Texas Law Review

THE WIRES GO TO WAR:
THE U.S. EXPERIMENT WITH GOVERNMENT OWNERSHIP OF THE TELEPHONE SYSTEM DURING WORLD WAR I
Michael A. Janson & Christopher S. Yoo

REMAPPING THE PATH FORWARD:
TOWARD A SYSTEMIC VIEW OF FORENSIC SCIENCE REFORM AND OVERSIGHT
Jennifer E. Laurin

BOOK REVIEWS
Akhil Reed Amar & Sanford Levinson
Barak Orbach
Michael Scaperlanda
Jane S. Schacter
Robert F. Williams

JURY UNANIMITY AND THE PROBLEM WITH SPECIFICITY:
TRYING TO UNDERSTAND WHAT JURORS MUST AGREE ABOUT BY EXAMINING THE PROBLEM OF PROSECUTING CHILD MOLESTERS

U-TURNS ON THE ONE-WAY STREET:
PUBLIC RIGHTS AND REPRESENTATION THEORY IN PATENT VALIDITY LITIGATION

PURSUING ACADEMIC FREEDOM AFTER *GARCETTI V. CEBALLOS*

APRIL 2013 | VOL. 91 NO. 5 | PAGES 983 TO 1282

Texas Law Review

A national journal published seven times a year

Recent and Forthcoming Articles of Interest

Visit www.texaslrev.com for more on recent articles

DREAM ON: THE OBAMA ADMINISTRATION'S NONENFORCEMENT OF IMMIGRATION LAWS, THE DREAM ACT, AND THE TAKE CARE CLAUSE
Robert J. Delahunty & John C. Yoo
March 2013

COERCION, COMPULSION, AND THE MEDICAID EXPANSION: A STUDY IN THE DOCTRINE OF UNCONSTITUTIONAL CONDITIONS
Mitchell N. Berman
May 2013

Individual issue rate: $15.00 per copy
Subscriptions: $47.00 (seven issues)

Order from:
School of Law Publications
University of Texas at Austin
727 East Dean Keeton Street
Austin, Texas USA 78705
(512) 232-1149
http://www.utexas.edu/law/publications

Texas Law Review *See Also*

Responses to articles and notes found in this and other issues are available at www.texaslrev.com/seealso

PHYSICIAN HEAL THYSELF: WHITHER THE POLICE AND PROSECUTOR IN THE TALE OF FORENSIC SCIENCE GONE WRONG?
Erin Murphy

Receive notifications of all *See Also* content—sign up at www.texaslrev.com.

TEXAS LAW REVIEW ASSOCIATION

OFFICERS

NINA CORTELL
President

ERIC NICHOLS
President-Elect

AMELIA A. FRIEDMAN
Executive Director

JAMES A. HEMPHILL
Treasurer

HON. DIANE P. WOOD
Immediate Past President

BOARD OF DIRECTORS

R. DOAK BISHOP
JAMES A. COX
ALISTAIR B. DAWSON
KARL G. DIAL
GARY L. EWELL
STEPHEN FINK
DIANA M. HUDSON
DEANNA E. KING
JEFFREY C. KUBIN
D. MCNEEL LANE
LEWIS T. LECLAIR
JOHN B. MCKNIGHT
ELLEN PRYOR
CHRIS REYNOLDS
DAVID M. RODI
REAGAN W. SIMPSON
HON. BEA ANN SMITH
STEPHEN L. TATUM
MARK L.D. WAWRO

SCOTT J. ATLAS, *ex officio Director*
PARTH S. GEJJI, *ex officio Director*

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

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

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

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

The *Texas Law Review* is pleased to consider unsolicited manuscripts for publication but regrets that it cannot return them. Please submit a single-spaced manuscript, printed on one side only, with footnotes rather than endnotes. Citations should conform with *The Greenbook: Texas Rules of Form* (12th ed. 2010) and *The Bluebook: A Uniform System of Citation* (19th ed. 2010). Except when content suggests otherwise, the *Texas Law Review* follows the guidelines set forth in the *Texas Law Review Manual on Usage & Style* (12th ed. 2011), *The Chicago Manual of Style* (16th ed. 2010), and Bryan A. Garner, *A Dictionary of Modern Legal Usage* (2d ed. 1995).

© Copyright 2013, Texas Law Review Association

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

THE UNIVERSITY OF TEXAS SCHOOL OF LAW

ADMINISTRATIVE OFFICERS

WARD FARNSWORTH, B.A., J.D.; *Dean, John Jeffers Research Chair in Law.*
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.*
STEFANIE A. LINDQUIST, B.A., J.D., Ph.D.; *Associate Dean for External Affairs, Charles Alan Wright Chair in Federal Courts.*
EDEN E. HARRINGTON, B.A., J.D.; *Associate Dean for Experiential Education, Dir. of William Wayne Justice Ctr. for Public Interest Law, Clinical Professor.*
KIMBERLY L. BIAR, B.B.A.; *Assistant Dean for Financial Affairs, Certified Public Accountant.*
MICHAEL J. ESPOSITO, B.A., J.D., M.B.A.; *Assistant Dean for Continuing Legal Education.*
KIRSTON FORTUNE, B.F.A.; *Assistant Dean for Communications.*
MICHAEL HARVEY, B.A., B.S.; *Assistant Dean for Technology.*
MONICA K. INGRAM, B.A., J.D.; *Assistant Dean for Admissions and Financial Aid.*
TIM KUBATZKY, B.A.; *Interim Assistant Dean for Development and Alumni Relations.*
DAVID A. MONTOYA, B.A., J.D.; *Assistant Dean for Career Services.*
BRANDI L. WELCH, B.A., J.D.; *Interim Assistant Dean for Student Affairs.*

FACULTY EMERITI

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

PROFESSORS

DAVID E. ADELMAN, B.A., Ph.D., J.D.; *Harry Reasoner Regents Chair in Law.*
DAVID A. ANDERSON, A.B., J.D.; *Fred & Emily Marshall Wulff Centennial Chair in Law.*
MARK L. ASCHER, B.A., M.A., J.D., LL.M.; *Joseph D. Jamail Centennial Chair in Law.*
RONEN AVRAHAM, M.B.A., LL.B., LL.M., S.J.D.; *Thomas Shelton Maxey Professor in Law.*
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.*
MITCHELL N. BERMAN, A.B., M.A., J.D.; *Richard Dale Endowed Chair in Law.*
BARBARA A. BINTLIFF, M.A., J.D.; *Joseph C. Hutcheson Professor in Law, Director of Tarlton Law Library & 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.*
J. BUDZISZEWSKI, B.A., M.A., Ph.D.; *Professor.*
NORMA V. CANTU, B.A., J.D.; *Professor of Law and Education.*
LOFTUS C. CARSON, II, B.S., M. Pub. Affrs., M.B.A., J.D.; *Ronald D. Krist Professor.*
MICHAEL J. CHURGIN, A.B., J.D.; *Raybourne Thompson Centennial Professor.*
JANE M. COHEN, B.A., J.D.; *Edward Clark Centennial Professor.*
FRANK B. CROSS, B.A., J.D.; *Herbert D. Kelleher Centennial Professor of Business Law.*
WILLIAM H. CUNNINGHAM, B.A., M.B.A., Ph.D.; *Professor.*
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 Law and Philosophy.*
MECHELE DICKERSON, B.A., J.D.; *Arthur L. Moller Chair in Bankruptcy Law and Practice.*
GEORGE E. DIX, B.A., J.D.; *George R. Killam, Jr. Chair of Criminal Law.*
JOHN S. DZIENKOWSKI, B.B.A., J.D.; *Dean John F. Sutton, Jr. Chair in Lawyering and the Legal Process.*
KAREN L. ENGLE, B.A., J.D.; *Minerva House Drysdale Regents Chair in Law, Co-Director of Bernard and Audre Rapoport Center for Human Rights and Justice.*
KENNETH FLAMM, Ph.D.; *Professor.*
JULIUS G. GETMAN, B.A., LL.B., LL.M.; *Earl E. Sheffield Regents Chair.*
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.*
CHARLES G. GROAT, B.A., M.S., Ph.D.; *Professor.*
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.*
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.; *Andrews & Kurth Centennial Professor.*
EMILY E. KADENS, B.A., M.A., Dipl., M.A., Ph.D., J.D.; *Baker and Botts Professor in Law.*
SUSAN R. KLEIN, B.A., J.D.; *Alice McKean Young Regents Chair in Law.*

SANFORD V. LEVINSON, A.B., Ph.D., J.D.; *W. St. John Garwood & W. St. John Garwood, Jr. Centennial Chair in Law, Professor of Government.*
VIJAY MAHAJAN, M.S.Ch.E., Ph.D.; *Professor.*
BASIL S. MARKESINIS, LL.B., LL.D., D.C.L., Ph.D.; *Jamail Regents Chair.*
INGA MARKOVITS, LL.M.; *"The Friends of Joe Jamail" Regents Chair.*
RICHARD S. MARKOVITS, B.A., LL.B., Ph.D.; *John B. Connally Chair.*
THOMAS O. MCGARITY, B.A., J.D.; *Joe R. & Teresa Lozano Long Endowed Chair in Administrative Law.*
STEVEN A. MOORE, B.A., Ph.D.; *Professor.*
LINDA S. MULLENIX, B.A., M. Phil., J.D., Ph.D.; *Morris & Rita Atlas Chair in Advocacy.*
STEVEN P. NICHOLS, B.S.M.E., M.S.M.E., J.D., Ph.D.; *Professor.*
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 Law and Government.*
LUCAS A. POWE, JR., B.A., J.D.; *Anne Green Regents Chair in Law, Professor of Government.*
WILLIAM C. POWERS, JR., B.A., J.D.; *President of The University of Texas at Austin, Hines H. Baker & Thelma Kelley Baker Chair, University Distinguished Teaching Professor.*
DAVID M. RABBAN, B.A., J.D.; *Dahr Jamail, Randall Hage Jamail & Robert Lee Jamail Regents Chair, University Distinguished Teaching Professor.*
ALAN S. RAU, B.A., LL.B.; *Mark G. & Judy G. Yudof Chair in Law.*
DAVID W. ROBERTSON, B.A., LL.B., LL.M., J.S.D.; *W. Page Keeton Chair in Tort Law, University Distinguished Teaching Professor.*
JOHN A. ROBERTSON, A.B., J.D.; *Vinson & Elkins Chair.*
WILLIAM M. SAGE, A.B., M.D., J.D.; *Vice Provost for Health Affairs, James R. Dougherty Chair for Faculty Excellence.*
LAWRENCE G. SAGER, B.A., LL.B.; *Alice Jane Drysdale Sheffield Regents Chair.*
JOHN J. SAMPSON, B.B.A., LL.B.; *William Benjamin Wynne Professor.*
CHARLES M. SILVER, B.A., M.A., J.D.; *Roy W. & Eugenia C. MacDonald 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.*
JAMES C. SPINDLER, B.A., M.A., J.D., Ph.D.; *The Sylvan Lang Professor.*
MATTHEW L. SPITZER, B.A., Ph.D., J.D.; *Hayden W. Head Regents Chair for Faculty Excellence.*
JANE STAPLETON, B.S., Ph.D., LL.B., D.C.L., D. Phil.; *Ernest E. Smith Professor.*
JORDAN M. STEIKER, B.A., J.D.; *Judge Robert M. Parker Endowed Chair in Law.*
MICHAEL F. STURLEY, B.A., J.D.; *Fannie Coplin Regents Chair.*
GERALD TORRES, A.B., J.D., LL.M.; *Bryant Smith Chair in Law.*
GREGORY J. VINCENT, B.A., J.D., Ed.D.; *Professor, Vice President for Diversity and Community Engagement.*
WENDY E. WAGNER, B.A., M.E.S., J.D.; *Joe A. Worsham Centennial Professor.*
LOUISE WEINBERG, A.B., J.D., LL.M.; *William B. Bates Chair for the Administration of Justice.*
OLIN G. WELLBORN, A.B., J.D.; *William C. Liedtke, Sr. Professor.*
JAY L. WESTBROOK, B.A., J.D.; *Benno C. Schmidt Chair of Business Law.*
ABRAHAM L. WICKELGREN, A.B., Ph.D., J.D.; *Bernard J. Ward Professor in Law.*
ZIPPORAH B. WISEMAN, B.A., M.A., LL.B.; *Thos. H. Law Centennial Professor.*
PATRICK WOOLLEY, A.B., J.D.; *Beck, Redden & Secrest Professor in Law.*

ASSISTANT PROFESSORS

MARILYN ARMOUR, B.A., M.S.W., Ph.D.
DANIEL M. BRINKS, A.B., J.D., Ph.D.
JUSTIN DRIVER, B.A., M.A., M.A., J.D.
ZACHARY S. ELKINS, B.A., M.A., Ph.D.
JOSEPH R. FISHKIN, B.A., M. Phil., D. Phil., J.D.
CARY C. FRANKLIN, B.A., M.S.T., D. Phil., J.D.
MIRA GANOR, B.A., M.B.A., LL.B., LL.M., J.S.D.
JENNIFER E. LAURIN, B.A., J.D.
ANGELA K. LITTWIN, B.A., J.D.
MARY ROSE, A.B., M.A., Ph.D.
SEAN H. WILLIAMS, B.A., J.D.

SENIOR LECTURERS, WRITING LECTURERS, AND CLINICAL PROFESSORS

ALEXANDRA W. ALBRIGHT, B.A., J.D.; *Senior Lecturer.*
WILLIAM P. ALLISON, B.A., J.D.; *Clinical Professor, Director of Criminal Defense Clinic.*
MARJORIE I. BACHMAN, B.S., J.D.; *Clinical Instructor.*
PHILIP C. BOBBITT, A.B., J.D., Ph.D.; *Distinguished Senior Lecturer.*
KAMELA S. BRIDGES, B.A., B.J., J.D.; *Lecturer.*
CYNTHIA L. BRYANT, B.A., J.D.; *Clinical Professor, Director of Mediation Clinic.*
JOHN C. BUTLER, B.B.A., Ph.D.; *Clinical Associate Professor.*
MARY R. CROUTER, A.B., J.D.; *Lecturer, Assistant Director of William Wayne Justice Center for Public Interest Law.*
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.*
ELANA S. EINHORN, B.A., J.D.; *Lecturer.*
TINA V. FERNANDEZ, A.B., J.D.; *Lecturer, Director of Pro Bono Program.*
LYNDA E. FROST, B.A., M.Ed., J.D., Ph.D.; *Clinical Associate Professor.*
DENISE L. GILMAN, B.A., J.D.; *Clinical Professor, Co-Director of Immigration Clinic.*
KELLY L. HARAGAN, B.A., J.D.; *Lecturer, Director of Environmental Law Clinic.*
BARBARA HINES, B.A., J.D.; *Clinical Professor, Co-Director of Immigration Clinic.*
HARRISON KELLER, B.A., M.A., Ph.D.; *Vice Provost for Higher Education Policy, Senior Lecturer.*
JEANA A. LUNGWITZ, B.A., J.D.; *Clinical Professor, Director of Domestic Violence Clinic.*
TRACY W. MCCORMACK, B.A., J.D.; *Lecturer, Director of Advocacy Programs.*
ROBIN B. MEYER, B.A., M.A., J.D.; *Lecturer.*
RANJANA NATARAJAN, B.A., J.D.; *Clinical Professor, Director of National Security Clinic.*

Jane A. O'Connell, B.A., M.S., J.D.; *Lecturer, Deputy Director of Tarlton Law Library Public Services.*

Robert C. Owen, A.B., M.A., J.D.; *Clinical Professor.*

Sean J. Petrie, B.A., J.D.; *Lecturer.*

Wayne Schiess, B.A., J.D.; *Senior Lecturer, Director of Legal Writing.*

Stacy Rogers Sharp, B.S., J.D.; *Lecturer.*

Pamela J. Sigman, B.A., J.D.; *Adjunct Professor, Director of Juvenile Justice Clinic.*

David S. Sokolow, B.A., M.A., J.D., M.B.A.; *Distinguished Senior Lecturer, Director of Student Life.*

Leslie L. Strauch, B.A., J.D.; *Clinical Professor.*

Gretchen S. Sween, B.A., M.A., Ph.D., J.D.; *Lecturer.*

Melinda E. Taylor, B.A., J.D.; *Senior Lecturer, Executive Director of Center for Global Energy, International Arbitration, & Environmental Law.*

Heather K. Way, B.A., B.J., J.D.; *Lecturer, Director of Community Development Clinic.*

