without international *legal* protection.

When the U.N. Charter was drafted and adopted, few—if any—human rights were protected by international law, so it is largely in this broader sense that the purposes of the United Nations include the promotion and encouragement of respect for human rights. The United Nations itself was intended, at least by the Allied leaders who planned and created it, as a security framework that would balance great powers against one another and thereby promote international peace and security as well as ensure the sovereign equality of all nations.[182]

Perhaps it is unsurprising, then, that the United Nations' track record in some aspects of protecting human rights has been poor. Although the United Nations served as the institutional forum for the negotiation and conclusion of many multilateral human rights agreements,[183] the mechanisms of international legal enforcement of human rights have not been a clear success for human rights themselves. Moreover, efforts by the United Nations to enforce human rights, especially international human rights law, appear to have detracted from international cooperation as a whole by contributing to the organization's loss of credibility, as discussed in subpart II(C), and by contributing to polarization, as discussed below.

1. Human Rights Council.—The unfortunate failure of the Human Rights Commission, its dismantling by the United Nations in 2006, and its reincarnation as the Human Rights Council[184] illustrate some of the difficulties of the U.N.'s efforts to enforce human rights through international

180. Daphna Shraga, *The Security Council and Human Rights—from Discretion to Promote to Obligation to Protect*, *in* SECURING HUMAN RIGHTS?: ACHIEVEMENTS AND CHALLENGES OF THE UN SECURITY COUNCIL 8, 9–10 (Bardo Fassbender ed., 2011).

181. Bardo Fassbender, *Introduction* to SECURING HUMAN RIGHTS?: ACHIEVEMENTS AND CHALLENGES OF THE UN SECURITY COUNCIL, *supra* note 180, at 1, 2.

182. Koskenniemi, *supra* note 179, at 335; *see also* C. L. Lim, *The Great Power Balance, the United Nations and What the Framers Intended: In Partial Response to Hans Köchler*, 6 CHINESE J. INT'L L. 307, 309–14 (2007).

183. *See, e.g.*, Convention Against Torture and Other Cruel, Inhuman or Degrading Treatment or Punishment, Dec. 10, 1984, 1465 U.N.T.S. 85; G.A. Res. 217 (III) A, Universal Declaration of Human Rights (Dec. 10, 1948); World Conference on Human Rights, Vienna Declaration and Programme of Action, U.N. Doc. A/CONF.157/23 (June 25, 1993).

184. G.A. Res. 60/251, ¶ 1 (Mar. 15, 2006).

law. The Human Rights Commission was created in 1946.[185] Its loss of legitimacy and the reasons it was shuttered are generally summarized in these terms:

> The Commission was notoriously impotent, farcical even. Countries with egregious human rights records, from Libya to Sudan, managed to become Commission members, affording them a platform to deflect criticism, obstruct meaningful action against flagrant atrocities, and, to the chagrin of the United States, disproportionately bash Israel. The Commission's work often pitted the West against non-Western countries, which tended to vote in blocs along the lines of the Organization for Islamic Cooperation (OIC) and the Non-Aligned Movement (NAM). By the time it was shuttered in 2006, the Commission was discredited and disgraced.[186]

Early reports on the Council suggest that its work is also dominated by regional and political bloc voting, including a disproportionate focus on Israel. As Professor Rosa Freedman describes:

> Despite warnings about selectivity, bias, double standards, and loss of credibility, from the outset Council discussions were dominated by states seeking to vilify Israel and to keep the spotlight on that region. A large number of OIC states were able to express, and use their votes to achieve, collective positions. The OIC sought to retain focus on the OPT as part of national and regional foreign policies including political, religious, cultural, and regional ties with the Palestinians and with affected neighbouring states. OIC states also used the situation to divert attention away from other gross and systemic violations within the Middle East or within influential OIC Council members such as Pakistan, Algeria, and Egypt.[187]

As a result, the Human Rights Council, like the Human Rights Commission before it, is of questionable value, despite its new tools such as Universal Periodic Review, which requires all states to appear before the Council to report on human rights in their country.[188] To be sure, it is possible

---

185. *United Nations Commission on Human Rights*, U.N. OFF. HIGH COMMISSIONER, http://www.ohchr.org/EN/HRBodies/CHR/Pages/CommissionOnHumanRights.aspx [https://perma.cc/F2NR-N74U].

186. Daniel Chardell, *Gaining Ground at the UN Human Rights Council*, COUNCIL ON FOREIGN REL. (Oct. 3, 2014), http://blogs.cfr.org/patrick/2014/10/03/gaining-ground-at-the-un-human-rights-council/ [https://perma.cc/K88C-U4CH].

187. Rosa Freedman, *The United Nations Human Rights Council: More of the Same?*, 31 WIS. INT'L L.J. 208, 226 (2013) [hereinafter Freedman, *More of the Same?*]; *see also* SUBEDI, *supra* note 14 at 129–39.

188. Vincent Chetail, *The Human Rights Council and the Challenges of the United Nations System on Human Rights: Towards a Cultural Revolution?*, *in* INTERNATIONAL LAW AND THE QUEST FOR ITS IMPLEMENTATION 222–23 (Laurence Boisson de Chazournes & Marcelo Kohen eds., 2010).

that the Council's work is improving,[189] or that it has a positive overall impact on human rights in indirect ways, for example by increasing the effectiveness of other enforcement mechanisms such as domestic courts and regional courts and tribunals. Both points highlight the difficulty in measuring human rights outcomes and underscore the need for caution in proposing solutions.

The questionable efficacy of the U.N.'s human rights work does mean, however, that at a minimum we should take seriously the potential costs of that work to the other objectives of the United Nations. The Human Rights Commission and Council may have had a negative effect on the work of the United Nations as a whole. First, as human rights have become more central to the work of the United Nations, the failures of the Commission and Council may become more closely tied to the success—or not—of the entire institution.[190] These "credibility costs" are discussed below in subpart II(C). Second, divisions within the Human Rights Commission and Council, as well as human rights issues generally, appear to contribute to regional and political divisions among states, which hinder the work of the General Assembly and the Security Council on a wide variety of issues. These divisions are referred to in this Article as "polarization."[191]

The work of the U.N. General Assembly is, according to many observers, undercut by polarization.[192] The Cold War politics within the General Assembly—which often prevented it from acting constructively—illustrate the point historically.[193] Even in the period after 1990, U.N. General Assembly voting has been dominated by a divide between the West and non-Western countries. Indeed, in the years 1990–2011, "[m]ore than 90 percent of every single vote casted ('yes' or 'no') can be classified correctly by the placement of states on one single dimension": the West and the rest.[194]

---

189. *See* Alston, *supra* note 9 (noting that "the Human Rights Council has been 'operating in a way that is surprisingly balanced in the last few years, especially if the issue of Israel and the occupied Palestinian territories is put to one side,'" but also noting that China, Russia, and growing populism are likely to change this dynamic).

190. *See The Shame of the United Nations*, N.Y. TIMES (Feb. 26, 2006), http://www.nytimes.com/2006/02/26/opinion/the-shame-of-the-united-nations.html [https://perma.cc/8HUS-ZEB6]; George W. Bush, President of the U.S., Address at the UN High-Level Plenary Meeting (Sept. 14, 2005) (transcript available at https://2001-2009.state.gov/p/io/rls/rm/53111.htm [https://perma.cc/N8JZ-645V]).

191. *See* Nicholas O. Stephanopoulos, *Race, Place, and Power*, 68 STAN. L. REV. 1323, 1349 (2016) (defining "polarization" in terms of group political cohesion and bloc voting). A related problem is that of "politicization," sometimes defined as deliberation or decision-making based on political objectives unrelated to the issue being debated or decided. *See* Freedman, *More of the Same?*, *supra* note 187, at 210–11 for a discussion of politicization in the human rights context.

192. *See, e.g.*, LINDA FASULO, AN INSIDER'S GUIDE TO THE UN 88–90 (3d ed. 2015).

193. *See* Frederick H. Gareau, *Cold-War Cleavages as Seen from the United Nations General Assembly: 1947–1967*, 32 J. POL. 929, 931–32 (1970); Amy L. Sayward, *International Institutions*, *in* THE OXFORD HANDBOOK OF THE COLD WAR 384–85 (Richard H. Immerman & Petra Goedde eds., 2013).

194. Nicolas Burmester & Michael Jankowski, Comparing Regional Organizations in the United Nations General Assembly—Is There a Shift to Regionalism? 13 (Mar. 26, 2014)

General Assembly votes on human rights resolutions have had a significant role in generating the polarization found in General Assembly voting patterns. One analysis finds, for example, that human rights resolutions within the General Assembly have a very high "discrimination parameter," meaning they "weigh heavily" in the conflict between liberal and nonliberal states.[195] That study showed human rights votes with a higher discrimination parameter than colonialism, economic issues, disarmament, Middle East, and nuclear issues.[196] To be clear, this data does not show that removing human rights from the U.N. agenda would improve cooperation in other areas;[197] it shows only that voting on human rights plays a large role in the measures of polarization within the U.N.

A connection between bloc voting and the U.N.'s human rights agenda is also suggested by the data on voting within the Human Rights Council (HRC). A study focusing just on human rights has found that in both the General Assembly and the HRC, "China, Russia and developing countries pass regular resolutions undercutting Western human rights agendas."[198] Other studies found greater polarization in the HRC than in the Commission[199] and that measures proposed by Cuba (one of the most frequent proposers) "strongly polarize the member states of the UNHRC in their voting."[200]

Polarization leads to several problems, including lengthy procedural discussions and many repetitive resolutions.[201] More importantly, however, the voting blocs impede the ability of the General Assembly to make meaningful progress on other issues.[202] For example, Russia's use of force to

---

(unpublished Paper Presented at the 55th Annual Convention of the International Studies Association), https://pure.au.dk/ws/files/81269013/Burmester_Jankowski_ISA_2014.pdf [https://perma.cc/QV2R-BRQE]; *see also* Erik Voeten, *Data and Analyses of Voting in the United Nations General Assembly*, *in* ROUTLEDGE HANDBOOK OF INTERNATIONAL ORGANIZATION 54, 60–61 (2013) (providing statistical analysis of the "West–Rest" divide).

195. Michael A. Bailey, Anton Strezhnev & Erik Voeten, *Estimating Dynamic State Preferences from United Nations Voting Data*, 61 J. CONFLICT RESOL. 430, 444 (2017).

196. *Id.* For more discussion of states' voting patterns on human rights resolutions, see generally Bernhard Boockmann & Axel Dreher, *Do Human Rights Offenders Oppose Human Rights Resolutions in the United Nations?*, 146 PUB. CHOICE 443 (2011).

197. *Cf.* JOSEPH E. SCHWARTZBERG, TRANSFORMING THE UNITED NATIONS SYSTEM: DESIGNS FOR A WORKABLE WORLD 26–27 (2013) (suggesting that almost all human rights issues be taken off the U.N. General Assembly agenda).

198. Richard Gowan, *Who is Winning on Human Rights at the UN?*, EUR. COUNCIL ON FOREIGN REL. (Sept. 24, 2012), http://www.ecfr.eu/article/commentary_who_is_winning_on_human_rights_at_the_un [https://perma.cc/C9CA-SDEK].

199. Simon Hug, *Dealing with Human Rights in International Organizations*, 15 J. HUM. RTS. 21, 22 (2016).

200. Simon Hug & Richard Lukás, *Preferences or Blocs? Voting in the United Nations Human Rights Council*, 9 REV. INT'L ORG. 83, 103 (2014).

201. Freedman, *More of the Same?*, *supra* note 187, at 216.

202. *See* FREEDMAN, FAILING TO PROTECT, *supra* note 14, at 20–21 (2014); *see also* FASULO, *supra* note 192, at 88; ARYEH NEIER, THE INTERNATIONAL HUMAN RIGHTS MOVEMENT: A

annex Crimea was a violation of international law that undermines peace and security in Eastern Europe. Yet when the U.N. General Assembly voted to condemn Russia's actions, a large number of countries abstained, showing that "many countries see this as a struggle between power blocs rather than as a fundamental question of international order and do not accept the West's self-identification as the guardian of liberal order."[203]

The polarization costs of international human rights also appear to effect the work of the U.N. Security Council. As described above, the Security Council was not originally intended to address human rights issues. The structure of the Security Council is ill-suited to that task. As Martii Koskenniemi puts the point, the Security Council safeguards security as the "technician of peace," but when it comes to human rights, "[i]ts composition, procedures and practices are completely indefensible,"[204] in part because it is controlled by the "Great Powers"—the five veto-wielding permanent members (the P-5).[205] Decisions about human rights enforcement are especially prone to the perception of selection bias and to the perception that human rights serve as a "smokescreen for intervention or regime change,"[206] in part because human rights violations are so widespread and yet go largely unaddressed. These perceptions (or realities) exacerbate the "West v. rest" divide. Human rights also are likely to be issues for which the Security Council has difficulty reaching principled agreement, in part because Russia and China hold different views about human rights enforcement than do the other members of the P-5.[207]

2. U.N. Treaty Bodies.—Each of the ten core human rights treaties establishes a U.N. human rights body. Although they differ in some particulars, especially in their competence to investigate or visit countries under their mandate, the human rights treaty bodies share many features and tasks. Members of the treaty bodies are not states themselves (distinguishing these bodies from the HRC) but are instead independent human rights experts who are nominated and elected by the state parties. State parties are required to submit periodic reports to the treaty bodies, which then review the states'

HISTORY 110–12 (2012); THOMAS G. WEISS, WHAT'S WRONG WITH THE UNITED NATIONS AND HOW TO FIX IT 53 (2d ed. 2012).

203. Piotr Buras et al., *Ten Global Consequences of the Ukraine Crisis*, EUR. COUNCIL ON FOREIGN REL. (June 16, 2014), http://www.ecfr.eu/article/commentary_ten_global_consequences_of_ukraine272 [https://perma.cc/SP5T-YU3L].

204. Koskenniemi, *supra* note 179, at 344.

205. *Id.* at 338.

206. U.N., SCOR, 66th Sess., 6531st mtg. at 11, U.N. Doc. S/PV.6531 (May 10, 2011); *see also* RHOADS, *supra* note 17, at 2 (describing a shift at the U.N. towards "[t]he realization, promotion, and protection of human rights," which has "translated into forms of international engagement that are less consensual and more compulsory and coercive, justified by upholding human rights").

207. *See infra* subpart III(C).

domestic legislation and policies and make recommendations to the states on how to better comply with their obligations under the treaty. Treaty bodies sometimes issue general comments on how to best interpret the treaties that they are charged with enforcing. None of the decisions, recommendations, comments, or other output of any treaty body is binding on any state.[208]

The primary goal of the treaty bodies is to promote compliance with the substantive obligations of the human rights treaties for which they are responsible. Their main tools are the reporting and monitoring requirements, coupled with the recommendations made by the treaty bodies and dialogue with the state parties which results from this process.[209] The written and oral presentation of data and viewpoints that are part of this process can clarify to states what the treaty requires, and it may motivate and assist states to achieve greater compliance.

Measuring the effectiveness of the treaty bodies (and the treaties themselves) is notoriously difficult.[210] It is hard to assess compliance with the substantive norms of the treaties, and it is even more difficult to determine whether changes in norm compliance are caused by the treaty and its monitoring body, as opposed to other factors.[211] Also, the treaty monitoring bodies could make other enforcement mechanisms, such as domestic and regional legal regimes, more successful. However, what data is available suggests the treaty monitoring bodies are not especially effective. In 2010 and 2011, only 16% of the reports due to the treaty bodies were submitted on time.[212] Under some treaties, such as the Convention Against Torture, around 20% of state parties have never submitted a report, and for the Convention on the Rights of Persons with Disabilities and some other treaties, the figure is even higher.[213] There was a relative decrease in reporting compliance between 2000 and 2012.[214]

The low level of reporting has a positive side. The treaty bodies lack the resources to adequately evaluate even the reports they do receive so that there

208. *See* Kerstin Mechlem, *Treaty Bodies and the Interpretation of Human Rights*, 42 VAND. J. TRANSNAT'L L. 905, 907–08 (2009). For a detailed introduction, see HELLEN KELLER & GEIR ULFSTEIN, UN HUMAN RIGHTS TREATY BODIES: LAW AND LEGITIMACY 1–2 (2012).

209. *See* Mechlem, *supra* note 208, at 907; Office of the High Commissioner for Human Rights, *Report on the Second Consultation of States on Treaty Body Strengthening*, ¶ 9 (Feb. 7–8, 2012); Navanethem Pillay, *Rep. of U.N. High Commissioner for Human Rights on Strengthening the Human Rights Treaty Body System*, at 12, U.N. Doc. A/66/86 (June 26, 2012); G.A. Res. 68/268, ¶ 2 (Apr. 21, 2014).

210. *See* Adam Chilton & Dustin Tingley, *Why the Study of International Law Needs Experiments*, 52 COLUM. J. TRANSNAT'L L. 173, 175–76 (2013); Yuval Shany, *Assessing the Effectiveness of International Courts: A Goal-Based Approach*, 106 AM. J. INT'L L. 225, 239 (2012).

211. *See* Shany, *supra* note 210, at 239.

212. Pillay, *supra* note 209, at 9.

213. *Id.* at 22.

214. *Id.*

is already a growing backlog of reports awaiting consideration.[215] The time lag between submission and consideration is costly because the drafters of the reports may no longer be available when the report is considered, undermining the potential for constructive dialogue on compliance.[216] Scholars are generally skeptical about the effectiveness of the treaty monitoring bodies,[217] although perhaps there is reason for some optimism in countries with a strong civil society that engages with monitoring bodies on an ongoing basis.[218]

The questionable effectiveness of the treaty bodies may also be inferred from the repeated efforts to reform the system in order to improve their work.[219] The most recent reform effort began in 2009, when the U.N. High Commissioner for Human Rights began a review of treaty monitoring bodies designed to make them more effective as a whole.[220] Professor Yuval Shany evaluated the reform proposals and concluded that the report "refus[ed] to acknowledge the 'elephant in the room'—namely, what appears to be a conscious decision by a significant number of state-parties to maintain the treaty bodies under permanent conditions of under-effectiveness."[221] He continues:

> The fact that many states-parties have opposed, by and large, past attempts to seriously explore such fundamental changes suggests, however, that they do not wish to strengthen the treaty body system. The unhappy situation of the UN treaty bodies may thus be explained in large part by a tension between a superficial commitment by many state-parties to the goal of human rights promotion and a *realpolitik* aversion to actual treaty implementation.[222]

---

215. *See* Pillay, *supra* note 209, at 23.

216. JASPER KROMMENDIJK, THE DOMESTIC IMPACT AND EFFECTIVENESS OF THE PROCESS OF STATE REPORTING UNDER UN HUMAN RIGHTS TREATIES IN THE NETHERLANDS, NEW ZEALAND, AND FINLAND: PAPER-PUSHING OR POLICY PROMPTING? 14 (2014).

217. *See, e.g.*, *id.* at 9–24; SUBEDI, *supra* note 14 at 88-94; EMILIE HAFNER-BURTON, MAKING HUMAN RIGHTS A REALITY 86–115 (2013); Christof Heyns & Frank Viljoen, *The Impact of the United Nations Human Rights Treaties on the Domestic Level*, 23 HUM. R.Q. 483, 488, 511 (2001).

218. *See* Creamer & Simmons, *supra* note 8, at 12 (finding participation in the review process by a treaty monitoring body may provide an opening for constructive engagement and "small" but "incremental" improvement in human rights protections).

219. *See* Philip Alston, *Final Report on Enhancing the Long-Term Effectiveness of the United Nations Human Rights Treaty System*, transmitted by the Secretary-General to the Comm'n on Human Rights, ¶ 98, UN Doc. E/CN.4/1997/74 (Mar. 27, 1997); U.N. High Comm'r for Human Rights, *The OHCHR Plan of Action: Protection and Empowerment*, ¶¶ 145–50 (May 2005), http://www2.ohchr.org/english/planaction.pdf [https://perma.cc/U89E-ZDKA].

220. Pillay, *supra* note 209, at 9.

221. Yuval Shany, *The Effectiveness of the Human Rights Committee and the Treaty Body Reform*, *in* DER STAAT IM RECHT: FESTSCHRIFT FÜR ECKART KLEIN ZUM 70 (Marten Breuer et al. eds., 2013).

222. *Id.*; *see also* COURTNEY HILLEBRECHT, DOMESTIC POLITICS AND INTERNATIONAL HUMAN RIGHTS TRIBUNALS: THE PROBLEM OF COMPLIANCE 140–41 (2014); KROMMENDIJK, *supra* note 216, at 12; SUBEDI, *supra* note 14, at 226–29, 255–56.

That is the key point. States are not committed enough to the international legal enforcement of human rights norms to establish a successful system. This unwillingness helps explain why human rights are understood in political terms. Most states lack a true commitment to the meaningful enforcement of human rights norms as international law. Instead, they are willing to enforce if doing so aligns with or furthers their other political interests, which leads to selective enforcement. Yet the language and the aspiration of international human rights law is of *universal* rights and obligations.[223]

If efforts to enforce human rights through the United Nations have no costs on the other work of the United Nations or on international law as a whole, then perhaps their apparently limited impact on human rights is of little moment. But, as this section has argued, the selective enforcement and other problems with human rights may make the other work of the United Nations less successful. Those potential costs are not part of the contemporary debate about human rights and international law, but they should be.

D. Conclusion

Human rights, and especially their transformation of sovereignty and their enforcement through international law, are in a period of stasis and decline. One potential objection is that any such decline is temporary or cyclical. After all, states remain bound by substantive human rights treaties and the customary international law protecting human rights. In a few years, the international legal enforcement of human rights could be back on an upswing. Maybe. But several factors suggest that the decline is longer-term, including the relative decline in Western economic and political power, broader global trends towards populism, and the selective and half-hearted efforts by Western countries to promote human rights through international law. Human rights themselves may, let us hope, improve over time, and the drive and motivation to secure human rights will certainly continue, but neither means that international human rights law should or must be the preferred vehicle for doing so.

A second potential objection is that there is no decline because there was no heyday: states never took the international legal enforcement of human rights law seriously, even if activists did. As the doctrinal discussions above illustrate, however, state practice in many areas changed around the turn of the century, even if those doctrines were never fully implemented and their on-the-ground effectiveness is contested.

---

223. G.A. Res. 217 (III) A, Universal Declaration of Human Rights (Dec. 10, 1948).

II. Human Rights and the Cost of Expansion

Human rights have not just purportedly transformed sovereignty and the work of the United Nations; they have also changed what international lawyers call the "doctrine of sources" as part of what one scholar has termed the "expansionist project" of international human rights law.[224] The resulting changes to the two primary sources of international law—treaties and custom—have successfully ensured that a wide range of human rights norms are legally protected by international law. But this elasticity in the doctrine of sources has come at the price of widespread non- and under-compliance with international human rights law. There has been a parallel development in the U.N. Security Council. The work of the Security Council has expanded to include international human rights, perhaps leading to diminished credibility and effectiveness.

This Part considers whether greater non- and under-compliance with human rights norms makes international law as a whole less credible and less effective. Other authors have suggested this possibility, but with little supporting analysis.[225] This Part applies the most important theories of why states comply with international law, including rational choice, constructivism, and organizational sociology. They, along with studies of social psychology and domestic law, all suggest that widespread noncompliance with human rights will make the rest of international law less effective and more costly to enforce. In other words, for *whatever* reason(s) international law is able to generate compliance, that compliance is more difficult to secure in the context of greater overall noncompliance. Analogizing to criminal law, we might call this a "broken windows" theory of international law. If it is correct, a decline in international human rights law might ultimately result in benefits to international law as a whole, but only if international law does not continue to shoulder the burden of enforcing human rights as legally binding norms.

A. Expanded Sources

International human rights law has expanded the two traditional sources of international law: treaties and custom. Both sources have been stretched to ensure that a broad array of human rights is governed by international law.

---

224. D'Aspremont, *supra* note 3, at 224; *see also* Cohen, *supra* note 3, at 65 (noting that human rights (along with other forces) have changed the sources of international law).

225. *See, e.g.*, Cohen, *supra* note 3, at 67 ("For some, the patterns of noncompliance are proof that international law is 'law' in name only."); Carlos M. Vázquez, *Direct vs. Indirect Obligations of Corporations Under International Law*, 43 COLUM. J. TRANSNAT'L L. 927, 958 (2005) (contending that an international human rights regime, without an enforcement mechanism, would trivialize international law); *cf.* KROMMENDIJK, *supra* note 216, at 43 (noting that state compliance depends on qualities and legitimacy of international norms).

1. Treaties.—Treaty law has changed in order to accommodate international human rights instruments. Treaty law now permits reservations to multilateral agreements without the consent of all contracting states (unless the treaty provides otherwise), a change that has allowed states, in effect, to undermine or even vitiate their consent to a treaty through far-reaching reservations.[226] Based on an analogy to contracts, international law traditionally required that, in order to be valid, a reservation must receive the consent of all contracting parties to a treaty.[227] This rule changed during the middle of the twentieth century, initially in order to accommodate reservations to the Convention on the Prevention and Punishment of the Crime of Genocide.[228] The International Court of Justice was asked whether states which filed reservations to the Genocide Convention could nonetheless be parties to that Convention.[229] The Court's opinions abandoned the unanimity rule in a 7-to-5 vote, with the majority emphasizing that the Genocide Convention was "adopted for a purely humanitarian and civilizing purpose."[230] In this kind of convention, "the contracting States do not have any interests of their own; they merely have, one and all, a common interest, namely, the accomplishment of those high purposes which are the *raison d'être* of the convention."[231] The object and purpose of the Convention accordingly implied that as many states as possible should participate, so reservations should be permitted as long as they did not "sacrifice the very object of the Convention."[232]

The general approach of the International Court of Justice was eventually adopted by the International Law Commission in the draft articles of the Vienna Convention on the Law of Treaties (VCLT), but only after two Special Rapporteurs sought to preserve the unanimity rule.[233] As finally drafted, the VCLT rules apply not just to human rights agreements but to all treaties.[234] Some treaties prohibit reservations, but many do not, and states

---

226. FRANK HORN, RESERVATIONS AND INTERPRETATIVE DECLARATIONS TO MULTILATERAL TREATIES 156 (1988) ("A reservation resembles a breach in that it also constitutes a derogation from an obligation. The derogation is however a legitimate one.").

227. *Id.* at 22–23.

228. *Id.* at 17, 21.

229. Reservations to the Convention on the Prevention and Punishment of the Crime of Genocide, Advisory Opinion, 1951 I.C.J. Rep. 15 (May 28).

230. *Id.* at 23.

231. *Id.*

232. *Id.* at 24.

233. *See* HORN, *supra* note 226, at 20–21 (documenting the change of the unanimity rule after discussions at the 5th session of the U.N. General Assembly and the ICJ decision, while notifying that Special Rapporteurs Lauterpacht and Fitzmaurice insisted on maintaining the unanimity rule); *see also* SHABTAI ROSENNE, DEVELOPMENTS IN THE LAW OF TREATIES 1945–1986, at 424–36 (1989) (surveying discussions of reservation rules before the International Law Commission and U.N. General Assembly).

234. Vienna Convention on the Law of Treaties, arts. 19–21, May 23, 1969, 1155 U.N.T.S. 331.

have made significant reservations to treaties governing a diverse set of topics from whaling to private international law.[235] Reservations to all of the treaties (human rights or otherwise) can be seen as a direct byproduct of the controversial change to treaty law in order to accommodate international human rights lawmaking.

It is true that the VCLT only sets default rules. States can and do contract out of those rules through specific treaty language, such as language that prohibits reservations entirely. The default language remains significant, however, as the example of the International Convention on the Regulation of Whaling (Whaling Convention) demonstrates. When the Whaling Convention came into force, the old unanimity rule applied, pursuant to which Denmark's proposed reservation was rejected.[236] Denmark abandoned its reservation and became a party.[237] Since the new default rule has become part of customary international law, states have made a variety of reservations to the Whaling Convention, including Iceland, which made a reservation to its basic terms and yet ultimately became a party despite the opposition of many states.[238] It is also true that the ability to use reservations may induce some states to become parties to treaties, encouraging broader participation, as the International Court of Justice reasoned in the language quoted above.[239] Whatever advantages there are to reservations, however, they may also be costly to international law as a whole.

Reservations to human rights treaties are generally more common than reservations to other kinds of treaties.[240] States have also made especially far-reaching reservations to human rights treaties. Some parties to the Convention for Elimination of Discrimination against Women, for example,

---

235. *See, e.g.*, United Nations Convention Against Corruption, Oct. 31, 2003, 2349 U.N.T.S. 41 (India reservation at 2753 U.N.T.S. 412); Convention on International Trade in Endangered Species of Wild Fauna and Flora, Mar. 3, 1973, 993 U.N.T.S. 243 (Ireland reservation at 2213 U.N.T.S. 122); Vienna Convention on Consular Relations, Apr. 24, 1963, 596 U.N.T.S. 261; Convention on the Privileges and Immunities of the United Nations, Feb. 13, 1946, 1 U.N.T.S. 15 (although the Convention includes no explicit language on reservations, reservations have been made, Turkey reservation at 70 U.N.T.S. 266); *see also* Alexander Gillespie, *Iceland's Reservation at the International Whaling Commission*, 14 EUR. J. INT'L L. 977, 977–78 (2003) (chronicling Iceland's three attempts to modify its reservation for readmission to the International Whaling Commission); Ulrich G. Schroeter, *Reservations and the CISG: The Borderland of Uniform International Sales Law and Treaty Law After Thirty-Five Years*, 41 BROOK. J. INT'L L. 203, 206–07 (2015) (discussing the reservations made in the 1980 United Nations Convention on Contracts for the International Sale of Goods).

236. Gillespie, *supra* note 235, at 982.

237. *Id.* at 981–82.

238. *Id.* at 977–78. There was also a dispute about whether the International Whaling Commission had the competence to determine the validity of the reservation. *Id.* at 993–96.

239. *See also* Edward T. Swaine, *Reserving*, 31 YALE J. INT'L L. 307, 331–40 (2006) (defending a default rule favoring reservations because reservations may increase treaty depth and because they yield valuable information).

240. Eric Neumayer, *Qualified Ratification: Explaining Reservations to International Human Rights Treaties*, 36 J. LEGAL STUD. 397, 397–98 (2007).

have made sweeping reservations to its substantive terms. Saudi Arabia included a general reservation to the Convention that provides: "In case of contradiction between any term of the Convention and the norms of islamic [sic] law, the Kingdom is not under obligation to observe the contradictory terms of the Convention."[241] Saudi Arabia has a truly dismal record on the rights of women and did not, to name just one example, permit women to drive cars until very recently.[242]

Wide-ranging and sometimes crippling reservations to human rights treaties have led treaty-monitoring bodies to declare such reservations invalid and also that invalid reservations are severable so that the treaty remains in force, but absent the reservation.[243] Although this practice by treaty-monitoring bodies is controversial,[244] if accepted it would permit human rights treaties to govern conduct which the contracting state intended to exclude from the treaty entirely. That conduct is likely to violate the treaty—otherwise there is little reason to make the reservation in the first place—leading to even more conduct that is inconsistent with the terms of the treaty.

When a state engages in conduct that would be prohibited by a treaty but for that state's reservation, the state is formally compliant with the treaty (so long as the reservation is permissible), but the conduct is "non-conforming"—because it does not correspond to the terms of the treaty. The practice of reservations, designed to encourage widespread ratification of human rights treaties, has thus resulted in widespread nonconformity with and violations of the terms of treaties, especially (but not exclusively) human rights treaties.

2. Custom.—The rules governing customary international law have also become more flexible so as to incorporate human rights into international law. Today, customary international law norms can be generated based on what states say, even if their actual conduct does not conform to the norm.[245]

---

241. *See* Linda M. Keller, *The Convention on the Elimination of Discrimination Against Women: Evolution and (Non)implementation Worldwide*, 27 T. JEFFERSON L. REV. 35, 39 (2004) (quoting Saudi Arabia's reservation made at 2121 U.N.T.S. 342).

242. Ben Hubbard, *Saudi Arabia Agrees to Let Women Drive*, N.Y. TIMES (Sept. 26, 2017), https://www.nytimes.com/2017/09/26/world/middleeast/saudi-arabia-women-drive.html [https://perma/cc.M8AD-J559].

243. Ryan Goodman, *Human Rights Treaties, Invalid Reservations, and State Consent*, 96 AM. J. INT'L L. 531, 531 (2002). These changes in how human rights treaties are interpreted and applied are apparently leading to changes in treaty law as a whole. *See generally* Marko Milanovic & Linos-Alexander Sicilianos, *Reservations to Treaties: An Introduction*, 24 EUR. J. INT'L L. 1055 (2013).

244. Goodman, *supra* note 243, at 531.

245. *See* Military and Paramilitary Activities in and Against Nicaragua (Nicar. v. U.S.), Judgment, 1986 I.C.J. Rep. 14, ¶ 186 (June 27); Jan Wouters & Cedric Ryngaert, *Impact on the Process of the Formation of Customary International Law*, *in* THE IMPACT OF HUMAN RIGHTS LAW ON GENERAL INTERNATIONAL LAW, *supra* note 33, at 111, 111–31; Anthea Elizabeth Roberts, *Traditional and Modern Approaches to Customary International Law: A Reconciliation*, 95 AM. J. INT'L L. 757, 758 (2001).

This change means that customary international law corresponds less to the actual conduct of states, which in turn means there will be more violations at the point in time when the norm crystalizes into customary international law.

For example, customary international law prohibits torture.[246] The primary sources relied upon to demonstrate the customary international prohibition on torture include the following: U.N. General Assembly resolutions condemning torture; the prohibition on torture found in domestic constitutions; states' universal condemnation of torture; and treaties such as the Geneva Conventions, the International Covenant on Civil and Political Rights, and the Convention Against Torture. The prohibition on torture is even a *jus cogens* norm of international law, meaning that it is understood as absolute and nonderogable.[247] Torture itself is "a direct attack on the core of the dignity and integrity of human beings."[248] Yet torture is widely practiced—even if it is also widely denounced as violating international law. A well-placed observer recently concluded "torture is practiced in more than 90 percent of all countries and constitutes a widespread practice in more than 50 percent of all countries."[249] As another example, customary international law is said to include a right to food, or at least a right to be free from hunger.[250] Yet many people around the world are not free from hunger and many governments fail to remedy or are complicit in the scarcity of food in their countries.[251]

Like changes to the law of treaties, change in the rules governing the formation of customary international law is justified on the grounds that it permits the widespread adoption of international human rights law, which is considered a valuable step forward despite the violations it accepts.[252] The elements needed to show the formation of custom are in general contested, and the actual content of customary international law tends to be vague so that the application of customary norms to particular facts is frequently contested.[253] As well, basing custom on state declarations rather than on their actions is an issue that extends beyond human rights. Nevertheless, human rights have unmistakably pushed customary international law towards what

246. RESTATEMENT (THIRD) OF FOREIGN RELATIONS § 702(d) (AM. L. INST. 1987).

247. Manfred Nowak, *What's in a Name? The Prohibitions on Torture and Ill Treatment Today*, *in* THE CAMBRIDGE COMPANION TO HUMAN RIGHTS LAW 307, 307 (Conor Gearty & Costas Douzinas eds., 2012).

248. *Id.*

249. *Id.*

250. *See* Smita Narula, *The Right to Food: Holding Global Actors Accountable Under International Law*, 44 COLUM. J. TRANSNAT'L L. 691, 705 (2006).

251. *See* SIMONE HUTTER, STARVATION AS A WEAPON: DOMESTIC POLICIES OF DELIBERATE STARVATION AS A MEANS TO AN END UNDER INTERNATIONAL LAW 2–3 (2015).

252. *See* Roberts, *supra* note 245, at 764–65.

253. Larry R. Helfer & Ingrid Wuerth, *Customary International Law: An Instrument Choice Perspective*, 37 MICH. J. INT'L L. 563, 576 (2016).

some call a "tremendous implementation gap."[254] As one observer puts it, international human rights got "sidetracked off into an international arena of pious declarations and unenforceable agreements."[255] These problems have generally been considered, if at all, only in terms of their costs for international human rights itself. But various theories of compliance with international law suggest that widespread noncompliance with international human rights law will tend to make the enforcement of other international law more difficult and costly.

B. A "Broken Windows" Effect?

Changes in the doctrine of sources to accommodate human rights have made international law more elastic; it now permits the adoption of norms despite greater noncompliance and more nonconforming behavior. And beyond the changes to the formal sources of international law, there are other widespread violations of international human rights law, such as a failure to abide by even those treaty norms to which no reservation was made. Such failures extend even to ministerial tasks, such as filing required reports to treaty-monitoring bodies.[256] There is a parallel development in the mandate of the U.N. Security Council, which has grown to include many human rights issues that the Council cannot remedy or prevent.