Elizabeth M. Youngdale, B.A., M.L.I.S., J.D.; *Lecturer.*

ADJUNCT PROFESSORS AND OTHER LECTURERS

Elizabeth Aebersold, B.A., M.S.
William R. Allensworth, B.A., J.D.
Craig D. Ball, B.A., J.D.
Sharon C. Baxter, B.S., J.D.
Karl O. Bayer, B.A., M.S., J.D.
William H. Beardall, Jr., B.A., J.D.
Jerry A. Bell, B.A., J.D.
Allison H. Benesch, B.A., M.S.W., J.D.
Craig R. Bennett, B.S., J.D.
James B. Bennett, B.B.A., J.D.
Melissa J. Bernstein, B.A., M.L.S., J.D.
Raymond D. Bishop, B.A., J.D.
Murff F. Bledsoe, B.A., J.D.
William P. Bowers, B.B.A., J.D., LL.M.
Hugh L. Brady, B.A., J.D.
Stacy L. Brainin, B.A., J.D.
Anthony W. Brown, B.A., J.D.
James E. Brown, B.A., LL.B.
Tommy L. Broyles, B.A., J.D.
Paul J. Burka, B.A., LL.B.
W.A. Burton, Jr., B.A., M.A., LL.B.
Erin G. Busby, 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.
Jeff Civins, A.B., M.S., J.D.
Leif M. Clark, B.A., J.D.
Elizabeth Cohen, B.A., M.S.W., J.D.
James W. Collins, B.S., J.D.
Patricia J. Cummings, B.A., J.D.
Keith B. Davis, B.S., J.D.
Dick Deguerin, B.A., LL.B.
Richard D. Deutsch, B.A., B.A., J.D.
Steven K. Dewolf, B.A., J.D, LL.M.
Rebecca H. Diffen, B.A., J.D.
Philip Durst, B.A., M.A., J.D.
Billie J. Ellis, Jr., B.A., M.B.A., J.D.
Jay D. Ellwanger, B.A., J.D.
Edward Z. Fair, B.A., M.S.W., J.D.
John C. Fleming, B.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.
Fred J. Fuchs, B.A., J.D.
Charles E. Gholz, B.S., Ph.D.
Michael J. Golden, A.B., J.D.
David Halpern, B.A., J.D.
Elizabeth Haluska-Rausch, B.A., M.A., M.S., Ph.D.
Jett L. Hanna, B.B.A., J.D.
Clint A. Harbour, B.A., J.D., LL.M.
Robert L. Hargett, B.B.A., J.D.
Mary L. Harrell, B.S., J.D.
James C. Harrington, B.A., M.A., J.D.
Christopher S. Harrison, Ph.D., J.D.
John R. Hays, Jr., B.A., J.D.
P. Michael Hebert, A.B., J.D.
Steven L. Highlander, B.A., Ph.D., J.D.
Susan J. Hightower, B.A., M.A., J.D.
Kenneth E. Houp, Jr., J.D.
Randy R. Howry, B.J., J.D.
Monty G. Humble, B.A., J.D.
Jeff Jury, B.A., J.D.
Patrick O. Keel, B.A., J.D.
Douglas L. Keene, B.A., M.Ed., Ph.D.
Chari L. Kelly, B.A., J.D.
Robert N. Kepple, B.A., J.D.
Mark L. Kincaid, B.B.A., J.D.
Ami L. Larson, B.A., J.D.
Jodi R. Lazar, B.A., J.D.
Kevin L. Leahy, B.A., J.D.
David P. Lein, B.A., M.P.A., J.D.
Maurie A. Levin, B.A., J.D.
Andres J. Linetzky, LL.M.
James-Lloyd Loftis, B.B.A., J.D.
Jim Marcus, B.A., J.D.
Harry S. Martin, A.B., M.L.S., J.D.
Frances L. Martinez, B.A., J.D.
Laura A. Martinez, B.A., J.D.
Ray Martinez, III, B.A., J.D.
Lisa M. McClain, B.A., J.D., LL.M.
Barry F. McNeil, B.A., J.D.
Angela T. Melinaraab, B.F.A., J.D.
Margaret M. Menicucci, B.A., J.D.
Jo A. Merica, B.A., J.D.
Ranelle M. Meroney, B.A., J.D.
Elizabeth N. Miller, B.A., J.D.
Jonathan F. Mitchell, B.A., J.D.
Daryl L. Moore, B.A., M.L.A., J.D.
Edwin G. Morris, B.S., J.D.
Sarah J. Munson, B.A., J.D.
Manuel H. Newburger, B.A., J.D.
David G. Nix, B.S.E., LL.M., J.D.
Patrick L. O'Daniel, B.B.A., J.D.
M.A. Payan, B.A., J.D.
Mark L. Perlmutter, B.S., J.D.
Eliza T. Platts-Mills, B.A., J.D.
Jonathan Pratter, B.A., M.L.S., J.D.
Velva L. Price, B.A., J.D.
Brian C. Rider, B.A., J.D.
Robert M. Roach, Jr., B.A., J.D.
Brian J. Roark, B.A., J.D.
Betty E. Rodriguez, B.S.W., J.D.
James D. Rowe, B.A., J.D.
Matthew C. Ryan, B.A., J.D.
Karen R. Sage, B.A., J.D.
Mark A. Santos, B.A., J.D.
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.
David A. Sheppard, B.A., J.D.
Hon. Eric M. Shepperd, 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.
Hon. Bea A. Smith, B.A., M.A., J.D.
Lydia N. Soliz, B.B.A., J.D.
Stephen M. Sonnenberg, A.B., M.D.

JAMES. M. SPELLINGS, JR., B.S., J.D.
DAVID B. SPENCE, B.A., J.D., M.A., Ph.D.
KACIE L. STARR, B.A., J.D.
WILLIAM F. STUTTS, B.A., J.D.
MATTHEW J. SULLIVAN, B.S., J.D.
JEREMY S. SYLESTINE, B.A., J.D.
BRADLEY P. TEMPLE, B.A., J.D.
SHERINE E. THOMAS, B.A., J.D.
TERRY O. TOTTENHAM, B.S., LL.M., J.D.
MICHAEL S. TRUESDALE, B.A., M.A., J.D.
JEFFREY K. TULIS, B.A., M.A., Ph.D.
TIMOTHY J. TYLER, B.A., J.D.
SUSAN S. VANCE, B.B.A., J.D.
LANA K. VARNEY, B.J., J.D.
SRIRAM VISHWANATH, B.S., M.S., Ph.D.
DEBORAH M. WAGNER, B.A., M.A., J.D.
CLARK C. WATTS, B.A., M.D., M.A., M.S., J.D.
WARE V. WENDELL, A.B., J.D.
RODERICK E. WETSEL, B.A., J.D.
THEA WHALEN, B.A., J.D.
DARA J. WHITEHEAD, B.A., M.S.
RANDALL B. WILHITE, B.B.A., J.D.
TIMOTHY A. WILKINS, B.A., M.P.P., J.D.
DAVID G. WILLE, B.S.E.E., M.S.E.E., J.D.
ANDREW M. WILLIAMS, B.A., J.D.
MARK B. WILSON, B.A., M.A., J.D.
HON. PAUL L. WOMACK, B.S., J.D.
LUCILLE D. WOOD, B.A., J.D.
DENNEY L. WRIGHT, B.B.A., J.D., LL.M.
LARRY F. YORK, B.B.A., LL.B.
DANIEL J. YOUNG, B.A., J.D.

VISITING PROFESSORS

OWEN L. ANDERSON, B.A., J.D.
ANTONIO H. BENJAMIN, LL.B., LL.M.
PETER F. CANE, B.A., LL.B., D.C.L.
JOSHUA DRESSLER, B.A., J.D.
ROBIN J. EFFRON, B.A., J.D.
VICTOR FERRERES, J.D., LL.M., J.S.D.
PETER M. GERHART, B.A., J.D.
LARRY LAUDAN, B.A., M.A., Ph.D.
GRAHAM B. STRONG, B.A., J.D., LL.M.

* * *

Texas Law Review

Volume 91 Number 5 April 2013

PARTH S. GEJJI
Editor in Chief

BENJAMIN S. MORGAN
Managing Editor

MOLLY M. BARRON
Chief Articles Editor

AMELIA A. FRIEDMAN
Administrative Editor

LAUREN K. ROSS
Chief Notes Editor

RALPH C. MAYRELL
Book Review Editor

LISA D. KINZER
Chief Online Content Editor

TYSON M. LIES
Research Editor

BRITTANY R. ARTIMEZ
ALESE L. BAGDOL
WILLIAM P. COURTNEY
MONICA E. GAUDIOSO
Articles Editors

ALEXANDER G. HUGHES
Managing Online Content Editor

MONICA R. HUGHES
ROSS M. MACDONALD
MICHAEL N. SELKIRK
Notes Editors

WILLIAM J. MCKINNON
MARTHA L. TODD
COLIN M. WATTERSON
COLLIN R. WHITE
Articles Editors

MICHAEL ABRAMS
BRIAN J. BAH
JULIA C. BARRETT
BRADEN A. BEARD
DAWSON A. BROTEMARKLE
ERIN L. GAINES
DANIEL D. GRAVER
JONATHAN LEVY
NATHANIEL H. LIPANOVICH
KYLE E. MITCHELL
CORBIN D. PAGE
CHRISTOPHER S. PATTERSON
ADAM R. PERKINS
KATHRYN G. RAWLINGS
STEPHEN STECKER
JAMES T. WEISS
Associate Editors

Members

MICHELLE K. ARISHITA
KATHRYN W. BAILEY
CECILIA BERNSTEIN
MICHAEL R. BERNSTEIN
MATTHEW J. BRICKER
CAITLIN A. BUBAR
KRISTIN C. BURNETT
CHARLES D. CASSIDY
SAMANTHA CHEN
CHASE E. COOLEY
JASON A. DANOWSKY
MICHAEL C. DEANE
MARIE E. DELAHOUSSAYE
DAVID D. DOAK
ALLISON L. FULLER
REBECCA L. GIBSON
JOSHUA S. GOLD
SEAN M. HILL
ALEXANDRA C. HOLMES
ROBERT P. HUGHES
LAURA C. INGRAM
SAMUEL F. JACOBSON
HANNAH L. JENKINS
COURTNEY H. JOHNSON
ELIZABETH M. JOHNSON
MICHAEL C. KELSO
MELANIE M. KISER
JEFFREY P. KITCHEN
KELSIE A. KRUEGER
ARIELLE K. LINSEY
JONATHAN D. LIROFF
ROCCO F. MAGNI
YANIV M. MAMAN
THOMAS K. MATHEW
DINA W. MCKENNEY
RYAN E. MELTZER
JOHN K. MORRIS
JACOB MOSS
DAVID A. NIEDRAUER
MARTIN OBERST
MATTHEW M. OLSON
JACKSON A. O'MALEY
SPENCER P. PATTON
JAMES D. PETERS
CHRISTA G. POWERS
JAMES R. POWERS
JENNIFER N. RAINEY
ALEZA S. REMIS
AMANDA D. ROBERTS
A. ELIZABETH ROMEFELT
BRETT S. ROSENTHAL
BRENT M. RUBIN
JONATHAN E. SARNA
JOHN W. STRIBLING
WILLIAM C. VAUGHN
VINCENT M. WAGNER
LECH K. WILKIEWICZ
JAMIE L. YARBROUGH
E. ALEXINE ZACARIAS

PAUL N. GOLDMAN
Business Manager

MITCHELL N. BERMAN
JOHN S. DZIENKOWSKI
Faculty Advisors

TERI GAUS
Editorial Assistant

* * *

Texas Law Review

Volume 91, Number 5, April 2013

ARTICLES

The Wires Go to War: The U.S. Experiment with Government Ownership of the Telephone System During World War I
Michael A. Janson & Christopher S. Yoo 983

Remapping the Path Forward: Toward a Systemic View of Forensic Science Reform and Oversight
Jennifer E. Laurin 1051

BOOK REVIEWS

What Do We Talk About When We Talk About the Constitution?
Akhil Reed Amar & Sanford Levinson 1119

reviewing Akhil Reed Amar's
AMERICA'S UNWRITTEN CONSTITUTION: THE PRECEDENTS AND PRINCIPLES WE LIVE BY

and Sanford Levinson's
FRAMED: AMERICA'S 51 CONSTITUTIONS AND THE CRISIS OF GOVERNANCE

Unsettling the Settled: Challenging the Great and Not-So-Great Compromises in the Constitution
Robert F. Williams 1149

also reviewing Sanford Levinson's
FRAMED: AMERICA'S 51 CONSTITUTIONS AND THE CRISIS OF GOVERNANCE

The Limits of Antitrust Scholarship
Barak Orbach 1165

reviewing THE GLOBAL LIMITS OF COMPETITION LAW

Stirring the Melting Pot: A Recipe for Immigrant Acceptance
Michael Scaperlanda 1171

reviewing Philip Kretsedemas's
THE IMMIGRATION CRUCIBLE: TRANSFORMING RACE, NATION, AND THE LIMITS OF THE LAW

Making Sense of the Marriage Debate
Jane S. Schacter 1185

reviewing Michael J. Klarman's
FROM THE CLOSET TO THE ALTAR: COURTS, BACKLASH, AND THE STRUGGLE FOR SAME-SEX MARRIAGE

NOTES

Jury Unanimity and the Problem with Specificity: Trying to Understand What Jurors Must Agree About by Examining the Problem of Prosecuting Child Molesters
Brian Bah 1203

U-Turns on the One-Way Street: Public Rights and Representation Theory in Patent Validity Litigation
Brett Rosenthal 1227

Pursuing Academic Freedom After *Garcetti v. Ceballos*
Lauren K. Ross 1253

Articles

The Wires Go to War: The U.S. Experiment with Government Ownership of the Telephone System During World War I

Michael A. Janson* & Christopher S. Yoo**

One of the most distinctive characteristics of the U.S. telephone system is that it has always been privately owned, in stark contrast to the pattern of government ownership followed by virtually every other nation. What is not widely known is how close the United States came to falling in line with the rest of the world. For the one-year period following July 31, 1918, the exigencies of World War I led the federal government to take over the U.S. telephone system. A close examination of this episode sheds new light into a number of current policy issues. The history confirms that natural monopoly was not solely responsible for AT&T's return to dominance and reveals that the Kingsbury Commitment was more effective in deterring monopoly than generally believed. Instead, a significant force driving the re-monopolization of the telephone system was the U.S. Postmaster General, Albert Burleson—not Theodore Vail, President of AT&T. It also demonstrates that universal service was the result of government-imposed emulation of the postal system, not, as some have claimed, a post hoc rationalization for maintaining monopoly. The most remarkable question is, having once obtained control over the telephone system, why did the federal government ever let it go? The dynamics surrounding this decision reveal the inherent limits of relying on war to justify extraordinary actions. More importantly, it shows the difficulties that governments face in overseeing industries that are undergoing dynamic technological change and that require significant capital investments.

Introduction 984

I. Setting the Stage 988

* Attorney Advisor, Federal Communications Commission, Washington, D.C. J.D., 2009, University of Pennsylvania Law School; Ph.D., 2007, University of Pennsylvania.

** John H. Chestnut Professor of Law, Communication, and Computer & Information Science and Founding Director of the Center for Technology, Innovation and Competition, University of Pennsylvania.

The opinions expressed herein are the authors' alone and do not necessarily represent the views of the Federal Communications Commission, its Commissioners, or its staff. The authors would like to thank the participants in the 2010 Annual Meeting of the American Political Science Association, 2010 Wharton Colloquium on Media and Communications Law, the 40th Annual Telecommunications Policy Research Conference (TPRC), and Stefan Heumann for their comments on earlier drafts, and Julie Goldemberg, Daniel Janovitz, Benjamin Meltzer, Rebecca Payne, and David Shifren for their research assistance. Professor Yoo would like to thank the Milton and Miriam Handler Foundation, the New York Bar Foundation, and the University of Pennsylvania's Center for Technology, Innovation and Competition for their financial support.

A. Proposals for Government Ownership of the U.S. Telegraph System 988
B. The Nationalization of the British Telephone System 993
C. The Government Takeovers of the U.S. Railroad and Radio Systems 994
D. Labor Conflict 998
E. Progressive Sentiment and the Kingsbury Commitment 998
F. Early Debates over Nationalizing the U.S. Telephone System 1002
II. The History of the Government Takeover 1006
A. Enacting of the Takeover 1006
B. Running the Telephone System 1010
1. *Mandating the Return of Monopoly* 1013
2. *Labor Unrest* 1017
3. *Rate Increases* 1019
C. Returning the Wires 1022
III. Implications of the Government Takeover 1030
A. The Reemergence of Monopoly 1030
1. *Natural Monopoly* 1031
a. *Scale Economies* 1031
b. *Network Economic Effects* 1033
2. *The Supposedly Lax Enforcement of the Antitrust Laws* 1037
3. *Government-Mandated Consolidation as a Missing Consideration* 1039
B. The Origins of Universal Service 1042
C. Federal–State Relations 1044
D. The Limits of Government 1047
Conclusion 1049

Introduction

One of the characteristics of the U.S. telephone system generally thought to distinguish it from all others is that it has always been privately owned. In all other major countries, telephone systems have generally been owned and operated by the government, most commonly through an organization known as a Post, Telephone, and Telegraph (PTT).[1] The United States took a notably different course, having private ownership of telephone and telegraph systems.[2] Indeed, the American emphasis on individualism

1. Philip J. Weiser, *The Ghost of Telecommunications Past*, 103 MICH. L. REV. 101, 103 (2005) (reviewing PAUL STARR, THE CREATION OF THE MEDIA: POLITICAL ORIGINS OF MODERN COMMUNICATION (2004)).

2. *See* ROBERT MILLWARD, PRIVATE AND PUBLIC ENTERPRISE IN EUROPE 245 (2005) ("[P]ublic ownership by a single enterprise of a national network was the rule by 1950 and reflected, in part, the unwillingness of governments . . . to use arm's-length regulation of private

and lack of a legacy of strong sovereign states has led some to regard government ownership of the telephone in the United States as unthinkable.[3] The wave of privatizations that began worldwide in the 1980s is widely regarded as an implicit endorsement of the American approach.[4]

What is not widely known is how close the United States came to falling in line with the rest of the world. For the roughly one-year period following July 31, 1918, the federal government took over the U.S. telephone system.[5] This period of history is important for many reasons. It provides a fascinating insight into the dynamics of institutional change, particularly regarding the role of individuals, political processes, and technology.

The episode also sheds light on many central issues of telecommunications policy today. For example, the analysis reveals that the reassertion of the Bell System's monopoly, long blamed on natural monopoly,[6] or the Antitrust Division's failure to curb the ambitions of AT&T President Theodore Vail,[7] was assisted and encouraged by the deliberate policies of the Postmaster General to consolidate the industry.[8] Moreover, the Kingsbury Commitment of 1913 may have been more effective at preventing consolidation than generally realized.[9] Further, contrary to the criticism that universal service was a concept that arose during the 1960s to rationalize the Bell monopoly after the fact,[10] history reveals that universal service has its roots during the government takeover, much earlier than previously thought.[11] The episode marked a nascent revolution in federal–state relations that would ultimately collapse due to the unpopularity of rate increases.[12] Perhaps most revealing is the government's surprising decision, after having taken over the telephone system, to once again return it to

---

monopolies . . . ."); RICHARD B. KIELBOWICZ, POSTAL ENTERPRISE: POST OFFICE INNOVATIONS WITH CONGRESSIONAL CONSTRAINTS, 1789–1970, at 51 (2000), *available at* http://www.prc.gov/prc-docs/library/refdesk/techpapers/Kielbowicz/enterprise.pdf ("Except for the United States, virtually every nation regarded the telegraph and telephone as natural extensions of the state's mail monopoly and operated them under a postal ministry.").

3. ALAN STONE, PUBLIC SERVICE LIBERALISM: TELECOMMUNICATIONS AND TRANSITIONS IN PUBLIC POLICY 42 (1991).

4. *See* Wei Li & Lixin Colin Xu, *The Impact of Privatization and Competition in the Telecommunications Sector Around the World*, 47 J.L. & ECON. 395, 395–96 (2004) (documenting the dramatic increase in privately owned telephone systems during the 1980s and 1990s and showing that privatization improved industry performance).

5. Comment, *The Telegraph Industry: Monopoly or Competition*, 51 YALE L.J. 629, 633 (1942).

6. *See* GERALD R. FAULHABER, TELECOMMUNICATIONS IN TURMOIL: TECHNOLOGY AND PUBLIC POLICY 107 (1987) ("Indeed, until the late 1960s few questioned that the telephone industry was a natural monopoly."); PETER W. HUBER ET AL., FEDERAL TELECOMMUNICATION LAW § 2.1.2, at 86 (2d ed. 1999) ("Is the telephone industry (or any part of it) a natural monopoly? Until the 1960s, the answer was generally presumed to be yes, from end to end.").

7. TIM WU, THE MASTER SWITCH: THE RISE AND FALL OF INFORMATION EMPIRES 56 (2010).

8. *See infra* section II(B)(1).

9. *See infra* section III(A)(2).

10. MILTON L. MUELLER, UNIVERSAL SERVICE 6, 150–64 (1997).

11. *See infra* subpart III(B).

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

private control.[13] The government's reasons for doing so are quite revealing about the realities of management and ownership in an industry characterized by dynamic technological change.

Despite the importance of this episode in history, it has been largely overlooked by the scholarly community. Many histories discussing the competitive dynamics of this period fail to mention it at all.[14] Other accounts offer a passing reference to it[15] or devote a few pages to it.[16] Indeed, only a handful of published works examine the history of the government takeover at any length,[17] and these accounts focus on the political consequences of this

---

13. *See infra* subpart III(D).

14. *E.g.*, FAULHABER, *supra* note 6; HUBER ET AL., *supra* note 6; KENNETH LIPARTITO, THE BELL SYSTEM AND REGIONAL BUSINESS: THE TELEPHONE IN THE SOUTH, 1877–1920 (1989); ALAN STONE, WRONG NUMBER: THE BREAKUP OF AT&T (1989); Glen O. Robinson, *The Federal Communications Act: An Essay on Origins and Regulatory Purpose*, *in* A LEGISLATIVE HISTORY OF THE COMMUNICATIONS ACT OF 1934, at 3 (Max D. Paglin ed., 1989).

15. *See, e.g.*, GERALD W. BROCK, THE TELECOMMUNICATIONS INDUSTRY: THE DYNAMICS OF MARKET STRUCTURE 156 (1981) (mentioning the government's control of telephone systems during World War I); JEFFREY E. COHEN, THE POLITICS OF TELECOMMUNICATIONS REGULATION 38 (1992) (referencing the government's experimentation with nationalization during World War I); CLAUDE S. FISCHER, AMERICA CALLING: A SOCIAL HISTORY OF THE TELEPHONE TO 1940, at 50 (1992) (making mention of the government's one-year takeover of the telephone industry); AMY FRIEDLANDER, NATURAL MONOPOLY AND UNIVERSAL SERVICE 77 (1995) (citing the telephone system's brief nationalization during World War I); 1 LEONARD S. HYMAN ET AL., THE NEW TELECOMMUNICATIONS INDUSTRY: EVOLUTION AND ORGANIZATION 81 (1987) (noting the Post Office's control of AT&T from August 1, 1918 to August 1, 1919 as part of the war effort); SUSAN E. MCMASTER, THE TELECOMMUNICATIONS INDUSTRY 47 (2002) (same); MUELLER, *supra* note 10, at 133 (mentioning centralization); PETER TEMIN, THE FALL OF THE BELL SYSTEM 11 n.3 (1987) (referencing the Postmaster General's control over the telephone system); William P. Barnett & Glenn R. Carroll, *How Institutional Constraints Affected the Organization of Early U.S. Telephony*, 9 J.L. ECON. & ORG. 98, 112 (1993) (indicating the government's brief period of control over the telephone industry); Kenneth A. Cox & William J. Byrnes, *The Common Carrier Provisions—A Product of Evolutionary Development*, *in* A LEGISLATIVE HISTORY OF THE COMMUNICATIONS ACT OF 1934, *supra* note 14, at 29 (mentioning the Post Office's control of the telephone and telegraph companies as a wartime measure); Geoffrey M. Peters, *Is the Third Time the Charm? A Comparison of the Government's Major Antitrust Settlements with AT&T This Century*, 15 SETON HALL L. REV. 252, 257 (1985) (pointing out the government's operation of the telephone systems during World War I).

16. *E.g.*, JOHN BROOKS, TELEPHONE: THE FIRST HUNDRED YEARS 150–53, 157–59 (1975); ROBERT BRITT HORWITZ, THE IRONY OF REGULATORY REFORM: THE DEREGULATION OF AMERICAN TELECOMMUNICATIONS 101–02 (1989); GEORGE P. OSLIN, THE STORY OF TELECOMMUNICATIONS 278–79 (1992); J. WARREN STEHMAN, THE FINANCIAL HISTORY OF THE AMERICAN TELEPHONE AND TELEGRAPH COMPANY 175–81 (Augustus M. Kelley Publishers 1967) (1925); STONE, *supra* note 3, at 197–99; RICHARD H.K. VIETOR, CONTRIVED COMPETITION 172–73 (1994); Adam D. Thierer, *Unnatural Monopoly: Critical Moments in the Development of the Bell System Monopoly*, 14 CATO J. 267, 275–76 (1994); *The Telegraph Industry: Monopoly or Competition*, *supra* note 5, at 633–37.

17. N.R. DANIELIAN, A.T.&T.: THE STORY OF INDUSTRIAL CONQUEST 243–70 (1939); RICHARD R. JOHN, NETWORK NATION: INVENTING AMERICAN TELECOMMUNICATIONS 395–406 (2010); CHRISTOPHER N. MAY, IN THE NAME OF WAR: JUDICIAL REVIEW AND THE WAR POWERS SINCE 1918, at 26–59 (1989). The only unpublished discussions of any significance of which we are aware are a dissertation by political scientist Kenneth Bickers and a brief note by an FCC economist. Kenneth N. Bickers, The Politics of Regulatory Design: Telecommunications in Historical and Theoretical Perspective 134–56 (1988) (unpublished Ph.D. dissertation, University of

episode without discussing its consequences for the telecommunications system in general or its role in paving the way for AT&T's return to monopoly and the establishment of universal service in particular. The omission is rendered all the more curious by the recent heightening of interest in government ownership of communications networks, reflected in the support for municipal WiFi,[18] the Dutch government's efforts to promote the buildout of municipal broadband networks,[19] the inclusion of government funds for U.S. broadband deployment in the 2009 stimulus package,[20] and the Australian government's decision to fund more than three quarters of the cost to build fiber optic cable to the home.[21] In addition, some scholars have either advocated government funding of broadband networks[22] or proposed giving the postal system a greater role in the Internet.[23] These calls for the postalization of the Internet would do well to take into account the lessons from our nation's past experience with the postalization of telecommunications.

This Article is organized as follows: Part I discusses certain developments that set the stage for the takeover. These include the proposed takeover of the U.S. telegraph system, the nationalization of the British Telephone System, the wartime takeover of the U.S. railroad system, the early debates about nationalizing the U.S. telephone system, and the antitrust

---

Wisconsin-Madison) (on file with authors); Douglas Galbi, *Government Takeover of All Telephone Systems*, PURPLE MOTES (Apr. 4, 2010), http://purplemotes.net/2010/04/04/government-takeover-of-all-telephone-systems/.

18. *See* François Bar & Namkee Park, *Municipal Wi-Fi Networks: The Goals, Practices, and Policy Implications of the U.S. Case*, 61 COMM. & STRATEGIES 107, 107 (2006) (commenting on the growing number of municipal Wi-Fi networks in the United States and abroad).

19. Willem van Winden & Paulus Woets, *Urban Broadband Internet Policies in Europe: A Critical Review*, 41 URB. STUD. 2043, 2046, 2049–51 (2004).

20. *See* Lynne Holt & Mark Jamison, *Broadband and Contributions to Economic Growth: Lessons from the US Experience*, 33 TELECOMM. POL'Y 575, 575 (2009) (examining the connection between information and communications technologies and economic growth, and noting that Congress approved $7.2 billion in funding for broadband planning and deployment initiatives as part of the American Recovery and Reinvestment Act of 2009).

21. *See* Roland Montagne & Valérie Chaillou, *Public Funding & FTTx: Assessing the Impact of Public Action*, 80 COMM. & STRATEGIES 153, 161 (2010) (noting that Australia represents a prime example of a national project to build a neutral, national FTTH network and reporting that Australia has invested €30 billion in the construction of an open national network); *see also National Broadband Network—Overview*, DEP'T BROADBAND, COMM. & THE DIGITAL ECON., AUSTL. GOV'T, http://www.dbcde.gov.au/broadband/national_broadband_network/nbn_overview (last modified Dec. 18, 2012) (describing Australia's National Broadband Network).

22. *E.g.*, SUSAN P. CRAWFORD, CAPTIVE AUDIENCE: THE TELECOM INDUSTRY AND MONOPOLY POWER IN THE NEW GILDED AGE 263–67 (2013); LAWRENCE LESSIG, THE FUTURE OF IDEAS 244 (2002); BARBARA VAN SCHEWICK, INTERNET ARCHITECTURE AND INNOVATION 370 (2010).

23. *See* PRESIDENT'S COMM'N ON THE U.S. POSTAL SERV., EMBRACING THE FUTURE: MAKING THE TOUGH CHOICES TO PRESERVE UNIVERSAL MAIL SERVICES 143–58 (2003), *available at* http://www.treasury.gov/press-center/press-releases/Documents/pcusps_report.pdf (positing that "a digital postal network will enhance the value of the mail as a 21st century communications mode and improve virtually every aspect of the nation's postal service").

scrutiny of AT&T that led to the Kingsbury Commitment. Part II lays out the history of the takeover, analyzing its enactment by Congress, its operation by the Postmaster General, and the decision to return the wires. Part III examines the lessons of the takeover, discussing the Postmaster General's active promotion of AT&T's return to monopoly, the origins of universal service, the transformation of federal–state relations, and the acknowledgement of the limits of government control implicit in the decision to return the wires.

I. Setting the Stage

The failure of the Soviet bloc's pattern of state-owned enterprises and the wide-scale privatization of telephone systems in recent years make it all too easy to reject government ownership as a viable policy option in modern history. During the Progressive Era, however, "the specter of nationalization was present and gaining momentum," a "fact [that] is often lost on historians of telephony during this era."[24] In fact, nationalization of the telephone system was not the exclusive province of socialists: A wide range of respectable voices, including many conservatives, supported government ownership.[25] A better appreciation for key aspects of the historical context, including proposals for government ownership of the U.S. telegraph system, the 1911 nationalization of the British telephone system, Progressive hostility toward large enterprises that led to the Kingsbury Commitment, and early debates over nationalization, helps put the debates over government ownership into perspective.

A. *Proposals for Government Ownership of the U.S. Telegraph System*

The telegraph preceded the telephone as the dominant means of telecommunications, and policy makers debated the merits of public ownership since its earliest days.[26] The history of the electromagnetic telegraph in the United States began on September 4, 1837, when Samuel Morse made a successful transmission across 1,700 feet of wire arranged in his classroom.[27] Ill suited to commercializing the invention himself, he

---

24. COHEN, *supra* note 15, at 38.

25. JOHN, *supra* note 17, at 363–65, 372–74; STONE, *supra* note 3, at 141, 195.

26. A report submitted by the Post Office to Congress in 1914 provides a useful overview of the early advocacy for government ownership of telecommunications. POSTMASTER GEN., GOVERNMENT OWNERSHIP OF ELECTRICAL MEANS OF COMMUNICATION, S. DOC. NO. 63-399, at 19–36 (2d Sess. 1914).

27. OSLIN, *supra* note 16, at 19. Morse was neither the first nor the only inventor working on telegraphy. Beginning in 1793, France deployed an optical telegraph system that used a series of towers topped by a set of movable arms that could send signals in a semaphore-like manner. In 1809, a German inventor developed a telegraph that used electrochemical processes connected by thirty-five wires to communicate. European inventors were also independently experimenting with electromagnetic telegraphs at more or less the same time as (indeed, perhaps slightly before) Morse. A.N. HOLCOMBE, PUBLIC OWNERSHIP OF TELEPHONES ON THE CONTINENT OF EUROPE 3–8 (1911). The U.S. Supreme Court would ultimately rule that Morse's invention came first. O'Reilly

convinced Congress to appropriate $30,000 to establish a telegraph connection between Washington, D.C. and Baltimore,[28] through which Morse successfully sent a message on May 24, 1844.[29]

At the time, many observers thought that the telegraph network should be owned and operated by the government. For example, Henry Clay found that the telegraph "is destined to exert great influence on the business affairs of society. In the hands of private individuals they will be able to monopolize intelligence and to perform the greatest operations in commerce . . . . I think such an engine ought to be exclusively under the control of the [G]overnment."[30] Postmaster General Cave Johnson's 1845 and 1846 Reports similarly supported government ownership of the entire telegraph system.[31]

In 1866, Congress enacted legislation that gave the government a five-year right to purchase all the telegraph lines at a value appraised by five disinterested arbitrators.[32] The following year, Andrew Johnson's Postmaster General, Alexander Randall, urged Congress to study the possibility of a postal takeover of the telegraph system.[33] In 1871, shortly after the government's option to purchase the telegraph system expired, Postmaster General John Creswell endorsed the idea of a postal telegraph, pointing to the fact that Great Britain had nationalized its telegraph system in 1870.[34] This recommendation drew the approbation of President Ulysses S. Grant, who "recommend[ed] favorable consideration of the plan for uniting the telegraphic system of the United States with the postal system."[35] Not only would public ownership reduce rates while rendering the same level of service, if not better[36]: "It would secure the further advantage of extending

---

v. Morse, 56 U.S. (15 How.) 62, 108 (1853); *see also* Smith v. Downing, 22 F. Cas. 511, 513 (C.C.D. Mass. 1850) (No. 13,036) (holding that Morse had a right to patent the new method).

28. Act of Mar. 3, 1843, ch. 84, § 1, 5 Stat. 618; OSLIN, *supra* note 16, at 32.

29. OSLIN, *supra* note 17, at 32–33.

30. Letter from Henry Clay to Alfred Vail (Sept. 10, 1844), *reprinted in* Frank G. Carpenter, *Henry Clay on Nationalizing the Telegraph*, 154 N. AM. REV. 380, 382 (1892).

31. CAVE JOHNSON, REPORT OF THE POSTMASTER GENERAL, S. DOC. NO. 29-1, at 861 (1st Sess. 1845); CAVE JOHNSON, REPORT OF THE POSTMASTER GENERAL, S. DOC. NO. 29-1, at 688–89 (2d Sess. 1846). Morse even offered to sell the patent to the federal government for $100,000, see S. DOC. NO. 63-399, at 19, and apparently attempted to give the patent to the Republic of Texas in 1838. OSLIN, *supra* note 16, at 23.

32. Act of July 24, 1866, ch. 230, § 3, 14 Stat. 221, 221–22.

33. ALEXANDER RANDALL, REPORT OF THE POSTMASTER GENERAL, H.R. EXEC. DOC. NO. 40-1, pt. 4, at 29 (2d Sess. 1867).

34. JOHN CRESWELL, REPORT OF THE POSTMASTER GENERAL, H.R. EXEC. DOC. NO. 42-1, pt. 4, at 28–9 (2d Sess. 1871).

35. President Ulysses S. Grant, Third Annual Message (Dec. 4, 1871), *in* 7 A COMPILATION OF THE MESSAGES & PAPERS OF THE PRESIDENTS 1789–1897, at 149–50 (James D. Richardson ed., 1898).

36. "[B]y such a course the cost of telegraphing could be much reduced, and the service as well, if not better, rendered." *Id.* at 150.

the telegraph through portions of the country where private enterprise will not construct it," as well as promote commerce and education.[37]

During 1871 and 1872, Congress seriously debated government ownership of the telegraph system, dividing between one proposal (endorsed by the President and the Postmaster General) under which the federal government would take possession of the entire telegraph system and merge it with the post office,[38] and another proposal (backed by Gardiner Hubbard, who would eventually become President of the Bell System as well as Alexander Graham Bell's father-in-law) that would place the entire industry in the hands of a single private company that was granted special privileges by the government and give the government preferential terms.[39] Congress deadlocked over these proposals and failed to enact either of them.[40] Creswell would repeat his call for government ownership of the telegraph system in 1872 and 1873 to no avail.[41]

The matter lay quiescent until 1880, when a visit to the British post office prompted Postmaster General Horace Maynard to ask whether the federal government should once again take up the issue of public ownership of the telegraph system.[42] These calls were renewed in 1882 and 1883 by Postmasters General Timothy Howe and Walter Gresham.[43] Minority political party platforms in the 1880s echoed these sentiments. The Greenback Party platform of 1884 demanded "the establishment of a

---

37. *Id.*

38. Gardiner Hubbard, *The Proposed Changes in the Telegraph System*, 117 N. AM. REV. 80, 102–03 (1873).

39. *Id.* at 103–04.