These changes to international law may encourage noncompliance with other international legal norms, not just those governing human rights. The intuition here is an imperfect analogy to the "broken windows" theory of crime prevention: widespread violations of human rights law may be a symbol of unaccountability,[257] a signal that "no one cares" about violations of international law and that "no one is in charge."[258] Accountability is a central concern of public international law. The system lacks a centralized enforcement mechanism, and as a result, compliance and effectiveness pose important—some would say fundamental—challenges to the relevance of

---

254. Nowak, *supra* note 247, at 307.

255. Conor Gearty, *Spoils for Which Victor? Human Rights Within the Democratic State*, *in* THE CAMBRIDGE COMPANION TO HUMAN RIGHTS LAW, *supra* note 247, at 214. As with treaties, the change to the rules of the formation of custom are not limited to the customary international law of human rights, instead they are changes to customary international law generally. *See, e.g.*, Geoffrey R. Watson, *The Humanitarian Law of the Yugoslavia War Crimes Tribunal: Jurisdiction in* Prosecutor v. Tadic, 36 VA. J. INT'L L. 687, 708–09 (describing a customary norm of international humanitarian law). See Arnold N. Pronto, *"Human-Rightism" and the Development of General International Law*, 20 LEIDEN J. INT'L L. 753 (2007), for a discussion of the significant influence of human rights norms and imperatives over the development of general international law.

256. *See supra* text accompanying notes 214–16.

257. George L. Kelling & James Q. Wilson, *Broken Windows: the Police and Neighborhood Safety*, THE ATLANTIC (March 1982).

258. Bernard E. Harcourt, *Reflecting on the Subject: A Critique of the Social Influence Conception of Deterrence, the Broken Windows Theory, and Order-Maintenance Policing New York Style*, 97 MICH. L. REV. 291, 303, 305–06 (1998).

public international law.[259] In this context, behavior that signals a lack of accountability may be especially damaging to the enforcement and deterrence of international law writ large. To some extent, this intuition has already been voiced within the human rights discourse.[260]

The domestic broken windows theory argues that "mere" "disorder" or "victimless" crimes lead to more serious violations of the law.[261] Human rights violations are not victimless and can impose significant costs to human lives and dignity. As well, the broken windows theory of domestic law enforcement may depend upon and itself create certain subjects or categories of people, such as the "honest" versus the "disorderly," upon which social influences operate differently.[262] Although the analogy is thus imperfect—and despite great controversy over the domestic broken windows theory and its relationship to domestic policing[263]—the question remains: if states are in widespread violation of, or noncompliance with, international human rights law, are they (and other states) more likely to violate *other* norms of international law? Models of state compliance with international law—rational choice, constructivism, the sociology of international organizations—answer this question affirmatively,[264] and so do empirical studies of social psychology and domestic law.

---

259. *See, e.g.*, Harlan Grant Cohen, *Can International Law Work? A Constructivist Expansion*, 27 BERKELEY J. INT'L L. 636, 641 (2009) (book review); JEFFREY L. DUNOFF & MARK A. POLLACK, *International Law and International Relations: Introducing an Interdisciplinary Dialogue*, *in* INTERDISCIPLINARY PERSPECTIVES ON INTERNATIONAL LAW AND INTERNATIONAL RELATIONS 3, 5–21 (2013) [hereinafter INTERDISCIPLINARY PERSPECTIVES].

260. NEIER, *supra* note 202, at 259; Philip Alston, *Conjuring Up New Human Rights: A Proposal for Quality Control*, 78 AM. J. INT'L L. 607, 609 (1984); *see also supra* note 231 and accompanying text.

261. *See generally* George L. Kelling & William J. Bratton, *Declining Crime Rates: Insiders' Views of the New York City Story*, 88 J. CRIM. L. & CRIMINOLOGY 1217 (1998) (applying the broken windows theory to police common, victimless, and minor offenses).

262. Harcourt, *supra* note 258, at 297.

263. *See, e.g.*, Sam Roberts, *Author of "Broken Windows" Policing Defends His Theory*, N.Y. TIMES (Aug. 10, 2014), https://www.nytimes.com/2014/08/11/nyregion/author-of-broken-windows-policing-defends-his-theory.html?mcubz=3 [https://perma.cc/4PBN-J63R].

264. Some rational choice scholars argue that international law does little to change state behavior and that what appears as compliance is usually coercion or coincidence of interest. *See* JACK L. GOLDSMITH & ERIC A. POSNER, THE LIMITS OF INTERNATIONAL LAW 108–18 (2005). That work is not directly relevant to this section, which argues that the reasons to think that international law *does* generate compliance also suggest that widespread non- or underenforcement of some international law will make compliance with other international law more difficult to secure. Some liberal international relations scholarship is also not directly relevant. Generally, liberal theorists do not take the state as a unitary actor, and they view compliance as at least partially a function of domestic processes. *See* Andrew Moravscik, *Liberal Theories of International Law*, *in* INTERDISCIPLINARY PERSPECTIVES, *supra* note 259, at 83, 84. Beth Simmons, for example, argues that international human rights treaties can be effective because the commitment to the treaty empowers domestic political groups or parts of the state. BETH SIMMONS, MOBILIZING FOR HUMAN RIGHTS: INTERNATIONAL LAW IN DOMESTIC POLITICS 125–29 (2009). Her work is limited, however, to compliance with international human rights treaties and does not discuss compliance generally. Similarly, Harold Koh describes a transnational process that leads states to incorporate

1. Rational Choice.—Rational choice theories of state behavior support the broken windows analogy. They assume that states are rational and self-interested.[265] They take state preferences as exogenous and fixed, and they assume states have "no innate preference for complying with international law."[266] Scholars have focused on several mechanisms that states use to enforce their international commitments, including reputation, retaliation, and reciprocity.[267] Beginning with reputation, widespread violations of international human rights norms by states mean that, as a whole, states will have a poorer reputation for compliance and that the (many) noncomplying states themselves will have a poorer reputation for compliance. These two drops in reputation—one in states' overall reputation for noncompliance and one in the reputation of particular noncomplying states—should produce two effects.

First, rational choice scholars argue that, on the margin, states with little reputation for compliance may decide it is "too costly to build a good reputation."[268] A diminished reputation for compliance by many states as a result of widespread violations of international human rights norms means that more states have an overall reputation of noncompliance. As a result, overall noncompliance with international law will increase because more states will simply give up on compliance.

Second, and more importantly, if states as a whole tend to expect noncompliance from each other, the costs of entering into treaties or developing norms of customary international law become higher for all states. A baseline reputation of noncompliance among states generally means that states will have to do more in a treaty agreement to generate trustworthy commitments (such as monitoring noncompliance), making some agreements not worth the time or effort.[269] Similarly, if entities tasked with

---

international norms to become internalized through domestic law and domestic actors. Harold Hongju Koh, *Why Do Nations Obey International Law?*, 106 YALE L.J. 2599, 2651–58 (1997). Koh uses human rights as his main example but leaves open the question of how the relevant issue networks and epistemic communities form and whether they are influenced by developments in other areas of international law. In a more recent article, Koh suggests that his theory of compliance is drawn from the work of Tom Tyler. Koh, *supra* note 128, at 973 n.2. Tyler's work is discussed *infra* at text accompanying notes 307–12.

265. ANDREW T. GUZMAN, HOW INTERNATIONAL LAW WORKS: A RATIONAL CHOICE THEORY 17 (2008).

266. *Id.*

267. *Id.* at 33–34; ROBERT E. SCOTT & PAUL B. STEPHAN, THE LIMITS OF LEVIATHAN: CONTRACT THEORY AND THE ENFORCEMENT OF INTERNATIONAL LAW 10, 115–27 (2006); Rachel Brewster, *Reputation in International Relations and International Law Theory*, *in* INTERDISCIPLINARY PERSPECTIVES, *supra* note 259, at 524, 533.

268. Andrew T. Guzman, *Reputation and International Law*, 34 GA. J. INT'L & COMP. L. 379, 382 (2006).

269. *See* ANDREW H. KYDD, TRUST AND MISTRUST IN INTERNATIONAL RELATIONS 26, 119 (2005); Andrew T. Guzman, *A Compliance-Based Theory of International Law*, 90 CALIF. L. REV. 1823, 1849 (2002).

formal enforcement—such as the treaty monitoring bodies discussed in the next section—fail to ensure compliance, states may be deterred from making additional international commitments,[270] while the possibility of formal enforcement may at the same time make informal enforcement by the parties less likely.[271]

Rational choice scholars also argue that retaliation and reciprocity generate compliance with international law.[272] These mechanisms, too, are undermined by widespread violations of international human rights law. First, as mentioned above, a widespread belief that states do not comply with international obligations makes it more difficult to generate trustworthy commitments, even if those commitments might be enforced through retaliation or reciprocity rather than directly through reputation. Second, states benefit from having a reputation for using reciprocity or retaliatory sanctions, which can be costly to impose. A state contemplating a violation of its international legal obligations might be deterred from doing so if the state (or states) that would be aggrieved by the breach has (or have) a general reputation for imposing sanctions.[273] The widespread under- and noncompliance with international human rights law can lead states to believe there is general unwillingness to impose retaliatory sanctions for violations of international law.

Building on rational choice models, behavioral law and economics scholars hypothesize that states' willingness to engage in retaliation is partly a function of their perception of fairness and bias.[274] If correct, this suggests another problem with widespread violations of international human rights law: the view that international human rights law is selectively enforced. If only some states are "punished" for human rights violations, the result may be perceptions of bias and unfairness in the international legal system as a whole. Those perceptions may, in turn, make states generally less willing to impose sanctions on other states, in particular those which they perceive as receiving unfair treatment in the human rights context. One possible example of this dynamic is some states' reluctance to condemn Russia's annexation of Crimea, which observers attribute in part to non-Western states' "conviction that the West enjoys an unjustified position of privilege in the international system."[275]

---

270. SCOTT & STEPHAN, *supra* note 267, at 178–79, 181.

271. *Id.* at 26–27, 84–109.

272. GUZMAN, *supra* note 265, at 33–34.

273. *Id.* at 47–48; SCOTT & STEPHAN, *supra* note 267, at 115–17.

274. Anne van Aaken, *Behavioral International Law and Economics*, 55 HARV. INT'L L. J. 421, 434–35 (2014).

275. *See* Buras et al., *supra* note 203.

One objection to the arguments advanced in this section is that states' reputations may be compartmentalized and issue-specific.[276] If so, widespread violations of international human rights agreements will not create reputational losses to individual states or to states as a whole in other issue areas such as trade or security. Similarly, making a reservation might signal to other states that the reserving state is especially compliant, because rather than simply violating the treaty, the state instead made the reservation and became under- rather than noncompliant.[277] But significant reservations may instead lead to less regard for international legal rules as a whole by signaling that states get to pick and choose their commitments or that their consent to the treaty is not genuine,[278] thereby decreasing general reputations for compliance. In any event, in the context of retaliatory sanctions for human rights violations, there is no reason to assume that states imposing such sanctions do not see their general reputational capital for imposing sanctions grow as a result.[279] In the end, these are empirical questions to which there is no definitive answer. But if reputation does not cross into other issue areas, then it is not "an important cause of compliance with a wide range of agreements."[280] Most rational choice and human rights scholars thus assume or argue that reputation is not entirely issue-specific.[281]

2. Constructivism.—Constructivist theories of state behavior also support the broken window analogy and suggest that widespread violations of human rights norms will result in reduced effectiveness for the rest of international law. Constructivist accounts of international relations focus on the social construction of identity.[282] Under this view, international norms, including international legal norms, have a constitutive function in that they

276. *See* George W. Downs & Michael Jones, *Reputation, Compliance, and International Law*, 31 J. LEGAL STUD. 95, 109–10 (2002); van Aaken, *supra* note 274, at 476–77.

277. Swaine, *supra* note 239, at 338.

278. Ineke Boerfijn, *Impact on the Law on Treaty Reservations*, *in* THE IMPACT OF HUMAN RIGHTS LAW ON GENERAL INTERNATIONAL LAW, *supra* note 33, at 63, 65; *see* Laurence R. Helfer, *Exiting Treaties*, 91 VA. L. REV. 1579, 1641 (2005) (noting that reservations may signal that a state is uncooperative, even if compliant); *see also* Rachel Brewster, *Unpacking the State's Reputation*, 50 HARV. INT'L L.J. 231, 263–64 (2009) (describing arguments about the impact of treaty reservations on state reputation).

279. *See* GUZMAN, *supra* note 265, at 46–47 (citing President Clinton's statement that the United States has "an interest in standing up against the principle of ethnic cleansing" as an example of an effort to bolster the reputational capital of the United States).

280. Brewster, *supra* note 278, at 261.

281. *E.g.*, *id.*; GUZMAN, *supra* note 265, at 100–11 ("[I]t is likely that states have different reputations in different issue areas, but these reputations are related to one another."); SCOTT & STEPHAN, *supra* note 267, at 118 ("Reputations survive particular treaties and similar arrangements.").

282. Benedict Kingsbury, *The Concept of Compliance as a Function of Competing Conceptions of International Law*, 19 MICH. J. INT'L L. 345, 358 (1998).

help create the desires and preferences of states.[283] Thus, ideas and beliefs (not just material interests), which are formed through interaction and communication, help determine how states behave, including their compliance with international law.[284]

Drawing on the work of legal philosopher Lon Fuller, influential constructivist scholars Jutta Brunnée and Stephen Toope argue that international legal obligations arise from communities of practice which have shared understandings and which generate norms with specific characteristics of legality.[285] These characteristics of legality include generality, consistency, and alignment between legal norms and the actions of officials.[286] A community of practice, grounded in shared understanding, that generates and maintains such norms results in a "practice of legality."[287] It is not the formal characteristics of instruments and norms that give rise to legal obligation, but instead this practice of legality.[288]

A communities-of-practice analysis suggests that states' obvious non- and under-compliance with legal norms inhibits the development of a meaningful community of practice with shared values. Indeed, Brunnée and Toope question whether the prohibition on torture qualifies as a norm of international law at all because of the misalignment between the norm and actual official conduct.[289] They reason that "a widespread failure to uphold the law as formally enunciated leads to a sense of hypocrisy which undermines fidelity to law."[290] Their account is telling. In a horizontal, decentralized legal system, whether it relies on reputation or on shared normative values, widespread "violations" may be especially costly, for they suggest that states do not care about their reputations for compliance or that the shared normative values are not so shared after all.

3. Organizational Sociology.—Theories of state behavior based on organizational sociology support the broken windows thesis by suggesting that systematic violations of international human rights law will result in lower compliance with other international law. This work, too, could be characterized as constructivist, but it focuses on a different process of social

283. *See id.* (giving a constructivist reading of international law as an account of "legal relations," "patterns of interactive behavior," and "particularizing society's universal purposes").

284. *See* Alexander Wendt, *Collective Identity Formation and the International State*, 88 AM. POL. SCI. REV. 384, 385 (1994).

285. JUTTA BRUNNÉE & STEPHEN J. TOOPE, LEGITIMACY AND LEGALITY IN INTERNATIONAL LAW: AN INTERACTIONAL ACCOUNT 53–54 (2010). Brunnée & Toope compare their work to that of other constructivists. *Id.* at 52–65.

286. *Id.* at 6.

287. *Id.* at 46–52, 65–77, 350–52.

288. *Id.* at 46–52, 100.

289. *Id.* at 250–68.

290. *Id.* at 232.

interaction to form state preferences: acculturation.[291] A state's or individual's identification with a group can lead to cognitive and social pressures to conform to the group's behavior. Here, the mechanisms of influence on state behavior are social expectations and cultural identity, not necessarily acceptance of the legitimacy of the norm. Organizational sociology predicts acculturation will shape the formal structure of institutions, which are embedded in wider institutional environments, but that convergence across institutions (called isomorphism) may not actually result in functional changes and results. States appear to behave in just this way in the global community: there is remarkable isomorphism in structural organization and policy commitments at the global level, but this convergence is decoupled from any significant changes within states themselves.[292] Or, put more succinctly, the sociology of organizations predicts "cross-national isomorphism irrespective of local circumstance," which is precisely what the data shows.[293]

The acculturation model of state behavior suggests widespread noncompliance with human rights norms will lead to noncompliance with other international legal obligations. Pervasive noncompliance means the "global scripts as members of world society"[294] does not include actual compliance with international norms but instead mere commitment to them. The name organizational sociologists give this dynamic is "decoupling."[295] Decoupling appears itself to be a global script for behavior. That is, both the prosocial pressures applied by groups and the psychological benefits of conforming to group norms operate on the global level to push states toward making international commitments, but not to successfully implement those commitments.[296] This particular global script may have special purchase in international human rights law, but it also operates in other contexts: environmental policy, education curricula, and militarization.[297] Ryan Goodman and Derek Jinks present decoupling in positive terms because it facilitates the global adoption of human rights norms.[298] Other scholars have challenged whether acculturation is normatively defensible in large part

---

291. RYAN GOODMAN & DEREK JINKS, SOCIALIZING STATES: PROMOTING HUMAN RIGHTS THROUGH INTERNATIONAL LAW 41 (2013) [hereinafter GOODMAN & JINKS, SOCIALIZING STATES]; Ryan Goodman & Derek Jinks, *How to Influence States: Socialization and International Human Rights Law*, 54 DUKE L.J. 621, 626 (2004) [hereinafter Goodman & Jinks, *How to Influence States*].

292. Goodman & Jinks, *How to Influence States*, *supra* note 291, at 646–49.

293. *See id.* at 647–55 (surveying empirical studies).

294. *Id.*

295. *Id.* at 649.

296. *Id.* at 643.

297. *Id.* at 648; *see also* Ryan Goodman & Derek Jinks, *Incomplete Internationalization and Compliance with Human Rights Law: A Rejoinder to Roda Mushkat*, 20 EUR. J. INT'L L. 443, 444 (2009).

298. *See* Goodman & Jinks, *How to Influence States*, *supra* note 291, at 670.

because decoupling simply accepts widespread violations of international law.[299]

Two specific aspects of the acculturation model suggest that widespread non- and under-compliance with human rights law will lead to noncompliance with international law as a whole. First, Goodman and Jinks argue that there is a *world* polity with *global* scripts of behavior and *global* cultural models pursuant to which states are "defined" and "legitimated."[300] The claim of worldwide scripts and worldwide cultural norms undercuts the possibility that widespread noncompliance with and violation of international human rights norms have an isolated impact only on human rights norms and only within a specified human rights community. Second, the evidence of decoupling suggests there are already powerful, global behavioral scripts pursuant to which states ignore problems with on-the-ground enforcement of and compliance with international norms.[301] Expansive human rights norms that accept widespread noncompliance may have initially generated what has now become a global script. Even if not, they may have made a powerful contribution to an existing, global nonchalance about violations of international law.

4. Social Psychology.—Research from social psychology also suggests the broken windows analogy is correct. Social psychologists writing on domestic legal systems have developed a causal theory of legitimacy and measured its impact on actual compliance rates.[302] In this work, "value-based" legitimacy is defined as an individual's sense of obligation and willingness to obey, as measured by questions about whether individuals feel

---

299. *See, e.g.*, José E. Alvarez, *Do States Socialize?*, 54 DUKE L.J. 961, 971 (2005); Harold Hongju Koh, *Internalization Through Socialization*, 54 DUKE L.J. 975, 978–81 (2005). *But cf.* GOODMAN & JINKS, SOCIALIZING STATES, *supra* note 291, at 135–65 (defending the normative value of acculturation).

300. Ryan Goodman & Derek Jinks, *Toward an Institutional Theory of Sovereignty*, 55 STAN. L. REV. 1749, 1756–57 (2003).

301. Goodman and Jinks argue that decoupling is evidence of acculturation but that acculturation can occur without decoupling, that decoupling is not necessarily inconsistent with deep reform, and that even decoupling that initially hinders compliance may eventually lead to the progression of more meaningful change over time. GOODMAN & JINKS, SOCIALIZING STATES, *supra* note 291, at 135–65. Their account nevertheless presents evidence of widespread decoupling and describes a model of state behavior pursuant to which states are socialized to conform their behavior at the international level without actually complying with the legal norm through their behavior domestically.

302. TOM R. TYLER, WHY PEOPLE OBEY THE LAW 3–5, 19–39 (Princeton Univ. Press 2006); Margaret Levi et al., *The Reasons for Compliance with Law*, *in* UNDERSTANDING SOCIAL ACTION, PROMOTING HUMAN RIGHTS 70, 70–71 (Ryan Goodman et al. eds., 2012). The international legal theorist Thomas Franck also argued that the "legitimacy" of international law generates compliance. THOMAS M. FRANCK, THE POWER OF LEGITIMACY AMONG NATIONS 3, 25–26 (1990). Franck did not, however, offer any well-developed causal mechanism to explain why legitimacy causes compliance.

obligated to obey the law.[303] This kind of legitimacy is linked to actual compliance with the law ("behavioral legitimacy"), as measured by precinct-level police data and self-reporting from respondents.[304] The factors that tend to generate "value-based" legitimacy include "trustworthiness in government," which depends on the government's general ability to solve problems and to enforce laws.[305] The perception of widespread noncompliance with the law is correlated with a lower trustworthiness in government,[306] and accordingly with lower legitimacy and compliance. The general result is supported by data from both the United States and Africa.[307]

States are obviously not individuals, and the determinants of individual behavior are not necessarily those of state behavior.[308] Decisions about state behavior are, however, made by individuals or groups of individuals. Moreover, empirical evidence suggests states' decision-making about treaty commitments is susceptible to some of the same cognitive biases that impair individual decision-making, including salience effects, status-quo bias, and peer effects.[309] Along the same lines, the acculturation research described above argues that "[s]tate socialization" is empirically measurable and that it is a process explained by the "beliefs, conduct, and social relations of individuals."[310] Research on social psychology and domestic law shows that the propensity of individuals to comply with law is an indirect function of their belief that the government is good at solving problems and good at enforcing laws. Noncompliance with law undermines both beliefs. If we can extrapolate from individual to state behavior, this research from social psychology suggests the perception of widespread noncompliance with international law across issue areas will tend to undermine people's sense of obligation and willingness to obey.

C. Expanded Mandate: Credibility and the United Nations

International human rights law has also expanded the work of the U.N. Security Council. The Security Council is charged with maintaining international peace and security, and it has the power to impose coercive

303. Levi et al., *supra* note 302, at 70–72, 74–75, 82.

304. *Id.* at 82–83, 89.

305. *Id.* at 73, 76–77.

306. *Id.* at 73.

307. *Id.* at 89–90.

308. *See* van Aaken, *supra* note 274, at 441–49 (discussing the relationship between individual and state behavior and its implications for both rational choice and behavioral law and economics analysis).

309. *See* Jean Galbraith, *Treaty Options: Towards a Behavioral Understanding of Treaty Design*, 53 VA. J. INT'L L. 309, 349–55 (2013); *see also* van Aaken, *supra* note 274, at 449–80 (describing insights from behavioral law and economics research about individuals that help to explain the behavior of states).

310. *See* GOODMAN & JINKS, SOCIALIZING STATES, *supra* note 291, at 41.

measures including the use of force.[311] Its functions were originally viewed as separate from those of the General Assembly, and it was the General Assembly that had the responsibility to promote voluntary respect for human rights through its supervision of the Economic and Social Council.[312] Until the end of the Cold War, the Security Council played "a negligible role in the protection of human rights," but since then it has shown a "greater readiness to invoke human rights provisions"[313]—at least sometimes.

Even before the end of the Cold War, the U.N. Security Council condemned the "illegal" and "racist minority" in Southern Rhodesia and eventually imposed mandatory sanctions based on human rights violations by Southern Rhodesia.[314] But the Security Council expanded its powers in the decades since the end of the Cold War by adopting a broad definition of what constitutes a threat to international peace and security.[315] Humanitarian crises and refugee flows, repression of civilian populations, serious human rights violations, impunity for violation of international humanitarian law, and the overthrow of a democratically elected regime have all warranted action under Chapter VII, at least in the view of the Security Council itself.[316] In some of these situations, the cross-border effects of human rights violations were clear, so that the link to international (rather than internal) peace and security was straightforward.[317] In other situations, the link between the two was more tenuous.[318] The Security Council's increased activity in the field of human rights has gone so far that some have said "[b]y

311. U.N. Charter arts. 39, 41–42.

312. *See supra* text accompanying note 182; *see also* Vera Gowland-Debbas, *The Security Council and Human Rights—from Discretion to Promote to Obligation to Protect in* SECURING HUMAN RIGHTS?: ACHIEVEMENTS AND CHALLENGES OF THE UN SECURITY COUNCIL, *supra* note 180, at 36–37.

313. Bertrand G. Ramcharan, *Norms and Machinery*, *in* THE OXFORD HANDBOOK ON THE UNITED NATIONS 439, 454 (Thomas G. Weiss & Sam Davis eds., 2007).

314. S.C. Res. 455, ¶ 4 (Nov. 23, 1979); S.C. Res. 246 (Mar. 14, 1968); S.C. Res. 245, ¶ 2 (Jan. 25, 1968).

315. Gowland-Debbas, *supra* note 312, at 38–39; Fen O. Hampson & Christopher K. Penny, *Human Security*, *in* THE OXFORD HANDBOOK ON THE UNITED NATIONS, *supra* note 313, at 539, 554; Shraga, *supra* note 180, at 11–13.

316. Koskenniemi, *supra* note 179, at 341–42; Shraga, *supra* note 180, at 12–13.

317. The harms to international peace and security are clear when the humanitarian crises involve serious flows of refugees. *See, e.g.*, S.C. Res. 794 (Dec. 3, 1992) (addressing armed conflict in Somalia); S.C. Res. 827 (May 25, 1993) (establishing the International Criminal Tribunal for the Former Yugoslavia); S.C. Res. 1244 (June 10, 1999) (addressing the grave humanitarian situation in Kosovo).

318. *See* Mehrdad Payandeh, *The United Nations, Military Intervention, and Regime Change in Libya*, 52 VA. J. INT'L L. 355, 365–66, 366 n.55 (2012) (listing cases where the Security Council found a threat to international peace and security relying on the significant "magnitude of the human tragedy" in domestic conflicts). The decisions of the U.N. Security Council as to the scope of its own powers are controlling, at least as a practical matter.

the end of the 20th century" the Security Council was "at the centre of the human rights protection system."[319]

Broadening the Council's tasks to include the protection of human rights expanded the room for failure in performing those tasks, potentially resulting in costs to the organization's credibility. Failure to act or to follow through on a broad set of promises, implicit or explicit, makes the Security Council look ineffective and weak. The Council will thus, as one scholar put it, "have to face up to the consequences of its inability to make reality of its inflated promises."[320] A broader mandate also has the potential to further increase the perception of bias and selectivity,[321] which may lead to polarization and diminished effectiveness of the United Nations as a whole.[322] The change in U.N. Security Council-authorized peacekeeping serves as an example. As peacekeeping mandates expanded over the past two decades to focus on human rights and the protection of civilians, its work has been viewed as more political, less impartial, and in many respects, less successful.[323]

Credibility and polarization costs have also resulted from the U.N. Security Council's response to Libya and its subsequent failure to act in Syria. Russia and China abstained from (but did not veto) Security Council Resolution 1973, which authorized the use of force in Libya to protect civilians from the imminent threat of massive human rights violations.[324] Whether that Resolution is best read to have authorized the use of air strikes to facilitate the rebel ousting of Qaddafi (in addition to merely protecting civilians) is disputed.[325] But it is clear that Russia and China became skeptical

---

319. Shraga, *supra* note 180, at 13.

320. Koskenniemi, *supra* note 179, at 346; *see also* SUBEDI, *supra* note 14 at 27–28; JEAN-MARIE GUÉHENNO, THE FOG OF PEACE: A MEMOIR OF INTERNATIONAL PEACEKEEPING IN THE 21ST CENTURY 305 (2015). To be sure, there are other aspects of the United Nation's human rights work which also undermine the credibility and effectiveness of the organization as whole, including human rights violations resulting from the actions of the Security Council, the oil-for-food scandal, sexual abuse and gross misconduct by U.N. peacekeepers, and the tragic cholera epidemic in Haiti. *See, e.g.*, Jonathan M. Katz, *In the Time of Cholera*, FOREIGN POL'Y (Jan. 10, 2013), http://foreignpolicy.com/2013/01/10/in-t\he-time-of-cholera [https://perma.cc/3MMQ-TRPK].

321. *See* Payandeh, *supra* note 318, at 398 ("As the Security Council becomes more active, its selective course of action becomes ever more apparent. And with the increasing awareness of this selectivity the international community's acceptance of the Security Council practice might decline significantly."); *see also* Joanna Weschler, *Acting on Human Rights* (arguing that the Security Council is a "political body" and is likely to be guided by "national interests"), *in* THE UN SECURITY COUNCIL IN THE 21ST CENTURY, *supra* note 139, at 259, 273.

322. *See supra* subpart I(C).

323. *See* RHOADS, *supra* note 17, at 64–80, 172–91.

324. Press Release, Security Council, Security Council Approves "No-Fly Zone" over Libya, Authorizing "All Necessary Measures" to Protect Civilians, by Vote of 10 in Favour with 5 Abstentions, U.N. Press Release SC/10200 (Mar. 17, 2011) [hereinafter Security Council Press Release 2011].

325. Payandeh, *supra* note 318, at 383–86.

of the Resolution and its use to assist in the ousting of Qaddafi.[326] The regime change aspect of the intervention appeared to many countries as the use of human rights and humanitarian issues as a smokescreen for the removal of Qaddafi, a result explicitly desired by the West.[327] As one writer puts it, the use of force in Libya has "fueled speculations as to which other countries are also likely candidates for intervention" by Western countries.[328] Cooperation between Russia and Western countries on other issues became more difficult. These are polarization costs.

The United Nations' actions with respect to Libya also led to credibility costs. As the intervention in Libya proceeded, the Syrian government used increasingly violent measures to quell domestic unrest.[329] The conflict spiraled into a civil war, killing hundreds of thousands of people.[330] The Syrian rebels have hoped for years for a U.N. Security Council resolution authorizing the use of force to assist them, yet the Council has taken no meaningful action.[331] The broader mandate of the Security Council over mass atrocities and other human rights violations leads to a loss of credibility when the Council is hamstrung by political differences and accordingly cannot act in response to massive atrocities in violation of human rights law.[332]

---

326. *See, e.g.*, *Foreign Minister Yang Jiechi Meets with the Libya Government Envoy* (外交部长杨洁篪8日会见来访的利比亚政府特使), TODAY'S CHINA (June 8, 2011), http://www.gov.cn/jrzg/2011-06/08/content_1879973.htm [https://perma.cc/5V39-VTC9] (declaring China's desire for a prompt cease-fire and respect for the sovereignty of Libya and the right to self-determination of the Libyan people) (translation available upon request); Nabi Abdullaev, *Lavrov Eases Libya Stance, but NATO Rift Stays*, MOSCOW TIMES (Apr. 28, 2011), http://oldtmt.vedomosti.ru/news/article/tmt/435194.html [http://perma.cc/VYU5-WLMU] (stating Russia's view that the Resolution did not sanction regime change in Libya).

327. *See, e.g.*, Security Council Press Release 2011, *supra* note 324 (reporting that the German Ambassador intended the Security Council to send a message to Colonel Qaddafi that "their time is over [and] they must relinquish power immediately"); *see generally* U.N. SCOR, 66th Sess., 6531st mtg. at 11, 17–18, 20, 34, U.N. Doc. S/PV.6531 (May 10, 2011) (noting that representatives of Brazil, South Africa, China, and Nicaragua voiced concern that the protection of civilians could be used as a smokescreen for intervention or regime change).

328. *See* Payandeh, *supra* note 318, at 397.

329. HUMAN RIGHTS WATCH: WORLD REPORT 2012, at 624–31 (2012).

330. *Id.* at 624.

331. *See* Press Release, Security Council, Security Council Fails to Adopt Resolution Condemning Chemical Weapons Use in Syria, Following Veto by Russian Federation, U.N. Press Release SC/12791 (Apr. 12, 2017); Press Release, Security Council, Security Council Fails to Adopt Two Draft Resolutions on Syria Despite Appeals for Action Preventing Impending Humanitarian Catastrophe in Aleppo, U.N. Press Release SC/12545 (Oct. 8, 2016); *see also supra* note 147.

332. *See* SUBEDI, *supra* note 14, at 22–28; Einsiedel et al., *Conclusion: The Security Council and a World in Crisis*, *in* THE UN SECURITY COUNCIL IN THE 21ST CENTURY, *supra* note 139, at 827, 870; Roland Paris, *Is It Possible to Meet the "Responsibility to Protect"?*, WASH. POST (Dec. 9, 2014), https://www.washingtonpost.com/news/monkey-cage/wp/2014/12/09/is-it-possible-to-meet-the-responsibility-to-protect/?utm_term=.ac089fd96769 [https://perma.cc/UZ2L-N6BM]; *cf.* Ganesh Sitaraman, *Credibility and War Powers*, 27 HARV. L. REV. F. 123, 123 (2014) (arguing in the context of war powers that past inaction does not necessarily result in a lack of credibility).

D. *Conclusion*

As international human rights law expanded the doctrine of sources and the work of the United Nations to include an ever-growing set of legal prohibitions and demands, there has been little discussion of what impact this growth may have on the rest of international law. Yet the most important and relevant theories of compliance and effectiveness in international law, from rational choice to constructivist, posit that widespread noncompliance with human rights norms will make cooperation and compliance more difficult and costly in other areas.[333]

III. Reframing the Debate: A Post-Human Rights Agenda for International Law

International human rights law, as described above, imposes costs on international law as a whole. This Part considers a range of possible responses. It suggests that international law and the United Nations should not double down on the decaying human rights enforcement architecture, but should instead turn to developing a strong core of sovereignty-protecting international legal norms devoted to protecting international peace and cooperation. Doing so may promote international cooperation and protect the territorial peace. It would also reduce broken window costs by relaxing the claim that human rights must be enforced as binding international law and by focusing on a smaller set of more vigorously enforced norms. The benefits of such an approach depend in part, however, on the unresolved debate about whether and how international law effectively advances human rights. The arguments proffered here are accordingly tentative and are not intended to be dispositive.

A. *Protecting the "Territorial Peace"*

International law might, for a variety of reasons described in Parts I and II, be best used not to secure a broad set of poorly enforced human rights but instead to promote territorial stability. Doing so might in turn help secure the territorial peace. There are several important potential counterarguments, including that the literature on the "territorial peace" only discusses *interstate* conflicts, not civil wars and other internal or cross-border conflicts, and so the relationship between international law and peace is not fully understood. That is true, and more empirical work on the causes of noninterstate armed conflict would be helpful, but in any event interstate conflict remains an important potential threat which international law may help diminish.

---

333. This Part does not address the potential benefits of noncompliance. *See* Jacob Katz Cogan, *Noncompliance and the International Rule of Law*, 31 YALE J. INT'L L. 189, 193 (2006) (defending noncompliance based on its role in developing new norms and enforcing current law). *See generally* Timothy Meyer, *Shifting Sands: Power, Uncertainty and the Form of International Law*, 27 EUR. J. INT'L L. 161 (2016) (explaining that states use noncompliance as a renegotiation strategy).

A second counterargument is that the "democratic peace" suggests that in order to promote peace international law should promote human rights.[334] If international human rights law makes countries more "democratic" as that term is defined in the empirical literature, it may promote interstate peace. One problem with this argument is that democratic countries are apparently more peaceful only in their relations with other democracies, but not with nondemocratic countries.[335] Still, if all countries were democracies, the data suggests the world would be more peaceful. To this extent, promoting democracy also promotes peace. A more intractable difficulty, however, is the weak relationship between human rights and "democracy" as defined in the democratic peace literature.

A "democracy" as defined by democratic peace literature is not the same as a human rights-respecting regime. The empirical measures of democracy used in the democratic peace literature are provided by the "Polity IV" database.[336] The Polity IV dataset users' manual explains that "democracy" has three elements: the ability of citizens to express effective policy preferences, institutional constraints on executive power, and civil liberties.[337] But the Polity IV database only includes information on the first two elements—not on civil liberties, which would have the clearest relationship to human rights. A human right to democracy, as discussed above in section I(A)(3), would have a close relationship to the "democracy" of the "democratic peace." Intervention (military or otherwise) to ensure a democratic form of government would thus appear to have long-term peace benefits—if the intervenors succeed in their mission. But military interventions to promote democracy have a dismal track record, from the ongoing failures in Libya and Iraq to those in Haiti.[338]

There is also evidence that for pairs of states that both respect human rights, military conflict with each other is less likely, even controlling for democracy. The "human rights" peace has not been tested as extensively as either the territorial or the democratic peace. One study shows, however, that pairs of countries that protect physical integrity rights (torture, political imprisonment, extrajudicial killing, and disappearance) and empowerment rights (freedom of speech, assembly and association, worker rights, freedom

334. *See supra* note 151 and accompanying text.

335. SOBEK, *supra* note 154, at 51, 84–85; *see also* POSNER, *supra* note 14, at 126.

336. *See, e.g.*, Johann Park, *Forward to the Future? The Democratic Peace After the Cold War*, 30 CONFLICT MGMT. & PEACE SCI. 178, 182 (2013) (describing studies that use Polity IV); Douglas M. Gibler, *Bordering on Peace: Democracy, Territorial Issues, and Conflict*, 51 INT'L STUD. Q. 509, 521 (2007) (using Polity IV); *see also* David Sobek et al., *The Human Rights Peace: How the Respect for Human Rights at Home Leads to Peace Abroad*, 68 J. POL. 519, 519 (2006) (noting the difference between human rights and democracy scores on the Polity scale).