40. *Cf. id.* at 104 (stating that time ran out before the proposal could be considered that session, and history shows us that they ultimately were not successful).

41. JOHN CRESWELL, REPORT OF THE POSTMASTER GENERAL, H.R. EXEC. DOC. NO. 42-1, pt. 4, at 21–35 (3d Sess. 1872); JOHN CRESWELL, REPORT OF THE POSTMASTER GENERAL, H.R. EXEC. DOC. NO. 43-1, pt. 4, at xxxiii–xxxvii (1st Sess. 1873). Congressional consideration of government ownership occurred contemporaneously with national political parties advocating for increased government regulation of the telegraph. The Labor Reform Party's platform of 1872 resolved that "it is the duty of the government to so exercise its power over railroads and telegraph corporations that they shall not in any case be privileged to exact such rates . . . as may bear unduly or inequitably upon either producer or consumer." LABOR REFORM PLATFORM OF 1872, *in* 1 NATIONAL PARTY PLATFORMS, 1840–1956, at 43 (Donald B. Johnson ed., 1978). Likewise, the Prohibition Party's platform of 1872 called for reduction of telegraph rates "to the lowest practical point, by force of laws wisely and justly framed, with reference not only to the interest of capital employed but to the higher claim of the general good." PROHIBITION PLATFORM OF 1872, *in* 1 NATIONAL PARTY PLATFORMS, 1840–1956, *supra*, at 46. In 1876, the Prohibition Party also called for the "reduction of the rates of inland and ocean postage of telegraphic communication." PROHIBITION REFORM PLATFORM OF 1876, *in* 1 NATIONAL PARTY PLATFORMS, 1840–1956, *supra*, at 52.

42. HORACE MAYNARD, REPORT OF THE POSTMASTER GENERAL, H.R. EXEC. DOC. NO. 46-1, pt. 4, at 42 (3d Sess. 1880).

43. TIMOTHY HOWE, REPORT OF THE POSTMASTER GENERAL, H.R. EXEC. DOC. NO. 47-1, pt. 4, at xxvii–xxx (2d Sess. 1882); WALTER GRESHAM, REPORT OF THE POSTMASTER GENERAL, H.R. EXEC. DOC. NO. 48-1, pt. 4, at 33–37 (1st Sess. 1883).

government postal telegraph system."[44] The Union Labor Party of 1888 declared, "The means of communication and transportation shall be owned by the people, as is the United States postal system."[45] In accord with these calls for government action, between 1871 and 1884, Congress considered over two dozen proposals to nationalize the telegraph system, three quarters of which were apparently reported favorably out of committee.[46] None, however, was ever enacted.[47]

In 1901, the Industrial Commission heard testimony from Professor Frank Parsons advocating government ownership.[48] The Postal Service Appropriations Act of 1901 contained a provision directing the Postmaster General "to report to Congress the probable cost of connecting a telegraph and telephone system with the postal service by some feasible plan,"[49] although it does not appear that the Postmaster General ever did so.[50] Aside from a passing mention by George Cortelyou in 1906 including the postal telephone in a laundry list of future improvements to the postal system,[51] no further action was taken for more than a decade despite continuing support from minority parties.[52]

Interest returned in 1912, when Postmaster General Frank Hitchcock once again proposed, "The telegraph lines in the United States should be made a part of the postal system,"[53] only to see that recommendation specifically disavowed by President Taft's message transmitting this report. Taft "believe[d] that the true principle is that private enterprise should be permitted to carry on such public utilities under due regulation as to rates by proper authority rather than that the Government should itself conduct them."[54] Taft thought it would be bad public policy "greatly to increase the

---

44. GREENBACK NATIONAL PLATFORM OF 1884, *in* 1 NATIONAL PARTY PLATFORMS, 1840–1956, *supra* note 41, at 69–70.

45. UNION LABOR PLATFORM OF 1888, *in* 1 NATIONAL PARTY PLATFORMS, 1840–1956, *supra* note 41, at 83.

46. *See* POSTMASTER GEN., GOVERNMENT OWNERSHIP OF ELECTRICAL MEANS OF COMMUNICATION, S. DOC. NO. 63-399, at 30 (2d Sess. 1914) (explaining that over seventy bills had been introduced to Congress for the purpose of establishing a postal telegraph and sixteen times the House and Senate Committees had reported favorably on the issue).

47. *Id.*; COHEN, *supra* note 15, at 37.

48. S. DOC. NO. 63-399, at 33.

49. Postal Service Appropriations Act of 1901, ch. 851, 31 Stat. 1099, 1104.

50. S. DOC. NO. 63-399, at 35.

51. GEORGE B. CORTELYOU, ANNUAL REPORT OF THE POSTMASTER-GENERAL FOR THE FISCAL YEAR ENDED JUNE 30, 1906, H.R. DOC. NO. 59-4, at 81 (2d Sess. 1906).

52. The People's Party platform of 1908 stated: "To perfect the postal service, the Government should own and operate the general telegraph and telephone systems and provide a parcels post." PEOPLE'S PLATFORM OF 1908, *in* 1 NATIONAL PARTY PLATFORMS, 1840–1956, *supra* note 41, at 155.

53. FRANK H. HITCHCOCK, ANNUAL REPORT OF THE POSTMASTER GENERAL FOR THE FISCAL YEAR ENDED JUNE 30, 1911, H.R. DOC. NO. 62-118, at 14 (2d Sess. 1912).

54. PRESIDENT WILLIAM H. TAFT, MESSAGE OF THE PRESIDENT OF THE UNITED STATES, H.R. DOC. NO. 62-559, at 8 (2d Sess. 1912).

body of public servants."[55] Although the argument for government ownership would be strong if government could operate the system "at a less price . . . and with equal efficiency," Taft was

> not satisfied from any evidence that if these properties were taken over by the Government they could be managed any more economically or any more efficiently or that this would enable the Government to furnish service at any smaller rate than the public are now required to pay by private companies.[56]

In any event, Taft believed that any such initiatives should be postponed until after the Post Office had established a postal savings bank and a parcel post.[57]

Still, at this point, AT&T was sufficiently optimistic to predict in its annual report that "[t]he discussion of the government ownership of wire companies is not likely to become anything more than academic, at least for the present."[58] The company was sufficiently concerned, however, to devote four additional pages to laying out arguments against government ownership of the telegraph system.[59] The report concluded, "The facts are, that there is hardly a telegraph or telephone system in the world now operated by any government which shows a profit, even under accounting methods employed, and not one that would not show a deficit under accounting methods obligatory upon private enterprise."[60]

Undeterred by the President's opposition, Postmaster General Hitchcock's next report in 1912 did not back down, arguing that government ownership of the telegraph lines would lower rates, and that the successful creation of the postal savings system and the parcel post justified renewing attention on the proposal.[61]

The advent of the Wilson Administration brought in a new Postmaster General, Albert S. Burleson, who would play a pivotal role in the debates over government ownership. Indeed, Burleson would advocate government ownership of the telephone system with a zeal that strained the limits of even the most ardent Progressive.[62] On December 1, 1913, Burleson submitted his initial annual report, which adhered to Hitchcock's position: "The

---

55. *Id.*

56. *Id.*

57. *Id.*

58. AT&T CO., ANNUAL REPORT OF THE DIRECTORS OF AMERICAN TELEPHONE & TELEGRAPH COMPANY TO THE STOCKHOLDERS FOR THE YEAR ENDING DECEMBER 31, 1911, at 36 (1912).

59. *Id.* at 36–40.

60. *Id.* at 39–40.

61. FRANK H. HITCHCOCK, ANNUAL REPORT OF THE POSTMASTER GENERAL FOR THE FISCAL YEAR ENDED JUNE 30, 1912, H.R. DOC. NO. 62-931, at 13 (3d Sess. 1913).

62. *See* Adrian Anderson, *President Wilson's Politician: Albert Sidney Burleson of Texas*, 77 SW. HIST. Q. 339, 345 (1974) (noting that Burleson's advocacy of government ownership of the telephone system was "a goal that was really a little too radical even for most Progressives").

monopolistic nature of the telegraph business makes it of vital importance to the people that it be conducted by unselfish interests, and this can be accomplished only through Government ownership."[63] Expanding his point to include the telephone system, Burleson wrote, "Every argument in favor of Government ownership of telegraph lines may be advanced with equal logic and force in favor of the Government ownership of telephone lines."[64] The report indicated that the Post Office Department was conducting an investigation and promised to submit a legislative proposal shortly.[65]

B. *The Nationalization of the British Telephone System*

Another consideration that framed and colored debates over nationalization was the global trend toward governmental ownership of telephone systems. Manitoba had nationalized its telephone system in 1907.[66] Even more importantly, Great Britain had nationalized its telephone system in 1911.[67] Indeed, by 1913, the United States was the only major country whose telephone system was not publicly owned.[68]

The fact that Britain implemented a highly successful rate cut immediately following the government takeover made government ownership seem alluring to many.[69] The demand for nationalization in the United Kingdom also came from the postal service's fear of loss of revenue.[70] As Harper explains it: "The reason [for nationalization] was fear that [private telegrams] would damage the revenues of the postal service, coupled with serious dissatisfaction among the business community about the service being given by competing private interests."[71]

---

63. ALBERT S. BURLESON, ANNUAL REPORT OF THE POSTMASTER GENERAL FOR THE FISCAL YEAR ENDED JUNE 30, 1913, H.R. DOC. NO. 63-712, at 15 (2d Sess. 1914).

64. *Id.* at 16.

65. *Id.*

66. JAMES MAVOR, GOVERNMENT TELEPHONES: THE EXPERIENCE OF MANITOBA CANADA 26–28 (1917).

67. STONE, *supra* note 3, at 141. For the classic study on public ownership of telephone systems in Europe, see generally HOLCOMBE, *supra* note 27.

68. *See supra* note 2 and accompanying text.

69. *See* STONE, *supra* note 3, at 41–42 ("Earlier regulatory statutes had failed to bring rates down, leading to the drive to nationalize a business that was widely conceived as a public service."); *Revision of Telephone Rates*, TIMES (London), Jan. 27, 1912, at 13 (quoting Postmaster General Herbert Samuel as stating that "it would be necessary before long to revise the rates of telephone users" and explaining that the rates "were at present unequal, in some cases not wholly equitable"); *Telephone Trunk Calls: Introduction of Lower Rates*, TIMES (London), Aug. 8, 1912, at 2 (reporting the announcement of Postmaster General Samuel that "he proposes to introduce lower rates for the use of telephone trunk lines during the less busy hours of the day," with reductions ranging from one-quarter to three-quarters of the ordinary rates).

70. JOHN HARPER, MONOPOLY AND COMPETITION IN BRITISH TELECOMMUNICATIONS: THE PAST, THE PRESENT AND THE FUTURE 5 (1997).

71. *Id.*

C. *The Government Takeovers of the U.S. Railroad and Radio Systems*

The takeover of the telephone system also took place in the shadow of the federal government's decision to take over the railroad system.[72] Unlike the telephone system, after years of corporate mismanagement and restrictive rate regulation by the ICC, the railroad industry was in a state of financial and operational disarray.[73] The flood of traffic to the Atlantic ports pushed the rail network to the brink of collapse.[74] Moreover, the industry had long sought coordination of the entire industry by a single entity to curb what it viewed as the excesses of competition.[75]

Congress had anticipated the need for the government to take control of the railroads by including a provision in the Army Appropriations Act of 1916 authorizing the President to do so in the event of war.[76] Wilson issued the proclamation taking over the railroad system on December 28, 1917.[77] Congress ratified his decision by enacting the Federal Control Act on March 21, 1918.[78] The takeover was supported both by the industry, which welcomed cartelization as a sanctuary from unbridled competition, and the U.S. Chamber of Commerce, which embraced the idea of "Scientific Management."[79]

Government operation of the rail system under the direction of the U.S. Railway Administration (USRA) (headed by William Gibbs McAdoo, Treasury Secretary and Wilson's son-in-law) proved controversial. Most controversial was the approval of a 28% across-the-board increase in rates.[80] McAdoo's determination to maintain labor peace and avoid strikes led him to order a series of wage increases, extend the eight-hour day to all rail employees, promote union membership by encouraging collective bargaining, and create an elaborate system of job classifications.[81] The increased costs caused the railroads to operate at a substantial deficit.[82]

---

72. *See* BROOKS, *supra* note 16, at 150 ("With the coming of war, agitation for government ownership [of the telephone system] greatly increased . . . . Government takeover of the railroads [in December 1917] fanned the flames; thereafter, advocates of a telephone takeover argued that government-run railroads and privately run wire communications constituted a logical inconsistency.").

73. JAMES W. ELY, JR., RAILROADS AND AMERICAN LAW 241 (2001).

74. *Id.*; Landon H. Rowland, The Last Hurrah for the Gilded Age: The 1917 Nationalization of U.S. Railways, Remarks at the World War I Museum 5 (Nov. 29, 2008), *available at* http://www.landonrowland.com/RailroadSpeechFINAL.pdf.

75. Rowland, *supra* note 74, at 5–6.

76. Army Appropriations Act of 1916, ch. 418, 39 Stat. 619, 645.

77. President Woodrow Wilson, Proclamation (Dec. 26, 1917), *in* 17 A COMPILATION OF THE MESSAGES AND PAPERS OF THE PRESIDENTS 8409, 8410 (James D. Richardson ed., 1921).

78. Federal Control Act, ch. 25, 40 Stat. 451 (1918).

79. Rowland, *supra* note 74, at 9.

80. *Id.* at 10; *see also* ELY, *supra* note 15, at 244 (discussing rates increases generally); GABRIEL KOLKO, RAILROADS AND REGULATION 1877–1916, at 228 (1965) (discussing specific rate increases).

81. ELY, *supra* note 73, at 244–45.

82. *Id.* at 245.

The cessation of hostilities raised the question of what to do next. The Federal Railroad Control Act of 1918 only authorized government control through twenty-one months after the end of hostilities,[83] which Wilson noted in his Annual Message expired in January 1921.[84] The fact that the USRA had been more generous in rates than the ICC made the railroads in no hurry to reclaim control.[85] The generous wage increases and work rules made the labor unions supportive as well.[86] The shippers who had borne the burden of the rate and wage increases disagreed.[87]

Senator Albert Cummins (R-Iowa) introduced legislation on September 2, 1919, that would have consolidated the industry into a single entity, outlawed strikes, and based rates on a "fair" return on capital.[88] Although this bill passed the Senate, it faced opposition in the House, led by shippers complaining about the rate increases and who preferred the more shipper-friendly ICC.[89] Pressured by Wilson's announced intention to end federal control on March 1, 1920,[90] Congress enacted compromise legislation.[91] The effect of the legislation was to reinstate the prewar status quo, while protecting the industry from competition by authorizing pooling arrangements (subject to ICC approval), authorizing the ICC to set minimum as well as maximum rates, and forcing the most profitable lines to subsidize weaker lines.[92] It also gave railroads a two-year guarantee of 5.5% return on investment and established a Railroad Labor Board to settle labor disputes.[93]

The government's contemporaneous takeover of the U.S. radio system has drawn less attention. The Navy had argued for government control over radio communications even before the war broke out.[94] Representative Joshua W. Alexander, Chairman of the committee with jurisdiction over the issue, introduced legislation in December 1916 that would have authorized

83. Federal Railroad Control Act, ch. 25, § 14, 40 Stat. 451, 458 (1918).

84. President Woodrow Wilson, Sixth Annual Address (Dec. 2, 1918), *in* 18 A COMPILATION OF THE MESSAGES AND PAPERS OF THE PRESIDENTS, *supra* note 77, at 8645.

85. Rowland, *supra* note 74, at 12.

86. ELY, *supra* note 73, at 245.

87. *Id.*

88. S. 2906, 66th Cong. (1919).

89. Rowland, *supra* note 74, at 13.

90. President Woodrow Wilson, A Proclamation (Dec. 24, 1919), *in* 18 A COMPILATION OF THE MESSAGES AND PAPERS OF THE PRESIDENTS, *supra* note 77, at 8804, 8804–05.

91. Transportation Act of 1920, ch. 91, 41 Stat. 456, 457.

92. ELY, *supra* note 73, at 246–47.

93. Rowland, *supra* note 74, at 13.

94. SUSAN J. DOUGLAS, INVENTING AMERICAN BROADCASTING, 1899–1922, at 258 (1987); PHILIP T. ROSEN, THE MODERN STENTORS: RADIO BROADCASTERS AND THE FEDERAL GOVERNMENT, 1920–1934, at 21–22 (1980); Ronald E. Sutton, *The Nationalization of the United States Radio System in 1917*, 10 J. VISUAL LITERACY 8, 9, 12 (1990). Some lower Navy officials disagreed. ROSEN, *supra*, at 22.

the Navy to take control over the entire radio system.[95] After hearings conducted in January 1917,[96] these proposals were allowed to die in committee.[97] The outbreak of war allowed the President to invoke the provision of the Radio Act of 1912 authorizing him to close or take control of all radio stations during times of war,[98] which Wilson asserted on the very day the Senate ratified his declaration of war on Germany.[99]

The Navy ran the radio system with an iron fist. It incorporated more than fifty commercial stations into its network and closed all of the others.[100] It shut down all amateur operators, requiring that they certify that they had lowered their antennae and disconnected and sealed all of their transmitting and receiving equipment.[101] Then, Undersecretary of the Navy Franklin D. Roosevelt broke the logjam caused by blocking patents[102] by indemnifying all companies from liability for patent infringement.[103] The Navy also acquired radio companies both to consolidate patents and the industry structure.[104] As we shall see, during this time, the Navy testified in support of the federal takeover of the telephone system.[105]

95. H.R. 19350, 64th Cong. § 6 (2d Sess. 1916); *see also Wireless Bill Introduced*, N.Y. TIMES, Dec. 20, 1916, at 12, http://query.nytimes.com/mem/archive-free/pdf?res=9A06EEDD153BE633A25753C2A9649D946796D6CF (noting the date of introduction).

96. *Radio Communication: Hearings Before the H. Comm. on the Merch. Marine & Fisheries on H.R. 19350*, 64th Cong. 3 (1917) (statement of Joshua W. Alexander, Chairman).

97. DOUGLAS, *supra* note 94, at 282; Sutton, *supra* note 94, at 12–13.

98. Act of Aug. 13, 1912, ch. 287, § 2, 37 Stat. 302, 303.

99. President Woodrow Wilson, Executive Order 2582 (Apr. 6, 1917), *in* 17 A COMPILATION OF THE MESSAGES AND PAPERS OF THE PRESIDENTS, *supra* note 77, at 8241.

100. JOSEPHUS DANIELS, ANNUAL REPORT OF THE SECRETARY OF THE NAVY FOR THE FISCAL YEAR 1917, H.R. DOC. NO. 65-618, at 44 (2d Sess. 1917).

101. Sutton, *supra* note 94, at 14.

102. In 1917, the Second Circuit upheld the validity of a Marconi-held patent that was essential to the vacuum tube. *See* Marconi Wireless Tel. Co. of Am. v. De Forest Radio Tel. & Tel. Co., 243 F. 560, 565–67 (2d Cir. 1917) (upholding the lower court's finding of a valid Marconi-held patent and rejecting De Forest's counterclaim for patent infringement). Other key patents needed by vacuum tubes were held by AT&T and Columbia student Edwin H. Armstrong. 1 ERIK BARNOUW, A TOWER IN BABEL: A HISTORY OF BROADCASTING IN THE UNITED STATES 47 (1966). General Electric held the patent on another key vacuum-tube technology known as the Alexanderson alternator. *Id.* at 48–49.

103. GLEASON L. ARCHER, HISTORY OF RADIO TO 1926, at 137, 138 n.12 (1938); Sutton, *supra* note 94, at 15. The government's initial position was that the Act of June 25, 1910, Pub. L. No. 61-305, ch. 423, 36 Stat. 851, transferred any liability for patent infringement from the government contractors to the government. OFFICE OF NAVAL RECORDS & LIBRARY, HISTORY OF THE BUREAU OF ENGINEERING OF THE NAVY DURING THE WORLD WAR 128–29 (1922). The Supreme Court rejected this conclusion. Marconi Wireless Tel. Co. of Am. v. Simon, 246 U.S. 46, 55–56 (1918). Congress subsequently enacted legislation establishing that the only remedy for patents infringed by government contractors would be against the United States. Naval Appropriations Act, Pub. L. No. 65-182, ch. 114, 40 Stat. 704, 705 (1918).

104. OFFICE OF NAVAL RECORDS & LIBRARY, *supra* note 103, at 113–14; ROSEN, *supra* note 94, at 23; Sutton, *supra* note 94, at 15–16.

105. *See infra* note 197 and accompanying text. Support within the Navy was not universal. Chief of Naval Operations Captain David Todd argued that unlike radio communications, wireline

On November 21, 1918, ten days after the end of the war, Chairman Alexander submitted a bill supported by both Wilson and Secretary of the Navy Josephus Daniels that would have given the Navy permanent control.[106] Daniels's Annual Report of December 1 opined that the Navy's successful operation of the commercial radio system "presages the way for making this service entirely governmental."[107] Alexander's committee conducted hearings on the proposal from December 12–19.[108] The Navy found little industry support, having tactlessly alienated both the large industry players and amateur enthusiasts.[109] It also dissipated its energies jousting with other federal departments such as the Post Office, the Commerce Department, and the Army, which were also attempting to assert control over the radio industry, and failed to marshal popular and congressional support.[110] The new Republican Congress proved less amenable than its Democratic predecessor.[111] The Committee was particularly angered by the Navy's acquisition of radio companies, which led to the tabling of the bill,[112] the enactment of an appropriations rider prohibiting further acquisitions,[113] and calls for the divestiture of the acquired properties[114] and even Daniels's impeachment.[115] The Navy tried again in July 1919,[116] only to face similar opposition.[117]

The Navy's belief that the American Marconi Company was controlled by British interests led it to view returning the radio industry to its prior owners as unacceptable.[118] It approached Owen D. Young, who was General Counsel to General Electric, to form a new company known as the Radio

communications were less subject to interference and were provided by companies that were more cooperative. Sutton, *supra* note 94, at 16–17.

106. H.R. 13159, 65th Cong. (2d Sess. 1918).

107. JOSEPHUS DANIELS, ANNUAL REPORT OF THE SECRETARY OF THE NAVY, H.R. DOC. NO. 65-1450, at 22 (3d Sess. 1918).

108. *Government Control of Radio Communication: Hearings Before the H. Comm. on the Merch. Marine & Fisheries on H.R. 13159*, 65th Cong. (1918).

109. Sutton, *supra* note 94, at 19.

110. *Id.* at 10, 19.

111. DOUGLAS, *supra* note 94, at 282.

112. *Id.* at 283.

113. *Blocks Purchase of Radio Systems: House Amends Naval Bill to Prevent Diversion of Steam Engineering Bureau Funds*, N.Y. TIMES, Feb. 8, 1919, http://query.nytimes.com/mem/archive-free/pdf?res=F0061FF73B5D147A93CAA91789D85F4D8185F9.

114. *Wants Radios Returned: Congressman Rowe Also Thinks $3,000,000 Should Be Recovered*, N.Y. TIMES, Jan. 18, 1919, http://query.nytimes.com/mem/archive-free/pdf?res=F50E1FF6385D147A93CAA8178AD85F4D8185F9.

115. *Wants Daniels Ousted: Mann Says He Should Be Impeached for Radio Purchases*, N.Y. TIMES, Jan. 30, 1919, http://query.nytimes.com/mem/archive-free/pdf?res=F20A14FB3B5D147A93C2AA178AD85F4D8185F9.

116. S. 3399, 66th Cong. (1st Sess. 1919).

117. DOUGLAS, *supra* note 94, at 284.

118. *Id.*

Company of America (RCA) guaranteed to be under American control.[119] American Marconi transferred all of its assets into the company in exchange for RCA stock.[120] RCA, General Electric, and AT&T entered into cross-licensing agreements that neutralized the patent thicket obstructing U.S. development.[121] The effect was to create a government-sanctioned monopoly.[122]

D. *Labor Conflict*

These debates took place in the shadow of a looming telegraph operators' strike designed to force Western Union to unionize.[123] The Commercial Telegraphers' Union initially called the strike for April 9, 1918, but postponed it to permit arbitration by the War Labor Board.[124] After that failed, the union called for another strike on July 8, only to relent once again at the request of the Secretary of Labor.[125] The war footing made the strike particularly controversial. As one contemporary editorialist put it, labor strife and labor rights "must be instantly swept aside if they in the slightest degree threaten the country's efforts to win the war."[126] Some advocates saw government control as a means of maintaining service levels despite rising labor militancy and the corresponding increased threat of strikes.[127]

E. *Progressive Sentiment and the Kingsbury Commitment*

Debates over government ownership of the telephone system were also framed by the rise of the Progressive movement. Some Progressives expressed strong distrust for large organizations and advocated strong antitrust enforcement to return to an economy dominated by small businesses.[128] Although they were suspicious of big government, they generally distrusted corporations more.[129] Others accepted corporations as a

---

119. RCA's charter stipulated that only U.S. citizens could serve as officers or directors and required that foreigners own no more than 20% of the stock. BARNOUW, *supra* note 102, at 59.

120. *Id.*

121. *Id.* at 60.

122. DOUGLAS, *supra* note 94, at 288.

123. MAY, *supra* note 17, at 28–30; OSLIN, *supra* note 16, at 278.

124. OSLIN, *supra* note 16, at 278.

125. *Washington Plea Prevents Strike on Western Union: Operators' Chief Yields to the Appeals of Secretary Wilson and Gompers*, N.Y. TIMES, July 8, 1918, http://query.nytimes.com/mem/archive-free/pdf?res=F60911FB355F157A93CAA9178CD85F4C8185F9.

126. George Harvey, *The Postal and the Western Union*, N. AM. REV.'S WAR WKLY., June 22, 1918, at 7.

127. *See* OSLIN, *supra* note 16, at 278 (indicating that President Wilson's reaction to the increased risk of strikes was to announce that Postmaster General Burleson would take over the telegraph and telephone systems for the government).

128. For a classic statement, see Louis D. Brandeis, *A Curse of Bigness*, HARPER'S WKLY., Jan. 10, 1914, at 21, *available at* http://hdl.handle.net/2027/mdp.39015022184223 (recommending implementing legislation with an antitrust focus in order to "remedy the evils" of railroad monopolies).

129. RICHARD HOFSTADTER, THE AGE OF REFORM 229, 231, 233 (1955).

part of modern life and instead favored curbing them through technocratic regulation.[130] To these Progressives, the goal of returning to the old competitive order was a chimera.[131]

This ambivalence created a large gap between rhetoric and reality.[132] Notwithstanding his reputation as the preeminent trustbuster in history, Theodore Roosevelt expressed reservations about aggressive antitrust enforcement, criticizing "the impossible task of restoring flintlock conditions of business sixty years ago."[133] He therefore brought surprisingly few antitrust cases and limited the Antitrust Division to five attorneys and an annual budget of $100,000.[134] Woodrow Wilson similarly regarded the emergence of large enterprises as "characteristic of our time" and "normal and inevitable" and stated that "we shall never return to the old order of individual competition."[135] To them, bigness was not bad per se; instead, it was culpable only when it crossed certain lines.[136] Ironically, it was the non-Progressive William Howard Taft who asserted the antitrust laws most vigorously.[137]

This ambivalence became apparent in the government's policies with respect to AT&T. When competition first emerged in 1894, AT&T's initial reaction was to attempt to outbuild the independents.[138] The result was a boon to consumers. The number of telephone connections, which had been growing at the somewhat languid annual rate of 6% prior to 1894, jumped to 20%.[139] Initially, the independents focused on areas that the Bell System had ignored, such as rural areas, small towns, and the suburbs of major cities.[140] Over time, they began to enter into direct competition with Bell. By 1902, competition existed in more than half of all cities with populations of greater than five thousand people.[141] Consumers who purchased both connections

---

130. HERBERT CROLY, THE PROMISE OF AMERICAN LIFE 357 (1909); G. EDWARD WHITE, PATTERNS OF AMERICAN LEGAL THOUGHT 104 (1978).

131. CROLY, *supra* note 130, at 358–59.

132. HOFSTADTER, *supra* note 129, at 252–53.

133. President Theodore Roosevelt, Theodore Roosevelt's Confession of Faith at the National Convention of the Progressive Party 27 (Aug. 6, 1912).

134. HOFSTADTER, *supra* note 129, at 245.

135. WOODROW WILSON, THE NEW FREEDOM 163 (1913); Woodrow Wilson, Response of Woodrow Wilson to Notification Address at the Democratic National Convention (Aug. 7, 1912), *in* OFFICIAL REPORT OF THE PROCEEDINGS OF THE DEMOCRATIC NATIONAL CONVENTION OF 1912, at 400, 407 (Urey Woodson ed., 1912).

136. HOFSTADTER, *supra* note 129, at 248.

137. Bickers, *supra* note 17, at 108.

138. Richard Gabel, *The Early Competitive Era in Telephone Communication, 1893–1920*, 34 LAW & CONTEMP. PROBS. 340, 354 (1969); Stanley M. Besen & Joseph Farrell, *Choosing How to Compete: Strategies and Tactics in Standardization*, 8 J. ECON. PERSP. 117, 122, 124 (1994).

139. Gabel, *supra* note 138, at 350 tbl.4.

140. *See id.* at 343–44 (arguing that Bell's method of providing service prevented it from developing residential, suburban, and rural service went largely undeveloped, which restricted Bell's growth).

141. VIETOR, *supra* note 16, at 170; Gabel, *supra* note 138, at 344.

could typically connect to five to ten times the number of other customers while paying less than the cost of buying a single connection during the monopoly period.[142]

From AT&T's standpoint, this strategy turned out to be a dismal failure. AT&T's prices, profits, and stock price plummeted, and the capital requirements strained the company's ability to raise capital.[143] By 1907, AT&T's market share had fallen below 50%.[144] The financial markets had had enough. The Morgan banking interests took over the company and forced a change in management, installing Theodore Vail as president.[145] Under Vail's leadership, the company stopped competing directly with the independents and instead began pursuing two classic anticompetitive strategies. As an initial matter, they attempted to merge to monopoly by offering to buy out independents with whom they competed directly.[146] If the independent refused to sell, they pursued a classic division of markets by offering to withdraw from direct competition in return for a promise from the independent that it would not expand its territory and would interconnect with AT&T's long-distance network.[147]

Vail justified the consolidation of all telephone companies into a single system with his "belie[f] that the telephone system should be universal, interdependent and intercommunicating, affording opportunity for any subscriber of any exchange to communicate with any other subscriber of any other exchange."[148] AT&T backed its strategy of withdrawing from competition with what has been described as the first major corporate public-relations campaign in history decrying the cost and inconvenience of having to maintain two separate connections, each with its own lines and handsets (known as dual service).[149] To compensate for the lack of price discipline

142. MUELLER, *supra* note 10, at 94; Robert Bornholz & David S. Evans, *The Early History of Competition in the Telephone Industry*, *in* BREAKING UP BELL 7, 30 (David S. Evans ed., 1983); David F. Weiman & Richard C. Levin, *Preying for Monopoly? The Case of Southern Bell Telephone Company, 1894–1912*, 102 J. POL. ECON. 103, 123–24 (1994); *see also* G. JOHNSTON, SOME COMMENTS ON THE 1907 ANNUAL REPORT OF AT&T 15–16 (1908) (describing the dramatic drop in the Bell System's rates).

143. Gabel, *supra* note 138, at 345–46; Weiman & Levin, *supra* note 142, at 109–23; MUELLER, *supra* note 10, at 70.

144. U.S. CENSUS BUREAU, CENSUS OF ELECTRICAL INDUSTRIES: 1917—TELEPHONES 11 (1920), *available at* http://www2.census.gov/prod2/decennial/documents/1917telephones.pdf (noting that Bell reported 51.2% of the market share in 1907).

145. Gabel, *supra* note 138, at 345.

146. *See, e.g.*, AT&T CO., ANNUAL REPORT OF THE DIRECTORS OF THE AMERICAN TELEPHONE & TELEGRAPH COMPANY TO THE STOCKHOLDERS FOR THE YEAR ENDING DECEMBER 31, 1910, at 21 (1911) ("Wherever it could be legally done, and done with the acquiescence of the public, opposition companies have been acquired and merged into the Bell System.").

147. Daniel F. Spulber & Christopher S. Yoo, *Toward a Unified Theory of Access to Local Telephone Networks*, 61 FED. COMM. L.J. 43, 71 (2008).

148. AT&T CO., *supra* note 146, at 22–23.

149. Spulber & Yoo, *supra* note 147, at 71.

resulting from the elimination of competition, AT&T dropped its long-standing opposition to government oversight and willingly submitted to rate regulation.[150]

The result was an abrupt end to the erosion of AT&T's market share.[151] Some competitors began to complain that the mergers represented a violation of the antitrust laws.[152] These complaints did not prompt any immediate action by the Taft Administration, which despite its willingness to use the antitrust laws to break up Standard Oil and American Tobacco,[153] viewed each telephone merger as an independent event instead of evaluating them as part of a systematic campaign.[154] Independents warned that although each individual acquisition involved purely intrastate commerce, "[t]he avowed purpose of the Bell Company is to buy or crowd out the independent companies, which in the end will give them a complete monopoly of the telephone."[155] The Attorney General referred the matter to the ICC, which declined to act and eventually dropped the investigation.[156]

Toward the end of the Taft Administration, however, the Justice Department began to view AT&T's acquisition campaign with greater skepticism. Concerned about acting too hastily and giving the appearance of political grandstanding on the eve of a presidential election, the Attorney General simply asked AT&T not to consummate any pending transactions until after the election,[157] a request with which AT&T complied.[158] The Wilson Administration successfully settled the case on December 13, 1913, when AT&T agreed to the so-called Kingsbury Commitment, named after the AT&T Vice President, Nathan Kingsbury, who brokered it.[159] According

150. AT&T CO., ANNUAL REPORT OF THE DIRECTORS OF AMERICAN TELEPHONE & TELEGRAPH COMPANY TO THE STOCKHOLDERS FOR THE YEAR ENDING DECEMBER 31, 1907, at 18 (1908).

151. Spulber & Yoo, *supra* note 147, at 73.

152. *See, e.g.*, Letter from Edward F. Murray, President, Murray's Line, to George W. Wickersham, Att'y Gen. (Nov. 13, 1912), *cited in* Bickers, *supra* note 17, at 116 n.64. *See generally* Letter from George W. Wickersham, Att'y Gen., to Charles A. Prouty, Chairman, Interstate Commerce Comm'n (Jan. 7, 1913), *quoted in* Bickers, *supra* note 17, at 113–14 (summarizing these complaints).

153. United States v. Standard Oil Co., 221 U.S. 1, 81–82 (1911); United States v. Am. Tobacco Co., 221 U.S. 106, 188 (1911).