337. MONTY G. MARSHALL ET AL., POLITY IV PROJECT (POLITICAL REGIME CHARACTERISTICS AND TRANSITIONS, 1800-2015): DATASET USERS' MANUAL 14 (2016).

338. *See* Philippe R. Girard, *Peacekeeping, Politics, and the 1994 US Intervention in Haiti*, 24 J. CONFLICT STUD., no. 1, 2004, at 20, 31–34 .

of religion, and political participation) are less likely to engage in military conflict with each other, whether or not they are democratic.[339] This study has some anomalies—it does not appear to show democratic peace, for example.[340] Another study has shown that there is also an "abusers' peace"—states with the worst human rights records are "relatively more peaceful with similarly abusive states."[341]

There is, however, an additional problem with the claim that international human rights law has helped create either the democratic or the human rights peace. Even if *human rights* are correlated with peace between some pairs of states, the extent to which *international* human rights *law* generated or sustained those human rights would be difficult to measure in most situations. The protection of human rights is provided for in an overlapping set of regional human rights systems as well as domestic statutory and constitutional law, in addition to binding and nonbinding international legal instruments. The same is not true of international law, which limits conflicts over territory. The basis of that system—indeed a key basis for post-World War II international law—is the prohibition on the use of force against the territorial integrity of another state, which is not an issue meaningfully regulated by domestic law or by regional court systems.

B. Refocusing the Work of the United Nations

A second response would be for the principal organs of the United Nations to narrow and focus their work, so as to reduce credibility and polarization costs, thereby improving the organization's overall effectiveness and reputation. Since the end of the Cold War human rights have become an important aspect of the United Nations' overall mission. Even the U.N. bodies explicitly tasked only with human rights are, however, generally unable to make clear progress.[342] Moreover, human rights appear to be divisive, leading to polarization and lack of progress on other substantive issues.[343] And the enlarged mandate to include human rights has also generated credibility costs for the United Nations as whole, and the Security Council in particular.[344]

To reduce or eliminate these problems, the United Nations might forgo efforts to develop or enforce international human rights law. To be sure, it may be difficult to distinguish the promotion of human rights generally from

339. Sobek et al., *supra* note 336, at 519–25; *cf.* Mary Caprioli & Peter F. Trumbore, *Human Rights Rogues in Interstate Disputes, 1980–2001*, 43 J. PEACE RES. 131 (2006).

340. Sobek et al., *supra* note 336, at 525. The authors hypothesize that the democratic peace was really just a "Cold War peace." A recent study suggests otherwise. *See* Park, *supra* note 336.

341. Timothy M. Peterson & Leah Graham, *Shared Human Rights Norms and Military Conflict*, 55 J. CONFLICT RES. 248, 249 (2011).

342. *See supra* subpart I(C).

343. *See supra* subpart I(C).

344. *See supra* subpart II(C).

the application of international human rights law. To solve the broken windows problem, the focus should be on *legal* norms which go unenforced. The distinction is less significant, however, when it comes to credibility costs, polarization, and lack of cooperation which could result from the promotion of human rights generally, not just from its enforcement through law. Separating out legal and nonlegal enforcement mechanisms is also difficult. Even if human rights are not enforced through international law, their enforcement through domestic law and in other ways might be enhanced by having them included in major human rights treaties. Thus, this Part proposes leaving the treaties in place but acknowledging that their commitments are soft and refocusing the work of the United Nations away from human rights and their legal enforcement.

Disbanding the treaty bodies and the Human Rights Council may not be possible under current human rights treaties, but those bodies could, for example, focus less of their attention on human rights as an international *legal* obligation, reducing the broken windows problems associated with legal norms that are widely violated. The idea would be to take the potential positive effects of iterative engagement with these bodies without the claim that each human rights commitment is a binding legal obligation.

The Paris Agreement on climate is structured in this way. The Agreement itself is binding, as are the reporting requirements, but states select their own nationally determined contributions to reducing greenhouse gasses, which they then "aim" to achieve.[345] Research suggests that the iterative process of reviewing and reporting on human rights before an international body improves human rights practices.[346] The "Paris Model" would preserve and enhance these processes, while acknowledging that not all countries are going to meet all human rights obligations at the same pace. The Paris Agreement itself was structured to convince as many states as possible to join.[347] Human rights agreements have a different problem: countries join the agreements, but the challenge lies in enforcement. Yet the Paris Agreements' benefits in terms of "transparency and accountability" might provide on-the-ground benefits for efforts to improve human rights, even if the legal character of the obligations themselves have changed. The Paris Agreement was designed in part to "drive deeper roots into the domestic policy making processes that will be so key to the success of the kinds of legal, social, and economic transformations that will be necessary to achieve

345. Multinational Agreement on Climate Change, art. 4, *opened for signature* Apr. 22, 2016, T.I.A.S. No. 16-1104 (entered in force Nov. 4, 2016) [hereinafter Paris Agreement].

346. Creamer & Simmons, *supra* note 8; de Búrca, *supra* note 8.

347. JACOB WERKSMAN, INTERNATIONAL LEGAL CHARACTER OF THE PARIS AGREEMENT 7 (Feb. 9, 2016), http://www.law.ed.ac.uk/other_areas_of_interest/events/event_documents/BrodiesLectureontheLegalCharacteroftheParisAgreementFinalBICCLEdinburgh.pdf [https://perma.cc/TJT4-Z8X7].

the Agreement's ambitious goals."[348] Exactly the same kinds of transformation are necessary to the protection and promotion of human rights.

The U.N. General Assembly and Security Council could focus their attention on the many pressing issues other than human rights—either their legal enforcement or their promotion more generally—to avoid credibility and polarization costs. Efforts to protect human rights through other institutions and mechanisms could be redoubled. The result might be a more credible, less polarized United Nations, and a more effective protection of human rights.

Other proposals to reform the United Nations are consistent with the foregoing in important ways. In order to increase the General Assembly's declining prestige, for example, others have suggested that it needs to have greater authority over a circumscribed, more limited agenda—one that includes few, if any, human rights issues.[349] Similarly, many view the human rights work of the specialized U.N. bodies as politicized and arbitrary; going so far as to argue that it "has sullied the UN's reputation, cast doubt on its legitimacy," and "led to diminished national support and popular support" for the organization as a whole.[350] The author suggests doubling down on human rights enforcement, an option addressed below.[351] Even if the proposed solution is different, however, the problems with the current system provide important common ground.

Consider how reform along the lines proposed here would improve the work of the United Nations Security Council. As an initial matter, note that the Security Council is well-positioned to play a vital role in global affairs. Although critics have charged that the composition of the Security Council no longer represents the distribution of global power, in fact the five permanent members of the Council still reflect a large share of global economic and military power, based on their percentage of global GDP,[352] on

348. *Id.*

349. *See, e.g.*, SCHWARTZBERG, *supra* note 197, at 26–27 (2013); *see also* G.A. Res. 68/307 (Sept. 10, 2014) (emphasizing "the need to further enhance the role, authority, effectiveness and efficiency of the General Assembly"); U.N. Secretary-General, *In Larger Freedom: Towards Development, Security and Human Rights for All*, ¶¶ 158–64, annex ¶ 8, U.N. DOC. A/59/2005 (Mar. 21, 2005) (noting that the General Assembly's agenda is too broad and should be focused on issues like migration and terrorism).

350. SCHWARTZBERG, *supra* note 197, at 114–15, 124.

351. *See infra* text accompanying notes 385-93.

352. In 1950, the P-5 share of global GDP was 52%. In 2009, it was 38%, and it rose to 41% in 2014. *See* Bart M. J. Szewczyk, *Variable Multipolarity and U.N. Security Council Reform*, 53 HARV. INT'L L.J. 450, 459 tbl.3, 460 tbl.4 (2012); *see also GDP as Share of World GDP at PPP by Country*, WORLD BANK, https://data.worldbank.org/indicator/NY.GDP.MKTP.PP.CD?end=2014&locations=CN-US-GB-FR-RU-1W&start=2009&view=chart [https://perma.cc/86WP-NBWJ].

GDP per capita of P-5 countries as compared to other countries,[353] on military spending,[354] and on the share of the world's military personnel.[355] What *has* changed substantially in the recent past is Russia's willingness to use force, especially in the Middle East, increasing the significance of P-5 military power.[356] Today, the road to solving global problems runs right through China, Russia, the United States, and Europe (represented in the P-5 by France and the U.K.).

These geopolitical facts create an opportunity for the Security Council to act as an important forum for international legal cooperation. But the opportunity is generated in part by the global rise in power of Russia and China, countries which have made increasingly clear in the last few years—by deed and word—that they disagree with the Western human rights agenda.[357] If the Security Council becomes more active and more effective, it will probably not be around a common aim of providing strong and effective protections for human rights.

Instead, common ground lies in the core purpose of the Security Council itself: ensuring international peace and security.[358] Failure to appreciate this point with respect to Syria is part of what led to the eventual marginalization of both the United States and the Security Council in belated efforts to broker a ceasefire in 2016. The Obama administration's focus on removing Bashar al-Assad from power, in part because of his poor human rights record, generated conflict with Russia and stymied Security Council efforts to

---

353. In 2010, the average GDP per capita, PPP, of the P-5 was 30,072. The world average was 12,828. In 2015 the numbers were 35,221 and 15,546 respectively. *GDP Per Capita, PPP (Current International $)*, WORLD BANK, https://data.worldbank.org/indicator/NY.GDP.PCAP.PP.CD?end=2015&locations=CN-US-RU-GB-FR-1W&start=2010 [https://perma.cc/MVX4-EW2J].

354. The P-5 share of global military spending was 59% in 2015. *See* Sam Perlo-Freeman et al., *SIPRI Fact Sheet: Trends in World Military Expenditure, 2015*, STOCKHOLM INT'L PEACE RES. INST., Apr. 2016, at 2, https://www.sipri.org/sites/default/files/EMBARGO%20FS1604%20Milex%202015.pdf [https://perma.cc/3TSQ-SBLP]. It was 61% in 2010 and 78% in 1950. *See* Szewczyk, *supra* note 352.

355. The only indicator that declined significantly is military personnel: the P-5 had 22% of the world's military personnel in 2014, 26% in 2010, and 63% in 1950. *See* Szewczyk, *supra* note 352; *Armed Forces Personnel, Total*, WORLD BANK, https://data.worldbank.org/indicator/MS.MIL.TOTL.P1?end=2014&locations=FR-CN-US-RU-GB-1W&start=2010 [https://perma.cc/R5MA-8XTM].

356. *See* Itamar Rabinovich, *The Russian-U.S. Relationship in the Middle East: A Five-Year Projection*, CARNEGIE ENDOWMENT INT'L PEACE (Apr. 5, 2016), http://carnegieendowment.org/2016/04/05/russian-u.s.-relationship-in-middle-east-five-year-projection-pub-63243 [https://perma.cc/C87S-6QW2].

357. HUMAN RIGHTS WATCH: WORLD REPORT 2016, *supra* note 18, at 476; *see Relations with Russia*, *supra* note 175; *Russia and China Veto UN Resolution Against Syrian Regime*, *supra* note 146.

358. *See* Bruce Jones, *The Security Council and Changing Distribution of Power* (arguing for strengthening the Security Council's role in maintaining state integrity), *in* THE UN SECURITY COUNCIL IN THE 21ST CENTURY, *supra* note 139, at 793, 798.

resolve the conflict.[359] The U.S.-led agenda of ousting Qaddafi, in part for human rights reasons, angered Russia and China, ultimately impeding U.N. Security Council action in Syria.[360] Of course, in both situations, the United States also had strategic reasons to favor regime change,[361] but that fact only strengthens the view that human rights norms are selectively enforced to serve other interests.

Critics argue that the Security Council's biggest post-Cold War failures have been those of omission: failures to act in Bosnia, Rwanda, Kosovo, and now Syria.[362] Indeed, the failure to act is the problem that R2P was supposed to solve, so it seems counterintuitive to argue that by abandoning humanitarian-based intervention, the Security Council will become more effective and relevant. However, these criticisms of the Security Council overlook its success at maintaining a great powers peace and peace among nation states more generally, risks which are sometimes discounted.[363] But as Robert Kagan recently argued, interstate territorial conflict is a real danger: the decline of U.S. power, the growth of Chinese and Russian power, and territorial ambitions of these two powers which aim to shake up the existing global order, is a recipe for large-scale armed conflict.[364] International law that diminishes such risks remains vitally important today, as it has since the U.N. Charter was drafted.

C. Human Rights: Is International Law Necessary?

An alternative response to the decline in human rights and the costs of human rights to the rest of international law would be to enforce—or at least take seriously the enforcement of—international human rights law.

---

359. *See, e.g.*, Shaik & Roberts, *supra* note 146, at 723–25, 733; Jason Ukman & Liz Sly, *Obama: Syrian President Assad Must Step Down*, WASH. POST (Aug. 18, 2011), https://www.washingtonpost.com/blogs/checkpoint-washington/post/obama-syrian-president-assad-must-stepdown/2011/08/18/gIQAM75UNJ_blog.html?-utm_term=.6e4e60573757 [https://perma.cc/AG3T-LATT]; Macon Phillips, *President Obama: "The Future of Syria Must Be Determined by Its People, but President Bashar al-Assad Is Standing in Their Way"*, WHITE HOUSE BLOG (Aug. 18, 2011), https://www.whitehouse.gov/blog/2011/08/18/president-obama-future-syria-must-be-determined-its-people-president-bashar-al-assad [https://perma.cc/3FGS-4BQ5]; Steven Mufson, *"Assad Must Go": These Three Little Words Are Huge Obstacles for Obama on Syria*, WASH. POST (Oct. 19, 2015), https://www.washingtonpost.com/business/economy/assad-must-go-these-three-little-words-present-a-huge-obstacle-for-obama-on-syria/2015/10/19/6a76baba-71ec-11e5-9cbb-790369643cf9_story.html?utm_term=.849f34ba1c4e [https://perma.cc/WRH6-YZ7B].

360. *See supra* text accompanying notes 147, 335.

361. *See* Christopher Fermor, *NATO's Decision to Intervene in Libya (2011): Realist Principles or Humanitarian Norms?*, 8 J. POL. & INT'L STUD., no. 2, 2012, at 323.

362. *See* Szewczyk, *supra* note 352, at 454.

363. *Id.*

364. Robert Kagan, *Backing into World War III*, FOREIGN POL'Y (Feb. 6, 2017), https://foreignpolicy.com/2017/02/06/backing-into-world-war-iii-russia-china-trump-obama/ [https://perma.cc/W4JS-FZ7F]; *see also* GRAHAM ALLISON, DESTINED FOR WAR: CAN AMERICA AND CHINA ESCAPE THUCYDIDES'S TRAP? (2017) (describing risk of conflict with China).

Enforcing "binding" international human rights law would solve the broken windows problem. It would make the United Nations more credible, reduce polarization, and arguably merits whatever expenditures necessary, assuming that human rights themselves would improve. But there is disagreement about the effectiveness of international law in protecting human rights. And serious efforts at enforcing human rights through international law are politically infeasible.

Empirical scholarship on the effectiveness of international human rights law has focused primarily on treaty ratification, an easily measured variable. There is very little work, by contrast, on whether the doctrinal innovations to enforce human rights had a meaningful impact. The ratification of certain treaties has had modest positive effects on human rights practices of some countries, while ratification of others has not, and some studies even show a correlation between the ratification of certain human rights treaties and more human rights violations.[365] A recent overview of the large body of relevant empirical work concludes, for example, that the Convention on the Elimination of Discrimination Against Women has shown the most impressive results, with small gains for women's political rights and the education of girls in some countries.[366] For other treaties, the correlation between ratification and improved human rights outcomes is weaker or even nonexistent.[367]

A consistent finding across issue areas and across studies is that positive effects of ratifying human rights treaties are correlated with particular domestic conditions within ratifying countries.[368] Beth Simmons has developed this claim in the most detail.[369] Ratification can influence domestic politics under certain circumstances by influencing elite agendas, providing

365. *E.g.*, Daniel W. Hill Jr., *Estimating the Effects of Human Rights Treaties on State Behavior*, 72 J. POL. 1161, 1172 (2010) (finding that "the ratification of [the Convention Against Torture] is associated with worse torture practice than would be expected in the absence of ratification").

366. Kevin L. Cope & Cosette D. Creamer, *Disaggregating the Human Rights Treaty Regime*, 56 VA. J. INT'L L. 459, 467–68 (citing Hill Jr., *supra* note 365; Yonantan Lupu, *The Information Power of Treaty Commitment: Using the Spatial Model to Address Selection Effects*, 57 AM. J. POL. SCI. 912 (2013)).

367. *See* Hathaway, *supra* note 14, at 1940; Hill Jr., *supra* note 365, at 1161, 1172; Cope & Creamer, *supra* note 366, at 467.

368. *See, e.g.*, Wade M. Cole, *Mind the Gap: State Capacity and the Implementation of Human Rights Treaties*, 69 INT'L ORG. 405, 433–34 (2015) (showing a correlation between effective bureaucracies and improved rights protections after ratification of the International Covenant on Civil and Political Rights); *see also* Cope & Creamer, *supra* note 366, at 476–78 (summarizing a number of studies that suggest the post-ratification process within a state matters to its compliance with the treaty).

369. SIMMONS, *supra* note 264, at 125–48.

a focal point for national lawmaking, enabling litigation in national courts, and sparking mass political mobilization.[370]

The significance of domestic politics and law to the enforcement of international human rights law suggests that today—after human rights treaties have been widely ratified—human rights can perhaps be enforced just as well through domestic and transnational legal work as they can through international law. To consider this claim, note that the positive effects of treaty ratification on rights practices can be separated into two groups. First, there are the benefits conferred by ratification itself. For example, elites, policy makers, and domestic interest groups may be mobilized by the process of treaty ratification.[371] The ratification process itself might also create longer term effects by generating changes in domestic statutes, regulations, bureaucratic structure, and even civil society.[372] All of these benefits have already been conferred through the process of ratification. A new approach to enforcement and implementation of human rights treaties would have no effect, except perhaps on the few states which have not ratified human rights treaties.

Second, however, treaty ratification might generate positive human rights outcomes through the domestic legal and political effect of *ongoing* efforts to enforce the treaty as international law, as distinct from the benefits conferred by ratification itself. Far less work attempts to measure these second effects specifically, and the results are generally mixed.[373] One forthcoming study measures the effect of both self-reporting to the Committee Against Torture and of the ensuing review process. Controlling for a wide variety of factors, it concludes with "moderate confidence" that some countries which go through the reporting and review process more than once have a small reduction in the incidence of torture.[374] It also shows that the reporting and review process tends to be covered in the press in Latin

370. *Id.* at 138–39, 147; *see also* Yonatan Luptu, *Best Evidence: The Role of Information in Domestic Judicial Enforcement of International Human Rights Agreements*, 67 INT'L ORG. 469 (2013); Emilia Justyna Powell & Jeffrey K. Staton, *Domestic Judicial Institutions and Human Rights Treaty Violation*, 53 INT'L STUD. Q. 149 (2009).

371. *See* Cope & Creamer, *supra* note 366, at 475–77 (observing the attention that the negotiation and ratification process receives and recognizing that after a state's ratification of an international treaty, these new regulations may take years to "diffuse into . . . local law and custom").

372. *Id.*

373. One study uses the decision to accept a treaty's individual and interstate complaint procedures as an independent variable and finds that it is associated with better rights practices, worse rights practices, or has no effect, depending on the treaty in question. Wade M. Cole, *Human Rights as Myth and Ceremony/Reevaluating the Effectiveness of Human Rights Treaties*, 117 AM. J. SOC. 1331, 1337–42, 1363 (2012).

374. Creamer & Simmons, *supra* note 8, at 15.

America, suggesting that the positive effect of the review process may take place through domestic political mobilization.[375]

Leaving aside the moderate confidence level in a modest effect, and assuming that the results would hold for other treaty regimes,[376] it may be possible to achieve the measured effects even if human rights commitments were not characterized as legally binding *international* law. One way of maximizing benefits and minimizing costs might be to acknowledge that international human rights are—in some senses—soft international legal obligations, although they are often included in binding domestic law. Domestic enforcement mechanisms may be effective without ongoing enforcement through international law. In order to evaluate the possibility empirically, we would need to distinguish between the domestic pressures that result from binding international commitment and those that result from nonbinding international instruments. Some experimental work supports this conclusion by finding that soft international human rights norms have the same effect on people's perceptions as hard international law.[377]

Most substantive human rights obligations are today imposed by an overlapping set of instruments that includes nonbinding international norms like the Universal Declaration of Human Rights; treaty obligations; regional human rights agreements and tribunals; and domestic constitutions, statutes, and common law.[378] Relaxing the claim that international human rights treaties and custom are formally enforceable as binding international law would accordingly leave open many enforcement options including "naming and shaming," transnational human rights advocacy movements, soft or nonbinding norms, active civil society, iterative engagement with international review bodies, domestic enforcement procedures, diplomatic pressure, regional human rights enforcement, conditioning of development aid, and so on. We also know that some of the most important victories for worldwide human rights have been achieved through these methods, including the role of the nonbinding Helsinki Accords in bringing about the

375. *Id.* at 15–23.

376. *But cf.* Cope & Creamer, *supra* note 366, at 479 (showing that many results vary from treaty to treaty).

377. Geoffrey P. R. Wallace, *International Law and Public Attitudes Toward Torture: An Experimental Study*, 67 INT'L ORG. 105, 105, 127–29 (2013); *see also* Stephen Meili, *The Human Rights of Non-Citizens: Constitutionalized Treaty Law in Ecuador*, 31 GEO. IMMIGR. L.J. 347, 385 n.180 (2017) (describing the effect of the nonbinding Cartagena Declaration on domestic legal protections of noncitizens). *But cf.* Adam S. Chilton, *The Influence of International Human Rights Agreements on Public Opinion: An Experimental Study*, 15 CHI. J. INT'L L. 110, 127–28 (2014) (describing an experimental study on adults in the United States that showed less support for solitary confinement among those who were told that the practice violated international law).

378. *See generally* STEVEN R. RATNER ET AL., ACCOUNTABILITY FOR HUMAN RIGHTS ATROCITIES IN INTERNATIONAL LAW: BEYOND THE NUREMBERG LEGACY (3d ed. 2009); FRANCISCO FORREST MARTIN ET AL., INTERNATIONAL HUMAN RIGHTS AND HUMANITARIAN LAW: TREATIES, CASES, AND ANALYSIS (2005).

end of the Cold War.[379] It is a great achievement of the human rights movement that so many human rights are protected in so many different, overlapping ways—direct enforcement of the global treaties may no longer be necessary.

Many readers will disagree with even the suggestion of turning away from the international legal enforcement of human rights. At a minimum, however, it work on human rights outcomes should be accompanied by more and better work on the costs of international human rights enforcement, including cooperation and broken windows costs. A striking feature of the scholarship on human rights is the assumption that human rights treaties and their enforcement regimes have no costs, except possibly to human rights themselves. Even very modest, statistically weak, or otherwise unclear gains in human rights protections are defensible if there are no costs. But as David Kennedy cautions: "[w]e should be on guard when someone seeks to recruit us to a project that only has upsides."[380]

There is a second problem with doubling down on the enforcement of human rights through international law—it is not politically feasible. China, Russia, and what seems to be a growing number of other countries increasingly reject the enforcement of human rights through international law. Indeed, a largely unnoticed aspect of the shift towards a multipolar international order is an increasingly pointed conflict about international law generally and about human rights in particular.[381] A 2016 joint Chinese–Russian declaration on international law illustrates the point.[382] Unlike earlier declarations, the one from 2016 speaks not in terms of "universal" norms in international law, nor does it mention human rights.[383] China and Russia also

379. *See* SARAH B. SNYDER, HUMAN RIGHTS ACTIVISM AND THE END OF THE COLD WAR: A TRANSNATIONAL HISTORY OF THE HELSINKI NETWORK 1–2 (2011); *see generally* Zachary Elkins et al., *Getting to Rights: Treaty Ratification, Constitutional Convergence, and Human Rights Practice*, 54 HARV. INT'L L.J. 61 (2013) (finding that the nonbinding Universal Declaration of Human Rights played an important role in the spread of human rights into domestic constitutions).

380. David Kennedy, *The International Human Rights Regime: Still Part of the Problem*, *in* EXAMINING CRITICAL PERSPECTIVES ON HUMAN RIGHTS 19, 27 (Rob Dickinson et al. eds., 2012); *see also* David Kennedy, *International Human Rights Movement: Part of the Problem?*, 15 HARV. HUM. RTS. J. 101, 102 (2002).

381. *See* William W. Burke-White, *Power Shifts in International Law: Structural Realignment and Substantive Pluralism*, 56 HARV. INT'L L.J. 1, 4 (2015) (noting that few authors have considered "how changes in the distribution of power influence the processes and substance of international law").

382. Ministry of Foreign Affairs of the Russ. Federation, The Declaration of the Russian Federation and the People's Republic of China on the Promotion of International Law (June 25, 2016), http://www.mid.ru/en/foreign_policy/position_word_order/-/asset_publisher/6S4RuXfeYlKr/content/id/2331698 [https://perma.cc/6CN7-N3J8] [hereinafter Joint Declaration].

383. *Cf.* Communiques, Ministry of Foreign Affairs of China, Joint Statement of the People's Republic of China and Russian Federation on Major International Issues (May 23, 2008), http://www.fmprc.gov.cn/mfa_eng/wjdt_665385/2649_665393/t465821.shtml [https://perma.cc/6PBT-N7T5] (reiterating China and Russia's respect for human rights); Communiques, Ministry of Foreign Affairs of China, China–Russia Joint Statement on 21st Century

have a growing influence on the approach that many countries take to international law, including Brazil, India, and South Africa.[384] Turkey illustrates the newfound power of China and Russia—and the centrality of human rights to their growing control over the content of international law. Turkey has become increasingly autocratic since the government thwarted a coup attempt in 2016.[385] Human rights have long been a major stumbling block to Turkey's admission to the European Union; in response to the post-coup human rights violations, Europe has threatened to end the (recently re-opened) process of admitting Turkey.[386] Turkey, on the other hand, has explicitly threatened to join the Shanghai Cooperation Organization, a security and economic bloc led by China and Russia.[387] Whatever the ultimate outcome of this particular dispute—and Turkey has a long history of violent conflict with Russia—the growing power of Russian and Chinese approaches to international law is clear.

Doubling down on the international legal protection of human rights is not feasible for other reasons. A serious enforcement effort might involve the doctrines described in Part I, except on a far broader scale than previous efforts. The true cost of human rights enforcement would involve potentially great costs to the friendly relations of states and even interstate peace. It could mean stepping up civil and criminal universal jurisdiction prosecutions and eliminating immunity even in cases involving U.S., Chinese, or Israeli defendants tried in foreign national courts. Taking seriously a right of secession as a meaningful tool to improve human rights would involve secessions from human rights-violating states, potentially including Russia, Turkey, China, and others. The high costs to interstate relations are clear. Unlike international human rights law, there is not an overlapping set of

---

World Order (中俄关于21世纪国际秩序的联合声明) (July 1, 2005), http://www.mfa.gov.cn/chn//gxh/zlb/smgg/t201988.htm [https://perma.cc/4KFB-KDUA] (emphasizing respect for the Universal Declaration of Human Rights and state sovereignty); Russian-Chinese Joint Declaration on a Multipolar World and the Establishment of a New International Order, in letter dated May 15, 1997 from the Permanent Representatives of China and the Russian Federation to the United Nations, Addressed to the Secretary-General, U.N. Doc. A/52/153-s/1997/384 (May 20, 1997) (mentioning and respecting "universally recognized principles of international law").

384. Burke-White, *supra* note 381, at 2, 4, 7; *see also* GIDEON RACHMAN, EASTERNIZATION: ASIA'S RISE AND AMERICA'S DECLINE FROM OBAMA TO TRUMP AND BEYOND 14, 212–24 (2016).

385. Tim Arango & Ceylan Yeginsu, *Turkish President Returns to Istanbul in Sign Military Coup Is Faltering*, N.Y. TIMES (July 15, 2016), http://www.nytimes.com/2016/07/16/world/europe/military-attempts-coup-in-turkey-prime-minister-says.html [https://perma.cc/MK29-SEAH].

386. Rod Nordland & James Kanter, *Turkey and E.U. Near Breaking Point in Membership Talks*, N.Y. TIMES (Nov. 23, 2016), https://www.nytimes.com/2016/11/23/world/europe/turkey-eu-membership-talks.html [https://perma.cc/BZ77-T3LM].

387. Samuel Osborne, *Turkey Could Join Bloc with Russia and China Instead of EU, President Erdogan Says*, INDEPENDENT (Nov. 20, 2016), http://www.independent.co.uk/news/world/europe/turkey-president-erdogan-eu-russia-china-a7428391.html [https://perma.cc/896F-TK65].

international and domestic prohibitions on interstate war. Only international law can do that work.

IV. Conclusion

International human rights law, despite its many historical successes, is no longer in its golden age. The beneficiary may be international law as a whole. But the benefits will be fully realized only if those who create and shape international law consider carefully the dangers and opportunities that the decline presents. At a minimum, scholars should focus not only on the decades-old debate about international human rights law and human rights, but instead on international human rights law and international law as a system. As this Article has shown, empirical work and models drawn from political science, social psychology, and sociology all suggest that human rights have imposed costs on international law and on international peace and security. Examples bear out what the models predict.

Of particular significance, because the stakes involved are so high, is empirical work showing a strong correlation between territorial disputes and armed conflict. Some aspects of the human rights transformation of international law have undermined international legal norms and institutions designed to limit such conflicts. It is easy to think that large-scale interstate war is simply impossible under contemporary conditions, due to globalization, international economic ties, and the deadly weapons available. That same view was prominent just over a century ago, on the eve of World War I.[388] The remarkable contribution that international law has apparently made to interstate peace may be *the* great success of post-World War II international law. It is well worth preserving today through vigorous and credible international institutions such as the United Nations and through international legal norms, which are as strong and robust as possible.

388. *See, e.g.*, FLORIAN ILLIES, 1913: DER SOMMER DES JAHRHUNDERTS (2013) (discussing the work of Norman Angell, including his book *The Great Illusion*).

* * *

Law, Organizing, and Status Quo Vulnerability

Benjamin I. Sachs*

This is an era of striking economic and political inequality. The statistics are familiar, but a few numbers are worth highlighting. Income disparities in the United States are now the highest they have been since the Gilded Age,[1] with the top 1% of earners taking home 20.1% of national income.[2] In fact, the twenty wealthiest people in the United States own more wealth than the bottom half of the population put together: twenty individuals with more wealth than another 152 million people.[3] On the political front, the picture is no better. Recent studies by political scientists Larry Bartels and Martin Gilens reveal that the government's policies reflect the views of the wealthy but not of the poor or middle class.[4] As Gilens puts it, "when preferences between the well-off and the poor diverge, government policy bears absolutely no relationship to the degree of support or opposition among the poor."[5]

There are multiple causes for the rise of economic and political inequality in the United States, but one such cause is the decline of unions. Between 1973 and 2007, union membership rates fell from 34% of the male private-sector workforce to 8% of that workforce.[6] The private-sector rate is now below 7%.[7] This dramatic decline is causally related to the increase in both economic and political inequality. Jake Rosenfeld and Bruce Western

---

*Kestnbaum Professor of Labor and Industry, Harvard Law School. Thanks to Kate Andrias and Andrew Strom for helpful comments and suggestions, and to K. Sabeel Rahman for helpful discussion. Alexander Miller and Maia Usui provided superb research assistance.

1. *See, e.g.*, ESTELLE SOMMEILLER, MARK PRICE & ELLIS WAZETER, ECON. POLICY INST., INCOME INEQUALITY IN THE U.S. BY STATE, METROPOLITAN AREA, AND COUNTY 25–26 (2016), http://www.epi.org/files/pdf/107100.pdf [https://perma.cc/HSK5-9W26] (documenting the share of all income held by the top 1% of families in the United States between 1917 and 2013).

2. *Id.* at 4. In some areas of the country, the numbers are even more stark. In New York, for example, the average family in the top 1% of earners has forty-five times more income than an average family in the bottom 99%. *Id.* at 7.

3. CHUCK COLLINS & JOSH HOXIE, INST. FOR POLICY STUDIES, BILLIONAIRE BONANZA: THE FORBES 400 AND THE REST OF US 1 (2015), http://www.ips-dc.org/billionaire-bonanza/ [https://perma.cc/YMB5-SVHZ].

4. LARRY M. BARTELS, UNEQUAL DEMOCRACY: THE POLITICAL ECONOMY OF THE NEW GILDED AGE 5–6 (2d ed. 2016); MARTIN GILENS, AFFLUENCE AND INFLUENCE: ECONOMIC INEQUALITY AND POLITICAL POWER IN AMERICA 81 (2012).

5. GILENS, *supra* note 4, at 81.

6. Bruce Western & Jake Rosenfeld, *Unions, Norms, and the Rise in U.S. Wage Inequality*, 76 AM. SOC. REV. 513, 513 (2011).

7. BUREAU OF LABOR STATISTICS, U.S. DEP'T OF LABOR, UNION MEMBERS—2016 (2017), http://www.bls.gov/news.release/pdf/union2.pdf [https://perma.cc/3RBU-JKBV].

estimate that up to one-third of the climb in income inequality for men, and one-fifth for women, across these decades is attributable to the decline in unionization rates.[8] The decline in union strength is also routinely identified as a critical factor in explaining the rise in political inequality.[9]

The link between union decline and inequality has led to a burgeoning interest in how the union movement might be reinvigorated, with scholars and popular commentators from a wide range of fields exploring the question anew.[10] For labor lawyers and legal academics, the primary area of focus is how law can better enable workers to organize unions when those workers wish to organize them. This is a notoriously difficult question because of an interrelated set of facts about labor in the United States. First, workplaces are nonunion by default, so unionization requires workers to take a series of difficult steps to change the status quo.[11] Second, unionization is a collective good and thus organizing a union poses all the difficulties and dilemmas of collective action.[12] Third, management is nearly uniformly opposed to unionization and has at its disposal a wide range of tools that are highly

---

8. Western & Rosenfeld, *supra* note 6, at 514.

9. JACOB S. HACKER & PAUL PIERSON, WINNER-TAKE-ALL POLITICS: HOW WASHINGTON MADE THE RICH RICHER—AND TURNED ITS BACK ON THE MIDDLE CLASS 142 (2010); KAY LEHMAN SCHLOZMAN, SIDNEY VERBA & HENRY E. BRADY, THE UNHEAVENLY CHORUS: UNEQUAL POLITICAL VOICE AND THE BROKEN PROMISE OF AMERICAN DEMOCRACY 325–26 (2012).

10. On the scholarship side, see, e.g., Daryl J. Levinson, *Foreword: Looking for Power in Public Law*, 130 HARV. L. REV. 31, 137 (2016) ("Recognizing the importance of unions as a vehicle for mobilizing and empowering the nonwealthy and reducing political and economic inequality, commentators have suggested reforms designed to reinvigorate unions as political organizations."); Kate Andrias, *The New Labor Law*, 126 YALE L.J. 2 (2016); Benjamin I. Sachs, *The Unbundled Union: Politics Without Collective Bargaining*, 123 YALE L.J. 148 (2013); GILENS, *supra* note 4, at 158 (arguing that unions "would appear to be among the most promising interest group bases for strengthening the policy influence of America's poor and middle class," although a decline in unionization rates suggests that "unions' success in these efforts is likely to be fairly limited"); HACKER & PIERSON, *supra* note 9, at 303 ("[T]he organizations that traditionally bolstered middle-class democracy have declined. Nowhere is this clearer or more fateful than with regard to American labor."). For a sample of popular writing on the subject, see, e.g., Teresa Ghilarducci, *Farewell to America's Middle Class: Unions Are Basically Dead*, ATLANTIC (Oct. 28, 2015), http://www.theatlantic.com/business/archive/2015/10/unions-are-basically-dead/412831/ [https://perma.cc/D9FT-4ZHY]; Nicholas Kristof, *The Cost of a Decline in Unions*, N.Y. TIMES (Feb. 19, 2015), http://www.nytimes.com/2015/02/19/opinion/nicholas-kristof-the-cost-of-a-decline-in-unions.html?_r=0 [https://perma.cc/FG74-7X38]; Sean McElwee, *One Big Reason for Voter Turnout Decline and Income Inequality: Smaller Unions*, AM. PROSPECT (Jan. 30, 2015), http://prospect.org/article/one-big-reason-voter-turnout-decline-and-income-inequality-smaller-unions [https://perma.cc/FY2D-TZJN]; Robert Reich, *Unions Can Save the Middle Class*, SALON (June 6, 2015), http://www.salon.com/2015/06/06/robert_reich_unions_can_save_the_middle_class_partner/ [https://perma.cc/YVU2-BBYL].

11. *See, e.g.*, Mark Barenberg, *Democracy and Domination in the Law of Workplace Cooperation: From Bureaucratic to Flexible Production*, 94 COLUM. L. REV. 753, 959–61 (1994) (reviewing barriers to unionization).