154. Bickers, *supra* note 17, at 115.

155. Letter from Edward F. Murray, President, Murray's Line, to George W. Wickersham, Att'y Gen. (Nov. 23, 1912), *quoted in* Bickers, *supra* note 17, at 116.

156. Press Release, Interstate Commerce Comm'n, Investigation of Telephone and Telegraph Companies, Docket No. 5462 (Apr. 15, 1914), *cited in* Bickers, *supra* note 17, at 115 n.63.

157. Memorandum from George W. Wickersham, Att'y Gen., to J.A. Fowler, Assistant to the Att'y Gen. (Aug. 29, 1912), *quoted in* Bickers, *supra* note 17, at 117.

158. Letter from Theodore Vail, President, AT&T, to the presidents of all associated Bell telephone companies (Aug. 6, 1912), *cited in* Bickers, *supra* note 17, at 118 n.69.

159. Burleson stated, "If the efficient management and direction is given the telegraph and telephone that has been given the Postal Service, the probability is that they never will be returned to private control." DANIELIAN, *supra* note 17, at 246; *see also* Bickers, *supra* note 17, at 147

to the terms of the agreement, AT&T pledged to stop acquiring directly competing companies.[160] AT&T also promised to divest its ownership stake in Western Union and to permit the independents to interconnect with its long-distance network.[161]

Commentators have not been kind to the Kingsbury Commitment. Some complain that by allowing the Bell System to keep the properties instead of breaking it up as it did with Standard Oil, the antitrust authorities effectively condoned monopoly by refusing to undo the existing acquisitions and leaving them intact.[162] Others have complained that the Kingsbury Commitment was toothless. For example, the Commitment did not prevent AT&T from acquiring independent local telephone companies with which it did not directly compete.[163] Even where companies competed directly, other scholars claim that the Justice Department permitted mergers so long as AT&T divested an equal number of lines elsewhere.[164] Others are somewhat less critical, insisting that the Kingsbury Commitment was successful in slowing down mergers temporarily.[165] As discussed below, the Kingsbury Commitment was more effective than commonly thought at slowing consolidation.[166] In any event, the Kingsbury Commitment was an important progressive policy preceding the nationalization of the wires.

F. *Early Debates over Nationalizing the U.S. Telephone System*

Interest in government ownership of the telegraph system began to extend to the telephone system as well in late 1913. The *New York Times* reported in October of that year, "Notwithstanding efforts at profound secrecy, it has become known here that the Wilson Administration is

---

(characterizing the temporary nationalization of the telephone lines as an experiment in government control as a permanent policy); STEHMAN, *supra* note 16, at 177 (same).

160. DANIELIAN, *supra* note 17, at 76.

161. *Id.* Before the Kingsbury Commitment, courts had been reluctant to rely on the antitrust laws to justify mandating interconnection with the long-distance network. *See* U.S. Tel. Co. v. Cent. Union Tel. Co., 202 F. 66, 72 (6th Cir. 1913) (declining to discuss whether an exchange could be compelled to provide long-distance service); Pac. Tel. & Tel. Co. v. Anderson, 196 F. 699, 705 (E.D. Wash. 1912) (holding that a company had no right to demand a physical connection with another line).

162. WU, *supra* note 7, at 56; *see also* Dean Burch, *Common Carrier Communications by Wire and Radio: A Retrospective*, 37 FED. COMM. L.J. 85, 87 (1985) (a former FCC Chairman noting that "by the time of the so-called Kingsbury Commitment in 1913, . . . AT&T's monopolization of the telephone industry was well on its way to becoming an accomplished fact"); Harry M. Trebing, *Common Carrier Regulation—The Silent Crisis*, 34 LAW & CONTEMP. PROBS. 299, 305 (1969) (calling the reestablishment of monopoly "a *fait accompli* by 1913").

163. COHEN, *supra* note 15, at 48; VIETOR, *supra* note 16, at 172; Gabel, *supra* note 138, at 352–53.

164. BROCK, *supra* note 15, at 155–56; Thierer, *supra* note 16, at 272.

165. HUBER ET AL., *supra* note 6, § 4.4.2; MUELLER, *supra* note 10, at 134.

166. *See infra* section III(A)(2).

engaged in preparing the groundwork" for nationalization of the nation's telegraph and telephone lines.[167]

On December 1, 1913, Postmaster General Burleson's second annual report confirmed these suspicions, revealing that the Post Office had been studying the possible acquisition of the telegraph and telephone systems since the previous June.[168] On December 20, Representative David J. Lewis (D-Md.) introduced a resolution directing the relevant committees to consider a bill providing for the postalization of the telephone network.[169] He followed that with an extended defense of the merits of postalizing the telephone system that occupied thirty-five pages of the *Congressional Record*.[170] A December 23 meeting between Burleson and President Wilson left Burleson reluctant to press the issue, suggesting that the Administration was not unified in its support of Burleson's proposal.[171]

The debate continued into early 1914. On January 29, the Senate passed a resolution directing the Postmaster General to send the results of his investigation to the Senate.[172] Burleson complied on January 31, submitting a nearly 150-page report laying out the case for government ownership of both the telephone and telegraph system[173] that apparently drew the support of the Navy.[174] The House Committee on the Post Office and Post Roads would subsequently conduct hearings on "The Postalization of the Telephone" on January 15, 1915.[175] At these new hearings, Representative Lewis made a speech that was quite similar to his speech of December 1913, emphasizing the public benefits of government control.[176] The only other witness at the hearing was an officer of the second largest telegraph company

167. *Federal Wires New Wilson Plan: Policy Afoot to Control Nation's Telephones as Key to Government Telegraph*, N.Y. TIMES, Oct. 2, 1913, http://query.nytimes.com/mem/archive-free/pdf?res=FA0D11F63D5913738DDDAB0894D8415B838DF1D3.

168. ALBERT S. BURLESON, ANNUAL REPORTS FOR THE FISCAL YEAR ENDED JUNE 30, 1913, H.R. DOC. NO. 63-712, at 16 (2d Sess. 1914).

169. 51 CONG. REC. 1377 (1913).

170. *Id.* at 1377–412.

171. *Lewis Opens Fight for U.S. Telephones*, N.Y. TIMES, Dec. 23, 1913, http://query.nytimes.com/gst/abstract.html?res=FB0910F63A5813738DDDAA0A94DA415B838DF1D3; *accord* BROOKS, *supra* note 16, at 149 (noting that "the government advocates of nationalization seemed to hang back awaiting their opening").

172. Senate Resolution 242 was submitted on January 12 and was initially passed by unanimous consent. 51 CONG. REC. 1503 (1914). The next day, passage of the resolution was reconsidered, and the resolution was referred to the Committee on Post Offices and Post Roads. *Id.* at 1569. The Committee reported the resolution favorably January 29, and the Senate passed it. *Id.* at 2503–04.

173. POSTMASTER GEN., GOVERNMENT OWNERSHIP OF ELECTRICAL MEANS OF COMMUNICATION, S. DOC. NO. 63-399 (2d Sess. 1914).

174. ROSEN, *supra* note 94, at 21–22.

175. *The Postalization of the Telephone: Hearing Before the H. Comm. on the Post Office and Post Roads on H.R. 20471*, 63d Cong. (1915).

176. *Id.* at 3–143.

who testified in support of postalization in order to curb supposed abuses of his chief rival, Western Union.[177]

AT&T responded with a vigorous campaign against postalization that attempted to stake out a middle ground in favor of a private monopoly subject to government regulation.[178] Its most extensive statement was a thirty-four-page discussion in its annual report on 1913, which asserted "no government owned telephone system in the world is giving as cheap and efficient service as the American public is getting from all its telephone companies."[179] Other prominent examples of AT&T efforts to counter the rising sentiment in favor of nationalization include publishing a point-by-point rebuttal of Representative Lewis's floor statement,[180] public speeches made by Vail and other Bell officials,[181] as well as language in its annual reports.[182]

Although Progressives were willing to use the antitrust laws and regulation to curb monopolies, they were far more ambivalent about government ownership. Although some commentators have simplistically seen Progressivism as favoring nationalization,[183] Progressives' attitudes were much more complex. Specifically, government ownership pitted Progressives' faith in scientific administration and centralized control against their intuitive distrust of uncontrolled economic power.[184] Wilson's scholarly work placed him in the camp of the government-ownership skeptics. In a book published within two years of his inauguration, Wilson argued that, although natural monopolies can harm the public interest, in

---

177. *Id.* at 145–56.

178. Bickers, *supra* note 17, at 142–44.

179. AT&T CO., *supra* note 159, at 28–62.

180. AT&T CO., GOVERNMENT AND PRIVATE TELEGRAPH AND TELEPHONE UTILITIES: AN ANALYSIS (1914), *reprinted in* SELECTED ARTICLES ON GOVERNMENT OWNERSHIP OF TELEGRAPH AND TELEPHONE 129–57 (Katharine B. Judson ed., 1914).

181. Theodore Newton Vail, Some Observations on Modern Tendencies (Oct. 1915), *in* VIEWS ON PUBLIC QUESTIONS: A COLLECTION OF PAPERS AND ADDRESSES OF THEODORE NEWTON VAIL 1907–1917, at 240, 258–63 (1917); F.H. Bethell, Some Comment on Government Ownership of Telephone Properties (Feb. 25, 1914), *reprinted in* SELECTED ARTICLES ON GOVERNMENT OWNERSHIP OF TELEGRAPH AND TELEPHONE, *supra* note 152, at 159.

182. AT&T CO., ANNUAL REPORT OF THE DIRECTORS OF AMERICAN TELEPHONE & TELEGRAPH COMPANY TO THE STOCKHOLDERS FOR THE YEAR ENDING DECEMBER 31, 1915, at 50 (1916); AT&T CO., ANNUAL REPORT OF THE DIRECTORS OF AMERICAN TELEPHONE & TELEGRAPH COMPANY TO THE STOCKHOLDERS FOR THE YEAR ENDING DECEMBER 31, 1916, at 49–51 (1917).

183. *See, e.g.*, COHEN, *supra* note 17, at 37 (highlighting the impact the rise of the Progressive movement had on proposals to nationalize the telegraph). *See generally* RICHARD EPSTEIN, HOW PROGRESSIVES REWROTE THE CONSTITUTION (2006) (characterizing Progressives as supporters of economic nationalism).

184. Bickers, *supra* note 17, at 88–89. On the conflicts within the Progressive movement, see HOFSTADTER, *supra* note 129, at 215–71 (describing the tension between business monopoly and political freedom in an era of increased reliance on government regulation); Michael J. Sandel, *Democracy's Discontent: America in Search of a Public Philosophy*, 85 GEO. L.J. 2073, 2076–77 (1996) (detailing the varied responses among the Progressives to the threat corporate power posed to self-government).

most cases government ownership would be inferior to government regulation.[185]

Public reaction was also largely critical of government ownership. A number of scholars criticized the methodology of Burleson's analysis.[186] The popular press was largely critical as well, raising concerns about efficiency of government operations as well as the potential abuse of patronage.[187] They also denigrated the performance of government-owned telephone systems in Europe, with one industry executive quipping, "And as to service—Government service would be a joke as compared with present service. If you don't believe it just try the Government service—telegraph and telephone—in Europe."[188]

The imposition of the Kingsbury Commitment in December 1913 diverted Wilson's interest in pursuing government ownership.[189] Burleson maintained a steady drumbeat in support of nationalization in his annual reports, focusing some of his energy on the more limited goal of nationalizing the telephone systems of Alaska, Hawaii, and Puerto Rico.[190] The Navy did conduct a successful three-day test mobilization during May 6–8, 1916, during which the Navy used AT&T's network for all communication between all naval facilities and ships.[191]

In January and February 1917, the House Committee on the District of Columbia conducted extensive hearings on the possibility of the federal government taking over the D.C. telephone system, well before the U.S.

---

185. WOODROW WILSON, THE STATE: ELEMENTS OF HISTORICAL AND PRACTICAL POLITICS ¶¶ 1524–26 (1913).

186. *E.g.*, A.N. Holcombe, *Public Ownership of Telegraphs and Telephones*, 28 Q.J. ECON. 581, 583–86 (1914).

187. COHEN, *supra* note 15, at 38; STONE, *supra* note 3, at 197.

188. STONE, *supra* note 3, at 197; *see also Wilson Gets Facts on Wire Control*, N.Y. TIMES, Oct. 3, 1913, http://query.nytimes.com/mem/archive-free/pdf?res=940CE1DF133FE633A25750C0A9669D946296D6CF (citing Vail as saying government help with long-distance service could be welcome); *C.H. Mackay Derides Federal Ownership*, N.Y. TIMES, Dec. 19, 1913, http://query.nytimes.com/mem/archive-free/pdf?res=F40F13F73B5B13738DDDA00994DA415B838DF1D3 (projecting that government ownership would result in operating at a loss as it had in the English context).

189. STONE, *supra* note 3, at 197–98.

190. *See* ALBERT S. BURLESON, ANNUAL REPORT OF THE POSTMASTER GENERAL FOR THE FISCAL YEAR ENDED JUNE 30, 1914, H.R. DOC. NO. 63-1387, at 14–16 (3d Sess. 1914) (emphasizing that nationalizing the telegraph and telephone systems remained desirable, particularly in Alaska, Hawaii, and Puerto Rico); ALBERT S. BURLESON, ANNUAL REPORT OF THE POSTMASTER GENERAL FOR THE FISCAL YEAR ENDED JUNE 30, 1915, H.R. DOC. NO. 64-358, at 51–52 (1916) (same); ALBERT S. BURLESON, ANNUAL REPORT OF THE POSTMASTER GENERAL FOR THE FISCAL YEAR ENDED JUNE 30, 1916, H.R. DOC. NO. 64-1728, at 46–48 (1917) (same); ALBERT S. BURLESON, ANNUAL REPORT OF THE POSTMASTER GENERAL FOR THE FISCAL YEAR ENDED JUNE 30, 1917, H.R. DOC. NO. 65-770, at 79 (1918) (reiterating, generally, the claims contained in previous reports but omitting the claims regarding Hawaii and Puerto Rico).

191. DEP'T OF THE NAVY, ANNUAL REPORT OF THE NAVY DEPARTMENT FOR THE FISCAL YEAR 1916, H.R. DOC. NO. 64-1480, at 29 (2d Sess. 1917); BROOKS, *supra* note 16, at 150.

entry into World War I on April 6, 1917.[192] On March 4, 1918, Burleson responded to a Senate request for information with a scathing criticism of the telephone service provided in Washington, D.C.[193] Ten days later, the House Committee on the District of Columbia favorably reported a bill authorizing the government takeover of the D.C. phone system.[194] Advocates clearly regarded the D.C. takeover more as an experiment in a permanent policy than a wartime measure.[195] Proposals for long-term government control enjoyed little support. Legislation authorizing permanent government operation of the entire telephone system submitted in January 1918 by the same Representative who would sponsor the successful temporary takeover legislation died in committee.[196]

II. The History of the Government Takeover

The federal government's decision to take control of the U.S. telephone system was part of a broader debate over the proper role of the government during times of both peace and war. In reviewing this history, it is important to keep in mind that the forces driving the decision both to take the telephone system over and to give it back are complex. Were it simply a matter of reflexive support for the state during times of armed conflict, one would expect the takeover to have occurred as soon as war was declared, as was done with respect to radio. Instead, Congress waited eight months to take over the railroads and another nine months to assume control of the telephone system, pointedly declining to take action on an earlier proposal until the measure was framed as a prophylactic, emergency measure and the President gave it his political support. Perhaps even more interesting is the manner in which the underlying technology and the way that the government ran the telephone network influenced the decision to return the wires a year later.

A. Enacting of the Takeover

The legislation that would lead to the government takeover was introduced on June 27, 1918. The House Committee on Interstate and Foreign Commerce conducted hearings on July 2, at which the only witnesses were three government officials who were widely recognized as advocates of a permanent takeover: Burleson, Secretary of War Newton

192. *Government Monopoly of Telephone Communication in the District of Columbia: Hearing Before the H. Comm. on the D.C. on H.R. 18723*, 64th Cong. (1917).

193. ALBERT S. BURLESON, POSTMASTER GENERAL'S STATEMENT, H.R. REP. NO. 65-379, at 23–27 (2d Sess. 1918).

194. H.R. REP. NO. 65-379.

195. STEHMAN, *supra* note 16, at 177.

196. H.J. Res. 206, 65th Cong. (1918).

Baker, and Secretary of the Navy Josephus Daniels.[197] The Committee issued a report supporting the bill.[198] The Senate Committee on Interstate Commerce declined to conduct full hearings,[199] choosing on July 9 only to hear from Western Union President Newcomb Carlton, who testified that he saw no necessity that would justify taking over the telegraph system and that even if that were to happen, there was even less justification for taking over the telephone system.[200] He did state publicly that he would prefer a government takeover to yielding to unionization.[201] No representative from AT&T participated in either hearing, although there is some ambiguity about whether AT&T actively opposed the measure.[202]

The war added a new dimension to the debate over nationalizing the telephone system. On June 28, Burleson wrote to Representative Thetus W. Sims (D-Tenn.) that government control was necessary "to prevent communications by spies and other public enemies" and "imperative to safeguard public interests."[203] Burleson said "paralysis of a large part of the system" was threatened, and there were "possible consequences prejudicial to our military preparations and other public activities that might prove serious or disastrous."[204] Comparing the American response to those of European states, Burleson concluded, "We are reminded that there is not a nation engaged in the war that [e]ntrusts its military or other communications to unofficial agencies."[205] Burleson as well as Secretary of War Baker and Secretary of the Navy Daniels all indicated that government ownership was

---

197. *Federal Control of Systems of Communication: Hearings Before the H. Comm. on Interstate and Foreign Commerce on H.J. Res. 309*, 65th Cong. (1918) [hereinafter *1918 House Hearings*]. On all three witnesses' established support for government ownership, see 56 CONG. REC. 8717 (1918) (statement of Rep. Martin Madden).