12. Benjamin I. Sachs, *Enabling Employee Choice: A Structural Approach to the Rules of Union Organizing*, 123 HARV. L. REV. 655, 680–85 (2016) (discussing impediments to unionization posed by collective action problems).

effective in deterring union-organizing efforts, greatly exacerbating the collective action problems that would hinder unionization even without managerial opposition.[13] These facts make effective legal protection for organizing critical. They also make such protection challenging to design and enforce.

This Essay endeavors to deepen our understanding of how a legal regime like the National Labor Relations Act can make unionization feasible despite these challenges. It does this by offering a different way of understanding how labor law rights and remedies enable union organizing. Because the difficulties involved in unionization are similar to the difficulties that inhere in collective action and social movement mobilization generally, the Essay draws on a prominent strand of social movement theory to make its argument. Known as the political process model, this theory emphasizes two basic preconditions for successful collective action: one, objective structural conditions that make mobilization feasible and, two, the existence of a subjective view among participants that collective action is likely to succeed.[14] Although the Essay relies on political process theory, these preconditions for mobilization are common across multiple approaches to social movements.[15] The argument in this Essay is that law can facilitate both the objective and subjective preconditions for collective action.

The notion that law can create objective structural conditions for successful collective action is relatively familiar, but the point is worth reiterating, and the Essay will spend some time showing how labor law functions in this way. The Essay focuses, however, on the subjective front. Many scholars, including labor law scholars, have made the point that workers are likely to form unions only when they believe that such an effort is likely to succeed. As Rick Fantasia puts it in his seminal sociology of union organizing, "[w]hat workers here (or anywhere) will fight for is largely a function of what they can reasonably expect to win."[16] Workers' views about

---

13. *See, e.g.*, *id.* at 684 (discussing employers' use of threats that unionization will result in closure of the business and their willingness to discharge employees as retaliation for supporting unionization).

14. *See, e.g.*, DOUG MCADAM, POLITICAL PROCESS AND THE DEVELOPMENT OF BLACK INSURGENCY 1930–1970, at 34 (1st ed. 1982) (discussing "objective social conditions" and subjective belief that status quo is "subject to change" through group action). As discussed below, a third precondition is what McAdam calls "indigenous organizational strength." *Id.* at 43. By this he means the existence of organizations that can (or can be built to) channel the collective action facilitated by the first two preconditions. *See id.* at 43–48 (analyzing the four critical resources provided by indigenous organizations); *infra* text accompanying notes 36–39 (discussing McAdam's political process theory for collective action).

15. For a similar view in a different theoretical tradition, see, e.g., Mark Granovetter, *Threshold Models of Collective Behavior*, 83 AM. J. SOC. 1420 (1978).

16. RICK FANTASIA, CULTURES OF SOLIDARITY: CONSCIOUSNESS, ACTION AND CONTEMPORARY AMERICAN WORKERS 170 (1988). In the labor law literature, Brishen Rogers's recent work provides an insightful analysis. *See, e.g.*, Brishen Rogers, *Passion and Reason in Labor Law*, 47 HARV. C.R.-C.L. L. REV. 313, 361 (2012) ("[E]mployees will be free to choose

likelihood of success, however, involve two sets of beliefs that theorists often merge. One is workers' belief about the union: how likely it is that a union will be organized and how powerful that union will be. This belief is informed, for example, by estimations of how many other workers are likely to participate in the union effort,[17] and how militant the union is likely to be at the bargaining table. But predictions of *union* strength are only one side of the coin. Workers' views on the likelihood of success depend just as heavily on their perceptions of *managerial* strength. That is, workers' views about the likelihood of union success depend critically on their views about how vulnerable management is to challenge. This observation is critical to understanding how law functions in the context of organizing for collective action.

The existing literature stresses the prospects for union strength and focuses on how law can and does (or does not) contribute to the building of that strength. Here, in contrast, I focus on the other side of the coin. I argue that labor law can facilitate unionization by convincing workers that management (and the nonunion system of workplace relations it supports) is susceptible to challenge. When labor law intervenes visibly in the workplace and renders management observably subject to an authority more powerful than itself, the law helps convince workers that management and its preferred system of authority relations is vulnerable. Persuading workers that management is vulnerable, in turn, improves workers' assessments of the likely success of unionization. It therefore increases workers' willingness to participate in unionization efforts and facilitates unionization among workers who desire it—labor law's central function.[18]

The vulnerability discussed here, it is important to stress, is vulnerability to law and to a form of workplace organization the law protects. Vulnerability as used in this Essay, then, implies only that management is functionally subject to a set of laws that clearly governs its conduct. Managers involved in union campaigns often vigorously contest such vulnerability, telling workers, for example, "[I don't] give . . . a damn about the law,"[19] or "I'm the total dictator,"[20] or simply "I am the law."[21] These statements are meant to scuttle unionization efforts by conveying an aura of *in*vincibility. But given

---

unionization only if unionization is a realistic option, and only if they perceive it as a realistic option.").

17. For an excellent discussion, see, e.g., Brishen Rogers, *"Acting Like a Union": Protecting Workers' Free Choice by Promoting Workers' Collective Action*, 123 HARV. L. REV. F. 38, 47–48 (2010) (explaining that "no one is inclined to participate unless everyone else . . . is also participating").

18. *See* Sachs, *supra* note 12, at 693 (discussing labor law's goal of maximizing employee choice on the question of unionization).

19. Baberton Manor, 252 N.L.R.B. 380, 390 (1980).

20. E.E.C., Inc., 297 N.L.R.B. 943, 947 (1990).

21. Caroline Gardens Apartment Corp., No. 2-CA-23514, 1990 WL 1222181, at *2 (N.L.R.B. Div. of Judges Jan. 17, 1990).

that institutions ought to be subject to the laws that govern them, the form of vulnerability discussed in this Essay is a normatively desirable one. And one that labor law should ensure.

In fact, a wide range of labor law rights and remedies can be understood as demonstrating to workers that management is vulnerable—on this meaning—to labor law and to unionization. One such remedy, known as notice reading, will help illustrate the point and motivate the discussion.[22] In a typical unfair labor practice proceeding, the NLRB orders the employer to comply with a set of affirmative commands—to reinstate a discharged worker, for example, or to cease and desist from surveilling union activity—and also to post a notice detailing the terms of the Board's order.[23] Notice postings themselves generally are uncontroversial and generate little employer resistance. On the other hand, in certain cases the Board will order the employer not only to post the notice but also to read that notice to its employees. These orders generate intense employer opposition and harsh judicial skepticism. In one prominent case, the employer characterized the notice-reading order as "design[ed] to humiliate and embarrass . . . ."[24] The D.C. Circuit has described the remedy as "humiliating and degrading to the employer" and as "incompatible with the democratic principles of the dignity of man."[25]

Notice readings, like most remedies, have various effects. Like a notice posting, a notice reading conveys facts to employees, perhaps more effectively than postings do.[26] But notice readings also have a separate effect, captured by the view that they are humiliating, embarrassing, and undignified. When management reads to employees a list of actions the law requires it to take and not take, management *enacts* its vulnerability to labor

22. For a recent discussion of the remedy, see Andrew Strom, *Class Bias on Display at the D.C. Circuit*, ONLABOR (May 26, 2016), https://onlabor.org/2016/05/26/class-bias-on-display-at-the-d-c-circuit/ [https://perma.cc/YC55-FLTM].

23. For one of countless examples, see, e.g., HTH Corp., 361 N.L.R.B. No. 65, at 3 (2014) (highlighting the NLRB's broad discretion to exercise remedial authority—including the notice posting requirement).

24. Teamsters Local 115 v. NLRB, 640 F.2d 392, 401 (D.C. Cir. 1981) (quoting employer's brief).

25. Int'l Union of Elec., Radio, & Mach. Workers, AFL-CIO v. NLRB, 383 F.2d 230, 233–34 (D.C. Cir. 1967).

26. Notice readings are often defended on the ground that in-person communication is more effective than written communication. *See, e.g.*, Thomas C. Barnes, Note, *Making the Bird Sing: Remedial Notice Reading Requirements and the Efficacy of NLRB Remedies*, 36 BERKELEY J. EMP. & LAB. L. 351, 360–64 (2015) (labeling spoken notice "a more effective remedy" and detailing how it better furthers restorative and deterrent purposes); John W. Teeter, Jr., *Fair Notice: Assuring Victims of Unfair Labor Practices That Their Rights Will Be Respected*, 63 UMKC L. REV. 1, 2 (1994) (advocating for notice reading in light of the greater effectiveness of oral statements). The NLRB itself advocates readings when workforce illiteracy rates are high. NLRB General Counsel Memorandum OM 16-21, Notice Reading In Cases Where Unit Employees Have Literacy Issues (June 21, 2016), https://www.nlrb.gov/reports-guidance/operations-management-memos [https://perma.cc/9BWF-QZGE].

law.[27] Such a public commitment to comply with the law's requirements may well be experienced by managers as humiliating and embarrassing. But this tells us more about management's views of the legitimacy of labor law than about the remedy of notice reading. More important than the embarrassment a manager might feel is the signal the reading sends to employees—a signal of the legal fact that employers are susceptible to labor law and to the unionization project labor law protects.

Numerous labor law rights and remedies not only function in this way but can also be defended on these grounds. The very basic right to discuss unionization at the workplace, for example, not only enables employees to communicate with each other about the virtues of unionization, but also signals that the employer's control over the workplace is not absolute.[28] The same was true, perhaps even more so, when union organizers had a partial right to enter employer property to discuss unionization with employees: their very presence signaled that managerial control had its limits and that managerial power might therefore be subject to challenge.[29] Reinstatement orders clearly have this impact as well, as does the right to have a coworker present at a disciplinary interview.[30]

The goal of this Essay is to show that by convincing workers of management's vulnerability, labor law can facilitate union organizing and thereby advance its statutory mission. Recognizing that labor law can function in this way, in turn, suggests some new, perhaps unorthodox, reforms to the labor law regime. In the final Part, the Essay proposes these reforms, including a requirement that all covered employers read a notice of labor law rights to their employees on a regular basis.[31] The conclusion suggests that the argument developed here in the context of labor law has implications for collective action and social movement mobilization more broadly.

I. The Political Process Model: An Overview

A large literature in sociology is dedicated to addressing the question of why and under what conditions social movements emerge.[32] And while the

27. For an example of an actual notice reading, see *infra* text accompanying note 89.

28. *See infra* text accompanying notes 110–12 (emphasizing that despite employers' vigorous objections to union discussions in the workplace, the NLRB's enforcement of the right to discuss unionization at work demonstrates management vulnerability).

29. *See infra* text accompanying notes 133–34 (recognizing the importance of vulnerability to the prospects for unionization).

30. *See infra* text accompanying notes 113–18 (demonstrating that reinstatement indicates that the employer is still susceptible to the law and discussing a union campaign study showing that union win rates were higher after union activists were discharged and later reinstated by the NLRB).

31. As discussed, such a requirement would probably require statutory amendment. *See infra* Part III.

32. For a review, see Doug McAdam, *Conceptual Origins, Current Problems, Future Directions*, *in* COMPARATIVE PERSPECTIVES ON SOCIAL MOVEMENTS 24–25 (Doug McAdam et al.

precise definition of "social movement" is contested,[33] the literature is at its core dedicated to examining "the origins of collective action."[34] One prominent theory of social movement emergence is the political process model. As elaborated by Doug McAdam, political process theory stresses three necessary preconditions for movement emergence, that is, for the successful generation of collective action. The first is the existence of a "structure of political opportunities" that offers potential movement participants an "objective 'structural potential' for collective . . . action."[35] What constitutes such a structural potential depends, of course, on the nature of the movement or the form of collective action being contemplated. The focus of McAdam's work was the American civil rights movement (broadly defined), and he points to a set of social and economic dynamics—including the collapse of cotton as the core of the southern economy, migration, the increase in black electoral strength, and World War II—as constituting the relevant opportunities that made the civil rights movement possible.[36] But, in general terms, political opportunities are the objective structural conditions necessary to make a specific form of collective action possible.

Second, and most important for present purposes, McAdam argues that the successful emergence of collective action depends on the "collective assessment [by potential participants] of the prospects for *successful* insurgency."[37] This assessment itself turns on participants believing two related sets of things: one, participants must be optimistic about the potential strength of the organizations they are forming; and two, they must believe that the status quo regime is "vulnerable to challenge."[38]

Vulnerability is thus central to the political process theory of mobilization; it is a *sine qua non* of collective action.[39] The idea is intuitively

---

eds., 1996) (observing that "[t]he concept of political opportunities has . . . proven to be a welcome addition to the analytic arsenal of movement scholars").

33. DAVID A. SNOW, SARAH A. SOULE & HANSPETER KRIESI, THE BLACKWELL COMPANION TO SOCIAL MOVEMENTS 6–11 (2004) (listing examples of the variety of definitions of social movements and noting that they may differ in what terms are emphasized or accented).

34. DOUG MCADAM, POLITICAL PROCESS AND THE DEVELOPMENT OF BLACK INSURGENCY, 1930–1970, at viii (2d ed. 1999).

35. *Id.* at 48.

36. *See id.* at 73–83; *see also* MICHAEL J. KLARMAN, FROM JIM CROW TO CIVIL RIGHTS 4 (2004) (analyzing "factors behind the nation's racial transformation").

37. MCADAM, *supra* note 14, at 40 (emphasis added).

38. *See id.* at 40, 49 (discussing the "importance of indigenous organizations" and perceptions of a vulnerable political system as factors in "cognitive liberation").

39. Although McAdam is credited with developing the theory, he was not the first—nor the last—to identify perceptions of vulnerability as a prerequisite for collective action. In discussing vulnerability, McAdam himself draws on Piven and Cloward, who wrote in their influential book, *Poor People's Movements*, that for a movement to emerge, "[t]he social arrangements that are ordinarily perceived . . . immutable must come to seem . . . *mutable*." FRANCES FOX PIVEN & RICHARD A. CLOWARD, POOR PEOPLE'S MOVEMENTS: WHY THEY SUCCEED, HOW THEY FAIL 12 (1977) (emphasis added). In a similar vein, Rick Fantasia describes the importance—to union organizing specifically—of demonstrating to workers that "a seemingly omnipotent authority [i.e.,

accessible. If individuals believe that the current regime is invincible, that there is no prospect for change, then they are unlikely to participate in a collective effort to make such change, assuming that participation will entail costs—as it always will. This dynamic also explains why rulers, especially authoritarian ones, attempt to convey invulnerability to potential challengers. The literature on competitive authoritarianism is thus replete with examples of regimes deploying elections to "generate a public image of invincibility."[40] Closer to home, the dynamic helps explain why managers often attempt to communicate that managerial control—and the default regime of individual bargaining it defends—is invulnerable in the face of a union-organizing campaign. NLRB and court decisions contain countless statements like: "I am the boss, and no God damn Union is coming in here . . . ."[41]

Third, and finally, political process theory posits that successful collective action depends on "indigenous organizational strength," or resources within the relevant community that permit prospective movement participants to take advantage of the structural and cognitive opportunities discussed above.[42] When there is objective structural potential for collective action and when participants believe that the existing regime is vulnerable to challenge, the emergence of sustained collective action still depends on an organization capable of channeling such activity. Here, McAdam has in mind both existing organizations within communities that can channel collective activity, as well as more informal networks within those communities that can facilitate the creation of such organizations.[43]

II. Labor Law and Political Process Theory

This Essay's primary goal is to show how labor law can contribute to workers' perception that management is vulnerable to unionization and thereby to satisfy political process theory's second prong. But before getting to that discussion, this Part briefly shows how law can also contribute to the other two preconditions identified by the political process school. This Part thus proceeds by first taking up law's contribution to objective structural opportunity and organizational viability before turning, in subpart B, to the core discussion of vulnerability.

---

management] could be overcome." FANTASIA, *supra* note 16, at 145. Steven Buechler, in reviewing the evolution of social movement theory, stresses the importance to movement emergence of the "established order becom[ing] vulnerable to the actions of contenders." Steven M. Buechler, *Social Strain, Structural Breakdown, Political Opportunity, and Collective Action*, 2 SOC. COMPASS 1031, 1039 (2008). His conception, though, is more objective than subjective.

40. BEATRIZ MAGALONI, VOTING FOR AUTOCRACY: HEGEMONIC PARTY SURVIVAL AND ITS DEMISE IN MEXICO 9 (2006).

41. Barbers Iron Foundry, 126 N.L.R.B. 30, 46 (1960).

42. MCADAM, *supra* note 14, at 43.

43. *See id.* at 43–48 (discussing the "significance of indigenous organizations," formal and informal).

A. Political Opportunities and Organizational Strength

Within the framework of political process theory, it should be easy to see how labor law rights and remedies might provide the objective structural potential for successful collective action and how they can constitute a kind of micro-political opportunity structure.[44] In fact, although the political process frame may be unfamiliar, the idea that labor law aims to provide workers with an objectively feasible opportunity to unionize is fairly traditional.[45] A very brief overview is in order.

The most basic way labor law functions as opportunity structure is by making it illegal for employers to terminate workers who engage in union-organizing activity. Without this threshold assurance, union organizing would be exceedingly difficult and likely impossible: if workers lose their jobs when they try to form unions, they will either not try to unionize or they will not succeed. But such terminations would be legally permissible absent labor law's prohibition.[46] It is worth pausing to emphasize the point, precisely because it may have become too familiar: without labor law's prohibition on union-related discharges, it would be legally permissible for employers to fire anyone who engages in union organizing.[47] At certain historical moments, workers have been able to overcome the brute force of anti-union discharges

---

44. Throughout the bulk of this discussion, I will ignore the question of whether labor law's protections are actually enforced. I do this because my goal is to offer a theoretical account of how labor law functions—when it does function—and so I set the critical issue of enforcement aside. As many, myself included, have described elsewhere, American labor law suffers from a profound enforcement problem in practice. *See, e.g.*, Benjamin I. Sachs, *Employment Law as Labor Law*, 29 CARDOZO L. REV. 2685, 2694–700 (2008) (describing the NLRA's "deeply inadequate remedial regime").

45. *See, e.g.*, Sachs, *supra* note 12, at 656 (discussing labor law's goal of enabling employee choice on the union question). In Part IV, I note that the lack of enforcement can undermine law's capacity to demonstrate to workers that management is vulnerable to the law of union organizing. *See infra* Part IV.

46. This is true because of the background legal rule of at-will employment. Jay M. Feinman, *The Development of the Employment at Will Rule*, 20 AM. J. LEG. HIST. 118 (1976); Richard A. Epstein, *In Defense of the Contract at Will*, 51 U. CHI. L. REV. 947, 953–54 (1984).

47. Put somewhat differently, absent some external restraint, employers facing unionization will generally engage in what Richard Posner famously calls "rational predatory action." Richard A. Posner, *Some Economics of Labor Law*, 51 U. CHI. L. REV. 988, 994 (1984). Because unionization imposes costs on firms and—though this was not Posner's point—because unionization deprives managers of some of their control over the workplace, Posner predicts that management's rational response to union organizing is the termination of union supporters. *See id.* In fact, in a leading study of union-organizing drives, researchers concluded that employers opposed the union 96% of the time. *See* KATE BRONFENBRENNER, ECON. POLICY INST., NO HOLDS BARRED: THE INTENSIFICATION OF EMPLOYER OPPOSITION TO ORGANIZING 10 tbl.3 (May 20, 2009), http://www.epi.org/publication/bp235/ [https://perma.cc/P6FG-3WMD]. More to the point, studies suggest that employers terminate up to 20% of active union supporters. *See, e.g.*, Sachs, *supra* note 12, at 684 (collecting sources). Such terminations remove union supporters from the workplace and thus from the organizing effort. Terminations also signal to other potential union supporters that the cost of union support is very high. They are thus highly effective at subverting organizing. *See, e.g.*, *id.* at 684–85 (citing BRONFENBRENNER, *supra* note 47, at 10–11 tbl.3).

by preventing other employees—through picketing and other forms of social pressure—from accepting jobs left vacant by those fired for union activity.[48] But when workers lack the social power to deter such discharges, law has the capacity to prohibit them and make unionization feasible.[49]

While necessary, this type of threshold protection is hardly sufficient to provide workers with a genuine opportunity to unionize.[50] Unionization, like collective action of all kinds, calls on workers to overcome a set of coordination and information hurdles. With respect to coordination, a union campaign requires organizers to garner the support of a majority of a given workforce.[51] This, in turn, requires that organizers contact and speak with a large number of workers, who can be dispersed across large geographical areas. In a number of ways, labor law can make this coordination possible. First, although background rules of property and contract would entitle employers to prohibit organizers from using the workplace as a centralized locus for union activity, labor law allows organizers to do just that. Thus, for example, the law permits employees to use nonworking areas of their workplace (during nonworking time) to solicit support for unionization and to exchange information about unionization.[52] For a time, moreover, labor law also permitted full-time union organizers to conduct union activity on

48. It is also possible for labor market conditions to make anti-union discharges too costly to some employers.

49. There is a loose parallel available in Professor Klarman's analysis of the civil rights movement. Klarman shows how ensuring a basic guarantee of physical safety and security was necessary to enable African Americans to engage in protest activity during the civil rights era. *See* KLARMAN, *supra* note 36, at 445 (discussing factors contributing to the physical security necessary for social protest among urban blacks). This is not to equate the protection of physical safety and the protection against discharge. It is only to say that just as physical safety constituted a threshold for participation in the civil rights movement, protection against discharge is a threshold for participation in union-organizing activity.

50. Another threshold protection is labor law's prohibition on the bargaining of individual employment contracts once a union is organized. *See* J.I. Case Co. v. NLRB, 321 U.S. 332, 338–39 (1944) (holding that individual contracts cannot survive collective contracts). Under background contract rules, an employer could subvert the effects of a successful union-organizing drive by offering individual employment agreements to its employees. If those individual employment agreements were superior to the collective deal that the union struck, many of the employees would likely accept them, and thereafter have much less reason to continue to support the union. If the individual employment agreements were inferior to the collective deal, the employees would have no incentive to accept them, unless the employer made rejecting the agreement grounds for sanction (demotion, discharge, etc.), which, absent legal prohibition, the employer likely would do. Labor law, however, flatly prohibits employers from bargaining individual contracts in the presence of a union-negotiated collective bargaining agreement (absent union consent). *Id.*

51. More specifically, of a given bargaining unit. *See* 29 U.S.C. § 159(a) (2012) ("Representatives designated or selected for the purposes of collective bargaining by the majority of the employees in a unit appropriate for such purposes . . . .").

52. *See, e.g.*, Republic Aviation Corp. v. NLRB, 324 U.S. 793, 803–04 n.10 (1945) ("It is therefore not within the province of an employer to promulgate and enforce a rule prohibiting union solicitation by an employee outside of working hours, although on company property.").

company property, within certain limits.[53] Second, labor law also provides organizers with valuable information about the workers whose support will ultimately be required, information that an employer would otherwise have no obligation to disclose. So, when a threshold of support is reached, the law requires the employer to provide a union with names and addresses of workers in the bargaining unit, allowing organizers to reach workers outside the workplace and before or after the work day.[54]

Unionization also requires the existence of viable organizations capable of channeling and facilitating workers' collective action, the need anticipated by political process theory's third precondition for collective action. Here, labor law plays an important role in making unions—and not just unionization—viable. As many have observed, one of the most difficult tasks in fostering social movement organization—including union organization—is a sustainable financing mechanism. Indeed, this task is the quintessential collective action problem. Unions provide public goods to workers in the form of, for example, higher wages, safer working conditions, and just-cause employment protections.[55] Because unions are required to offer these terms on an equal basis to all workers in a bargaining unit—and because their positive effects generally extend well beyond workers in the bargaining unit[56]—there is a free-rider problem inherent in the union project: all workers have access to the goods unions provide, and if no worker is required to pay for those goods, many will decide not to do so.[57] Labor law gives unions a mechanism to overcome this free-rider problem in the form of union security agreements and dues checkoff provisions.

Union security agreements are contractual clauses in collective bargaining agreements that require employees to pay dues to the union as a condition of employment.[58] Where such agreements are operative, if an employee fails to pay dues to the union, the employer is obligated to

---

53. *See* Babcock & Wilcox Co., 109 N.L.R.B. 485, 494 (1954) (permitting nonemployee union representatives to distribute union literature on an employer's parking lot).

54. *See* Excelsior Underwear, Inc., 156 N.L.R.B. 1236, 1239–40 (1966) (addressing standards of disclosure in representation elections).

55. *See, e.g.*, JAMES T. BENNETT & BRUCE E. KAUFMAN, WHAT DO UNIONS DO? A TWENTY-YEAR PERSPECTIVE (2007) (compiling scholarship regarding the social and economic effects of unions, including their impact on wages, benefits, workplace policies, job satisfaction, and conflict resolution).

56. This includes the so-called "threat effect" and ways in which unions affect wage-bargaining norms in the nonunion sector. *See* Western & Rosenfeld, *supra* note 6, at 514 (noting that nonunion employers may increase wages to avoid unionization).

57. Indeed, Mancur Olson relied on the union example to describe his logic of collective action. MANCUR OLSON, THE LOGIC OF COLLECTIVE ACTION: PUBLIC GOODS AND THE THEORY OF GROUPS 66–97 (1965).

58. Kenneth G. Dau-Schmidt, *Union Security Agreements Under the National Labor Relations Act: The Statute, the Constitution, and the Court's Opinion in* Beck, 27 HARV. J. ON LEGIS. 51, 57 (1990).

discharge the employee.[59] By making dues payments mandatory—or, at least, a condition of employment—union security agreements help solve the free-rider problem that would otherwise plague unions. Labor law plays a somewhat less prominent role here than in the previous examples, but it is still crucial. Without a labor statute—and assuming unions could get organized in the first place—employers and unions could contractually agree to condition employment on workers' paying union dues. But labor law goes further than permitting the negotiation of such clauses, as background contract principles would. Under the NLRA, employers have a duty to bargain in good faith with unions and this bargaining duty explicitly covers the negotiation of union security agreements.[60]

In addition to union security clauses—which make dues payments mandatory—labor law also facilitates the bargaining of dues checkoff clauses. These clauses are administrative: they allow unions to deduct dues directly from employees' paychecks, automatically and on a recurring basis.[61] But this administrative process can be highly significant by saving unions the substantial burden of manually collecting dues payments each month from thousands, or even hundreds of thousands, of workers. Labor law facilitates this process by making dues checkoff clauses a mandatory subject of bargaining,[62] and by treating an employer's failure to agree to such a clause as evidence of that employer's failure to bargain in good faith.[63]

Taken together, these labor law rights help make unionization a viable prospect for workers. They create the "objective structural potential" for unionization and help ensure the existence of organizations capable of channeling workers' collective efforts.

B. Vulnerability

Structural potential and organizational viability, however, are not enough to generate collective action. A subjective view among participants that their mobilization is likely to succeed is also necessary. Perceptions about the likelihood of success, moreover, depend on two sets of projections: one concerning the potential strength of the group attempting to organize and another concerning the vulnerability of the status quo regime that the group hopes to alter. In our context, workers' perceptions of the likelihood of success of a union campaign combine views about, first, how likely it is that a union can be organized and (if organized) how strong it will be, and second,

---

59. *Id.*

60. This is not to say that employers are legally obligated to agree to such clauses, but they must bargain over them. *See, e.g.*, NLRB v. Gen. Motors Corp., 373 U.S. 734, 737–38, 744–45 (1963).

61. *See* Bethlehem Steel Co., 136 N.L.R.B. 1500, 1502 (1962) (describing a checkoff obligation requiring the employer to deduct union dues from employees' pay).

62. Reed & Prince Mfg. Co., 96 N.L.R.B. 850, 851 (1951).

63. H.K. Porter Co., Inc., 153 N.L.R.B. 1370, 1372 (1965).

how vulnerable management is to a campaign designed to change the status quo.

The legal literature and the literature on law and social movements tends either to merge these two factors or to focus on the first. In his pivotal study of the relationship between law and the mobilization of the pay-equity campaign, for example, Michael McCann summarizes the law and social movement scholarship by quoting Piven and Cloward for the proposition that when participants "'begin to assert their "rights" that imply demands for change,' there develops a new sense of efficacy; people who ordinarily consider themselves helpless come to believe that they have some capacity to alter their lot."[64] McCann also argues that "once it became widely known that the courts weighed in favorably, the prospects for change appeared promising and commitments to political organizing seemed worth the effort."[65] Why? Because judicial victories "signaled to potential activists that they might be able to count on judicial support for the cause."[66] McCann's work thus deftly illustrates the importance of participants' views about likelihood of success—and he highlights the ways law can contribute to these perceptions—but his focus is on the efficacy of the movement, not on the vulnerability of the opposition.

Some scholars of the civil rights movement similarly attribute to legal victories this ability to increase participants' optimism about the prospects for successful organizing and thus to increase the likelihood of successful mobilization. Mark Tushnet, writing about *Brown v. Board of Education*[67] and the Montgomery Bus Boycott, puts it this way:

> The boycott succeeded . . . because it lasted so long. Historians are uncomfortable with counterfactuals. Still, we can wonder whether the participants would have been so persistent . . . had they not known that one of the nation's major governing institutions had endorsed the principle for which they were contending.[68]

Indeed, even Michael Klarman, perhaps the leading skeptic of law's relationship to social movement mobilization, acknowledges that law can inspire participants and change perceptions about the possible: "*Brown* also plainly inspired blacks. To have the Court declare segregation to be unconstitutional was symbolically important, and it furthered the hope and the conviction that fundamental racial change was possible."[69] Again, then,

---

64. MICHAEL W. MCCANN, RIGHTS AT WORK: PAY EQUITY REFORM AND THE POLITICS OF LEGAL MOBILIZATION 89–90 (1994) (quoting FRANCES FOX PIVEN & RICHARD A. CLOWARD, POOR PEOPLE'S MOVEMENTS: WHY THEY SUCCEED, HOW THEY FAIL 4 (1977)).

65. *Id.* at 64.

66. *Id.* at 89.

67. Brown v. Board of Education of Topeka, 347 U.S. 483 (1954).

68. Mark Tushnet, *The Significance of* Brown v. Board of Education, 80 VA. L. REV. 173, 178–79 (1994).

69. KLARMAN, *supra* note 36, at 463.

these scholars recognize the importance of both belief in the possibility for change through collective action and the ability of law to shape such beliefs. But they subsume within this belief participants' views of the vulnerability of the status quo regime.

Like these writers, my previous work has focused on the importance to organizing success of workers' belief in their own efficacy and has missed the other side of the predictive coin. Describing law's ability to generate collective action, I have argued that "successful experiences with small-scale forms of collective activity increase the likelihood that workers will undertake successive, and more difficult forms of collective action."[70] To make this point, I consolidated the work of labor sociologists, all of whom stress the importance of workers' views of their own capacities to organize for change. Rick Fantasia, for example, argues that when workers participate in successful "mini-insurrections" they gain a "courageousness" that can be a "crucial component of union formation."[71] Rachel Meyer writes that union organizing requires organizers to "make it possible for people to first succeed at small collective actions in order that they become aware of their power to make change."[72] And Mark Steven Freyberg, also discussing unionization, concludes:

> [C]ollective efficacy fuels successful action [and] reinforces feelings of efficacy in actors and encourages the spread of such feelings to non-participants. This in turn, encourages further, more widespread and intense collective acts.[73]

Workers' optimism about the prospects for their own organization are critical to the success of their organizational efforts. But a subjective belief in the likely success of collective action depends not only on optimism about the group being organized. It turns equally on the view that the regime in power is susceptible to challenge and change.[74] If workers believe that management, which supports nonunion governance, is invincible,[75] then workers will not attempt to unionize; if workers think management is susceptible to unionization, organizing becomes possible.

---

70. Sachs, *supra* note 44, at 2735.

71. FANTASIA, *supra* note 16, at 121, 137, 145.

72. Sachs, *supra* note 44, at 2736 (quoting Rachel Meyer, The Irony of Power: Collective Action and Efficacy in Working Class Struggle 2 (Aug. 12, 2005) (unpublished manuscript) (on file with author)).

73. Mark Steven Freyberg, Constructing the UAW Dodge Local 3: Collective Identity, Collective Efficacy, Collective Action 224 (1996) (unpublished Ph.D. dissertation, University of Michigan). Brishen Rogers's work is also illuminating. He shows that union-organizing success depends on employees believing that "unionization is a realistic option," a belief which itself turns on workers' experience "build[ing] collective power" during their organizing campaign. Brishen Rogers, *Passion and Reason in Labor Law*, 47 HARV. C.R.-C.L. L. REV. 313, 361 (2012).

74. MCADAM, *supra* note 14, at 49.

75. MAGALONI, *supra* note 40, at 9.

The literature to date has not addressed law's ability to influence perceptions of vulnerability.[76] Yet labor law has significant capacity to shape workers' perceptions of management in just this way. To see this we can start by observing how strenuously managers attempt to convince workers of management's *invulnerability* to union organizing, expressed as invulnerability to unions themselves, as invulnerability to the law that protects unionization, or as invulnerability to both. NLRB opinions thus routinely quote management asserting that unionization simply will never occur:

> "[T]here will never be a union in this shop."[77]
>
> "[The firm will] never be a union shop, never, never, never."[78]
>
> "[N]o son of a bitch [will] bring a union to Wellstream."[79]

Board opinions also contain repeated assertions that management is immune from labor law and thus that management need not, and will not, yield to its mandates:

> "[I don't] care about the law."[80]
>
> "I don't care what you have been promised by the administration, the government, or anybody else. I am running this department and I will do it my way."[81]
>
> "[N]o other law don't mean nothing, this is my law[:] You just do as I say . . . I am the law . . . . I don't care about the court . . . . I don't care about the federal law . . . ."[82]

---

76. In his history of the civil rights movement, Aldon Morris comments—unfortunately only in passing—that NAACP litigation had the effect of rendering the entrenched regime "vulnerable at some points." ALDON D. MORRIS, THE ORIGINS OF THE CIVIL RIGHTS MOVEMENT: BLACK COMMUNITIES ORGANIZING FOR CHANGE 26 (1984).

77. United States v. Electro-Voice, Inc., No. 3:94-CV-1037RM, 1995 WL 353146, at *3 (N.D. Ind. May 4, 1995), *rev'd sub nom.*, NLRB v. Electro-Voice, Inc., 83 F.3d 1559 (7th Cir. 1996).

78. Insta-Print, Inc., 343 N.L.R.B. 368, 374 (2004).

79. Wellstream Corp., 313 N.L.R.B. 698, 706 (1994); *see also* Sun Co., 103 N.L.R.B. 359, 373 (1953) ("I don't care how many of you men join the union. . . . Don't you know that this will never be a union shop?"); Ra-Rich Mfg. Corp., 120 N.L.R.B. 503, 506 (1958) ("There will never be a union in this place as long as I live."); Schoolman Transp. Sys., Inc., 319 N.L.R.B. 701, 703 (1995) (employer states that as long as he was with the firm, "there would never be a union here"); NLRB v. Vanguard Oil & Serv., Inc., No. 75-4222, 1980 WL 18736, at *4 (2d Cir. Sept. 2, 1980) ("Ain't no union going to come in here."), *report and recommendation adopted* (2d Cir. Oct. 10, 1980); Bon Hennings Logging Co., 132 N.L.R.B. 97, 104 (1961) ("There will never be a union man on my job."), *enforcement granted in part and denied in part by* Bon Hennings Logging Co. v. NLRB, 308 F.2d 548, 558 (9th Cir. 1962).

80. Pac. Techs., Inc., No. 13-CA-33518, 1996 WL 33321406, at *2 (N.L.R.B. Div. of Judges Apr. 9, 1996).

81. Acad. of Art Coll., 241 N.L.R.B. 839, 841 (1979).

82. Caroline Gardens Apartment Corp., No. 2-CA-23514, 1990 WL 1222181, at *2 (N.L.R.B. Div. of Judges Jan. 17, 1990) (internal quotations omitted).

> "I have the power here. I am your boss . . . . I do whatever I want . . . ."[83]

And, in a consolidation of these various examples, one general manager made the point this way to an employee he was about to fire:

> "I have news for you. You will never have a union in this shop . . . . This is my shop. I will run this . . . shop my way. I am the law. I am the . . . law in this place and you get that straight."[84]

Statements like these are designed to quash organizational efforts by convincing workers that management is invulnerable to unionization.[85] NLRA rights and remedies can be understood, in contrast, as signals to workers that management's assertions are untrue. In other words, in the "struggle for [workers'] hearts and minds" that shapes every organizing campaign,[86] law can counter the claim that management is impervious to unionization.

To illustrate the point, I start with the remedy of notice reading, a remedy that calls on management to proclaim its vulnerability to labor law in front of an audience of workers. As noted in the Introduction, as a remedy in unfair labor practice proceedings, the NLRB on occasion orders employers to assemble their employees and to read them a notice. The notice states that the NLRB has found the employer in violation of federal law and it lists the actions that the employer must take and not take to remedy its illegal actions.[87] In a 2014 decision, for example, the Board ordered the employer to:

> [H]old a meeting or meetings, scheduled to ensure the widest possible attendance, at which the attached notice . . . is to be read to the

---

83. Ardsley Bus Corp., 357 N.L.R.B. 1009, 1011 (2011) (internal quotations omitted); *see also* The Broker, 282 N.L.R.B. 1265, 1269 n.9 (1987) ("I don't care about the Supreme Court."); Barberton Manor, 252 N.L.R.B. 380, 390 (1980) ("[I don't] give . . . a damn about the law.").