198. H.R. REP. NO. 65-741 (1918).

199. *Votes 7 to 3 for Wire Control*, N.Y. TIMES, July 10, 1918, http://query.nytimes.com/mem/archive-free/pdf?res=FB0F17F8355F157A93C2A8178CD85F4C8185F9.

200. *1918 House Hearings*, *supra* note 197, at 7–8, 17.

201. *House Votes Wire Control; Senate Waits*, N.Y. TIMES, July 6, 1918, http://query.nytimes.com/mem/archive-free/pdf?res=F00D1FF83C5A11738DDDAF0894DF405B888DF1D3.

202. Some commentators point out that AT&T did not oppose the measure and indicate that "it was freely said that President Vail was in favor of government control." DANIELIAN, *supra* note 10, at 246. Company legend holds that Vail went to Wilson in early 1918 and stated, "As long as you've taken over the railroads, you might as well take us over, too." BROOKS, *supra* note 9, at 151. AT&T Vice President Kingsbury later denied that AT&T supported the takeover and said that he had attempted to gain admission to both the House and Senate Committee hearings, but was denied in both cases. *Return of the Wire Systems: Hearings Before the H. Comm. on Interstate and Foreign Commerce on H.R. 421*, 66th Cong. 22–23 (1919) [hereinafter *Return of the Wire Systems Hearings*] (statement of Nathan C. Kingsbury, Vice President, American Telephone & Telegraph Company).

203. Letter from A.S. Burleson, Postmaster Gen., to Thetus W. Sims, Chairman, H. Comm. on Interstate & Foreign Commerce (June 28, 1918), *in* 56 CONG. REC. 8719 (1918).

204. *Id.*

205. *Id.*

needed to prevent government secrets from falling into enemy hands.[206] They also suggested that a strike by telecommunications workers would be particularly debilitating to the war effort.[207]

Most importantly, President Wilson signaled his support for a temporary takeover by sending a letter endorsing Burleson's arguments in support of the bill.[208] His support was pivotal. During the House floor debate, Representative Martin B. Madden (R-Ill.) noted the absence of the President's explicit support.[209] At the same time, he conceded, "If the President says to the House that he wants any power that will enable him to successfully conduct the war, there is no man in the House who will not vote to give it to him."[210] Supporters of the bill thereupon produced the letter, and it passed the House by a vote of 222 to 4.[211]

The debate forced supporters of the government takeover to place important limitations on the bill. Although Burleson and Daniels clearly harbored ambitions to make the takeover permanent,[212] Congress had rejected earlier legislation that was not limited to wartime.[213] The bill carefully avoided this problem by limiting its effect to the duration of the war.[214] Indeed, the bill's sponsor, Representative James D. Aswell (D-La.), specifically disavowed any intention of making government ownership permanent.[215]

The absence of any emergency to justify the takeover as well as concerns that Burleson might use a wartime measure as a prelude to a more permanent takeover of the wires[216] led Wilson to assure that the power would only be used in case of a telegraph strike and to emphasize the importance of

---

206. *1918 House Hearings*, *supra* note 197, at 3 (statement of Newton D. Baker, Secretary of War), 19 (statement of Josephus Daniels, Secretary of the Navy), 45 (statement of Albert S. Burleson, Postmaster General).

207. *Id.* at 41 (statement of Albert S. Burleson, Postmaster General).

208. Letter from President Woodrow Wilson to Thetus W. Sims, House of Representatives (June 28, 1918), *in* 56 CONG. REC. 8718 (1918).

209. 56 CONG. REC. 8717 (1918) (statement of Rep. Madden).

210. *Id.* at 8718.

211. *Id.* at 8735.

212. Burleson stated, "If the efficient management and direction is given the telegraph and telephone that has been given the Postal Service, the probability is that they never will be returned to private control." DANIELIAN, *supra* note 17, at 246; *see also* Bickers, *supra* note 17, at 147 (arguing the advocates of government control viewed it as an experiment for permanent government control); MAY, *supra* note 17, at 31–32, 36–38 (same); STEHMAN, *supra* note 16, at 177 (same).

213. *See* 56 CONG. REC. 8719 (1918) (statement of Rep. Sims) ("[T]he resolution introduced by Mr. Aswell in January was not a war-time proposition. . . . It was not confined to the operation of the war, and therefore never considered by the committee.").

214. *Id.* at 8721 (statement of Rep. Esch).

215. *Id.* at 8720 (statement of Rep. Aswell).

216. *Id.* at 8719 (statement of Rep. Madden) (noting that "[e]verybody knows the Postmaster General is a 'bug' on Government ownership").

putting such authority in place before it was needed.[217] Burleson's statement in support of the proposal was similarly contingent, urging passage of the resolution "in order that the President may act, *if necessary*."[218]

Indeed, the tone of the debate suggested that the authority was a prophylactic measure. Although no exigency currently existed, Congress felt that giving the President the authority would allow him to act promptly should the need arise.[219] Representative Sims, who was the floor manager, similarly noted that "this power might be needed at any moment" and that the Administration simply asked that "the President be clothed with the power, so that he might exercise it *if the emergency arose*."[220] Aswell emphasized that the takeover authority was not permanent,[221] as did other members in the debate.[222]

After a rancorous Senate floor debate[223] that forced postponement of a planned recess[224] and despite several editorials opposing the move,[225] Congress subsequently adopted the resolution, and Wilson signed it into law on July 16, 1918.[226] The text of the proposal made clear that it was an emergency measure. The takeover was to be exercised only when the President "shall deem it necessary for the national security or defense."[227] Moreover, the statute explicitly provided that the takeover would end with the ratification of a peace treaty ending the war.[228]

Wilson, however, wasted little time and exercised this power via a proclamation on July 22, 1918, that gave the federal government control of

217. *President Asks Power to Control All Wire Systems*, N.Y. TIMES, July 2, 1918, http://query.nytimes.com/gst/abstract.html?res=F50713FB3C5F15738DDDAB0894DF405B888DF1D3.

218. 56 CONG. REC. 8719 (1918) (emphasis added).

219. *Id.* at 8716 (statement of Rep. Sims); *see also* Bickers, *supra* note 17, at 148–49 (referencing Representative Sims's statement concerning the need for Presidential authority if exigent circumstances arose).

220. 56 CONG. REC. 8716 (1918) (statement of Rep. Sims) (emphasis added).

221. *Id.* at 8720 (statement of Rep. Aswell) (calling the "fear of permanent Government ownership resulting from this legislation . . . ill founded").

222. *Id.* at 8717 (statement of Rep. Sims) (noting that unlike Aswell's previous proposal, the current proposal "continued the control only during the existence of war").

223. *Id.* at 8741–47, 8841–43, 8934, 8937, 8959–62, 9069–78.

224. *Congress Recess Held up by Fight on Wire Control*, N.Y. TIMES, July 7, 1918, http://query.nytimes.com/gst/abstract.html?res=F60C12F63B5F1B7A93C5A9178CD85F4C8185F9.

225. Editorial, *Government Control*, N.Y. TIMES, July 6, 1918, http://query.nytimes.com/gst/abstract.html?res=9A00E7D8173EE433A25755C0A9619C946996D6CF&scp=2&sq=government+control&st=p; *The Unrepresented Public*, N.Y. TIMES, July 11, 1918, http://query.nytimes.com/gst/abstract.html?res=9E07E5D7173EE433A25752C1A9619C946996D6CF.

226. H.R.J. Res. 309, 65th Cong., 40 Stat. 904 (1918).

227. *Id.*

228. *Id.*

the wires effective July 31.[229] In his proclamation, Wilson directed Burleson to undertake "the supervision, possession, control and operation of [the] telegraph and telephone systems."[230] Wilson assured stock and bondholders that their interests would not be jeopardized. He stated, "Regular dividends hitherto declared, and maturing interest upon bonds, debentures, and other obligations, may be paid in due course."[231] On November 2, Wilson issued a proclamation taking over the submarine cables as well.[232] Although hostilities ended on November 11, 1918,[233] and the Treaty of Versailles ending the war was signed on June 28, 1919,[234] the government would continue to operate the wires until midnight July 31, 1919, just slightly more than one year after taking them over.

B. *Running the Telephone System*

The Post Office that took over the telephone system on July 31, 1918 faced some seemingly insurmountable challenges.[235] Most basically, the Post Office lacked the experience and administrative capacity to manage a large telecommunications network.[236] In stark contrast to the takeover of the railroads, Congress had failed to provide any detailed guidance as to how the system should be run or what the terms of compensation should be.[237] Moving quickly to address his lack of experience running the wires, Burleson's first step was to issue Bulletin No. 1 on July 23 forming a Wire Control Board consisting of Burleson, two other members of the post office, and David J. Lewis, the once and future Congressman who supported nationalization so avidly and who was then serving as a U.S. Tariff Commissioner after a failed bid for the Senate.[238] Next, on July 29, Burleson

229. President Woodrow Wilson, A Proclamation (July 22, 1918), *in* 18 A COMPILATION OF THE MESSAGES AND PAPERS OF THE PRESIDENTS, *supra* note 77, at 8551–53.

230. *Id.* at 8552.

231. *Id.*

232. President Woodrow Wilson, A Proclamation (Nov. 2, 1918), *in* 18 A COMPILATION OF THE MESSAGES AND PAPERS OF THE PRESIDENTS, *supra* note 77, at 8630–31. Indications that the government may not have taken control of the undersea cables until after hostilities had ended provoked accusations that the government's actions were motivated not by military exigency, but rather by a desire to assert permanent control over the wireline communications system. MAY, *supra* note 17, at 38–42.

233. MARGARET MACMILLAN, PARIS 1919: SIX MONTHS THAT CHANGED THE WORLD 157–58 (2001).

234. *Id.* at 485.

235. U.S. POST OFFICE, OFFICE OF POSTMASTER GEN., RETURN OF THE WIRE SYSTEMS, ORDER NO. 3380 (July 30, 1919), *reprinted in* U.S. POST OFFICE, GOVERNMENT CONTROL AND OPERATION OF THE TELEGRAPH, TELEPHONE AND MARINE CABLE SYSTEMS, AUGUST 1, 1918, TO JULY 31, 1919, at 56, 92 (1921).

236. Bickers, *supra* note 17, at 151.

237. *See* DANIELIAN, *supra* note 17, at 250 (detailing that Postmaster General Burleson relied on Bell's executives in deciding how to operate the wires and how to compensate the companies).

238. The other members from the Post Office aside from Burleson were John Koons, First Assistant Postmaster General, and William Lamar, Solicitor for the Post Office. U.S. POST OFFICE,

summoned Vail to meet with him in Washington, D.C., amid wide speculation that the first order of business would be to fire Vail.[239] Given the breadth of discretion that had been granted to Burleson, Vail felt almost entirely at Burleson's mercy.[240] Vail pledged that all of AT&T's officers and employees would do everything in their power to support the war effort.[241] Indeed, Vail regarded it as an opportunity to see what could be accomplished when both telephone and telegraph systems were operated by the same management.[242] Vail was unconcerned with compensation, stating, "You fix it, and I'll be satisfied."[243]

Burleson told Vail's biographer that he had expected Vail to be "in a class with the average railroad president—an autocrat, interested only in the success of his road as shown by profits accruing to his stockholders, and also largely concerned as to the continuance of his salary."[244] Instead, he quickly grew to regard Vail as "a great, unselfish patriot" and "a warm and true friend" who "never made a suggestion . . . that was in the slightest degree tinged with selfishness and that was not prompted by the highest motive."[245] Thoroughly disarmed, Burleson reassured Vail that he did not plan to operate the telephone system permanently and eventually regarded him as a "confidential adviser and counselor in all matters pertaining to the telephone."[246]

Given the absence of administrative personnel within the Postal Service to run the telephone system, on August 1, Burleson issued Bulletin No. 2 ordering that "[u]ntil further notice, the telegraph and telephone companies shall continue operation in the ordinary course of business through regular channels."[247] In addition, "[a]ll officers, operators, and employees of the telegraph and telephone companies will continue in the performance of their present duties, reporting to the same officers as heretofore and on the same

---

OFFICE OF POSTMASTER GEN., BULL. NO. 1, WIRE CONTROL BOARD, ORDER NO. 1744 (July 23, 1918), *reprinted in* U.S. POST OFFICE, *supra* note 235, at 61, 61–62.

239. BROOKS, *supra* note 16, at 151–52.

240. Vail reportedly told the two Vice Presidents who were accompanying him:

> Well, I never in my life felt so helpless as I do at this moment. These people we are going up to see have got us entirely in their hands—they have taken our property and probably intend to keep it. They can do what they please with us, and we cannot help ourselves. For once in my life I am completely at sea.

ALBERT BIGELOW PAINE, IN ONE MAN'S LIFE: BEING CHAPTERS FROM THE PERSONAL & BUSINESS CAREER OF THEODORE N. VAIL 320 (1921).

241. *Id.* at 321.

242. *Id.* at 322.

243. *Id.* at 323.

244. *Id.* at 322.

245. *Id.* at 323–24.

246. *Id.* at 323.

247. U.S. POST OFFICE, OFFICE OF POSTMASTER GEN., BULL. NO. 2, ORDER ASSUMING POSSESSION AND CONTROL, ORDER NO. 1783 (Aug. 1, 1918), *reprinted in* U.S. POST OFFICE, *supra* note 235, at 62, 62.

terms of employment."[248] Thus, the takeover was more akin to a change in management, rather than a change in ownership. On December 13, Burleson gave industry executives an even larger role when he transferred operational authority from the Wire Control Board to a new operating board consisting entirely of industry executives.[249] The operating board was expanded to include additional industry executives on January 10 and March 6, 1919.[250]

Burleson also effectively ensured that AT&T would do quite well by the deal. On October 5, Burleson approved a contract that was quite generous from AT&T's perspective. The contract promised to operate the system at the same level of efficiency achieved in the past and to maintain the property in its current state of repair and gave AT&T the right to inspect the books at reasonable times.[251] The government agreed to cover all taxes, licensee fees, and charges.[252] The contract preserved the 4.5% license contract fee that the local operating companies had been paying to the Bell System's long-distance arm and included a fairly generous depreciation rate of 5.72%.[253] The government also agreed to maintain AT&T's stock dividend of eight dollars per share.[254] Finally, the government agreed to hold AT&T harmless for any injuries or expenses that were incurred.[255] In short, the government effectively guaranteed AT&T's previous rate of return while assuming all of the risks of operating the system.

In addition, the Post Office took several actions that would have a lasting impact on the telephone system. These included ordering the industry to resume merging to monopoly, quelling labor unrest, and ordering rate increases.

---

248. *Id.*

249. The Board consisted of Union N. Bethell and F.A. Stevenson of AT&T; G.M. Yorke of Western Union; and A.F. Adams to represent the independent telephone companies. U.S. POST OFFICE, OFFICE OF POSTMASTER GEN., APPOINTMENT OF OPERATING BOARD, ORDER NO. 2479 (Dec. 13, 1918), *reprinted in* U.S. POST OFFICE, *supra* note 238, at 74, 74; *see also* DANIELIAN, *supra* note 17, at 256–57 (describing each man's corporate affiliations). It took control on January 1, 1919. U.S. POST OFFICE, OFFICE OF POSTMASTER GEN., OPERATING BOARD TO ASSUME OPERATION, ORDER NO. 2534 (Dec. 23, 1918), *reprinted in* U.S. POST OFFICE, *supra* note 238, at 82, 82 n.3.

250. Specifically on January 10, the operating board appointed F.B. MacKinnon of the U.S. Independent Telephone Association to serve as liaison to the independents. On March 6, the operating board placed N.T. Guernsey (AT&T's general counsel) in charge of the board's legal department, named Bancroft Gherardi (AT&T's acting chief engineer) head of the engineering department, and designated W.S. Gifford (AT&T's Comptroller) as head of the accounting department. DANIELIAN, *supra* note 17, at 257.

251. U.S. POST OFFICE, PROPOSAL FOR COMPENSATION OF —— TELEPHONE CO. TO THE POSTMASTER GENERAL, *reprinted in* U.S. POST OFFICE, *supra* note 235, at 23, 24–25.

252. *Id.* at 26.

253. DANIELIAN, *supra* note 17, at 252.

254. *Id.* at 251.

255. U.S. POST OFFICE, *supra* note 251, at 23, 25, 28.

1. Mandating the Return of Monopoly.—Burleson moved quickly to declare the scope of his intentions and his perspective on the future of the system. In Bulletin No. 2 (issued the day after he took over the telephone system), Burleson made clear that the purpose of government control was "to coordinate and unify these services so that they may be operated as a national system."[256]

On August 7, Burleson issued Bulletin No. 3 on "Consolidation of Competing Telephone Systems," which noted, "The Governmental operation and control of the telephone systems of the country *will undoubtedly cause the coordination and consolidation of competing systems wherever possible*."[257] To encourage the unification of the service, Burleson indicated that "negotiations . . . already under way for the consolidation of a number of competing telephone systems at the time the Government assumed control" would not be disturbed.[258] Even where such negotiations were not yet underway, Burleson made clear that he had "no objection to the companies taking up such negotiations."[259]

Burleson backed up his rhetorical support for consolidation with directives to the operators. Bulletin No. 4, issued on August 15, ordered companies "[t]o proceed as expeditiously as possible with the plans heretofore instituted for consolidating and unifying the telephone plants and properties."[260] In areas where such plans were not yet underway, the Bulletin ordered that consolidation plans "should be formulated as soon as practicable" wherever consolidation "is manifestly desired by the public" and "can be effected on fair terms and in accordance with law."[261] Where two competing operators continued to operate, Burleson ordered them to "cooperate in making extensions and betterments," in order to promote "unification and the elimination of waste."[262] On the same day, Burleson issued another order creating a Committee on Solicitation of Telephone Systems consisting of AT&T Vice President Nathan C. Kingsbury and the president of one of the independents "for the purpose of making the necessary investigations, conducting negotiations, and arriving at agreements *for the unification and consolidation of the various telephone companies*

256. U.S. POST OFFICE, OFFICE OF POSTMASTER GEN., *supra* note 247, at 62, 62.

257. U.S. POST OFFICE, OFFICE OF POSTMASTER GEN., BULL. NO. 3, CONSOLIDATION OF COMPETING TELEPHONE SYSTEMS (Aug. 7, 1918), *reprinted in* U.S. POST OFFICE, *supra* note 235, at 62, 62 (emphasis added).