84. Cleveland Pressed Prods. Corp., 203 N.L.R.B. 290, 293 (1973) (internal quotations omitted).

85. *Cf.* MCADAM, *supra* note 14, at 49 (indicating an increased likelihood of unionization when employees perceive management's vulnerability). This is not to imply that conveying invincibility is the only strategy management uses to defeat union organizing. For example, Karen Brodkin and Cynthia Strathmann point out that management may attempt to avoid unionization by building—or at least conveying—a culture of managerial "paternalism" in which management is the head of a "benevolent corporate family." Karen Brodkin & Cynthia Strathmann, *The Struggle for Hearts and Minds: Organization, Ideology, and Emotion*, 29 LAB. STUD. J. 1, 3, 17 (2004).

86. Rogers, *supra* note 73, at 318 (quoting Brodkin & Strathmann, *supra* note 85, at 3).

87. *See, e.g.*, HTH Corp. v. NLRB, 823 F.3d 668, 672–74 (D.C. Cir. 2016); NLRB v. Homer D. Bronson Co., 273 F. App'x 32, 39–40 (2d Cir. 2008); Federated Logistics & Operations v. NLRB, 400 F.3d 920, 929–30 (D.C. Cir. 2005); NLRB v. S.E. Nichols, Inc., 862 F.2d 952, 956, 962 (2d Cir. 1988); United Food & Commercial Workers Int'l Union, AFL-CIO v. NLRB, 852 F.2d 1344, 1348–49 (D.C. Cir. 1988); Conair Corp. v. NLRB, 721 F.2d 1355, 1385–87 (D.C. Cir. 1983); Teamsters Local 115 v. NLRB, 640 F.2d 392, 400–04 (D.C. Cir. 1981); J. P. Stevens & Co. v. NLRB, 417 F.2d 533, 539–40 (5th Cir. 1969); NLRB v. Bush Hog, Inc., 405 F.2d 755, 757–59 (5th Cir. 1968); Textile Workers Union of America, AFL-CIO v. NLRB, 388 F.2d 896, 902–05 (2d Cir. 1967); J. P. Stevens & Co. v. NLRB, 380 F.2d 292, 303–05 (2d Cir. 1967).

> employees in both English and Spanish by the Respondent's chief executive officer or, at the Respondent's option, by a Board agent in that officer's presence.[88]

The notice that the employer was required to read to its employees stated, among other things:

> The National Labor Relations Board has found that we violated Federal labor law and has ordered us to post and obey this notice. FEDERAL LAW GIVES YOU THE RIGHT TO [f]orm, join, or assist a union[;] [c]hoose representatives to bargain with us on your behalf[;] [and] [a]ct together with other employees for your benefit and protection . . . . WE WILL NOT discharge or otherwise discriminate against any of you for supporting [the] union.
>
> . . . .
>
> WE WILL NOT coercively question you about your union support or activities.
>
> WE WILL NOT threaten you with unspecified reprisals because you engaged in union activity.
>
> WE WILL NOT in any like or related manner interfere with, restrain, or coerce you in the exercise of the rights guaranteed you by Section 7 of the Act.[89]

We have no direct way to measure the effect of such readings on employees' perceptions of management, or more particularly, on employees' perceptions of how vulnerable to challenge management is. But there is some helpful indirect evidence of this effect. For one, employers themselves characterize the remedy in stark terms that indicate what employers believe the remedy's impact on employees' views will be. Most tellingly, employers routinely characterize the remedy as *humiliating* and *embarrassing*. In *Teamsters Local 115 v. NLRB*,[90] for example, the employer argued that the notice-reading requirement "demonstrates that the Board is intent on its

88. Marquez Bros. Enters., Inc., 361 N.L.R.B. No. 150, at *2 (Dec. 16, 2014). In *Marquez*, as in most cases, the employer has the option of allowing a Board agent to read the notice to the employees, in the presence of the employer, rather than reading the notice itself. With respect to demonstrating vulnerability, whether the notice is read by the employer or by the board agent in the employer's presence seems unlikely to matter much. For another example, see HTH Corp., 361 N.L.R.B. No. 65 (2014), in which the Board orders the employer to:

> [C]onvene meetings at [the workplace] during working time, scheduled to ensure the widest possible attendance, at which the attached notice and Explanation of Rights are to be read to all employees, supervisors, and managers. The meetings shall be held in the presence of a Board agent, and the notice and Explanation of Rights shall be read by [a senior management official] or, at the [Employer's] option, by the Board agent in the presence of [the senior management official] . . . . At least two supervisors/managers must be present at each reading.

89. *Marquez Bros.*, 361 N.L.R.B. at App.

90. 640 F.2d 392 (D.C. Cir. 1981).

design to humiliate and embarrass [us]."[91] In *Conair Corp. v. NLRB*,[92] the employer similarly argued that a notice reading is "punitive, oppressive, and unwarranted."[93] More colorfully, in *J. P. Stevens & Co., Inc. v. NLRB*,[94] the employer attacked the order by asking of the NLRB, "Upon what meat doth this our Caesar feed?"[95] The Shakespeare quote suggests not only an intensely humiliating effect of the notice-reading requirement but a degrading one.[96] And one need not be a psychoanalyst to know that, if management is humiliated and embarrassed by an NLRB order, this could contribute to a perception among workers that management is not quite as invincible as management wishes to convey.[97]

Many judges called on to review the NLRB's notice-reading orders, moreover, share employers' views about the remedy. Thus, judicial opinions characterize notice reading as "humiliating and degrading,"[98] as "incompatible with the democratic principles of the dignity of man,"[99] and as tantamount to "ignominy."[100] Writing about a notice reading in a case in which a senior manager had responded to a preliminary board order by saying "[f]uck the judge,"[101] the D.C. Circuit describes its reaction to the notice-reading remedy as follows:

> What is the subtext communicated by the sort of scene the Board would mandate? What is communicated to the assembled workers and the perpetrator himself? "You see before you one of your managers, who normally has a responsibility to make important choices as to your work. But who is he? Not merely is he a lawbreaker, but he is a

---

91. *Id.* at 401.

92. 721 F.2d 1355 (D.C. Cir. 1983).

93. *Id.* at 1385.

94. J. P. Stevens & Co., Inc. v. NLRB, 380 F.2d 292 (2d Cir. 1967).

95. *Id.* at 304; *see also* Brief for Petitioner at 48, UNF West, Inc. v. NLRB, 844 F.3d 451 (5th Cir. 2016) (No. 16-60124) (arguing that the notice-reading remedy should "not [be] available to humiliate employers for run of the mill allegations"); Brief of Petitioners at 37–39, HTH Corp. v. NLRB, 823 F.3d 668 (D.C. Cir. 2016) (Nos. 14-1222, 14-1283) (characterizing the notice-reading remedy as a "humiliating and counterproductive experience," an "extreme 'ad hominem remedy,'" and an "affront to [the employer's] dignity"); Brief for Petitioners at 26, Three Sisters Sportswear Co. v. NLRB, 55 F.3d 684 (D.C. Cir. 1995) (Nos. 94-1154, 94-1169) (characterizing the notice-reading remedy as "demeaning").

96. *See, e.g.*, Marvin L. Vawter, *"Julius Caesar": Rupture in the Bond*, 72 J. ENG. & GERMANIC PHILOLOGY 311, 311, 322 (1973) (interpreting phrase to convey "the enormity of Caesar's tyranny" and construing "meat" as a metaphor for the "self-debased bodies" of Caesar's subjects).

97. *See* Jeff Elison & Elizabeth J. Dansie, *Humiliation*, *in* ENCYCLOPEDIA OF ADOLESCENCE at 1342, 1346–47 (Roger J.R. Levesque ed., 2011) (describing devaluation of a humiliated party in the eyes of the relevant audience).

98. Int'l Union of Elec., Radio, & Mach. Workers, AFL-CIO v. NLRB, 383 F.2d 230, 233 (D.C. Cir. 1967).

99. *Id.* at 234.

100. *Id.*

101. HTH Corp., 361 N.L.R.B. No. 65, 2 n.10 (2014).

> pathetic creature who can be forced to spout lines some government officials have put in his mouth. He is not even a parrot, who can choose when to speak; he is a puppet who speaks on command words that he may well abominate. We have successfully turned him into a pathetic semblance of a human being."[102]

For the D.C. Circuit, then, notice reading is likely to have a *profound* effect on workers' perception of management's strength. From a boss who could confidently cast off an NLRB order with a "fuck the judge" comment, the manager becomes a "pathetic creature" and a "semblance of a human being" who lowly NLRB officials can force to "spout lines" about union rights.

We need not accept the circuit court's robust view to agree that a notice reading might impact employees' subjective impressions of management's vulnerability. A notice reading, palpably and viscerally, makes clear to employees that management is not, as management would have it, the only law in the workplace. It helps establish that management is not, in fact, invincible. And if government officials—not especially lofty ones—can overcome managerial resistance to the legal rules of union organizing, perhaps that will contribute to workers' belief that they might overcome managerial opposition to unionization. Borrowing McAdam's more circumspect language, we could say that the order is likely to constitute a "cognitive cue" that management is now "vulnerable to challenge."[103]

Notice reading is a clear example of how labor law remedies can work in this way, but it is not the only NLRA right or remedy with this capacity. Indeed, all those rights and remedies that effect an observable incursion into management's prerogatives can contribute to a perception among workers that management is vulnerable to labor laws and labor unions. Take the right of employees to request the presence of a coworker during a disciplinary interview, the so-called *Weingarten* right. In *NLRB v. J. Weingarten, Inc.*,[104] the Supreme Court upheld the Board's rule that any employee facing an investigatory interview with management has a protected right to request the presence of a union representative at that interview.[105] *Weingarten* arose in the context of an already-unionized workplace, and the question of whether nonunion employees enjoy *Weingarten* rights has been a vigorously contested and oft-litigated one over the last four decades. The Board has flip-flopped on the question,[106] but it is clear that when *Weingarten* rights are

102. HTH Corp. v. NLRB, 823 F.3d 668, 677 (D.C. Cir. 2016).

103. MCADAM, *supra* note 14, at 49.

104. 420 U.S. 251 (1975).

105. *Id.* at 252–53.

106. *Compare* Materials Research Corp., 262 N.L.R.B. 1010 (1982) (extending the *Weingarten* right to unrepresented employees), *with* Sears, Roebuck, and Co., 274 N.L.R.B. 230 (1985) (finding *Weingarten* inapplicable absent a recognized union). *See also* Epilepsy Foundation, 331 N.L.R.B. 676 (2000) (returning "to the rule set forth in *Materials Research*" that "*Weingarten* rights are

available to nonunion employees, they—like notice readings—can help demonstrate employer vulnerability. Investigatory interviews, after all, are a site for the direct exercise of managerial power: employees are facing discipline or discharge and enjoy few to no due process rights. Such interviews can thus be experienced by workers as expressions of near total managerial authority. As the Board has put it, an employee in an investigatory interview is "alone in the locus of managerial authority."[107]

For just this reason, however, the right to insist upon a coworker's presence at an investigatory interview can have important effects on the workers involved. Professor Finkin, in his article on *Weingarten* rights, comments on the "substantial psychological benefit" that flows from having a coworker present at such an interview.[108] The *Weingarten* opinion itself speaks to the right's effect on worker perceptions when it holds that the presence of a coworker can "redress the perceived imbalance of economic power between labor and management."[109] And although neither Finkin nor the *Weingarten* Court had workers' perceptions of managerial vulnerability in mind, such perceptions are likely one of the psychological implications of the rule's legal mandate. In a locus of otherwise unchecked managerial authority, the law enables workers to experience managerial authority as checked.

Or take the right of workers to discuss unionization on company property, even in circumstances when the employer seeks to prohibit such discussions. Since the 1940s, the Board has maintained a rule that employees have a presumptive right to talk union at work, as long as they do so during nonwork time.[110] Employers routinely and vigorously object to such discussions among their workforces, asserting property rights to control what employees do and what they talk about while at work. Again, the NLRB reports are filled with examples:

> I own this company, I own this building, I own this land . . . . And if I have to . . . I'll block the driveway with my jeep and my shotgun and

applicable in the nonunionized workplace"); IBM Corp., 341 N.L.R.B. 1288 (2004) (overruling *Epilepsy Foundation* and holding that the *Weingarten* right does not apply to a nonunionized workplace).

107. NLRB v. Ill. Bell Tel. Co., 674 F.2d 618, 621 (7th Cir. 1982); *cf.* Matthew W. Finkin, *Labor Law by Boz – A Theory of* Meyers Industries, Inc., Sears, Roebuck and Co., *and* Bird Engineering, 71 IOWA L. REV. 155, 186–87 (1985) (highlighting that investigatory interviews often resemble criminal interrogations and require management to "remain in firm command of the interview at all times").

108. Finkin, *supra* note 107, at 187.

109. 420 U.S. at 262 (quoting Am. Ship Bldg. Co. v. NLRB, 380 U.S. 300, 316 (1965)).

110. Peyton Packing Co., 49 N.L.R.B. 828, 843 (1943), *cited in* Republic Aviation Corp. v. NLRB, 324 U.S. 793, 803–04 n.10 (1945).

> I hope that the ringleader [of the union organizing] is in the front lines so I can blow his head off.[111]

Against such assertions of employer control, the right to discuss unionization at work provides employees with another experience of managerial vulnerability to labor law. This may be particularly true in the many cases where employers attempt to enforce an illegal ban on union speech and are then required by the Board to allow such discussions.[112] Employees who engage in conversations about unions at work do so despite employer opposition and, thus, only because of their employer's susceptibility to labor law's dictates.

Finally, take the remedy of reinstatement, which is perhaps the most salient way the NLRB can demonstrate an employer's susceptibility to law. Employers facing union campaigns may discharge up to 20% of active union supporters.[113] The message such discharges send is unmistakable, and the impact of discharge on union success rates is profound.[114] Precisely because discharge is such a powerful managerial tool, however, an order of reinstatement can have a powerful countervailing impact on employee views of managerial control: reinstatement is a demonstration that even the employer's ultimate form of authority is constrained by the obligations of law.[115] And, here, we have a helpful bit of empirical evidence. In her study of union campaigns, Kate Bronfenbrenner identified a set of cases in which union activists were discharged during an organizing drive but then reinstated by NLRB order prior to the time the union election was held.[116] Bronfenbrenner shows that union success rates were higher in these cases than in cases where no union activist was discharged at all. In her sample, in cases where no union activist was discharged, the union won 45% of the time.[117] On the other hand, where a union activist was discharged but then reinstated before the union election, the win rate was 58%, or 13% *higher*

---

111. Hagerstown Kitchens, Inc., 244 N.L.R.B. 1037, 1038 (1979); *see also* Care Manor of Farmington, Inc., 314 N.L.R.B. 248, 250 (1994) ("You can organize on your own time from your own home, but I will not have it on my property."); A & T Mfg. Co., 265 N.L.R.B. 1560, 1570–72 (1982) ("It's your right to promote a union . . . [but] you can't do it on my property, you can't do it on my time, you can't do it on my job . . . ."), *enforced in part, vacated in part*, 738 F.2d 148 (6th Cir. 1984).

112. *See, e.g.*, Loparex LLC v. NLRB, 591 F.3d 540, 545, 552 (7th Cir. 2009); Stanford Hosp. & Clinics v. NLRB, 325 F.3d 334, 338–42, 346 (D.C. Cir. 2003); United Parcel Serv., Inc. v. NLRB, 228 F.3d 772, 775–78, 780–82 (6th Cir. 2000).

113. *See* JOHN SCHMITT & BEN ZIPPERER, CTR. FOR ECON. & POLICY RESEARCH, DROPPING THE AX: ILLEGAL FIRINGS DURING UNION ELECTION CAMPAIGNS 11 (2007).

114. *See supra* note 47.

115. *See* 29 U.S.C. § 160 (2012) (empowering the NLRB to prevent unfair labor practices).

116. KATE BRONFENBRENNER, UNEASY TERRAIN: THE IMPACT OF CAPITAL MOBILITY ON WORKERS, WAGES, AND UNION ORGANIZING 48–49 (2000), http://digitalcommons.ilr.cornell.edu/reports/3/ [https://perma.cc/YVZ2-9JGV].

117. *Id.* at tbls. 8–9.

than in campaigns where there had been no discharge.[118] Here, then, data suggestive of the dynamic at the heart of this Essay: where the NLRB intervenes in a manner that makes management visibly susceptible to the power of law, union organizing is more likely to succeed.

III. Implications

This Essay claims that employee perceptions of employer vulnerability are an important factor in the generation of employee collective action. Such perceptions are, in fact, a necessary precursor to such collective action. The Essay also endeavors to show that the enforcement of labor law, at least when such enforcement is directly experienced by workers, can contribute to these perceptions of managerial vulnerability.

To be clear, the Essay does not claim—as it would be foolish to do—that employee perceptions of vulnerability are sufficient for the generation of collective action. Far more than this belief is required to enable unionization.[119] Nor does the Essay claim that labor law is the sole means by which employer vulnerability can be demonstrated. To the contrary, much of labor history is defined by workers and unions that have established the employer's vulnerability through their own acts, often entirely unassisted by—or, more accurately, hindered by—the legal regime.

But if the claims in the Essay are correct—that perceptions of vulnerability matter and that law can contribute to them—then there are implications for labor law. Namely, labor law ought to be structured so that employees do in fact experience management as subject to the rules of union organizing. Moreover, such experiences should be available to all employees and not just those who happen to be parties to an NLRB proceeding and a Board-ordered remedy.

One way to start would be for the Board to require all employers to read to its employees a notice of union-organizing rights. Such readings could be required, for example, on an annual basis and irrespective of the presence of ongoing organizing activity. Indeed, requiring such readings before organizing begins makes conceptual sense, given that organizing might not begin unless and until employees believe the employer is vulnerable to a campaign. More modestly, a notice reading could be required as part of the Board's representation-election procedures. Thus, when a union demonstrates to the Board that it has support from 30% of a relevant

---

118. *See id.* The data do not allow us to rule out another causal route: when an employer perceives that a union drive is likely to succeed, it fires a union activist in a manner that generates the relatively rare NLRB reinstatement order. In this case, the data reflect employer perceptions of the union's chances of success, and not necessarily any impact on employee perceptions.

119. For recent discussions of the range of factors necessary to generate successful unionization efforts, see, e.g., Rogers, *supra* note 73, at 361; Sachs, *supra* note 12, at 680–84 (discussing the collective action problem and forms of managerial intervention that deter unionization efforts).

bargaining unit and invokes the Board's election machinery,[120] the Board could—in addition to requiring that the employer disclose names and addresses of bargaining-unit members—require the employer to read employees a notice of their NLRA rights.[121]

Whether such a notice-reading requirement could be enforced by the Board absent legislative amendment is an open question. The Board recently adopted a rule requiring all covered employers to post a notice detailing employees' rights under the NLRA.[122] Implementation of the rule, however, has been enjoined by the Fourth Circuit on the ground that the Board lacks statutory authority to go beyond its "reactive" role of "addressing unfair labor practice charges and conducting representation elections upon request."[123] In addition, the D.C. Circuit found that enforcement of the rule ran afoul of § 158(c) of the NLRA,[124] which prohibits the Board from finding an unfair labor practice based on the dissemination of noncoercive speech.[125] If the Fourth Circuit and D.C. Circuit opinions remain the governing standards, the Board likely could not—absent congressional action—order employers to engage in regular notice readings outside the context of unfair labor practice proceedings. On the other hand, the Board could perhaps adopt the more modest proposal sketched above and make such a reading part of the representation-election procedure, at least under the Fourth Circuit's view of the law.[126]

---

120. 29 U.S.C. § 159(e)(1) (2012).

121. John Teeter recommends notice reading as a remedy for violations of the Act, a proposal that is fully defensible on the grounds articulated here but that does not go far enough to communicate vulnerability to employees engaged in, or considering, organizing. *See* Teeter, *supra* note 26, at 2 (arguing that "employers should always be required to read notices aloud to their workers as a standard remedy for violations of the Act").

122. 29 C.F.R. § 104.202(a) (2012), *invalidated by* Chamber of Commerce of the U.S. v. NLRB, 721 F.3d 152, 154 (4th Cir. 2013).

123. Chamber of Commerce of the U.S. v. NLRB, 721 F.3d 152, 154 (4th Cir. 2013).

124. Nat'l Ass'n of Mfrs. v. NLRB, 717 F.3d 947, 959 (D.C. Cir. 2013), *overruled in part by* Am. Meat Inst. v. U.S. Dep't of Agric., 760 F.3d 18, 22–23 (D.C. Cir. 2014).

125. *See* 29 U.S.C. § 158(c) (2012) (barring evidence of unfair labor practices based on expression of views absent a "threat of reprisal or force or promise or benefit").

126. The D.C. Circuit's opinion in *National Association of Manufacturers* was overruled in part by *American Meat Institute v. U.S. Department of Agriculture*, 760 F.3d 18, 22–23 (D.C. Cir. 2014) (en banc). In any event, the fact that notice reading involves reading rather than posting could generate compelled-speech challenges under the First Amendment, in addition to any statutory limitations on the Board's authority to enforce such a rule. The constitutional validity of the notice-posting rule is relatively settled, at least to the extent that the requirement is limited to factual information—as would be the requirement proposed here. *See Am. Meat Inst.*, 760 F.3d at 21 (quoting Zauderer v. Office of Disciplinary Counsel of the Supreme Court of Ohio, 471 U.S. 626, 651 (1985)) (describing the mandated disclosure as "limited to 'purely factual and uncontroversial information'"). Whether the verbal requirement would make the reading more susceptible to compelled-speech attack is beyond the scope of this Essay. For a discussion, see Robert Post, *Compelled Commercial Speech*, 117 W. VA. L. REV. 867, 869 (2015) (exploring "the burgeoning doctrine of compelled commercial speech").

In addition to highlighting the desirability of a global notice-reading requirement, the Essay's argument about vulnerability points to new ways to think about—and new ways to defend—other contested rights. Two examples should be obvious from the above discussion. One, the extension of *Weingarten* rights to the nonunion workplace can be defended as a means of enabling nonunion workers to experience the employer's susceptibility to labor law.[127] Two, the rapid reinstatement of workers discharged during organizing campaigns—through increased use of the Board's authority to seek preliminary injunctive relief in cases of retaliatory discharge—can be defended on the same ground.[128]

More broadly, recognizing the importance of vulnerability to the prospects for unionization suggests another reason why workplace access rights for nonemployee union organizers are so critical. This issue arises in at least two ways. First, union organizers often wish to speak with employees on company property. Doing so is frequently the best way—and at times the only way—for organizers to overcome the considerable coordination costs that inhere in trying to reach employees in their homes, or elsewhere outside of work. While the Supreme Court has recognized that "[t]he right of self-organization depends in some measure on the ability of employees to learn the advantages of self-organization from others,"[129] the Court's current jurisprudence forecloses—for all practical purposes—any claim that nonemployee organizers have a right to enter employer property to convey those advantages to employees.[130] The Court's rule, announced in *Lechmere, Inc. v. NLRB*,[131] has been roundly and persuasively critiqued, most

127. For other arguments supportive of *Weingarten* rights in the nonunion workplace, see Richard Michael Fischl, *Self, Others, and Section 7: Mutualism and Protected Protest Activities Under the National Labor Relations Act*, 89 COLUM. L. REV. 789, 814–15, 815 n.91 (1989); Charles J. Morris, *NLRB Protection in the Nonunion Workplace: A Glimpse at a General Theory of Section 7 Conduct*, 137 U. PA. L. REV. 1673, 1749–50 (1989).

128. The Board's authority flows from § 10(j) of the Act. *See* 29 U.S.C. § 160(j) (2012) (describing the Board's power to petition any United States district court upon issuance of a complaint of an unfair labor practice); *see also* Sachs, *supra* note 44, at 2695, 2695–96 n.35 (mentioning the Board's power under § 10(j) to seek injunctive relief in cases of retaliatory discharge). Other defenses of rapid reinstatement can be found in Charles J. Morris, *A Tale of Two Statutes: Discrimination for Union Activity Under the NLRA and RLA*, 2 EMP. RTS. & EMP. POL'Y J. 317, 357–58 (1998); PAUL C. WEILER, GOVERNING THE WORKPLACE: THE FUTURE OF LABOR AND EMPLOYMENT LAW 243–44 (1990). Of course, even 10(j) injunctions can take time, and on the account offered here, ought to be accelerated to best serve their purpose. *See, e.g.*, NLRB General Counsel Memorandum GC 17-02, Report on the Midwinter Meeting of the ABA (Mar. 10, 2017), https://www.nlrb.gov/reports-guidance/general-counsel-memos [https://perma.cc/U6JG-26S3] (providing statistics related to the timing of § 10(j) injunctions).

129. NLRB v. Babcock & Wilcox Co., 351 U.S. 105, 113 (1956).

130. *See, e.g.*, Cynthia L. Estlund, *Labor, Property, and Sovereignty after* Lechmere, 46 STAN. L. REV. 305, 319 (1994) (discussing the *Lechmere* holding that nonemployee organizers "virtually never have the right to enter private property to communicate with unorganized employees").

131. 502 U.S. 527 (1992).

prominently by Cynthia Estlund.[132] But understanding the importance of vulnerability to law gives us another way of thinking about the problem with *Lechmere*.

As vigorous as employers are when it comes to attempts to prohibit their own employees from talking union at work, they are even more adamant when it comes to the ability of full-time union organizers to come onto employer property to discuss unionization. One employer put it this way: "If [a union organizer] comes onto my property, I'll fill his butt with lead."[133] This opposition stems in part from the fact that physical presence of union organizers on company property marks a major incursion into the employer's control over the workplace. Excluding union organizers thus reinforces a perception of employer control. By the same token, a legal requirement that employers admit union organizers conveys to workers management's susceptibility to the law of union organizing.[134] The union organizer's physical presence on company property is a direct instantiation of law's power to compel employers to respect the union-organizing right.

The second context in which access rights for nonemployee organizers is critical is the union's right of reply to employer captive audience meetings. In organizing campaigns, employers nearly always require employees to attend meetings during which managers express their opposition to unionization.[135] Labor law protects management's right to hold such meetings and to make them mandatory.[136] At one time, the NLRB offered unions a right of reply: if management held a captive audience meeting, the union was entitled to come onto company property and present its side of the story.[137] But this right was extinguished soon after it was established, and since 1953 unions have enjoyed no right to enter employer property in order to respond to management's anti-union captive audience meetings.[138] Again,

---

132. *See* Estlund, *supra* note 130, at 308 (denying that a "naked property right" should "trump the substantial statutory interests of organized employees").

133. Mid-State, Inc., 331 N.L.R.B. 1372 (2000).

134. A cinematic example of this dynamic is available from *Norma Rae*. NORMA RAE (Twentieth Century Fox Film Corp. 1979). In an early sequence, the union organizer (Reuben Warshawsky) is excluded from the factory by a locked fence: the image is unmistakably one of employer power and control. Later in the film, however, the labor board orders the employer to allow Warshawsky to inspect a company bulletin board. The scene of Warshawsky walking through the factory—through what has previously been a site of unquestioned managerial authority—and, more particularly, the workers' reaction to Warshawky's presence, unmistakably conveys the workers' emerging sense that the employer is not invincible.

135. According to Bronfenbrenner's work, captive audience meetings are held in more than 90% of union campaigns. BRONFENBRENNER, UNEASY TERRAIN, *supra* note 116, at tbl.8.

136. So long as they are held more than twenty-four hours before workers are scheduled to vote on unionization. Peerless Plywood Co., 107 N.L.R.B. 427, 429 (1953).

137. *See* Bonwit Teller, Inc., 96 N.L.R.B. 608, 612 (1951) (holding, at the time, that unions have a right to present their case to employees under the same circumstances as employers).

138. *See* Livingston Shirt Corp., 107 N.L.R.B. 400, 408–09 (1953) (ruling that an employer may lawfully deny a union's request to rebut speech on company premises).

the doctrine has been persuasively critiqued,[139] but the vulnerability thesis gives us a new way of doing so. Allowing unions to come into the workplace and speak to an assembled group of workers, and to counter management's negative message about unionization, would be a powerful message about the employer's susceptibility to the law of union organizing.

Finally, and most generally, the importance of vulnerability gives us another way to understand—and object to—the profound weaknesses in the NLRA's enforcement regime. As many have observed, the NLRA suffers from major enforcement problems.[140] The agency charged with enforcing the statute's substantive mandates is often slow and weak.[141] And even when the Board endeavors aggressively to enforce the law, it is restrained by the remedial arsenal available to it: among other stark examples, the NLRB lacks the authority to order any type of punitive damages.[142] This often means that it is economically rational for employers to violate the labor laws and suffer the rather paltry damage awards the Board can order; it means that the available remedies "simply are not effective deterrents to employers."[143] From the employees' perspective, remedial weakness contributes to a perception of employer *in*vulnerability. Accordingly, improving enforcement would—among other things—help establish for employees the type of employer vulnerability highlighted here.

IV. Conclusion

For students of social movements, law's contribution to collective action has long been a subject of intense interest. This Essay shows that law can perform such a role by shaping perceptions about the vulnerability of the status quo. In particular, the Essay shows that labor law can contribute to workers' perceptions that the existing regime of managerial control is susceptible to challenge and thereby to facilitate union organizing. In doing so, the Essay offers a new way to both understand and to defend a host of NLRA rights and remedies.

If the thesis here is correct, however, its implications extend beyond labor law and labor unions. Indeed, because all social movements depend on participants viewing the relevant status quo regime as subject to change, law's ability to shape perceptions of vulnerability may be relevant across social movement contexts. One obvious place to look is the civil rights movement, and in particular, to the role played by civil rights litigation in the

---

139. Paul M. Secunda, *Toward the Viability of State-Based Legislation to Address Workplace Captive Audience Meetings in the United States*, 29 COMP. LAB. L. & POL'Y J. 209, 215 (2008).

140. *See, e.g.*, Sachs, *supra* note 44, at 2694–700.

141. *See, e.g.*, *id.* at 2696.

142. Cynthia L. Estlund, *The Ossification of American Labor Law*, 102 COLUM. L. REV. 1527, 1552 (2002).

143. Paul Weiler, *Promises to Keep: Securing Workers' Rights to Self-Organization Under the NLRA*, 96 HARV. L. REV. 1769, 1788–89 (1983).

generation of that movement. For example, perhaps part of what cases like *Brown v. Board of Education* did was to convey that the structure of Jim Crow was becoming less invincible. Although he makes the point only in passing, Aldon Morris suggests as much in his history of the civil rights movement:

> The endless court battles and agitation of the NAACP kept pressure on the Southern white power structures to abolish racial domination. It would be misleading to present the courtroom battles in a narrowly legal light. Their importance was more in demonstrating to Southern blacks and the NAACP that the Southern white power structure was vulnerable at some points.[144]

In broad terms, when law demonstrates to challengers that those in power are subject to an authority greater than themselves, law has the potential to convey the vulnerability of the regime and thereby to open space for mobilization.

144. ALDON D. MORRIS, THE ORIGINS OF THE CIVIL RIGHTS MOVEMENT: BLACK COMMUNITIES ORGANIZING FOR CHANGE 26 (1984).

* * *

A Shortcut to Death: How the Texas Death-Penalty Statute Engages the Jury's Cognitive Heuristics in Favor of Death*

Introduction

It is no secret that the leveling of death sentences and the administration of the death penalty in the United States has rapidly declined in recent years.[1] In fact, the number of United States jurisdictions imposing a death sentence has declined by 55.7% in the last four years.[2] Notably, however, the state of Texas still leads the nation in executions,[3] which necessarily require the issuance of a death sentence. Texas has executed 543 inmates since 1976; for comparison, the nation's second-leading state, Virginia, has executed 113 inmates in that time period.[4] But even Texas's propensity to issue death sentences has dramatically declined.[5] In 2016, Texas issued only four death sentences,[6] compared with its apex of forty-eight death sentences in 1999.[7]

This decline in death sentences and resultant executions, however, is no reason to ignore the administration of the death penalty altogether. To the extent that the Texas death-penalty statute has increased the number of death sentences (and this Note argues that it has), the statute is partly responsible for the inmates *currently* on death row. And even in an era of dwindling death sentences, Texas in particular continues to lead the nation, issuing the second-most death sentences in 2016,[8] second only to California, a state long

---

* I am grateful to Professor Jordan Steiker for his helpful comments and guidance. Thanks also to Charles Fowler and Jonathan Jackson for insightful conversations regarding this Note—and to the entire *Texas Law Review* staff, especially Andrew Van Osselaer, for their editing prowess. All remaining errors are mine alone.

1. *See Death Sentences in 2016*, DEATH PENALTY INFO. CTR. (2017), https://deathpenaltyinfo.org/2016-sentencing [http://perma.cc/8LJN-WV4Z] (describing the significant decline in death sentences in the United States); *The Death Penalty in 2016*, DEATH PENALTY INFO. CTR. (2017), https://deathpenaltyinfo.org/YearEnd2016#graphic [http://perma.cc/NW84-EHZR] (illustrating the recent notable decline in executions in the United States).

2. *The Death Penalty in 2016*, *supra* note 1.

3. *Facts About the Death Penalty*, DEATH PENALTY INFO. CTR. (Apr. 21, 2017), https://deathpenaltyinfo.org/documents/FactSheet.pdf [http://perma.cc/HKJ3-252G].

4. *Id.*

5. *Texas Death Penalty Fact Sheet*, TEX. COALITION TO ABOLISH DEATH PENALTY (Jan. 1, 2011), http://www.tcadp.org/wp-content/uploads/2011/01/TXDPFactSheet01-11.pdf [https://perma.cc/C6LL-QMX3].

6. *Death Sentences in 2016*, *supra* note 1.

7. *Texas Death Penalty Fact Sheet*, *supra* note 5.

8. *Death Sentences in 2016*, *supra* note 1.

hailed as a "symbolic death-penalty state."[9] (California has not executed an inmate in over a decade.)[10] Many factors have contributed to Texas's long-standing propensity to issue and carry out death sentences: zealous prosecutors,[11] right-leaning politics,[12] and a tough stance on crime.[13] But Texas's disproportionate use of the death penalty may have an additional, unsung culprit—the Texas death-penalty statute itself. Enacted in a "hurried and somewhat confused process,"[14] the statute centers the jury's deliberations on considerations of the defendant's future dangerousness—with a nod to mitigation.[15]

This Note argues that the Texas death-penalty statute skews the sentencing decision toward death. In particular, the structure of Texas's statute encourages juries to return a death verdict by engaging two cognitive heuristics: the representativeness heuristic and the anchoring-and-adjustment heuristic. Part I begins with a brief background of the Texas death-penalty statute and the corresponding constitutional jurisprudence. Part II summarizes various empirical findings regarding death-penalty juries, namely, their perception that death is required, their inaccurate predictions of future dangerousness, their disregard for mitigating evidence, and their moral distance from the life-or-death decision. Part III offers an explanation for these empirical findings, arguing that the Texas statute draws on the representativeness and anchoring-and-adjustment heuristics to bias the jury in favor of death. Finally, this Note concludes by asserting that Texas should abandon its haphazardly adopted statute. Texas should reverse the anchoring effect of the statute—tying the anchor to mitigation instead of future dangerousness.

---

9. *California Could Finally Resume Executions Next Year*, L.A. TIMES (Apr. 23, 2017), http://www.latimes.com/local/lanow/la-me-california-executions-20170423-story.html [https://perma.cc/F7DJ-FA7A].

10. *Id.*

11. *America's Top Five Deadliest Prosecutors: How Overzealous Personalities Drive the Death Penalty*, FAIR PUNISHMENT PROJECT 3 (June 2016), http://fairpunishment.org/wp-content/uploads/2016/06/FPP-Top5Report_FINAL.pdf [https://perma.cc/DY3B-76WA] (featuring Harris County prosecutor Johnny Holmes, who oversaw the imposition of death sentences against 201 people during his term).

12. *See* Ross Ramsey, *Analysis: The Blue Dots in Texas' Red Political Sea*, TEX. TRIB. (Nov. 11, 2016), https://www.texastribune.org/2016/11/11/analysis-blue-dots-texas-red-political-sea/ [https://perma.cc/883N-DKQE] (describing the Republican Party's winning streak in state and federal elections in Texas since 1994).

13. Christian McPhate, *10 Notorious Unsolved Texas Murders*, DALL. OBSERVER (May 17, 2016), http://www.dallasobserver.com/content/printView/8305930n [https://perma.cc/8EQT-N27N].

14. Eric Citron, Note, *Sudden Death: The Legislative History of Future Dangerousness and the Texas Death Penalty*, 25 YALE L. & POL'Y REV. 143, 163 (2006).

15. *See* TEX. CODE CRIM. PROC. ANN. art. 37.071 (West Supp. 2016) (calling the jury to consider the "probability that the defendant would commit criminal acts of violence that would constitute a continuing threat to society").

I. Background of the Texas Death-Penalty Statute

In the modern death-penalty era—that is, the years following the Supreme Court's landmark decision in *Furman*[16]—the Texas death-penalty statute has centered upon a series of Special Issues.[17] The pre-1991 Texas statute included two Special Issues:[18] a question of the defendant's future dangerousness and a question of the defendant's deliberate commission of the killing.[19] The pre-1991 statute—which largely remains the framework for the current Texas capital-sentencing statute—was passed in a "hurried and somewhat confused process," spanning mere weeks.[20]

One scholar "summarized" the "legislative history of the future dangerousness standard" as follows:

> Beginning in January 1973, committees and subcommittees began hearing testimony and thinking about a new death penalty in Texas. On May 10th, the House gave its best interpretation of *Furman* and passed a mandatory death penalty bill. Two weeks later, the Senate debated between that mandatory bill and a more discretionary approach, finally opting for the latter. With only Memorial Day weekend to go before adjournment, the House called a conference committee to resolve the differences between the two bills. On the very last day, the conferees presented a scheme which appeared in neither the House nor the Senate bill, along with newly minted language about "a probability" that the defendant would be a "continuing threat." That same day, both houses passed the committee report by huge margins without specifically considering the new language on future dangerousness.[21]

And seemingly overnight, the Texas death-sentencing statute came to fruition. In a "hurried and somewhat confused" process, the Texas Legislature centered the question of a defendant's life or death around future dangerousness—a standard that seems plucked out of thin air—for the indefinite future. The Texas statute now revolves around a precarious prediction of the future, rather than the defendant's moral blameworthiness.

The Supreme Court upheld the constitutionality of the Texas statute (and the corresponding Special Issues) against a facial challenge in *Jurek v.*

---

16. Furman v. Georgia, 408 U.S. 238 (1972) (holding that carrying out the death penalty in the cases before the Court would violate the Eighth Amendment to the Constitution).

17. TEX. CODE CRIM. PROC. ANN. art. 37.071.

18. The Texas statute also included a third Special Issue, regarding provocation, but that instruction is only given in select cases. JON SORENSON & ROCKY LEANN PILGRIM, LETHAL INJECTION: CAPITAL PUNISHMENT IN TEXAS DURING THE MODERN ERA 53 (2006).

19. TEX. CODE CRIM. PROC. ANN. art. 37.071 (West Supp. 1973), *amended by* TEX. CODE CRIM. PROC. ANN. art. 37.071 (2016).

20. Citron, *supra* note 14, at 162–63.

21. *Id.* at 162–63.

Texas.[22] The Court subsequently decided several cases that elaborated on the essential role of mitigation evidence in capital-sentencing proceedings.[23] Unsurprisingly, the effect of the Texas statute's lack of reference to mitigating evidence was well-documented.[24] In *Penry v. Lynaugh*,[25] the Court declared the Texas Special Issues constitutionally inadequate where they failed to offer jurors an opportunity to express a "reasoned moral response" to mitigating evidence.[26]

In response, the Texas Legislature abandoned the Special Issue on deliberateness—and replaced it with a Special Issue regarding mitigation.[27] While the Texas statute's new Special Issue on mitigation was a step in the right direction, the amendment did not eliminate serious problems embedded in the Texas statute. The centerpiece of the statute remains the future-dangerousness inquiry.[28] The original focus on future dangerousness created a sort of path dependence—and this focus remains problematic for many of the reasons discussed below.[29] Moreover, the statute still lacks guidance on the specific mitigating evidence that the jury should consider, "a critical failure of the statute."[30] Regrettably, many of the empirical problems inherent in the pre-1991 statute persist in the amended Texas regime.[31]

II. Empirical Findings Regarding Death-Penalty Juries

Despite the "black box" of the jury room, Texas death-penalty juries—to the extent possible—have been studied at great length, both before and

---

22. 428 U.S. 262, 276 (1976).

23. *See, e.g.*, Eddings v. Oklahoma, 455 U.S. 104, 105–06, 109 (1982) (reviewing an Oklahoma death sentence and the trial judge's refusal to consider certain mitigating circumstances); Lockett v. Ohio, 438 U.S. 586, 589 (1978) (reviewing the constitutionality of the range of mitigating circumstances that may be considered under Ohio's death-penalty statute).

24. *See, e.g.*, Ursula Bentele & William J. Bowers, *How Jurors Decide on Death: Guilt Is Overwhelming; Aggravation Requires Death; and Mitigation Is No Excuse*, 66 BROOK. L. REV. 1011, 1032 (2001) (suggesting that, due to the statute's lack of reference to mitigating evidence, jurors perceived that death was the mandatory sentence under certain circumstances).

25. 492 U.S. 302 (1989), *abrogated by* Atkins v. Virginia 536 U.S. 304 (2002). The *Penry* court sanctioned the death penalty for intellectually disabled offenders but nonetheless held the Texas death penalty constitutionally inadequate because the statute did not allow a full "reasoned moral response" to mitigation. *Id.* at 328. But the *Atkins* court later held that executions of the intellectually disabled constituted impermissible cruel and unusual punishment. 536 U.S. at 321.

26. *Penry*, 492 U.S. at 328.

27. Elizabeth S. Vartkessian, *Dangerously Biased: How the Texas Capital Sentencing Statute Encourages Jurors to Be Unreceptive to Mitigation Evidence*, 29 QUINNIPIAC L. REV. 237, 246 (2011).

28. *Id.* at 247; *see also* Citron, *supra* note 14, at 155–56 (lamenting the centrality of future dangerousness in the Texas statute).

29. *See infra* Part II.

30. Vartkessian, *supra* note 27, at 247.

31. Some findings discussed in Part II, *infra*, are based on interviews conducted before the Texas statute was amended. But, given the focus of the Texas statute on future dangerousness and the lack of clarity regarding mitigation, these problems persist in the amended version.

after the 1991 amendment. Many of the empirical findings come from the Capital Jury Project, which has conducted personal interviews with 1,198 jurors from 353 capital trials in 14 states.[32] These findings, combined with other studies, have established a number of conclusions about jury beliefs and behavior. Several patterns emerge: many jurors perceive that death is required,[33] they often make inaccurate predictions of defendants' future dangerousness,[34] they often give limited consideration to mitigating factors,[35] and they often perceive moral distance between themselves and the life-or-death decision.[36]

A. Inaccurate Predictions of Future Dangerousness

First, jurors often predict the defendant's future dangerousness with little-to-no accuracy. This notion is hardly surprising, given that experts in the field of psychology are themselves ill-equipped to make an accurate future-dangerousness prediction. For example, in a study of 155 capital cases in which expert witnesses predicted that the defendant would be a future danger to society, the witnesses were wrong 95% of the time.[37] Only 8 of the 155 inmates later engaged in seriously assaultive behavior.[38] Moreover, the American Psychiatric Association has asserted that the "unreliability of psychiatric predictions of long-term future dangerousness is by now an established fact within the profession."[39] Given the professional community's difficulty with the future-dangerousness question, it is no surprise that juries also struggle to make an accurate determination.[40]

Nevertheless, expert testimony regarding future dangerousness in capital-sentencing proceedings is ubiquitous, both in the form of clinical and actuarial predictions.[41] Juries' reliance on expert testimony is well-

32. *What Is the Capital Jury Project?*, U. ALB.: SCH. CRIM. JUST., http://www.albany.edu/scj/13189.php [https://perma.cc/Y2PQ-E3MP]. For an overview of the Capital Jury Project's methodology, see William J. Bowers, *The Capital Jury Project: Rationale, Design, and Preview of Early Findings*, 70 IND. L.J. 1043, 1044 (1995).

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

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

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

36. *See infra* subpart II(D).

37. *Future Dangerousness Predictions Wrong 95% of the Time: New Study on Capital Trials Exposes Widespread Unreliable Testimony*, DEATH PENALTY INFO. CTR. (2004), https://deathpenaltyinfo.org/node/1099 [https://perma.cc/93WM-GBEE].

38. *Id.*

39. ROGER J.R. LEVESQUE, THE PSYCHOLOGY AND LAW OF CRIMINAL JUSTICE PROCESSES 500 (2006).

40. *See* Brian Sites, *The Danger of Future Dangerousness in Death Penalty Use*, 34 FLA. ST. U. L. REV. 959, 970 (2007) ("[J]urors spend more time discussing future dangerousness than any other factor save the facts of the crime.").

41. Erica Beecher-Monas, *The Epistemology of Prediction: Future Dangerousness Testimony and Intellectual Due Process*, 60 WASH. & LEE L. REV. 353, 362 (2003).

documented[42]—and hardly surprising.[43] But in the context of future dangerousness, perhaps this reliance is misplaced. Clinical and actuarial expert witnesses often ignore sample sizes and base rates.[44] Moreover, ignoring base rates is "a particular problem in predicting [future] violence when the base rate of violent behavior is low overall and varies among different population subgroups."[45] Regardless of the appropriate calculations, however, focus on the accuracy of predicting future dangerousness distracts jurors from the grave, life-or-death decision at hand.

Furthermore, the Texas statute itself gives relatively little guidance on what constitutes a sufficient finding of future dangerousness. The Texas statute simply asks the jury to evaluate "whether there is a probability that the defendant would commit criminal acts of violence that would constitute a continuing threat to society."[46] The statute does not define what probability is required: is it 10%, 51%, or 90%? Nor does the statute define dangerousness: is a defendant dangerous if he commits a felony? Shoplifts? Drives recklessly? Such questions are left for each individual jury to deliberate.[47]

But despite the confusion surrounding the future-dangerousness question, jurors still center their focus largely on the future-dangerousness

---

42. *See, e.g.*, Scott Brewer, *Scientific Expert Testimony and Intellectual Due Process*, 107 YALE L.J. 1535, 1538–39 (1998).

43. Stanley Milgram's electric-shock studies in the 1960s demonstrated that reliance on authority can lead people to obey "even the most abhorrent of orders." *See* Cari Romm, *Rethinking One of Psychology's Most Infamous Experiments*, ATLANTIC (Jan. 28, 2015), https://www.theatlantic.com/health/archive/2015/01/rethinking-one-of-psychologys-most-infamous-experiments/384913/ [https://perma.cc/A7UU-P7PH].

44. *See* Beecher-Monas, *supra* note 41, at 362.

45. *Id.*

46. TEX. CODE CRIM. PROC. ANN. art. 37.071, § 2(b)(1) (West Supp. 2016).

47. Citron, *supra* note 14, at 158 ("[I]t also offers absolutely no guidance as to the level of certainty required for an answer of 'yes.' We might consider 10% a reasonably high probability, but we might also require 51% (i.e., more likely than not) or 95% (something like 'beyond a reasonable doubt'). As it stands, that question . . . is left to the jury, and the evidence suggests that they are as confused about it as everybody else."). In *Jurek v. State*, Judge Odom of the Texas Court of Criminal Appeals expressed similar dissatisfaction with the ambiguity of "probability" in the Texas statute:

> What did the Legislature mean when it provided that a man's life or death shall rest upon whether there exists a "probability" that he will perform certain acts in the future? Did it mean, as the words read, is there a probability, some probability, any probability? We may say there is a twenty percent probability that it will rain tomorrow, or a ten or five percent probability. Though this be a small probability, yet it is some probability, a probability, and no one would say it is no probability or not a probability. It has been written: "It is probable that many things will happen contrary to probability," and "A thousand probabilities do not make one fact." The statute does not require a particular degree of probability but only directs that some probability need be found. The absence of a specification as to what degree of probability is required is itself a vagueness inherent in the term as used in this issue. Our common sense understanding of the term leaves the statute too vague to pass constitutional muster.

522 S.W.2d 934, 945 (Tex. Crim. App. 1975) (Odom, J., concurring in part and dissenting in part).

inquiry. Even in South Carolina, a state whose capital-sentencing statute is not centered exclusively upon future dangerousness, data shows that the "defendant's dangerousness should he ever return to society (including the possibility and timing of such a return) are second only to the crime itself in the attention they receive during the jury's penalty phase deliberations."[48] Moreover, prosecutors often guide what constitutes a probability of future dangerousness—and prosecutors are "consistently clear about the fact that they [do] not have to prove that the defendant [will] kill again."[49] For example, one prosecutor in a Texas capital-sentencing proceeding argued:

> I never have to prove to you he would kill again or that he would rape somebody or that he would stab somebody. It could be setting a fire in a prison cell. It could be threatening to assault a guard over and over. It could be tearing up the cell or tearing up facilities.[50]

Furthermore, given prosecutors' guidance, coupled with the misleading nature of the Texas statute, it is no wonder that Texas juries often inaccurately predict a defendant's future dangerousness. But, since future dangerousness is the central inquiry in Texas capital-sentencing proceedings, an accurate prediction of future dangerousness is essential to the legitimacy of the Texas capital-sentencing scheme.

B. Perception that Death Is Required

Second, empirical evidence suggests that the Texas statute fosters the perception among jurors that death is required. The Texas statute does not *require* the imposition of death for any offender convicted of capital murder;[51] in fact, the Supreme Court rejected the idea of a mandatory death penalty in *Woodson v. North Carolina*[52] as constitutionally problematic.[53] Empirical evidence suggests that as many as a third of jurors in capital proceedings nationwide believe that a showing of future dangerousness

48. John H. Blume, Stephen P. Garvey & Sheri Lynn Johnson, *Future Dangerousness in Capital Cases: Always "At Issue"*, 86 CORNELL L. REV. 397, 404 (2001).

49. Vartkessian, *supra* note 27, at 261.

50. *Id.*

51. TEX. CODE CRIM. PROC. ANN. art. 37.071, § 2(a)(1) (stating that life imprisonment without parole is an alternative sentence for capital offenses).

52. 428 U.S. 280 (1976).

53. *Id.* at 305 ("Death, in its finality, differs more from life imprisonment than a 100-year prison term differs from one of only a year or two. Because of that qualitative difference, there is a corresponding difference in the need for reliability in the determination that death is the appropriate punishment in a specific case.").

requires a sentence of death.[54] This problem is probably exacerbated in the Texas scheme, which centers upon the future-dangerousness inquiry.[55]

But the Texas capital-sentencing scheme furthers the perception that death is required in another crucial way: through the 10–2 rule. The Texas capital-sentencing statute requires the judge to instruct the jury that, in answering the Texas Special Issues (including the defendant's future dangerousness), the jury "may not answer the issue 'no' unless it agrees unanimously and may not answer the issue 'yes' unless 10 or more jurors agree."[56] Further, the statute adopts a so-called "gag rule," which mandates that the judge may not instruct the jury regarding the consequences of a deadlock.[57] Thus, a Texas capital-sentencing jury is commonly instructed: "The jury may not answer Special Issues Numbers 1 and 2 'No' unless ten or more jurors agree. . . . The jury may not discuss or consider the effect of failure of the jury to agree on the answer to an issue."[58] While the jury may not answer "no" to future dangerousness unless ten jurors or more agree, the statute also prohibits the jury from answering "yes" to mitigation unless ten or more jurors agree.[59] As a result, jurors are, in effect, instructed that the law requires a consensus of ten or more jurors to issue a life sentence. But the effect of a holdout juror is the same as a consensus of ten jurors under Texas law:

> If the jury returns an affirmative finding on each issue submitted under Subsection (b) and a negative finding on an issue submitted under

---

54. Theodore Eisenberg & Martin T. Wells, *Deadly Confusion: Juror Instructions in Capital Cases*, 79 CORNELL L. REV. 1, 7 (1993); *see also* William J. Bowers, *The Capital Jury: Is It Tilted Toward Death?*, 79 JUDICATURE 220, 221–22 (1996) ("Four out of 10 capital jurors wrongly believed that they were 'required' to impose the death penalty if they found that the crime was heinous, vile, or depraved, and nearly as many mistakenly thought the death penalty was 'required' if they found that the defendant would be dangerous in the future.").

55. Bentele & Bowers, *supra* note 24, at 1032 ("Not surprisingly, in light of the structure of its statute, the perception that death was the mandatory sentence under certain circumstances, particularly if jurors thought the defendant would be dangerous in the future, was most prominent under the directed statute in Texas.").

56. TEX. CODE CRIM. PROC. ANN. art. 37.071, § 2(f)(2).

57. *Id.* § 2(a)(1); Sorto v. State, 173 S.W.3d 469, 492 (Tex. Crim. App. 2005) ("We have repeatedly upheld the constitutionality of Article 37.071, Section 2(a)(1), which prohibits informing jurors of the effects of their failure to agree on the special issues."). Although the Supreme Court has held that the failure to instruct jurors on the consequences of their deadlock does not violate the Eighth Amendment, *Jones v. United States*, 527 U.S. 373, 381–82 (1999), some have argued that the Due Process Clause requires such a disclosure. *See* Robert Clary, *Texas's Capital-Sentencing Procedure Has a* Simmons *Problem: Its Gag Statute and 12-10 Rule Distort the Jury's Assessment of the Defendant's "Future Dangerousness"*, 54 AM. CRIM. L. REV. 57, 110–11 (2016) ("[T]he fact that the defendant will automatically receive a sentence of life in prison without possibility of parole if a Texas capital jury fails to achieve the consensus required by the 12-10 Rule is directly relevant to the jury's assessment of the defendant's future dangerousness.").

58. *See, e.g.*, Court's Charge on Punishment at 2, State v. Storey, No. 1042204D (Crim. Dist. Ct. 3, Tarrant Cty., Tex. Sept. 11, 2008), 2008 WL 8188280, at *1 (providing typical capital-sentencing jury instructions).

59. TEX. CODE CRIM. PROC. ANN. art. 37.071, § 2(d)(2), (f)(2).

> Subsection (e)(1), the court shall sentence the defendant to death. If the jury returns a negative finding on any issue submitted under Subsection (b) or an affirmative finding on an issue submitted under Subsection (e)(1) *or is unable to answer any issue submitted under Subsection (b) or (e)*, the court shall sentence the defendant to confinement in the Texas Department of Criminal Justice for life imprisonment without parole.[60]

Based on the Texas statute, as the Fifth Circuit has acknowledged, "[a] single juror thus has the power to prevent a death sentence based on his personal view of the mitigation evidence."[61]

The mandated Texas jury instructions lead jurors who favor a life sentence to believe that they must convince nine other jurors to vote for life to avoid issuing a death sentence—but one holdout juror produces the same effect. This misunderstanding is corroborated not only by empirical evidence[62] but also by powerful anecdotal evidence. For example, Sven Berger, who served as a juror in a 2008 capital murder trial, recently spoke out regarding his own misunderstanding of the Texas capital jury instructions.[63] Berger did not want to sentence the defendant to death—based on his impression that the defendant would not be "a future danger to society"—but a majority of the jury voted for death.[64] Believing that he could not sway the other jurors' votes, Berger reluctantly assented to the death sentence.[65] But what he didn't realize "in part because of the language in the jury instructions . . . was that his vote alone could have blocked the jury from handing down a death sentence and given [the defendant] life in prison without the possibility of parole."[66] Although Berger's "haunt[ing]" experience with the Texas death-penalty statute inspired him to speak out (which in turn inspired two Texas legislators to file bills to change the statute),[67] most jurors remain unaware of their misunderstanding.

In reality, death is not required in Texas—and neither is unanimity for life. Such a capital-sentencing structure would violate the Eighth Amendment prohibition of mandatory death-penalty statutes.[68] But most jurors are never informed of the consequences of their dissent.

---

60. *Id.* § (2)(g) (emphasis added).

61. Allen v. Stephens, 805 F.3d 617, 631 (5th Cir. 2015).

62. *See supra* notes 54–55 and accompanying text.

63. Jolie McCullough, *Texas Death Penalty Juror Hopes to Change Law as Execution Looms*, TEX. TRIB. (Mar. 28, 2017), https://www.texastribune.org/2017/03/28/texas-death-penalty-juror-hopes-change-law-execution-looms/ [https://perma.cc/DY8K-W4M4].

64. *Id.*

65. *Id.*

66. *Id.*

67. *Id.*

68. Woodson v. North Carolina, 428 U.S. 280, 305 (1976).

C. *Limited Consideration of Mitigating Factors*

Third, empirical evidence suggests that, despite the mitigation Special Issue in the Texas statute, capital-sentencing juries fail to give much weight to mitigating circumstances in their deliberations. In order for a capital-sentencing statute to survive Eighth Amendment scrutiny,[69] the statute must provide for consideration of mitigating factors that may persuade a jury to sentence the defendant with less than death.[70] In fact, the statute must allow for the consideration of "as a mitigating factor, any aspect of a defendant's character or record and any of the circumstances of the offense that the defendant proffers as a basis for a sentence less than death."[71] Moreover, the sentence must give *at least some* weight to each mitigating factor.[72] The Texas statute in particular asks the jury to consider

> Whether, taking into consideration all of the evidence, including the circumstances of the offense, the defendant's character and background, and the personal moral culpability of the defendant, there is a sufficient mitigating circumstance or circumstances to warrant that a sentence of life imprisonment without parole rather than a death sentence be imposed.[73]

Empirical evidence suggests that jurors fail to give much weight to these constitutionally mandated mitigating factors. In particular, Capital Jury Project "interviews reflect a pattern in which mitigating factors play a disturbingly minor role in jurors' deliberations about whether a defendant should be sentenced to death."[74] Moreover, "[e]ven when jurors do report a discussion of mitigating factors, their understanding of what the law defines as mitigation is extremely limited."[75] Admittedly, this lack of attention to

---

69. Under prevailing Eighth Amendment doctrine, the words of the Eighth Amendment are "not precise, and . . . their scope is not static." Trop v. Dulles, 356 U.S. 86, 100–01 (1958). Instead, "[t]he Amendment must draw its meaning from the evolving standards of decency that mark the progress of a maturing society." *Id.* at 101.

70. *See* Lockett v. Ohio, 438 U.S. 586, 608 (1978) ("To meet constitutional requirements, a death-penalty statute must not preclude consideration of relevant mitigating factors."); *Woodson*, 428 U.S. at 305 (striking down North Carolina's mandatory death-penalty statute as unconstitutional).

71. *Lockett*, 438 U.S. at 604. In *Lockett*, the Supreme Court struck down an Ohio capital-sentencing statute that limited the sentence to consideration of three enumerated mitigating factors. *Id.* at 608–09. Such a statute that prohibited "consideration of a defendant's comparatively minor role in the offense, or age" was constitutionally impermissible. *Id.* at 608.

72. Eddings v. Oklahoma, 455 U.S. 104, 112–14 (1982) ("By holding that the sentencer in capital cases must be permitted to consider any relevant mitigating factor, the rule in *Lockett* recognizes that a consistency produced by ignoring individual differences is a false consistency. . . . Just as the State may not by statute preclude the sentencer from considering any mitigating factor, neither may the sentencer refuse to consider, as a matter of law, any relevant mitigating evidence.").

73. TEX. CODE CRIM. PROC. ANN. art. 37.071, § 2(e)(1) (West Supp. 2016).

74. Bentele & Bowers, *supra* note 24, at 1041.

75. *Id.* at 1042.

mitigation could be partly attributed to jurors' misunderstanding of the niche term "mitigation." In one California study, for example:

> [L]ess than one-half of . . . subjects could provide even a partially correct definition for the term 'mitigation,' almost one-third provided definitions that bordered on being uninterpretable or incoherent, and slightly more than one subject in ten was still so mystified by the concept that he or she was unable to venture a guess about its meaning.[76]

The inattention to mitigation is likely also attributable to jury instructions. The presumption of death is included in *the mitigation question itself*: does the mitigation evidence warrant that a life sentence is imposed "rather than a death sentence"?[77] Empirical evidence suggests that jurors pick up on this "death default": one juror interviewed by the Capital Jury Project recalled that the jury "had no instructions or didn't ask as to what role childhood should play."[78] The juror thus "[d]idn't know if defendant's childhood was [a] valid" mitigating factor and later lamented that he "[s]hould have asked [the] judge if that was [a] valid reason to deny death."[79]

This is hardly surprising: the Texas statute asks a very specific, concrete question regarding future dangerousness. But the question regarding mitigation is open-ended: there is no guidance from the law, and there are no enumerated mitigating factors. This scheme "encourages the dismissal of mitigation evidence as being irrelevant to jurors' sentencing decision."[80] In particular, scholars have argued that the Texas statute allows prosecutors to "dismantle and reframe" the sentencing scheme to focus purely on future dangerousness and encourage jurors to largely ignore mitigating evidence.[81]

The tendency of juries to downplay mitigating evidence is particularly troubling. The Supreme Court has mandated that virtually *all* mitigating factors be considered in capital-sentencing proceedings.[82] In fact, the Court has asserted that, "in capital cases[,] the fundamental respect for humanity underlying the Eighth Amendment *requires* consideration of the character and record of the individual offender and the circumstances of the particular offense as a constitutionally indispensable part of the process of inflicting the

---

76. Craig Haney, *Taking Capital Jurors Seriously*, 70 IND. L.J. 1223, 1229 (1995).

77. *See* TEX. CODE CRIM. PROC. ANN. art. 37.071, § 2(e)(1).

78. Bentele & Bowers, *supra* note 24, at 1044.

79. *Id.*

80. Vartkessian, *supra* note 27, at 240.

81. *See, e.g.*, *id.* ("Due to the statute's focus on the defendant's future dangerousness and the ambiguity of the mitigation instruction, legal actors are able to dismantle and reframe the sentencing scheme in a way which advances the dismissal of much mitigating evidence.").

82. *See* Tennard v. Dretke, 542 U.S. 274, 285 (2004) ("Thus, a State cannot bar 'the consideration of . . . evidence if the sentencer could reasonably find that it warrants a sentence less than death.'").

penalty of death."[83] If juries fail to consider such evidence, however, the result falls short of the required "reasoned moral response."[84]

D. Moral Distance Between the Decision Maker and the Decision Itself

Fourth, empirical evidence suggests that the structure of the Texas death-penalty statute creates moral distance between the decision maker and the decision itself. In a Capital Jury Project interview, one juror admitted that the life-versus-death decision was "easier" than she suspected and that she "thought it would be harder than just answering questions."[85] Another Texas juror recounted how the judge explained that the jury was *not* charged with the life-or-death inquiry:

> [The judge] said that he wanted us to understand that we were not choosing whether somebody should get the death penalty or not as far as being responsible if he ends up dying as a result of getting the death penalty. That it was up to us to answer yes or no to, I think it was three, questions. And based on the way we answered those questions. The death penalty would be assigned or not assigned, according to Texas law. The defense tried to make us feel as though we would be responsible for [the defendant] dying if we gave him the death penalty so I think that the judge maybe took some of that sting away.[86]

In fact, in live interviews of 153 capital jurors, only 28% of jurors agreed that determining whether the defendant lived or died was "strictly the jury's responsibility and no one else's"—and only 21% of jurors who gave a death sentence agreed with that proposition.[87]

It is unsurprising that the Texas statute discourages jurors from taking responsibility for the life-or-death decision. Texas jurors are *never* explicitly instructed to consider whether the defendant should be given a death sentence or life in prison without parole; instead, they are asked to answer a series of "yes" or "no" questions.[88] In most cases (when the defendant was not found guilty by the law of parties), the jury considers only two such questions: future dangerousness and mitigation.[89] But those questions do, in fact, make a life-or-death determination.

---

83. Woodson v. North Carolina, 428 U.S. 280, 304 (1976) (plurality opinion) (emphasis added) (citations omitted).

84. Franklin v. Lynaugh, 487 U.S. 164, 185 (1988) (O'Connor, J., concurring); *see also id.* at 194 (Stevens, J., dissenting) (concluding that "the risk that the jury did not give full consideration to the mitigating evidence petitioner introduced" required the death sentence to be vacated).

85. Bentele & Bowers, *supra* note 24, at 1039.

86. *Id.* at 1040.

87. Theodore Eisenberg, Stephen P. Garvey & Martin T. Wells, *Jury Responsibility in Capital Sentencing: An Empirical Study*, 44 BUFF. L. REV. 339, 350, 353 tbl.1 (1996).

88. TEX. CODE CRIM. PROC. ANN. art. 37.071, § 2(c), (f)(1) (West Supp. 2016).

89. *See id.* § 2(b)(1) (requiring the jury to consider "whether there is a probability that the defendant would commit criminal acts of violence that would constitute a continuing threat to society"); *id.* § 2(d)(1) (requiring the jury to consider "evidence of the defendant's background or

The Supreme Court has mandated that jurors must be able express their "reasoned moral response"[90] to the evidence. But if jurors do not feel responsible for the life-or-death decision, the question arises: are their yes-or-no answers truly indicative of a "reasoned moral response"?

III. How the Framing of the Texas Statute Affects Juror Decision Making

Several theories have been advanced to explain the foregoing empirical capital-jury findings. These include the Story Model,[91] the groupthink model,[92] and agentic shift.[93] Moreover, scholars have repeatedly noted the tendency of jurors to vastly underestimate the time a defendant sentenced to life in prison will actually serve in prison.[94] Such misperceptions undoubtedly contribute to a greater likelihood that a jury will return a death sentence.

But this Note advances an additional consideration that may increase the likelihood of a death sentence: the structure of the Texas statute itself. Specifically, the format and language of the Texas statute invokes two cognitive heuristics—the representativeness heuristic and the anchoring-and-adjustment heuristic—that arguably increase the likelihood that any given jury will return a death sentence.

A. Representativeness Heuristic

One plausible explanation for why jurors assume that death is default[95] and inaccurately predict future dangerousness[96] is the Texas statute's tendency to play into the representativeness heuristic. The representativeness heuristic, first articulated by Amos Tversky and Daniel Kahneman, centers upon the question: "What is the probability that object A belongs to class

---

character or the circumstances of the offense that militates for or mitigates against the imposition of the death penalty"); *id.* § 2(f)(4) (requiring the jury to "consider mitigating evidence to be evidence that a juror might regard as reducing the defendant's moral blameworthiness").

90. Franklin v. Lynaugh, 487 U.S. 164, 184–85 (1988) (O'Connor, J., concurring).

91. REID HASTIE, STEVEN D. PENROD & NANCY PENNINGTON, INSIDE THE JURY 22–23 (2d prtg. 2004) ("The Story Model provides a complete psychological account of cognitive processing in juror decision making, and it receives support from jury research, political science analysis, jurors' accounts of their experiences during trials, and other work.").

92. *See* Bentele & Bowers, *supra* note 24, at 1056 ("Irving Janis's 'groupthink' model of group decision making helps to account for the reluctance of jurors to switch gears when they move from the guilt to the penalty phase of the trial—to explain why jurors become fixated on guilt and aggravation while paying little attention to mitigation.").

93. *Id.* at 1058–59 ("A juror making a life or death sentencing decision is in the kind of situation that might well be expected to induce an agentic shift.").

94. *See, e.g.*, William J. Bowers & Benjamin D. Steiner, *Death by Default: An Empirical Demonstration of False and Forced Choices in Capital Sentencing*, 77 TEXAS L. REV. 605, 670 (1999) ("In every state examined here, capital jurors vastly underestimate the time that convicted first-degree murderers not given the death penalty will stay in prison.").

95. *See supra* subpart II(B).

96. *See supra* subpart II(A).

B?"[97] When individuals rely on the representativeness heuristic, they assess the degree to which object A is "representative of, or similar to, the stereotype of a" member of class B.[98] Tversky and Kahneman provide the following example:

> "Steve is very shy and withdrawn, invariably helpful, but with little interest in people, or in the world of reality. A meek and tidy soul, he has a need for order and structure, and a passion for detail." How do people assess the probability that Steve is engaged in a particular occupation from a list of possibilities (for example, farmer, salesman, airline pilot, librarian, or physician)? How do people order these occupations from most to least likely? In the representativeness heuristic, the probability that Steve is a librarian, for example, is assessed by the degree to which he is representative of, or similar to, the stereotype of a librarian.[99]

This method of categorizing objects—or, in this case, people—"leads to serious errors, because similarity, or representativeness, is not influenced by several factors that should affect judgments of probability."[100] These factors include insensitivity to the prior probability of outcomes, sample size, and predictability, as well as the illusion of validity and misconceptions of regression.[101] Moreover, relying on representativeness is problematic because "humans simply cannot assign optimal weights to variables, and they are not consistent in applying their own weights."[102]

As a result, many have warned against overreliance on expert testimony—specifically in the context of a defendant's dangerousness—due to the distorting effects of heuristics.[103] The representativeness bias in particular "causes the clinicians to compare the individuals they are assessing for dangerousness with their stereotypical conceptualization of a dangerous individual and construct a prediction on the basis of similarity."[104] But unfortunately, these stereotypes "are likely to be inaccurate and contain many attributes that are not linked to future violence," and thus clinicians'

---

97. Amos Tversky & Daniel Kahneman, *Judgment Under Uncertainty: Heuristics and Biases*, 185 SCI. 1124, 1124 (1974).

98. *Id.*

99. *Id.*

100. *Id.*

101. *Id.* at 1124–26.

102. William M. Grove & Paul E. Meehl, *Comparative Efficiency of Informal (Subjective, Impressionistic) and Formal (Mechanical, Algorithmic) Prediction Procedures: The Clinical–Statistical Controversy*, 2 PSYCHOL., PUB. POL'Y & L. 293, 315 (1996).

103. *See, e.g.*, GARY B. MELTON, JOHN PETRILA, NORMAN G. POYTHRESS & CHRISTOPHER SLOBGIN, PSYCHOLOGICAL EVALUATIONS FOR THE COURTS: A HANDBOOK FOR MENTAL HEALTH PROFESSIONALS AND LAWYERS 301 (3d ed. 2007) (arguing that "[c]linicians' judgments may also be affected by cognitive heuristics that influence the selection and weighting given to particular predictor variables").

104. Daniel A. Krauss & Bruce D. Sales, *The Effects of Clinical and Scientific Expert Testimony on Juror Decision Making in Capital Sentencing*, 7 PSYCHOL., PUB. POL'Y & L. 267, 280 (2001).

"predictions of violence are likely to be based on poor correlates of future violent behavior."[105] In fact, "[a] steadily growing body of peer-reviewed literature directly undercuts the legitimacy of future dangerousness diagnoses"[106]—so much so that the American Psychiatric Association has asserted that "[p]sychiatrists should not be permitted to offer a prediction concerning the long-term future dangerousness of a defendant in a capital case."[107] Despite these warnings, "Texas routinely allows such testimony."[108]

Although prior research has centered upon clinical assessments of future dangerousness, there is no reason to believe that jurors do not fall prey to the same cognitive shortcuts as clinicians. In fact, the structure of Texas capital proceedings tees up juries to focus on a representativeness question: that is, "whether there is a probability that the defendant would commit criminal acts of violence that would constitute a continuing threat to society."[109] The capital-sentencing jury is the same trial jury that found the defendant guilty of capital murder—often immediately beforehand.[110] Then, the Texas scheme requires the same jury to consider the future dangerousness of the convicted defendant, the primary consideration in the capital-sentencing determination.[111] Unsurprisingly, given that the Texas statute focuses on "whether there is a probability that the defendant would commit criminal acts of violence that would constitute a continuing threat to society,"[112] many jurors continue to dwell on guilt.[113] The Texas statute merely underscores this preoccupation with death by *requiring* jurors to focus on the future-dangerousness inquiry, a concept intertwined with the guilt determination.[114] In fact, findings from the Capital Jury Project suggest that many jurors begin "taking a stand on what the defendant's punishment should be well before they [are] exposed to the statutory guidelines for making this decision"—an almost predictable result.[115]

In effect, the Texas scheme requires jurors to ask whether object A—a defendant who the same jury convicted of capital murder—belongs in class B—a class of defendants that "would commit criminal acts of violence." The likelihood that such a defendant would be incongruent with

105. *Id.*

106. Meghan Shapiro, *An Overdose of Dangerousness: How "Future Dangerousness" Catches the Least Culpable Capital Defendants and Undermines the Rationale for the Executions It Supports*, 35 AM. J. CRIM. L. 145, 160–61 (2008).

107. Amicus Brief of American Psychiatric Ass'n for Petitioner at *3, Barefoot v. Estelle, 463 U.S. 880 (1983) (No. 82-6080).

108. Shapiro, *supra* note 106, at 160 n.64.

109. TEX. CODE CRIM. PROC. ANN. art. 37.071, § 2(b)(1) (West Supp. 2016).

110. *Id.* § 2(a)(1).

111. *Id.* § 2(b)(1).

112. *Id.*

113. Bentele & Bowers, *supra* note 24, at 1021.

114. *Id.*

115. Bowers, *supra* note 54, at 221.

the jurors' stereotypes of those who "would commit criminal acts of violence" is slim, and we see the result of that inquiry in jury verdicts themselves.[116]

B. Anchoring & Adjustment Heuristic

A second cognitive shortcut that affects jurors' assessment of a defendant's future dangerousness is the anchoring-and-adjustment heuristic. This heuristic, also articulated by Tversky and Kahneman, relies on the hypothesis that "people make estimates by starting from an initial value that is adjusted to yield the final answer."[117] Typically, this heuristic is triggered when a numeric value is in play; the initial value, known as the anchor, is a number that is adjusted upward or downward based on a set of variables.[118] For example, Judge Mark Bennett has argued that, despite the advisory nature of the Federal Sentencing Guidelines, the Guidelines create a numeric anchor that subconsciously affects judicial sentencing.[119]

But new evidence suggests that the anchoring-and-adjustment heuristic can be triggered by phenomena other than numeric figures, specifically by the order and format in which information is presented. Research suggests that the initial information presented can create a sort of "anchor" that constrains individuals' adjustments in response.[120] For example, some researchers have shown that the order in which information is presented on an exam affects students' evaluation of their own performance.[121] The initial questions presented create a sort of anchor, and the students adjust their perception of performance based on that anchor.[122] If information is presented in a different order, then a different self-evaluation results.[123] This phenomenon is known as belief persistence.[124]

---

116. *See supra* notes 54–55 and accompanying text.

117. Tversky & Kahneman, *supra* note 97, at 1128.

118. *See id.* ("The initial value, or starting point, may be suggested by the formulation of the problem, or it may be the result of a partial computation. . . . [D]ifferent starting points yield different estimates, which are biased toward the initial values. We call this phenomenon anchoring.").

119. Mark W. Bennett, *Confronting Cognitive "Anchoring Effect" and "Blind Spot" Biases in Federal Sentencing: A Modest Solution for Reforming a Fundamental Flaw*, 104 J. CRIM. L. & CRIMINOLOGY 489, 491 (2014).

120. *See, e.g.*, Yana Weinstein & Henry L. Roediger III, *The Effect of Question Order on Evaluations of Test Performance: How Does This Bias Evolve?*, 40 MEMORY & COGNITION 727, 728 (2012) ("One possibility is that the difficulty of the questions at the beginning of a test sets an anchor that constrains participants in their evaluations of performance throughout the remainder of the test.").

121. *Id.* at 727–28.

122. *Id.*

123. *Id.*

124. Lee Ross & Craig A. Anderson, *Shortcomings in the Attribution Process: On the Origins and Maintenance of Erroneous Social Assessments*, *in* JUDGMENT UNDER UNCERTAINTY: HEURISTICS AND BIASES 144 (D. Kahneman, P. Slovic & A. Tversky eds., 1982).

In the legal context, scholars have argued that the anchoring-and-adjustment heuristic plays a role in plea-bargaining discussions,[125] damage calculations,[126] and divorce negotiations.[127] In the criminal context, researchers have used the anchoring-and-adjustment heuristic and belief persistence to explain varied outcomes in analyzing offender profiles.[128] In particular, in a study identifying the decision-making mechanisms employed by participants analyzing offender profiles, participants were first presented with either a description of the suspect or a description of the stereotypical profile of an offender.[129] The order of information presented impacted the participant's likelihood of a guilt determination, which the researchers attributed in part to a confirmation bias.[130]

Reliance on an anchor, however, can be problematic in decision making. First, the selection of the anchor is often biased and self-serving.[131] Second, anchoring "depends not so much on relevance as on recency," and "[e]xperiment subjects are most influenced by information that they receive just before they make judgments, even if that information is obviously useless."[132] Third, and perhaps most importantly, "people usually do not adjust away from their anchors enough," leading to skewed decision making.[133]

The structure of the Texas statute, however, encourages reliance on an anchor—namely, the anchor of future dangerousness. The Texas capital-sentencing jury instruction is presented as a series of questions to which the jury must answer "yes" or "no."[134] The future-dangerousness inquiry is often

---

125. *See* Stephanie Bibas, *Plea Bargaining Outside the Shadow of the Trial*, 117 HARV. L. REV. 2463, 2515–17 (2004) ("Anchoring also helps to explain the course of negotiation. Bargainers who lack inside information about an opponent's payoff matrix are more influenced by the opponent's initial offer than by later concessions."). *See generally* Colin Miller, *Anchors Away: Why the Anchoring Effect Suggests that Judges Should Be Able to Participate in Plea Discussions*, 54 B.C. L. REV. 1667 (2013) (noting that "[i]n the majority of [plea bargains], the prosecutor makes the initial plea offer, which is typically high" and proposing that judges be involved in plea discussions to reduce this anchoring effect).

126. *See* Jeffrey J. Rachlinski, Andrew J. Wistrich & Chris Guthrie, *Can Judges Make Reliable Numeric Judgments? Distorted Damages and Skewed Sentences*, 90 IND. L.J. 695, 706 (2015) ("Anchoring can influence a wide variety of judgments in legal contexts, especially civil damage awards and criminal sentences.").

127. Tess Wilkinson-Ryan & Deborah Small, *Negotiating Divorce: Gender and the Behavioral Dynamics of Divorce Bargaining*, 26 LAW & INEQ. 109, 127 (2008).

128. Benjamin C. Marshall & Lawrence J. Alison, *Stereotyping, Congruence and Presentation Order: Interpretive Biases in Utilizing Offender Profiles*, 13 PSYCHOL., CRIME & L. 285, 285–87 (2007).

129. *Id.* at 288–91.

130. *Id.* at 291, 296–97.

131. *See* Bibas, *supra* note 125, at 2516 ("Because assessments of fairness are self-serving, each side may choose a different anchor.").

132. *Id.*

133. *Id.*

134. TEX. CODE CRIM. PROC. ANN. art. 37.071, § 2(f)(1) (West Supp. 2016).

the only aggravating Special Issue presented to a Texas capital-sentencing jury.[135] For example, a Texas capital-sentencing jury is often instructed:

> You are instructed that a sentence of imprisonment in the Institutional Division of the Texas Department of Criminal Justice for life imprisonment without parole, or a sentence of death is mandatory upon conviction of capital murder. In order for the Court to assess the proper punishment, certain special issues are submitted to you. . . .
>
> SPECIAL ISSUE NUMBER 1:
>
> Do you find from the evidence beyond a reasonable doubt that there is a probability that the Defendant would commit criminal acts of violence that would constitute a continuing threat to society? ANSWER "YES" OR "NO" in the space provided. . . .
>
> SPECIAL ISSUE NUMBER 3:
>
> Taking into consideration all of the evidence, including the circumstances of the offense, the Defendant's character and background, and the personal moral culpability of the Defendant, do you find from the evidence that there is a sufficient mitigating circumstance or circumstances to warrant that a sentence of life imprisonment without parole rather than a death sentence be imposed? ANSWER "YES" OR "NO" in the space provided.[136]

Given the structure of the Texas statute and jury instructions, it is not surprising that jurors view their role as "just answering questions."[137] But more importantly, the framing of the statute creates a future-dangerousness anchor, leading to deliberations that are tied to future dangerousness—and future dangerousness only.

The statute creates this effect not only by placing future dangerousness first, as an anchor, but also by giving jurors a specific concept (i.e., future dangerousness) to reference in deliberations. The mitigation question, however, is much more open-ended. Such an amorphous instruction, especially given jurors' misunderstanding of the term mitigation,[138] does not give jurors a second anchor from which to adjust. "Due to the statute's focus on the defendant's future dangerousness and the ambiguity of the mitigation instruction, legal actors"—that is, prosecutors—"are able to dismantle and

---

135. *See id.* § 2(b) (listing only two aggravating Special Issues, one of which is only presented to the jury when "the jury charge at the guilt or innocence stage permitted the jury to find the defendant guilty as a party"). When a jury was permitted to find the defendant guilty as a party, then the jury is also asked to consider "whether the defendant actually caused the death of the deceased or did not actually cause the death of the deceased but intended to kill the deceased or another or anticipated that a human life would be taken." *Id.*

136. Court's Charge on Punishment at 2–3, State v. Storey, No. 1042204D, 2008 WL 8185556 (Crim. Dist. Ct. 3, Tarrant County, Tex. Sept. 15, 2008), 2008 WL 8188280, at *1–2 (providing typical capital-sentencing jury instructions).

137. *See supra* notes 85–86 and accompanying text.

138. Haney, *supra* note 76, at 1229.

reframe the sentencing scheme in a way which advances the dismissal of much mitigating evidence."[139] Thus, jurors are anchored to the future-dangerousness question—a consideration prescribed by the law—and left to adjust for open-ended mitigation factors on their own.

This anchoring phenomenon helps explain several of the empirical findings regarding capital juries. First, anchoring to the future dangerousness helps explain juries' inaccurate predictions of defendants' future dangerousness.[140] Certainly, if clinicians are prone to inaccurate predictions because of this heuristic, capital jurors are as well. Second, the anchoring effect helps explain jurors' tendency to ignore mitigation evidence.[141] Because the future-dangerousness question is both first and specific, jurors anchor their discussions to that question—and largely ignore mitigation. Finally, the tendency of jurors to cling to the series of yes-or-no questions as a sort of checklist creates moral distance between the decision and the decision makers. This helps explain why juries may not feel responsible for the death sentence at all.[142]

Conclusion

The current Texas death-penalty scheme is problematic—based on a hastily adopted statute that harmfully plays on jurors' cognitive heuristics. And even in a time of dwindling death sentences, this problem should be corrected. When so few offenders are sentenced to the ultimate penalty, they should not be sentenced based on a statute that clouds the ability of jurors to express their reasoned moral response.

But what can be done to correct the problems caused by the Texas statute? Texas could proceed in one of two ways. First, Texas could simply abandon the future-dangerousness, Special Issue framework and revert to the country's mean. After all, Texas is one of only two states (joined by Oregon) that centers the inquiry on the future-dangerousness question.[143] Second, and perhaps more effectively, given the argument advanced in Part III, Texas could alter the Special Issue framework. Instead of anchoring the jury's

139. Vartkessian, *supra* note 27, at 240.

140. *See supra* subpart II(A).

141. *See supra* subpart II(C).

142. *See supra* subpart II(D).

143. Williams W. Berry III, *Ending Death by Dangerousness: A Path to the De Facto Abolition of the Death Penalty*, 52 ARIZ. L. REV. 889, 894 (2010). In fact, other states do not even embrace the Special Issues framework, instead focusing on enumerated mitigating factors—and thus requiring the jury to engage with the moral culpability of the defendant. *See, e.g.*, J. Michael Brown, *Eighth Amendment—Capital Sentencing Instructions*, 84 J. CRIM. L. & CRIMINOLOGY 854, 872 (1994) ("In thirty of the thirty-six states that provide for the death penalty, the legislature has adopted a sentencing statute that enumerates certain mitigating circumstances which the sentencer must consider in making its determination. While the states vary greatly on the type of mitigating circumstances deemed important, each of the thirty state statutes requires the sentencer to consider the moral culpability of the defendant.").

attention on future dangerousness (and compromising the jury's reasoning based on the representativeness and anchoring-and-adjustment heuristics), simply anchor the question on mitigation. Ask the mitigation question first—and preferably with more specificity. Further, Texas could require that the jury make an independent life-or-death determination, with a bifurcated sentencing proceeding, *before* the jury answers the future-dangerousness question; then, the jury would proceed to future dangerousness only if it deems a death sentence appropriate. Finally, Texas could abandon the 10–2 rule and instead instruct the jury of the effect of a failure to reach an agreement. Whatever the solution, however, Texas should abandon its confusing, biased, and haphazardly drafted death-penalty statute in favor of a statute that allows the jury to express its reasoned moral response—as free from cognitive biases as possible.

Brittany Fowler

Insider Trading Enforcement & Link Prediction*

Introduction

The Securities and Exchange Commission (SEC) has taken a very aggressive stance towards insider trading enforcement, promising to be an unrelenting and omnipresent foe to those who seek to abuse nonpublic information at the expense of investors. But in reality, the SEC's resources are limited. Until recently, the SEC was forced to rely on outside tips as its primary means of detection and to triage its cases heavily—pursuing only the lowest hanging fruit. Today, however, technology provides a means for the SEC to get more bang for its buck. To date, the SEC's tech-enforcement efforts have been centered on the development of in-house, automated, market-data-analysis systems to proactively detect suspicious trades, thereby reducing its reliance on outside tips. But despite the increasing ease with which investigators are able to pinpoint illegal trades amidst a din of market data, in cases where no readily discernable insider connection exists, the SEC must continue its old triage practices, mothballing potentially viable cases due to expected investigation costs. In short, despite its new toys, the SEC continues to suffer one of its old bottlenecks.

To that end, link prediction and corollary algorithmic technologies may be just the shot in the arm the SEC needs to better manage its caseload. Link prediction works by using network structures and other information to predict connections that should or will exist—for example, between people or organizations. In theory, with enough information, link prediction systems could predict traders' connections to inside sources. Investigators may be able to use these systems as a means to conduct low-cost preliminary case audits and, should a link be predicted, build a roadmap for a full investigation.

But despite the allure of this powerful tool, its application faces practical and legal hurdles. First, as a practical matter, link prediction systems are limited in their ability to synthesize dissimilar data into a usable format and in their ability to filter out false-positives (i.e., irrelevant connections). Fortunately for the SEC, ongoing research into data synthesis and link-strength prediction is helping to mitigate these issues, and link prediction technology is not so limited that it cannot be used in its current form. Second, as a legal matter, it may be difficult to satisfy the Fourth Amendment when obtaining certain data, to square algorithmic enforcement with notions of due process, and to rely on statistical data to sustain a verdict. Nevertheless, thanks to the third-party doctrine and the Stored Communications Act, the SEC will have almost free access to a wealth of transactional data—the type

* I thank David Garrett for guiding me through this spelunk into network analysis.

of data most useful in link prediction. Thus, at a bare minimum, link prediction systems can serve as a tool to inexpensively vet cases and point investigators in the right direction, ameliorating the current triage bottleneck.

This Note briefly reviews insider trading liability before moving into the SEC's role vis-à-vis insider trading, the issues it faces, and its current tech-enforcement efforts. From there, this Note discusses briefly the basics of link prediction, data synthesis, and link-strength prediction, and then turns to the application of those technologies and the limits imposed on them by the Constitution. Finally, this Note concludes with a worst-case scenario for the SEC should it decide to incorporate link prediction analysis into its current enforcement scheme.

I. An Overview of Insider Trading Liability

Many are familiar with the concept of insider trading: an "insider" with information unknown to the public makes a trade in order to profit from the market's ignorance. However, not all trades based on nonpublic information are illegal. Illegal insider trading, which is referred to here simply as "insider trading," requires, essentially, the "buying or selling [of] a security, in breach of a fiduciary duty or other relationship of trust and confidence, while in possession of material, nonpublic information about the security."[1] In short, (1) did the trader trade on nonpublic information, and (2) did the trader have a duty to disclose the information before doing so? In the second question lies the rub.

17 C.F.R. § 240.10b–5 (Rule 10b–5),[2] which contains more pointed (albeit still ambiguous) antifraud provisions than the statute under which it was promulgated,[3] has made it "relatively easy" to prohibit company *insiders* from profiting off of nonpublic information because "such behavior fit[s] within traditional notions of fraud."[4] But trades based upon nonpublic

1. *Fast Answers: Insider Trading*, U.S. SEC. & EXCHANGE COMMISSION (Jan. 15, 2013), https://www.sec.gov/fast-answers/answersinsiderhtm.html [https://perma.cc/CP9Q-ULYX].

2. 17 C.F.R. § 240.10b–5 (2017) reads:
> It shall be unlawful for any person, directly or indirectly, by the use of any means or instrumentality of interstate commerce, or of the mails or of any facility of any national securities exchange, (a) To employ any device, scheme, or artifice to defraud, (b) To make any untrue statement of a material fact or to omit to state a material fact necessary in order to make the statements made, in the light of the circumstances under which they were made, not misleading, or (c) To engage in any act, practice, or course of business which operates or would operate as a fraud or deceit upon any person, in connection with the purchase or sale of any security.

3. 15 U.S.C. § 78j (2016) forbids traders
> to use or employ, in connection with [the sale of securities,] . . . any manipulative or deceptive device or contrivance in contravention of such rules and regulations as the [Securities and Exchange] Commission may prescribe as necessary or appropriate in the public interest or for the protection of investors.

4. Thomas C. Newkirk, Assoc. Dir., Div. Enf't, U.S. Sec. & Exch. Comm'n, & Melissa A. Robertson, Senior Counsel, Div. Enf't, SEC, Insider Trading—A U.S. Perspective, Remarks at the

information made by *noninsiders* too can bear the markings of fraud. To that end, courts have used Rule 10b–5 as a versatile host, onto which they engrafted two noninsider-insider trading doctrines: tippee and misappropriation liability. The two defining cases for tippee and misappropriation liability are *Dirks v. SEC*[5] and *United States v. O'Hagan*[6] respectively. In *Dirks*, the Supreme Court held that to hold a tippee liable for trading on nonpublic information, the tipper must "breach[] [a] fiduciary duty to the shareholders by disclosing the information to the tippee," and the tippee must "know[] or should know that there has been a breach."[7] Fourteen years later, in *O'Hagan*, the Supreme Court held that traders who normally owe no fiduciary duty not to trade on nonpublic information can be liable when they acquire nonpublic information through the "deception of those who entrusted [them] with access to [that] information."[8]

The takeaway here is that one does not have to be an employee of a company to be liable for insider trading. Liability depends on how one obtained nonpublic information. An insider trader could be a contractor, a tennis buddy, a distant relative, or a friend of a friend. Needless to say, the task of tracking down a trader's inside connection can be an arduous one.

II. The SEC's Enforcement Division and Its Use of Technology

The SEC is not only tasked with promulgating rules under the 1934 Act but also with the Act's enforcement. Accordingly, a driving force against insider trading is the SEC's Enforcement Division, which conducts detection efforts and investigations, and handles enforcement proceedings on behalf of the SEC.[9] Before turning to the SEC's tech-enforcement practices, it is important to start with the basics of SEC enforcement and the SEC's philosophy on insider trading.

16th International Symposium on Economic Crime, Jesus College, Cambridge, England (Sept. 19, 1998), https://www.sec.gov/news/speech/speecharchive/1998/spch221.htm [https://perma.cc/492C-FJEN].

5. 463 U.S. 646 (1983).

6. 521 U.S. 642 (1997).

7. *Dirks*, 463 U.S. at 660. Although the Court did not explain every way in which one might breach one's duty through tipping, the Court did discuss one particular manner: receiving a benefit in return for the tip. *Id.* at 662. The Court noted that there is an assumption of benefit where there is "an intention to benefit the particular recipient." *Id.* at 664. This is because a gift of inside information for the purposes of trading is functionally equivalent to committing insider trading oneself then giving the proceeds to the recipient. *Id.* This theory of breach was recently reinforced in *Salman v. United States*, 137 S. Ct. 420, 422–24 (2016).

8. *O'Hagan*, 521 U.S. at 652.

9. Chien-Chung Lin & Eric Hung, *U.S. Insider Trading Law Enforcement: Issues and Survey of SEC Actions from 2009 to 2013*, 11 NAT'L TAIWAN L. REV. 37, 46 (2016). Insider trading cases constitute nearly 8% of all yearly SEC enforcement actions. *Id.* at 65.

A. SEC Enforcement

As a general matter, enforcement actions can be divided into four stages: detection, preliminary investigation, formal investigation, and prosecution.

1. Detection.—The Enforcement Division gathers information regarding potential violations from public, private, and internal sources. Common sources include periodic filings; complaints from investors, whistleblowers, and competitors; the media; referrals from self-regulatory organizations like the Financial Industry Regulatory Authority (FINRA) (which total in the hundreds yearly);[10] and in-house market surveillance, which is discussed at length below.[11]

2. Preliminary Investigation.—The Division sifts through its repository of nascent cases to determine which matters should be designated Matters Under Inquiry (MUI).[12] MUIs receive an informal investigation. During an informal investigation, investigators attempt to develop the facts "through . . . interviewing witnesses, examining brokerage records, reviewing trading data, and other methods."[13] Following an informal investigation, the Division decides whether to seek a formal order of investigation from the Commission.[14]

3. Formal Investigation.—Upon obtaining a formal order of investigation, the Division may compel witnesses and demand the production of documents.[15] The primary focus of the investigation is to build the case. One internal memo states that investigators are to focus on the specific

---

10. *See* Christopher P. Montagano, *The Global Crackdown on Insider Trading: A Silver Lining to the "Great Recession"*, 19 IND. J. GLOBAL LEGAL STUD. 575, 595 (2012) (stating that "[i]n 2010, FINRA referred 244 insider trading cases to the SEC, the highest in the history of FINRA").

11. Lin & Hung, *supra* note 9, at 48.

12. Because the Division cannot follow up on every matter brought to its attention, it must triage its investigations. The Division uses MUI designation as a threshold filter "to help ensure efficient allocation of resources." SEC. & EXCH. COMM'N DIV. ENF'T, ENFORCEMENT MANUAL § 2.3.1 (Oct. 28, 2016), https://www.sec.gov/divisions/enforce/enforcementmanual.pdf [https://perma.cc/7C23-BZPX]. MUI analysis has only two prongs: "whether the facts underlying the MUI show that there is potential to address conduct that violates the federal securities laws" and "whether the assignment of a MUI to a particular office will be the best use of resources for the Division as a whole." *Id.* Naturally, this test favors those cases that appear to be slam dunks *ex ante* and disfavors those that do not—regardless of whether they are in fact viable. Complementary to MUI designation, Division staff also rank open investigations for further resource allocation, dubbing certain cases "National Priority Matters." *Id.* § 2.1.1. NPM status is reserved for those cases that offer an "opportunity to send a particularly strong and effective message of deterrence," involve "egregious or extensive misconduct" that threatens "widespread and extensive harm to investors," and feature perpetrators that "occupy[] positions of substantial authority or responsibility." *Id.*

13. *How Investigations Work*, U.S. SEC. & EXCHANGE COMMISSION (Jan. 27, 2017), https://www.sec.gov/enforce/how-investigations-work.html [https://perma.cc/RG6U-SUXR].

14. *Id.*

15. 15 U.S.C. § 78u(b) (2012).

elements of the case: possession of nonpublic information, materiality, scienter, and duty.[16] After the investigation is complete, the Commission reviews the findings and decides how to proceed.[17] It may go forward by filing in Federal Court or initiating an administrative proceeding.[18] It may also decide to refer the case to the Department of Justice for parallel criminal proceedings.[19]

4. Prosecution.—The Division prosecutes cases in the name of the SEC.[20]

B. The SEC's Philosophy

According to the SEC, rooting out insider trading is critical to maintaining the "highest level of confidence" in the market, which gives investors the confidence to "put their fortunes at risk."[21] That said, the SEC faces a wealth of potential investigations while operating with finite resources.

1. No Place to Hide.—Former SEC Chair Mary Jo White once urged that the SEC must play the part of the bold and unrelenting enforcer, "perceived to be . . . everywhere."[22] She continued,

> [I]nvestors . . . want to know that there is a strong cop on the beat—not just someone sitting in the station house waiting for a call, but patrolling the streets and checking on things.
>
> They want to know that would-be fraudsters are spending more time looking over their shoulders, and less time stepping over the line.[23]

Although White acknowledged the SEC does its best "not to disappoint," she also acknowledged that omnipresence is more difficult to achieve in "today's fast moving, complex and changing markets."[24]

---

16. L. HILTON FOSTER, U.S. SEC. & EXCH. COMM'N., INSIDER TRADING INVESTIGATIONS 4–5 (Sept. 2000), https://www.sec.gov/about/offices/oia/oia_enforce/foster.pdf [https://perma.cc/MXR2-B6KK].

17. *How Investigations Work*, *supra* note 13.

18. *Id.*

19. Lin & Hung, *supra* note 9, at 48.

20. *Id.* at 46–47.

21. Arthur Levitt, Chairman, U.S. Sec. & Exch. Comm'n, A Question of Investor Integrity: Promoting Investor Confidence by Fighting Insider Trading, Address Before the "SEC Speaks" Conference (Feb. 27, 1998), https://www.sec.gov/news/speech/speecharchive/1998/spch202.txt [https://perma.cc/ZD5A-ZUMC].

22. Mary Jo White, Chair, U.S. Sec. & Exch. Comm'n, Remarks at the Securities Enforcement Forum (Oct. 9, 2013), https://www.sec.gov/news/speech/spch100913mjw [https://perma.cc/JU3Q-6RH3].

23. *Id.*

24. *Id.*

2. No Case Too Small.—To engender investor confidence, the SEC seeks to "pursue all wrongdoers—individual and institutional, of whatever position or size."[25] "Investors," White has said, "do not want someone who ignores minor violations, and waits for the big one that brings media attention."[26] This no-case-too-small mentality has also been championed by former Chair Arthur Levitt, who said, "It's not as if insider traders wander innocently into gray areas near the boundaries of legality. They willfully stride across the bright line of the law."[27] Nevertheless, other SEC sources disclose that, "given the inherent difficulties in investigating and proving insider trading cases, the reality is that there is a significant amount of clearly illegal activity that goes undetected or unpunished."[28]

3. No Resources Too Finite?—The SEC, like all law enforcement agencies, is an entity of finite resources. In the words of Chair White, enforcement resources are "not nearly sufficient to the enormity and scope of the responsibility [the SEC has]."[29] As discussed below, this has created tremendous tension between the SEC's need for omnipresence and unrelenting enforcement, and its checkbook.

C. Traditional Enforcement Methods

Traditionally, the Enforcement Division (the Division) has taken a "security-based" approach to insider trading detection.[30] Under this approach, the Division either receives a tip about potential abuses or learns of a disclosure of material nonpublic information.[31] It then inquires into who fortuitously traded the security in question and their motivation for doing so.[32] Until recently, this was the go-to detection approach because detecting illicit trades as they occurred was extremely difficult.[33] Thus, despite being reactive and inefficient, it was the only practical means.[34]

---

25. Michael D. Trager et al., *Mary Jo White Confirmed to Chair SEC; "Bold" Enforcement Envisioned*, ARNOLD & PORTER LLP 2 (Apr. 2013), http://files.arnoldporter.com/adv514recentdevelopmentsinsecenforcementreviewofwhite.pdf [https://perma.cc/G9SV-CN5G].

26. White, *supra* note 22.

27. Levitt, *supra* note 21.

28. Newkirk & Robertson, *supra* note 4.

29. White, *supra* note 22 ("In the first instance, we of course recognize that our resources are not infinite.").

30. Jake Steele, *SEC Utilizes Big Data to Fight Insider Trading*, CONSUMERS' RES. (Nov. 1, 2016), http://consumersresearch.org/sec-utilizes-big-data-to-fight-insider-trading/ [https://perma.cc/QE5J-GTU6].

31. *Id.*

32. *Id.*

33. *Id.* Only large and extremely fortuitous trades would jump out at anyone reviewing market data. Lin & Hung, *supra* note 9, at 53.

34. Steele, *supra* note 30.

When suspicious trades are detected, unless a connection to an insider is readily apparent, "one of the most challenging issues [an investigation faces] is establishing the relationship [to an insider]."[35] To do so, the Division has to beat the proverbial bushes and interview witnesses and review documents in the hope of flushing out a link.[36] Even wiretapping, publicly showcased in the Rajaratnam trial as a powerful tool for investigating insider trading,[37] is exceptionally labor-intensive for investigators, both in terms of obtaining authorization from the court and in terms of running the tap operation itself.[38] The resource-intensive nature of traditional investigation methods forced the Division to adopt a policy of assessing the viability of cases *ex ante* based upon what little information was initially available, thereby ensuring a lowest-hanging-fruit method of enforcement.[39]

D. The Move to Tech-Enforcement

Because the Division's resources are limited, it increasingly turns to technology to realize new efficiencies. To date, SEC efforts have been focused primarily on building in-house market-analysis systems to detect suspicious trades.[40] This new mode of detection has been dubbed the "trader-based" approach to enforcement.[41] In recent years, the Division's Market Abuse Unit has unveiled three market-analysis programs, which together form an impressive toolbox/alphabet soup: ARTEMIS,[42] ABAP,[43] and

35. White, *supra* note 22.

36. For a list of the multitude of witnesses and documents useful in an insider trading investigation, see generally the Foster memo, *supra* note 16.

37. For more information on the Rajaratnam wiretapping, see Eyder Peralta, *The Wiretaps that Built the Case Against Galleon's Rajaratnam*, NPR: THE TWO-WAY (May 11, 2011), http://www.npr.org/sections/thetwo-way/2011/05/11/136206203/the-wiretaps-that-built-the-case-against-galleons-rajaratnam [https://perma.cc/H9Q6-WC33].

38. 18 U.S.C. § 2518(1) (2012) provides that investigators must file an application "in writing . . . to a judge of competent jurisdiction" and include, among other things,

> details as to the particular offense[,] . . . a particular description of the nature and location of the facilities from which . . . the communication is to be intercepted, . . . a particular description of the type of communications sought[,] . . . the identity of the person, if known, . . . whose communications are to be intercepted[,] . . . [a] statement as to whether or not other investigative procedures have been tried . . . or why they reasonably appear to be unlikely to succeed, . . . [and] a statement of the period of time for which the interception is required to be maintained.

39. *See supra* note 12 and accompanying text.

40. Trager et al., *supra* note 25, at 4.

41. Steele, *supra* note 30.

42. *Id.* (describing the Advanced Relational Trading Enforcement Metrics Investigation System, which is designed to detect suspicious trading patterns between traders over time).

43. White, *supra* note 22 (describing the Advanced Bluesheet Analysis Program, which is designed to analyze specific securities transactions to detect "suspicious trading before market-moving events" and flag coordinated trades).

NEAT.[44] These programs are designed to detect both suspiciously fortuitous trades and suspiciously coordinated trades.[45] This approach not only provides a comparatively proactive means of detection, it also provides a means of detecting insider trading rings.[46] In short, the SEC's new technological tools provide an inexpensive way to self-tip and, in the case of coordinated trades, inexpensively tie conspirators together. This has allowed the SEC to do some of "[w]hat used to take weeks . . . in a matter of hours."[47]

But until this year, the SEC did not have a direct feed to the various stock exchanges.[48] It pulled data on an ad hoc basis and cobbled together its own data pool for processing.[49] Despite the database containing "over 6 billion electronic equities and options trading records,"[50] this ad hoc process meant an unrealized analytics capability as the lack of complete market data left gaps in the record and skewed the data pool. As a solution, the SEC recently announced it would begin aggregating all equity- and option-market trading activity through a centralized system it calls the Consolidated Audit Trail (CAT).[51] The CAT system will compile every trade order, execution, and cancellation for processing, pulling both from national stock exchanges and FINRA databases.[52] By enabling the SEC to use its algorithmic detection systems on all national-market trading data, the CAT system stands to advance the SEC one step closer towards its goal of appearing to be everywhere at once.

---

44. Mary Jo White, Chair, Sec. & Exch. Comm'n, Opening Remarks at the 21st Annual International Institute for Securities Enforcement and Market Oversight (Nov. 2, 2015), https://www.sec.gov/news/statement/remarks-21st-international-institute-for-securities-enforcement.html [https://perma.cc/8AUE-9B5M] (describing the National Exam Analytics Tool, which is designed to allow investigators to rapidly access transactions from a massive index).

45. Steele, *supra* note 30. Studies testing the accuracy of these detection methods have so far been positive. *See, e.g.*, Acar Tamersoy et al., *Inside Insider Trading: Patterns & Discoveries from a Large Scale Exploratory Analysis*, GA. TECH C. COMPUTING (Aug. 25–28, 2013), http://www.cc.gatech.edu/~dchau/insider/asonam13_insider.pdf [https://perma.cc/8B39-Q324] (describing synced-transaction analytics as "a major step" in detecting insider trading networks). Interestingly, private companies have also taken to using this technology to detect their own leaks. *See, e.g.*, *Insider Trading Detection*, FINANCIAL TRACKING, http://www.financial-tracking.com/insider-trading-detection/ [https://perma.cc/HX4P-BZJG] (describing a detection system that "alert[s] if a trade occurred prior to a market moving issuer event").

46. Steele, *supra* note 30. Recently, "suspicious trading patterns" led the SEC to charge 34 people in connection to an insider trading scheme. White, *supra* note 22.

47. White, *supra* note 22.

48. Nate Raymond, *Newest Weapon in U.S. Hunt for Insider Traders Paying Off*, REUTERS (Nov. 1, 2016), http://www.reuters.com/article/usa-insidertrading/rpt-insight-newest-weapon-in-u-s-hunt-for-insider-traders-paying-off-idUSL1N1D200U [https://perma.cc/RRW4-BUWB].

49. *Id.*

50. White, *supra* note 44 (referring to establishing connections to insiders).

51. *Rule 613 (Consolidated Audit Trail)*, U.S. SEC. & EXCHANGE COMMISSION, https://www.sec.gov/divisions/marketreg/rule613-info.htm [https://perma.cc/USL8-28TZ].

52. *Id.*

E. So, Where Does the SEC Go from Here?

By honing its ability to detect illegal trades and the coordinated efforts of insider trading rings, the SEC frees detection and investigation resources. This, in turn, allows for more investigations, thereby reducing the load on its triage system. But for those cases where links are not readily apparent, the Division continues to face "one of the most challenging issues" in enforcement.[53] Accordingly, while the SEC has made important steps towards achieving omnipresence and resource efficiency, it continues to suffer from its bottleneck at the investigation stage.[54]

But the bottleneck may not exist for long. Advances in link prediction technology may offer a way to predict connections to insiders by relying on traders' social network data and similar information. If the SEC can incorporate this technology, it should conserve investigation resources, allowing for additional investigations. Moreover, it should redeem a number of cases that would otherwise be left on the cutting room floor in the current triage system. But before addressing the practical capabilities and legal ramifications of using link prediction technology, in the name of plenitude, the basics of link prediction and its corollaries must be explained.

III. Detecting Links: Link Prediction, Data Synthesis, and Link-Strength Prediction

A. Introduction to Network Analytics

Imagine your social network. It consists of connections to friends, family, and acquaintances. At only two degrees of separation, you are connected to every friend, family member, and acquaintance of your friends, family members, and acquaintances. Taking this to its logical conclusion, if one were not limited by degrees of separation, it would be odd not to find a connection between any two people on Earth. When one considers that meaningful connections may exist between people not only through social relationships but also through other connections—for example, two strangers working for the same company—the paths seem infinite.

Algorithms are adept at tracing the shortest connection between two points on a network using breadth-first bilateral analysis.[55] With enough information, a program could detect a trader's link to an insider with little difficulty. Unfortunately for the SEC, information does not come freely, nor is it guaranteed to come in a form that will mesh easily for analysis.

---

53. White, *supra* note 22.

54. *See supra* note 12 and accompanying text.

55. "Breadth first" refers to a search pattern that establishes all branches at one degree of separation before moving to the next degree. Prateek Garg, *Breadth First Search*, HACKEREARTH, https://www.hackerearth.com/practice/algorithms/graphs/breadth-first-search/tutorial/ [https://perma.cc/P2W2-32MY].

Moreover, even if the veil can be lifted enough to find a link, it may not be the correct link. Nevertheless, these barriers have recently started to crumble as researchers in link prediction, data synthesis, and link-strength prediction continue to hone a capable network-analysis toolkit that is well-suited to insider trading investigation.

B. Link Prediction

1. What Is Link Prediction?—Link prediction, in its most formal sense, refers to predictions as to where links will be made in the future. However, the term is also commonly used to refer to existing-link prediction—sometimes called inferring missing links.[56] For the purpose of this Note, I refer to existing-link prediction simply as link prediction, the form relevant to an insider trading investigation.

2. The Context in Which Link Prediction Is Normally Discussed.—Link prediction is often discussed in the context of social media, ostensibly because social media provides researchers with ample information with which to experiment.[57] Indeed, being able to better recommend friends on Facebook would be a very obvious use of link prediction algorithms. However, no link prediction study worth its salt misses the opportunity to propose a novel application for link prediction. Notable examples include filling in information gaps in web-crawling software sweeps,[58] finding

56. *See, e.g.*, Michael Fire et al., *Computationally Efficient Link Prediction in a Variety of Social Networks*, 5 ACM TRANSACTIONS INTELLIGENT SYS. & TECH. 10:1, 10:1 (2013) (using the terms interchangeably). Fortunately for the SEC, attempts to predict existent links based upon present data have fared far better than attempts to predict future links based upon present data. This is in part because existent link prediction relies on a mix of structural and nonstructural information, rather than only on information regarding structural evolution, and in part because as time progresses, actual links develop linearly while negative links (links that do not occur) grow quadratically. Ahmad Sadraei, *Link Prediction Algorithms: What Will Facebook Friendships Look Like Tomorrow?* (Spring 2014), http://be.amazd.com/link-prediction/ [https://perma.cc/KR5T-YZL4]; *see also* David Liben-Nowell & Jon Kleinberg, *The Link-Prediction Problem for Social Networks*, 58 J. AM. SOC'Y INFO. SCI. & TECH. 1019, 1030 (2007) (obtaining only 16% accuracy predicting future links).

57. For example, one study ran tests using the members of pet-enthusiast websites Dogster, Catster, and Hamsterster. *See* Elena Zheleva et al., *Using Friendship Ties and Family Circles for Link Prediction*) ("On these sites, profiles include photos, personal information, characteristics, as well as membership in community groups. Members also maintain links to friends and family members."), *in* ADVANCES IN SOCIAL NETWORK MINING AND ANALYSIS 97, 103 (Lee Giles, Marc Smith, John Yen & Haizheng Zhang eds., 2010).

58. Fire et al., *supra* note 56, at 10:2.

interactions between proteins,[59] and identifying members of terrorist organizations.[60]

Although link prediction papers often note a potential application in criminal network detection, there has been little in the way of actual studies.[61] Among those studies that have discussed the use of link prediction in criminal investigations, its application to insider trading has only been mused about in passing.[62] This is regrettable because unlike the criminal-investigation applications that garner attention—namely, filling out the margins of larger conspiracies or detecting terrorist networks—insider trading investigations are particularly poised to benefit from link prediction.[63] As discussed above, these investigations increasingly begin with nothing more than a suspicious trade. Consequently, the case will rise or fall on whether a connection to inside information can be established or, just as important in the case of the SEC's triage system, how much effort will be required to establish a connection.

3. How It Works.—This Note will not attempt to explain the intricacies of link prediction.[64] However, a basic understanding of link prediction is necessary to understand how to apply link prediction to an insider trading investigation. At the highest level of abstraction, link prediction systems

59. *Id.*

60. Emrah Budur et al., *Structural Analysis of Criminal Network and Predicting Hidden Links Using Machine Learning*, ARXIV (Sept. 21, 2015), https://arxiv.org/pdf/1507.05739.pdf [https://perma.cc/PGJ6-GPN3].

61. Giulia Berlusconi et al., *Link Prediction in Criminal Networks: A Tool for Criminal Intelligence Analysis*, 11 PLOS ONE, no. 4, 2016, at 1, 2–3, http://journals.plos.org/plosone/article/file?id=10.1371/journal.pone.0154244&type=printable [https://perma.cc/3A56-2GQD] ("This is surprising, not only given the significant growth of works on missing links in other fields [and] with the development of a number of different strategies, but also given that criminal investigations face the problem of missing links almost by definition, due to the scarcity of investigative resources and the antidetection strategies by criminals." (footnotes omitted)).

62. *See* Andrew Arnold & William W. Cohen, Information Extraction as Link Prediction: Using Curated Citation Networks to Improve Gene Detection 2 (Jan. 1, 2009) (unpublished manuscript) (on file with the National Center for Biotechnology Information), https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3018763/pdf/nihms191712.pdf [https://perma.cc/4G6N-H8M2] (noting a potential application in detecting "insider trading cabals").

63. Berlusconi et al., *supra* note 61, at 3 (noting that the role of link prediction in law enforcement is not only a means of detecting the otherwise undetectable, but also a means to conserve "scarce investigative resources").

64. I am as relieved as you are.

work by predicting the existence or nonexistence of links based upon the existence of "features" that are similar to those they have seen before.[65]

Network features that speak to the network structure itself are called topographical features.[66] We almost always envision a network as a topographical model of the connections it contains. A topographical model can conceptualize a few different features. First, the topographical features can be comprised of dyadic relationships (e.g., person-to-person links with each person representing a "node") or of spheres of group membership, or a combination of the two.[67] Looking for commonality between two nodes based on the topographical features of their individual networks is often referred to as a test of "structural equivalence." As a rudimentary example of structural equivalence, consider that criminal networks are often arranged into a hierarchy.[68] It might surprise no one to find that two known criminals who live in the same city and whose social-network topographies contain traces of a similar hierarchical structure are part of the same criminal organization.[69]

But beyond topography, link predictors also consider descriptive features: those without regard to network structure.[70] A rudimentary example of a descriptive feature might be the dog one prefers. Again, it might surprise no one that two bichon frise enthusiasts are both connected to each other through a bichon frise-centric social group.[71]

4. Study Results.—Despite some studies' very impressive results, it is not called the "link-prediction problem" for no reason. While existent link prediction at one degree of separation between nodes has been shown to be a

65. Machine learning has become the prevailing methodology for link prediction. Machine learning refers to systems that are designed to generalize from the information they have previously processed to make inferences based upon new data. *E.g.*, Fire et al., *supra* note 56, at 10:2 (training machine-learning classifiers "on a set of easy-to-compute topological features"); *cf. id.* (using backward-looking Bayesian methods as an alternative or complement to supervised machine learning). One machine-learning study suggests that for social-network prediction, "near-maximal [area under the curve (AUC)] can be obtained by using training sets with up to 50,000 examples." *Id.* at 10:13. AUC under the receiver operating characteristic (ROC) curve is the "de facto performance measure for link prediction tasks." *Id.* at 10:6.

66. For example, the number of people one knows is considered a topographical feature.

67. Zheleva et al., *supra* note 57, at 99.

68. Budur et al., *supra* note 60, at 2.

69. Link prediction algorithms actually consider more nuanced topographical features—for example, "Jaccard's coefficient," which reflects link density around a group of individuals. Fire et al., *supra* note 56, at 10:8–10:9.

70. *See* Zheleva et al., *supra* note 57, at 103–04 (considering pet breed as a feature).

71. However, link prediction algorithms are not limited to the most glaring of features, and they will often consider hundreds of descriptive features. *See, e.g.*, Ben Taskar et al., *Link Prediction in Relational Data*, *in* ADVANCES IN NEURAL INFORMATION PROCESSING SYSTEMS 16, at 659, 663 (2003) (utilizing almost 1,000 features).

simple feat,[72] accurately predicting multi-degree links (e.g., based on X's data, X should know Y, and Y should know Z) is far more difficult, although still possible.[73]

C. *Data Synthesis*

Today, we store an incredible amount of electronic data, and our stockpile grows exponentially. By one account it doubles in volume every two years at the current rate of five quintillion bytes produced every two days.[74] Practically hemorrhaging data in our daily lives, information from our phone calls, credit card purchases, and internet usage (emails, social media, web surfing, etc.) is stored in phone company, bank, and internet service provider (ISP) records, among other places.[75] The information stored contains not only content (e.g., the text of an email), it also contains transactional (non-content) metadata that describes dates, times, people, and locations.[76] Collectively, this information is referred to as "big data."[77] In this data there is a seemingly infinite amount of information about an individual's social network topography, as well as descriptive information. However, while all the information may be available, taking it as it comes would be akin to drinking from a fire hose. Link prediction system designers must perform a balancing act between capturing the most data possible and doing so in a way that will be both accurate and usable. As discussed below, the SEC has access to a tremendous wealth of transactional metadata. And therein lies the difficulty.

Fortunately, link prediction researchers are on the job. Data fusion is the process of synthesizing data from multiple sources so that the information is not only digestible, but also achieves synergistic effect.[78] This is done by compiling disparate data into matrices and assigning weight to data types and combinations in order to create a standardized, consolidated output.[79] This

---

72. *See* Indika Kahanda & Jennifer Neville, *Using Transactional Information to Predict Link Strength in Online Social Networks*, *in* PROCEEDINGS OF THE 3RD INTERNATIONAL AAAI CONFERENCE ON WEBLOGS AND SOCIAL MEDIA 74–81 (2009) (achieving 87% AUC).

73. Fire et al., *supra* note 56, at 23 (stating that even when people are "two hops from each other," predictions are able to distinguish between friends and nonfriends).

74. Andrew Guthrie Ferguson, *Big Data and Predictive Reasonable Suspicion*, 163 U. PA. L. REV. 327, 354 (2015).

75. *Id.* at 355–56.

76. *Id.*

77. *Id.* at 352.

78. Kahanda & Neville, *supra* note 72, at 75.

79. For example, imagine one is trying to detect who is the better friend of X, Y, or Z. Phone records may show two calls from X to Y and one from X to Z, while no emails have been sent between X and Y, but two have been sent between X and Z. Now assume that test data shows that not only are phone calls a stronger indication of friendship than emails, but that the presence of phone calls combined with an absence of emails indicates a very close friendship. If the information is weighted accordingly, we can conclude that X and Y have a higher degree of friendship than X

process not only provides a means of multi-source information management, but can also be used to render more refined factors that are tailored to the link prediction task.

D. Link-Strength Prediction

As rewarding as building a better link predictor might be, those who study link prediction recognize that finding a link does not alone guarantee a practical result.[80] This fact has led some legal-policy-focused writers to doubt the successful application of link prediction in criminal investigation contexts. "People are related to criminals without being criminals themselves," one author writes.[81] "In a wired world, people are more closely connected than we think."[82] This poses, he argues, severe risk of including false-positive links at only a few degrees of separation.[83] However, researchers have not been so easily discouraged by the well-known colloquialism that everyone is six degrees from everyone else. Studies also focus on link-strength prediction, which estimates link importance rather than existence.[84] Not coincidentally, just as links can be predicted through topographical and descriptive information, so too can that information be used to predict link strength.[85]

But while factors like frequency of contact between two people certainly indicate a strong connection in an objective sense, when the goal is to filter connections among co-conspirators from innocent friendships, a different set of criteria is required. One particular technique has been touted for its usefulness in flagging the sort of links one might expect in criminal networks. It is based on the theory that the features of criminal networks are naturally unlike the features of normal social networks. Thus, connections that would otherwise be irrelevant (for example, a connection to one's grandmother) can be used to infer *a contrario* criminal links.[86]

---

and Z. The comparative degree of friendship, then, can be used as a factor for link prediction in lieu of, or in addition to, the raw data.

80. *See* Kahanda & Neville, *supra* note 72, at 74 (pointing out the issue of "spurious relationships").

81. Ferguson, *supra* note 74, at 409.

82. *Id.* (citing, among other things, a study showing that among Facebook users the "average number of acquaintances separating any two people in the world was not six [as is often said colloquially] but 4.74").

83. *Id.* ("Thus, a link analysis . . . can cast a very wide net, accidentally capturing many people who are only suspicious by this loose, associative relationship.").

84. Kahanda & Neville, *supra* note 72, at 74.

85. *Id.* at 74–75.

86. Berlusconi et al., *supra* note 61, at 18.

IV. Where the Rubber Meets the Road

In an insider trading investigation, link prediction would naturally be used to predict a suspect's inside connections by analyzing information investigators could expect to easily obtain. Thus, given that the technological restraints preventing practical applications of link prediction are fading, one must now ask, how much information can investigators expect to easily obtain? The answer is an astronomical amount.

Today, the easiest way to learn about a stranger is through Google. People regularly leave a trail of personal information online—for example, in public postings on social media, in work profiles on company websites, and in other various and sundry locales across the internet. However, to flesh out a social network to the degree required for link prediction, more data is needed. Company records, phone-company and ISP records, and personal-computer and cell-phone memory contain a treasure trove of data—both transactional and content. Getting access to that information poses surprisingly few legal hurdles. First, under the third-party doctrine exception to the applicability of the Fourth Amendment, transactional data held by phone companies and ISPs can be accessed with little fanfare, often with only an administrative subpoena. Second, obtaining information housed in company records, personal computers, and cell phones requires only a warrant.

But despite an optimistic forecast, courts and the Constitution are somewhat hostile to the use of statistical data in certain contexts. First, there is some debate as to whether statistical evidence alone—such as a statistical probability that a trade was an illegal insider trade—can be the basis for a warrant. In addition, there are potential due process concerns that arise from being investigated and prosecuted using proprietary probabilistic algorithms. Finally, there is considerable opposition to using probabilistic evidence to support a verdict. In the end, these restrictions define the worst-case scenario for the SEC, should it decide to adopt link prediction technology.

A. Acquiring Transactional Data: Administrative Subpoenas, the Constitution, and the Stored Communications Act

When building the topographical map emanating from a suspected insider trader, transactional data is king. While transactional data is certainly contained on private devices, and the seizure of those devices would require a warrant (discussed in Section C), a subpoena to produce those records does not. Furthermore, a sea of transactional data is kept in third-party storage (e.g., phone company records), vulnerable to administrative subpoena per the terms of the Stored Communications Act (SCA).

1. The SEC's Power to Issue Administrative Subpoena.—The SEC is empowered to commence any investigation "it deems necessary to determine whether any person has violated, is violating, or is about to violate any

provision of this chapter, the rules or regulations thereunder, the rules of a national securities exchange[, etc.] . . . ."[87] To further its investigations, it is authorized to compel the production of documents through administrative subpoena.[88]

For a court to give a subpoena the weight of law, the Supreme Court has stated that if "the investigation [is] conducted pursuant to a legitimate purpose, . . . the inquiry [is] relevant to the purpose, . . . the information sought is not already within the [agency's] possession, and . . . the administrative steps required by [statute] have been followed," the court should defer to the agency notwithstanding evidence of abuse.[89] Although the Supreme Court last addressed the issue in 1946, lower courts have continued to adhere to this policy of wide deference for administrative subpoenas.[90] For example, the Second Circuit in *In re McVane* held that "[a]n affidavit from a government official is sufficient to establish a prima facie showing that [the] requirements [for subpoena] have been met."[91] The Second Circuit continued,

> The courts' role in a proceeding to enforce an administrative subpoena is "extremely limited." We defer to the agency's appraisal of relevancy, which "must be accepted so long as it is not obviously wrong." . . . The relevance of the sought-after information is measured against the general purposes of the agency's investigation, "which necessarily presupposes an inquiry into the permissible range of investigation under the statute." . . . We have interpreted relevance broadly.[92]

Under the Second Circuit's policy of deference, the requirement that the information sought be relevant to an investigation is a particularly low bar. Assuming investigators can accurately pinpoint trades too fortuitous to have occurred without reliance on nonpublic information, it is hard to imagine that traders' communications records would not be relevant to an investigation into a violation of the Securities Exchange Act.

---

87. 15 U.S.C. § 78u(a)(1) (2012).

88. § 78u(b) reads,

> For the purpose of any such investigation, or any other proceeding under this chapter, any member of the Commission or any officer designated by it is empowered to administer oaths and affirmations, [subpoena] witnesses, compel their attendance, take evidence, and require the production of any books, papers, correspondence, memoranda, or other records *which the Commission deems relevant or material to the inquiry*.

Id. § 78u(b) (emphasis added).

89. United States v. Powell, 379 U.S. 48, 57–58 (1964).

90. *See, e.g.*, *In re* McVane, 44 F.3d 1127, 1135–36 (2d Cir. 1995) (acknowledging the court's limited role in enforcing an administrative subpoena).

91. *Id.* at 1136.

92. *Id.* at 1135–36.

2. A Lack of Constitutional Constraint.—Regarding any Fourth Amendment constraint on administrative subpoena power, the Supreme Court stated in 1946 that demands to produce records "present no question of actual search and seizure, but raise only the question whether orders of court for the production of specified records have been validly made."[93] After all, in a demand for production, "[n]o officer or other person [seeks] to enter . . . premises against [a person's] will, to search them, or to seize or examine their books, records or papers without their assent . . . ."[94] Moreover, for those subpoenas that have been taken to court for enforcement, the producing party has "adequate opportunity to present objections."[95] Thus, at most, the Court concluded, "the Fourth, if applicable, . . . guards against abuse only by way of too much indefiniteness or breadth" or by not being "relevant" to a legitimate inquiry.[96]

As for protesting a subpoena issued for data stored on third-party servers (e.g., phone records and emails), the Fourth Amendment offers no protection at all. The third-party doctrine, which states that one gives up any reasonable expectation of privacy when one gives information to a third party, leaves the subject of the transactional records without a constitutional leg to stand on.[97] The originator and recipients of the communications are not even required to have notice of the subpoena.[98]

As for the Fifth Amendment's protections against self-incrimination, the Supreme Court has stated, "in so far as [subpoenas] apply merely to the production of . . . records . . . [it] seems to be that the Fifth Amendment affords no protection by virtue of the self-incrimination provision."[99]

3. Statutory Constraints on Administrative Subpoenas: the SCA.—In the name of privacy, Congress enacted the Stored Communications Act, introducing statutory restraints on access to electronically stored third-party data.[100] Accordingly, the SEC must comply with the terms of the SCA if it is

---

93. Okla. Press Pub. Co. v. Walling, 327 U.S. 186, 195 (1946).

94. *Id.*

95. *Id.*

96. *Id.* at 208.

97. SEC v. Jerry T. O'Brien, Inc., 467 U.S. 735, 743 (1984) states,

> It is established that, when a person communicates information to a third party even on the understanding that the communication is confidential, he cannot object if the third party conveys that information or records thereof to law enforcement authorities. . . . These rulings disable respondents from arguing that notice of subpoenas issued to third parties is necessary to allow a target to prevent an unconstitutional search or seizure of his papers.

98. *Id.* at 742–43.

99. *Walling*, 327 U.S. at 208. However, the viability of a Fifth Amendment *due process* claim is discussed in a later section.

100. 18 U.S.C. § 2702 provides in part,

> a person or entity providing an electronic communication service to the public shall not knowingly divulge to any person or entity the contents of a communication while

to gain access to the wealth of transactional data held by phone companies and ISPs in order to satiate its need for topographical data. Fortunately for the SEC, the SCA provides a few exceptions,[101] including a mandate that government entities be given access to non-content data when presented with an administrative subpoena.[102] Again, giving notice to the subject of the investigation is not required.[103]

4. Final Thoughts on Transactional Data Acquisition.—When it comes to suspects' transactional data, the data most valuable to the success of link prediction systems, the SCA opens the door to a veritable smorgasbord. Armed with only a red flag marking a suspicious trade and a stack of subpoenas, investigators should be able to build a substantial topographical map surrounding the suspected insider trader without the trader's knowledge. When that network is coupled with a list of company employees and business relationships, a well-trained link prediction system should be able to detect simple, and potentially complex, hidden links, telling investigators where to start digging.

B. Fourth Amendment Seizures, and Probabilistic Probable Cause

Unlike for transactional data, the SCA provides increased protection for stored content, such as the body of an email. But even then, protections are minimal. The statute requires a warrant to access content in storage for less than 180 days,[104] but content stored for longer may be requested through administrative subpoena with prior notice.[105] This sweeps older emails, voicemails, text messages, and social media communication into the same easy-access category as transactional data. Nevertheless, during an investigation, waiting 180 days may not be feasible. Moreover, other important information—for example, electronic address books—may only be stored locally on privately owned devices, requiring a warrant to access. With warrants come more robust protections from the Fourth Amendment. This in

---

> in electronic storage by that service; and . . . a provider of remote computing service or electronic communication service to the public shall not knowingly divulge *a record or other information pertaining to a subscriber to or customer* of such service . . . *to any governmental entity.*

18 U.S.C. § 2702 (2016) (emphasis added).

101. *Id.* § 2702(b)–(c).

102. *Id.* § 2703(c)(2) provides that companies storing communication data,

> shall disclose to a governmental entity . . . when the governmental entity uses an administrative subpoena . . . [a subscriber's] name; address; local and long distance telephone connection records, or records of session times and durations; length of service (including start date) and types of service utilized; telephone or instrument number or other subscriber number or identity . . . .

103. *Id.*

104. *Id.* § 2703(a).

105. *Id.* § 2703(b).

turn raises concerns about finding probable cause based solely on probabilistic statistical data, such as a statistical probability that a trade was illegal.

1. The Requirements of Probable Cause.—The crux of reasonable search and seizure analysis under the Fourth Amendment is probable cause.[106] The probable cause standard exists to accommodate both "individual liberty and public security/crime prevention."[107] The equivalent standard for law enforcement stops is "reasonable suspicion."[108] While the threshold required for reasonable suspicion is generally agreed to be less than probable cause, the considerations are the same. As the Constitution protects against unreasonable search and seizure, the inquiry turns on objective reasonableness.[109] This in turn requires a look at the "totality of the circumstances" of the particular case to see whether there is a prospective "particularized and objective basis" for a search or seizure.[110] A search that was not justified *ex ante* cannot be legitimated by what was found during the search.[111]

Although the Court describes a seemingly unrestricted totality of the circumstances test to satisfy the Fourth Amendment's reasonableness requirement, one particular type of evidence must be present: there must be some evidence of wrongdoing that is particular to the individual.[112] For example, in *Florida v. J.L.*,[113] an anonymous call "reported to the Miami-Dade Police that a young black male standing at a particular bus stop and

106. U.S. CONST. amend. IV ("[N]o Warrants shall issue, but upon probable cause."); *Katz v. United States*, 389 U.S. 347, 357 (1967) ("[S]earches conducted outside the judicial process, without prior approval by judge or magistrate, are per se unreasonable under the Fourth Amendment—subject only to a few specifically established and well-delineated exceptions . . . .").

107. Erica Goldberg, *Getting Beyond Intuition in the Probable Cause Inquiry*, 17 LEWIS & CLARK L. REV. 789, 795 (2013).

108. The phrase "reasonable suspicion," however, was first articulated in a dissent. Terry v. Ohio, 392 U.S. 1, 37 (1968) (Douglas, J., dissenting).

109. Ferguson, *supra* note 74, at 341–42.

110. United States v. Arvizu, 534 U.S. 266, 273 (2002).

111. Ferguson, *supra* note 74, at 339.

112. Interestingly, this is not true for stops made pursuant to a policy to stop all passersby—for example, a DWI checkpoint on a public road. In *Michigan Department of State Police v. Sitz*, the Supreme Court, without acknowledging an individualized-evidence requirement, held DWI checkpoints reasonable under the Fourth Amendment. 496 U.S. 444, 455 (1990). In doing so, it performed a balancing test, weighing the strong state interest in preventing drunk driving, the proceduralized nature of the checkpoint, and evidence of drunk drivers on Michigan roads against the minimally invasive nature of the stop. *Id.* Justice Brennan, however, took issue with the majority divorcing checkpoint stops from traditional *Terry*-stop requirements, stating that "individualized suspicion is a core component of the protection the Fourth Amendment provides against arbitrary government action." *Id.* at 457 (Brennan, J., dissenting). Here, however, we discuss individual investigations—not checkpoints. Accordingly, *Sitz* should be noted as a strange, but ultimately irrelevant, exception to the individualized-evidence requirement.

113. 529 U.S. 266 (2000).

wearing a plaid shirt was carrying a gun."[114] Based only on this tip, police searched a young black man who was wearing a plaid shirt.[115] Finding a gun, they arrested him.[116] The Supreme Court held that the anonymous tip, which contained no information particular to that individual—only that a black man in a plaid shirt had a gun, was insufficient for reasonable suspicion.[117] Conversely, in *Alabama v. White*,[118] the Supreme Court held that a tip that included the suspect's name, vehicle, and destination was sufficient to satisfy the Fourth Amendment.[119] The individualized-evidence requirement creates problems for the use of statistical information to clear Fourth Amendment constraints. This is discussed below.

2. Using Probabilities for Probable Cause.—There are grave concerns about the ability of probabilistic evidence to support a verdict (discussed below). However, the core inquiry behind the Fourth Amendment's protections against unreasonable search and seizure dovetails nicely with probabilistic information.[120] Under the probable cause or reasonable suspicion standard, the inquiry focuses on the likelihood of the existence of a fact,[121] a purely probabilistic notion. Nevertheless, many have argued the merits and strikes against using solely probabilistic evidence to satisfy the Fourth Amendment.[122] Moreover, scholars have questioned whether big-data predictions can satisfy the individualized-evidence requirement.[123] To date, there is no word from the Supreme Court as to whether probabilistic evidence alone can satisfy the Fourth Amendment.

a. The Argument for Probabilistic Probable Cause.—It is hard to imagine a decision regarding a future outcome that is not made based upon a probability. Recent research on decision-making suggests that humans come to conclusions by entertaining multiple narratives, reevaluating their likelihood based on information as it is received.[124] Even older theories of human decision-making were based around the idea that humans make decisions based on probabilities that are continuously recalculated using new

---

114. *Id.* at 268.

115. *Id.*

116. *Id.* at 268–69.

117. *Id.* at 271 (stating that there was no evidence specific to the defendant).

118. 496 U.S. 325 (1990).

119. *Id.* at 327, 331–32.

120. Max Minzner, *Putting Probability Back into Probable Cause*, 87 TEXAS L. REV. 913, 957–58 (2009) (discussing probability as being integral to probable cause).

121. Brinegar v. United States, 338 U.S. 160, 175–76 (1949) ("In dealing with probable cause . . . as the very name implies, we deal with probabilities.").

122. Andrew Guthrie Ferguson, *Predictive Policing and Reasonable Suspicion*, 62 EMORY L. J. 259, 298–99 (2012).

123. Ferguson, *supra* note 74, at 330.

124. Minzner, *supra* note 120, at 953–54.

information.[125] Some have argued that people's resistance to relying on purely probabilistic information to make legal determination stems not from the inferiority of probabilistic information; it stems from the fact that the chance of being incorrect is plainly apparent—not buried beneath layers of human intuition and judgment.[126] To that end, an argument for using probabilistic evidence is that one must confront error rates as they come, and if they are unsatisfactory, pursue a more accurate result.

Moreover, we routinely use probabilistic evidence to satisfy the Fourth Amendment without regarding it as such. For example, an arrest may be premised on a drug dog's sniff or a breathalyzer reading even though dogs identify false-positive smells and breathalyzers have margins of error.[127] Consequently, it is illogical to declare certain probabilities an affront to the Constitution while others remain widely used in practice.

Finally, the use of quantitative empirical data reduces reliance on intuition, which infuses the constitutionally mandated question of reasonable likelihood with subjective valuations. Even the standard of probable cause itself suffers from subjectivity issues. For example, in one study, a professor asked federal judges to quantify probable cause.[128] The results ranged from ten to ninety percent.[129] Adding objectively determined probabilities to the mix would help ameliorate what is, at its core, an error-prone inquiry.

b. The Argument Against Probabilistic Probable Cause.—The main strike against probabilistic probable cause is that the Supreme Court has emphatically stated that "[t]he probable-cause standard is incapable of precise definition or quantification into percentages"[130] because, at its core, "it deals with probabilities and depends on the totality of the circumstances."[131] In other words, it is to be based upon "the factual and practical considerations of everyday life on which reasonable and prudent

125. Dan Simon, *A Third View of the Black Box: Cognitive Coherence in Legal Decision Making*, 71 U. CHI. L. REV. 511, 538–39 (2004).

126. Daniel Shaviro, *Statistical-Probability Evidence and the Appearance of Justice*, 103 HARV. L. REV. 530, 533–34 (1989).

127. *See, e.g.*, Illinois v. Caballes, 543 U.S. 405, 410 (2005) (Souter, J., dissenting) (presuming that a drug dog's sniff can be the sole basis for a search); United States v. Waltzer, 682 F.2d 370, 372 (2d Cir. 1982) (holding that a drug dog's alert "itself establish[es] probable cause, enough for [an] arrest"); *see also* Goldberg, *supra* note 107, at 791 ("Whether or not probable cause exists to issue the warrant depends largely on [the dog's] reliability, which can be quantified based on [the dog's] error rate in detecting drugs."). Given holdings like *Caballes* and *Waltzer*, it is hard to believe courts would question arrests made on the basis of positive breath tests on the grounds that the test results are probabilistic.

128. Ric Simmons, *Quantifying Criminal Procedure: How to Unlock the Potential of Big Data in Our Criminal Justice System*, 2016 MICH. ST. L. REV. 947, 987–88 (2016).

129. *Id.*

130. Maryland v. Pringle, 540 U.S. 366, 371 (2003).

131. *Id.*

men, not legal technicians, act."[132] Again, in other words, the probability component is human-derived—meant to be filtered through the reasonable actor.[133] Only at that point is the inquiry reduced to probability alone and, even then, only of a sort that is unrestrained by a need to assign a particular percentage. For that reason, one might argue a probable cause inquiry based solely on probabilistic information undermines the role of the reasonable actor by reducing what was intended to be an unconstrained evaluation of the totality of the circumstances into a pure numbers game.

Nevertheless, the Supreme Court's refusal to constrain probable cause to a number does not necessarily mean that probabilistic information cannot be considered nor that it cannot be the only thing considered when it is all that is available. Statistical evidence, rather than being taken as gospel, should be evaluated by investigators by questioning the credibility of its source and its relative value to the investigation, among other things.

c. The Individualized-Evidence Requirement in this Context.—A frequent complaint about probabilistic probable cause is that it can be used as a means to lawfully conduct searches and seizures based purely upon innocent behavior correlated with illegal conduct under circumstances outside the control of the subject, creating scenarios in which one's privacy rights may be deprived simply for engaging in innocent conduct.[134] It could be argued that seeking a warrant based upon an unusual trade as determined by statistical analysis of market movements constitutes such an ill. Indeed, scholars have argued that big-data predictions are an insufficient independent basis to satisfy the Fourth Amendment.[135] This, however, is an oversimplification.

While the idea of using a sweeping generality—like the street corner example above—as carte blanche to search anyone who falls within its umbrella seems patently unreasonable, in other contexts, we are perfectly comfortable relying on correlations built upon information gathered from sources other than the suspect to infer illegal conduct. As mentioned above, there are few qualms about using a drug-sniffing dog as the basis for a lawful search. However, a dog is not trained by smelling a particular suspect with and without drugs. The dog alerts its officer to a smell that has, in the past, successfully been associated with drugs. In essence, the suspect is searched

---

132. Brinegar v. United States, 338 U.S. 160, 175 (1949).

133. *See* Ferguson, *supra* note 74, at 405 (noting the Supreme Court's refusal to "quantify" the "standard of reasonable suspicion" because police officers, operating as "reasonable and prudent" persons, assess the standard).

134. For example, police officers searching every person passing by a particular street corner at night based on evidence that 90% of all people on that particular street corner at night are carrying drugs.

135. Simmons, *supra* note 128, at 950 (stating in no uncertain terms that big data is unable to provide the type of individualized information required to satisfy the Fourth Amendment).

for smelling like other circumstances the dog has been exposed to. The same can be said for breathalyzers. Because they are not calibrated by measuring the alcohol content of a particular suspect's breath, they can only speak to the similarities between the suspect and the device's other subjects.

d. Final Thoughts on Probabilistic Probable Cause.—Given that we in fact do use probabilistic evidence to clear Fourth Amendment burdens, the rule cannot be that probabilistic evidence alone is insufficient. Thus, the permissible use of probabilistic evidence must turn on other factors. One factor separating the street corner example from a breath test is the degree of accuracy. Although 90% accuracy seems high in most circumstances, perhaps it is insufficient when invading one's privacy based purely on ostensibly innocent conduct. Another separating factor is how central the conduct being scrutinized is to the offense in question.[136] The alcohol on one's breath speaks directly to intoxication whereas one's presence on a street corner by itself does not implicate drug possession. Yet another, some have argued, is the amount of individualized information used as the basis for the prediction.[137] However, this does not explain the acceptance of breathalyzers, which only take a single sample.

If these differences are meaningful, there should be no reason why statistical evidence of a trade too fortuitous to have been legal cannot pass muster under the Fourth Amendment, allowing investigators to obtain a warrant to search the personal devices of suspected insider traders. Provided that market-monitoring systems are sufficiently accurate, the conduct being scrutinized is the trade suspected of yielding fraudulent gains, and it is being critiqued in part by the magnitude of those gains.[138]

C. Due Process Concerns

Generally, the Fourth Amendment is said to displace Fifth Amendment due process claims arising from searches or seizures because the Fifth Amendment is not intended to serve as a vortex for substantive rights housed elsewhere in the Constitution. But some have argued that the Fourth Amendment might not always dictate the constitutional propriety of pulling

136. *See* Ferguson, *supra* note 74, at 338–39 (noting that "[s]uspicion must be particularized," relating to the "current criminal activity").

137. *Id.* at 330.

138. Of course, this is not the only level of abstraction from which to view the inquiry. One might just as easily recast the inquiry as scrutinizing profit—a perfectly legal and desirable outcome. However, this does not serve to distinguish it from the breathalyzer example. Having alcohol on one's breath is not itself illegal. Illegality comes from the level of intoxication or the amount of alcohol in one's blood. Breathalyzers attempt to predict blood-alcohol content by using, as their sole criterion, the alcohol on a subject's breath. Similarly, market analysis uses profit as one of many factors to predict illegal trading behavior.

data.[139] There are some circumstances where employing algorithmic data in the investigation and prosecution of a crime would implicate conceptions of fairness rather than reasonableness. In those circumstances, where the right being protected is in fact the right to due process, investigators' efforts may be constrained, not by the Fourth Amendment, but by the Fifth.

The foremost due process concern is a lack of transparency, which may hide a multitude of sins. As one scholar decried, with an algorithm "[t]here is no notice, no opportunity to be heard, no confrontation with adverse evidence, and no reason given—only a result."[140] So too has Justice Ginsburg expressed concern, writing,

> Inaccuracies in expansive, interconnected collections of electronic information raise grave concerns for individual liberty. The offense to the dignity of the citizen who is arrested, handcuffed, and searched on a public street simply because some bureaucrat has failed to maintain an accurate computer data base is evocative of the use of general warrants that so outraged the authors of our Bill of Rights.[141]

Indeed, the questions of fairness are many: If it took a computer two hours to do what a team of investigators would have taken weeks to do, how can defendants hope to unwind the steps taken by an algorithm to mount a defense against the information it found valuable? Do defendants receive due process when the building blocks of their criminal cases are only rebuttable through a battle of experts? Moreover, police are not allowed to recklessly withhold exculpatory evidence from an affidavit for a warrant.[142] What if a program was recklessly designed to miss certain exculpatory evidence? How would anyone know that such evidence was not taken into account?

There are no answers to these questions.[143] However, each implicates the fundamental sense of fairness that undergirds so much of our criminal process. For that reason, the Fifth Amendment might provide a means to protest the use of such evidence despite the general rule that searches and seizures are best analyzed under the Fourth. As to the extent defendants should expect to find shelter in the Fifth Amendment, at this point, it is too early to tell.

---

139. Daniel J. Steinbock, *Data Matching, Data Mining, and Due Process*, 40 GA. L. REV. 1, 45–46 (2005).

140. *Id.* at 45.

141. Herring v. United States, 555 U.S. 135, 155–56 (2009) (Ginsburg, J., dissenting) (internal quotation marks omitted).

142. *See, e.g.*, Whitlock v. Brown, 596 F.3d 406, 410–11 (7th Cir. 2010) (examining whether a policeman recklessly withheld evidence from a particular warrant application).

143. Although, for a very thorough *Mathews v. Eldridge* analysis of big data analytics, see Steinbock, *supra* note 139, at 54–62.

D. Supporting a Verdict on Analytics Alone?

Although there has been no word from the United States Supreme Court, various courts have held that probabilistic evidence alone cannot support a verdict.[144] This somewhat-accepted rule is derived from the proposition that fact finders must determine the existence of facts, not conclude that facts have a probability of existing.[145] To some this may seem absurd—after all, the preponderance of the evidence standard is regularly described as a finding that something was "more likely than not." However, proponents of the rule argue that burden of proof standards exist only to explain the level of confidence one must have in one's conclusion—whether a fact has or has not been proven—not to create a probability threshold to stand in the place of an actual determination.[146] Admittedly, this sounds like splitting hairs. However, this debate makes it difficult to hope that insider trading prosecutors will be able to reliably secure a verdict by proffering only a prediction that a trade was based on insider information and a prediction that the information was obtained from a particular source to satisfy two of the necessary elements of the action.

Conclusion: The Worst-Case Scenario

Should the SEC adopt link prediction technology as a means to investigate suspicious trades that do not readily appear to have inside connections, the worst-case scenario is that its application will be limited to reconnoitering suspects' networks using only transactional third-party data and the data investigators can procure from willing companies. Investigators will be unable to use market analysis or link predictions as a means to secure a warrant, and prosecutors will be unable to use that data to sustain a verdict. But even with these limitations, the SEC is poised to benefit greatly from link prediction. When faced with a suspicious trade with no discernable connection to an insider, investigators will have a means to conduct a low-cost preliminary case audit before deciding to proceed further. Furthermore, if a link is predicted, investigators will have a direction in which to move. This will allow the SEC to pursue cases that might otherwise have been cut

144. For a discussion, see Minzner, *supra* note 120, at 956–57.

145. As an illustration, take for example the seminal case *Smith v. Rapid Transit Inc.*, 58 N.E.2d 754 (Mass. 1945). In that case, the plaintiff was run off the road by a bus. *Id.* at 754–55. Although the plaintiff could not identify the bus company by memory, she proffered a mathematical probability that given the scheduled bus routes, the bus was owned by the defendant. *Id.* Dismissing the case, the court stated, "A proposition is proved by a preponderance of the evidence if it is made to appear more likely or probable in the sense that actual belief in its truth, derived from the evidence, exists in the mind or minds of the tribunal notwithstanding any doubts that may still linger there." *Id.* (internal quotation marks omitted). For a pragmatic argument against this rule, see Shaviro, *supra* note 1266, at 539–40 (arguing, among other things, that requiring jurors to form beliefs introduces inaccuracy to the point that we might be better off with pure probabilities).

146. *See, e.g.*, *Smith*, 58 N.E.2d at 755 (explaining the preponderance of the evidence standard as referring to the jurors' actual belief in the truth of the evidence).

in an *ex ante* triage decision and save investigative resources that can be redirected towards pursuing additional cases. Accordingly, even if the legal bars discussed above prevent extensive reliance on link prediction as a means of enforcement, the SEC is nevertheless poised to take yet another step towards its goal of unrelenting omnipresence.

Andrew P. Van Osselaer

* * *

* * *

* * *

* * *

* * *

* * *

* * *


TEXAS LAW REVIEW
TLR
FOUNDED 1922