258. *Id.* at 63.

259. *Id.*; Office of Information, Post Office Department (Aug. 7, 1918). Papers of Albert Sydney Burleson, Manuscripts Collection, Library of Congress (ASB), Box 21.

260. U.S. POST OFFICE, OFFICE OF POSTMASTER GEN., BULL. NO. 4, EXTENSIONS AND BETTERMENTS CURTAILED, ORDER NO. 1858 (Aug. 15, 1918), *reprinted in* U.S. POST OFFICE, *supra* note 235, at 63, 63.

261. *Id.*

262. *Id.*

operating in the same communities."[263] In all, Burleson would approve thirty-four consolidations of competing telephone operations.[264] In addition, Burleson deviated from the established principle that the common carriage obligation to provide nondiscriminatory service applied only to consumers and not to business rivals[265] by issuing a Bulletin indicating that he would order long-distance companies to interconnect with any requesting local telephone companies that did not have long-distance facilities "if upon investigation it is found practicable to do so."[266]

Burleson's advocacy for integration and consolidation extended beyond just the telephones. On November 18, 1918, Burleson ordered that as of November 18, all of the telegraph systems "shall hereafter be operated as one" and as of December 1, "all telegraph offices shall accept for transmission all classes of messages now accepted by any one of them at the prescribed tariff rates."[267] The same day, Burleson issued an order taking control over the submarine cable system, using the same language contained in his order taking over the telephone and telegraph systems, indicating that his goal was "to coordinate and unify these services so that they may be operated as a national system."[268]

The next day, Burleson stated that an effective communication system required "intimate relations under which a continuous circuit can be established .... The effectiveness of the service is dependent upon the extent of the common control of circuits."[269] Burleson drew support for his conclusion from the fact that each of the telegraph systems had its own

---

263. U.S. POST OFFICE, OFFICE OF POSTMASTER GEN., COMMITTEE ON CONSOLIDATION OF TELEPHONE SYSTEM, ORDER NO. 1855 (Aug. 15, 1918), *reprinted in* U.S. POST OFFICE, *supra* note 235, at 63, 63–64.

264. U.S. POST OFFICE, OFFICE OF POSTMASTER GEN., REPORT OF THE POSTMASTER GENERAL ON THE SUPERVISION AND OPERATION OF THE TELEGRAPH, TELEPHONE, AND CABLE PROPERTIES (Oct. 31, 1919), *reprinted in* U.S. POST OFFICE, *supra* note 235, at 5, 11.

265. The seminal decision is *The Express Package Cases*, 117 U.S. 1 (1885), in which the Supreme Court held that while railroads were obligated to carry passengers, they were not obliged to carry business rivals (including express package services), reasoning that railroads were not obligated to be a "common carrier of common carriers." *Id.* at 21. For descriptions of decisions extending this principle to telephony and holding that long-distance companies need not interconnect with local telephone companies, see HUBER ET AL., *supra* note 6, § 1.3.1, at 15–16, § 5.1.1, at 407–08; MUELLER, *supra* note 10, at 48–50. For an early deviation from this principle, see MUELLER, *supra* note 10, at 116 (citing U.S. Tel. Co. v. Cent. Union Co., 171 F. 130, 143 (C.C.N.D. Ohio 1909)).

266. U.S. POST OFFICE, OFFICE OF POSTMASTER GEN., BULL. NO. 13, LONG DISTANCE CONNECTIONS FOR ALL SYSTEMS (Nov. 18, 1918), *reprinted in* U.S. POST OFFICE, *supra* note 235, at 67, 67.

267. U.S. POST OFFICE, OFFICE OF POSTMASTER GEN., BULL. NO. 16, TELEGRAPH SYSTEMS OPERATED AS ONE, ORDER NO. 2353 (Nov. 18, 1918), *reprinted in* U.S. POST OFFICE, *supra* note 235, at 70, 70.

268. U.S. POST OFFICE, OFFICE OF POSTMASTER GEN., BULL. NO. 14, ORDER ASSUMING POSSESSION AND CONTROL OF CABLES, ORDER NO. 2351 (Nov. 18, 1918), *reprinted in* U.S. POST OFFICE, *supra* note 238, at 68, 68.

269. OFFICE OF INFO., POST OFFICE DEP'T (Nov. 19, 1918), ASB, Box 22.

"independent cable systems."[270] Moreover, Burleson pointed to recent problems to make his case:

> The recent breakdown in connection with one of the cable systems has demonstrated the absolute necessity of being able to utilize at will the facilities of either cable system with all of the land line systems, in order that traffic may be adjusted in the same hands as it is on the land lines.[271]

Burleson laid out the multifaceted rationale for these moves in his letter of December 4, 1918, ordering Western Union to place its European submarine cables under the control of its chief rival, the Commercial Cable Company.[272] First, there was the notion that the war required greater unity. Burleson explained that the "present emergency" demanded "unification in operation to the fullest extent possible [of] the cable systems" and that it could only be accomplished "through the operation of the two systems under one management."[273] Second, and perhaps most importantly, Burleson had a fixed set of beliefs about the importance of consolidation. He envisioned a national economy linked by a common communications system. Burleson wrote:

> To do this efficiently and economically requires the combination of every kind of electrical transmission of intelligence into one system over which the most efficient service could be rendered through the development of new and useful services, and the wire plant and other facilities being utilized to their fullest extent.[274]

The public was demanding "one telephone system,"[275] and the only real barrier to development was the disunity in the current system. Burleson argued:

> The transmission of speech or electrical continuous signals is now practically from every commercial industrial or social community as a center, to the limits of *effective common control over a continuity of circuits*. Any limitations are wholly in the lack of continuity in the facilities—not in the "state of the art."[276]

Echoing statements made by defenders of monopoly, Burleson said that a long-distance system required "perfect co-ordination which can only come from one unified system."[277] Burleson allowed that interconnection could

270. *Id.*
271. *Id.*
272. Letter from Albert S. Burleson, U.S. Postmaster General, to Clarence H. Mackay, President, Commercial Cable Co. (Dec. 4, 1918), ASB, Box 22.
273. *Id.*
274. U.S. POST OFFICE, OFFICE OF POSTMASTER GEN., REPORT ON POST OFFICE DEPARTMENT WIRE SYSTEM (Dec. 6, 1918), ASB, Box 22.
275. *Id.*
276. *Id.*
277. *Id.*

work, but only with a reciprocity equaling "virtual subordination" during the time of control by the other system.[278]

Burleson's belief was further underscored by his anger when the Commercial Cable Company refused to unify its facilities with Western Union's. Burleson declared that "the present emergency necessitated the unification in operation to the fullest extent possible of the cable systems leading from this country to Europe."[279] Indeed, it was manifest that full utilization "could only be accomplished through the operation of the two systems under one management."[280] Burleson underscored the importance that

> the operation of the said cable systems be unified not only for improvement of service but also that important economies in operation may be effected during the period of Government control which can be accomplished only by placing such unified operation under the management of persons in complete accord with the ends desired[.][281]

When Mackay refused to interconnect his cables with Western Union's in December 1918, Burleson removed the leadership of the Commercial Cable Company from any management role in the marine cable system and transferred those responsibilities to Western Union, ordering the president of Western Union to "carry into effect directions which have been given for the unification of the operation."[282] When the Commercial Cable Company continued to resist unification, Burleson removed its officers, board of directors, and owners from any supervisory responsibility and placed operating board member A.F. Adams in charge of the company.[283]

Burleson backed up such strong measures with broad statements evincing his support for consolidation. For example, Burleson's belief in unification was trumpeted in his first report. He sought to promote "the coordination and unification of all service rendered by [the telephone and telegraph] properties" by promoting "consolidations for the purpose of getting rid of pernicious competition and wasteful operation" as well as through "a general standardization of rates and rules of operation."[284] These statements reflected Burleson's belief in the "potential economies under a

---

278. *Id.*

279. U.S. POST OFFICE, OFFICE OF POSTMASTER GEN., APPOINTMENT OF MANAGER OF CABLES, ORDER NO. 2474 (Dec. 12, 1918), *reprinted in* U.S. POST OFFICE, *supra* note 235, at 72, 73.

280. *Id.*

281. *Id.* at 74.

282. *Id.*

283. U.S. POST OFFICE, OFFICE OF POSTMASTER GEN., REMOVAL OF CERTAIN OFFICES OF POSTAL TELEGRAPH SYSTEM, ORDER NO. 2904 (Mar. 19, 1919), *reprinted in* U.S. POST OFFICE, *supra* note 235, at 84, 85.

284. U.S. POST OFFICE, OFFICE OF POSTMASTER GEN., REPORT OF THE POSTMASTER GENERAL ON THE SUPERVISION AND OPERATION OF THE TELEGRAPH, TELEPHONE, AND CABLE PROPERTIES (Oct. 31, 1919), *reprinted in* U.S. POST OFFICE, *supra* note 235, at 10.

national system of telegraphs and telephones, provided such a system were brought under an efficient and intelligent management."[285] Based in part on the "[i]nterchange in the use of wires," such a system was "further illustrated by the steps taken for the consolidation of competing properties."[286]

In this regard, Burleson's vision of the telephone network was remarkably similar to Vail's.[287] Both clearly thought that the telephone system should consist of a single system under unitary control. During the government takeover, Burleson had the opportunity to put that vision into practice, not merely through gradual consolidation of the industry, but through executive fiat.

2. *Labor Unrest.*—Despite (or perhaps due to) his best efforts, Burleson never shook the labor troubles that he inherited when the Post Office took over the wires. Even before taking over the wires, Burleson's testimony before the House Committee on Interstate and Foreign Commerce had already put him on record as stating his belief that if the government were to take over the telegraph and telephone systems, their employees "should not be affiliated with any outside organization; that their sole allegiance and loyalty should be to the Government, and that no outside organization should have a voice in shaping their action."[288] Apparently Burleson thought that "outside organizations" included unions.[289]

The mutual dislike between Burleson and the labor movement served as the backdrop for Burleson's attempts to accommodate the wage demands of the telephone and telegraph operators. In an attempt to buy some time, Burleson appointed a committee on September 14, 1918, to "investigate the working conditions of and wages paid to employees of the telegraph and telephone companies, and report as to what improvements, if any, should be made in the working conditions, the wages which should be paid the various classes of employees, and the feasibility of standardizing the same."[290] The committee consisted of Union N. Bethell of AT&T; F.B. MacKinnon of the United States Independent Telephone Association; William S. Ryan, Assistant Superintendent, Division of Post Office Service; John B. Colpoys, Special Agent of the Department of Labor, and Julia S. O'Connor, "representing the organized telephone workers of the country."[291]

---

285. *Id.* at 11.

286. *Id.*

287. *See supra* subpart I(E).

288. *1918 House Hearings*, *supra* note 197, at 41.

289. *Quotes Burleson as Opposing Union: Head of Western Union Says Government Would Not Permit Workers to Organize*, N.Y. TIMES, July 5, 1918, http://query.nytimes.com/gst/abstract.html?res=FA0611F83C5A11738DDDAC0894DF405B888DF1D3.

290. OFFICE OF INFO., POST OFFICE DEP'T, ORDER 2005 (Sept. 14, 1918), ASB, Box 21.

291. *Id.* O'Connor's title was not mentioned in the initial report announcing the committee, but she was reported to be President of the Boston Local 1A of the Telephone Operators' Division of the International Brotherhood of Electrical Workers in June, 1919. *Congress Moves to Return*

Shortly after the Committee's appointment, Burleson issued a bulletin attempting to dispel widespread rumors "that it is the desire of the Government that employees of the telegraph and telephone companies should join the Commercial Telegraphers' Union, the International Brotherhood of Electrical Workers, or other unions."[292] Burleson averred that "[t]hese representations have no foundation in fact whatever" and "the Post Office Department will not distinguish between nonunion and union employees."[293]

The conflict boiled over the following April, and a large strike swept portions of New England.[294] Not surprisingly, Burleson took a hard line against the strikers. He released a statement on April 16, 1919, saying that the government could not authorize wage increases "merely upon demand from the employees. A strike on the part of employees working for the Government is not permissible."[295] On April 19, Burleson telegraphed President Wilson to explain the situation. Seeing strikes against the government as entirely illegitimate, he wrote, "To yield means for the Government to surrender to a strike demand without an opportunity to pass on the question whether it is a just demand. . . . To do this in my opinion would be a fatal mistake and will result in multiplying and aggravating these troubles."[296]

On the actual economics involved, Burleson was convinced that settlement on the terms that the union demanded would be cost prohibitive to AT&T and the independent operators. To support his position, he telegraphed President Wilson the results of a study that he had commissioned. According to Burleson, if the union's demands in Boston were applied to all operators nationwide, "it would increase operating expenses of Bell Company alone by nearly forty million dollars and Independents by nearly twelve million dollars."[297] Government control had not, as some advocates had hoped, quieted or resolved the labor issues in the industry. The demands and frustrations of the operators and the cost issues remained the same as when the wires were under private control. The most salient difference was simply that labor now found itself in open conflict with the government, as the owner of the system and ostensibly representing the interests of all, rather than with executives, representing the interests of

---

Wires, N.Y. TIMES, June 17, 1919, at 8, http://query.nytimes.com/mem/archive-free/pdf?res=F20C1FFA3C5E157A93C5A8178DD85F4D8185F9.

292. U.S. POST OFFICE, OFFICE OF POSTMASTER GEN., BULL. NO. 9, EMPLOYEES JOINING UNIONS, ORDER NO. 2067 (Oct. 2, 1918), *reprinted in* U.S. POST OFFICE, *supra* note 235, at 66, 66.

293. *Id.*

294. *Telephone Strike Ties Up 5 States*, N.Y. TIMES, Apr. 16, 1919, at 1, http://query.nytimes.com/mem/archive-free/pdf?res=FB0615FA395C1B728DDDAF0994DC405B898DF1D3.

295. OFFICE OF INFO., POST OFFICE DEP'T (April 16, 1919), ASB, Box 23.

296. Telegram from Albert Sydney Burleson to President Woodrow Wilson 5 (Apr. 16, 1919) ASB, Box 23.

297. Telegram from Albert Sydney Burleson to Woodrow Wilson (undated, circa Apr. 19, 1919), ASB, Box 23.

shareholders. The result was one that neither labor nor the Wilson Administration had sought: continuing labor conflict in a context in which the government's interests in containing costs conflicted with the wage increases being demanded.

3. Rate Increases.—As noted earlier, the government takeover left Vail with the helpless feeling of being unable to control the financial health of his company. Indeed, Vail confided in Western Union President Newcomb Carlton that he feared that the government takeover would make it impossible for him to raise the capital needed to continue expanding.[298] Carlton soothed Vail, replying, "It's your salvation. The government will be able to raise your rates and get you new money."[299]

Carlton's words would turn out to be prophetic. On August 28, 1918, just four weeks after assuming control, Burleson issued Bulletin No. 5 authorizing telephone companies to begin charging installation fees, which had been one of AT&T's longstanding goals.[300] The size of the charge varied with the cost of service: the installation fee was $5 when the monthly rate was $2 or less, $10 when the monthly rate was between $2 and $4, and $15 when the monthly rate was more than $4.[301] Burleson said that the charge was due "to the necessity for conserving labor and material."[302] In his final report on running the wires, Burleson indicated that wartime shortages made it "essential that the telephone companies . . . curtail their normal expenditures for extensions and temporarily arrest the normal development of their business."[303] As such, these fees had "no reference to the cost of installation."[304] Instead, their "prime purpose . . . was military in character, to be justified as a war measure and not as the expression of a commercial purpose."[305] A few weeks after the fees were enacted, Burleson began to refer to them as "service connection charges" and provided that changes of name where no lapse of service occurs and relocations of equipment within the same premises would result in a charge of $3.00.[306]

On November 18, Burleson concluded that the end of "the necessity for conserving labor and material" associated with the cessation of hostilities

298. OSLIN, *supra* note 16, at 278.

299. *Id.*

300. U.S. POST OFFICE, OFFICE OF POSTMASTER GEN., BULL. NO. 5, SERVICE CONNECTION CHARGES, ORDER NO. 1931 (Aug. 28, 1918), *reprinted in* U.S. POST OFFICE, *supra* note 235, at 64, 64; *see also* AT&T CO., 1911 ANNUAL REPORT 10–11 (1912) (explaining that installations paid for by the company represent a "large expenditure" that is "a burden not only on the capital but on the net revenue of the telephone, from which other service companies are free").

301. *Id.*

302. *Id.*

303. U.S. POST OFFICE, OFFICE OF POSTMASTER GEN., *supra* note 264, at 5, 12.

304. *Id.*

305. *Id.*

306. U.S. POST OFFICE, OFFICE OF POSTMASTER GEN., BULL. NO. 8, SERVICE CONNECTION CHARGES (Sept. 14, 1918), *reprinted in* U.S. POST OFFICE, *supra* note 235, at 65, 65–66.

justified revising the service connection charge in line with "the average cost of the initial expense of establishing service for new subscribers."[307] Consequently, he reduced the charge for new installations to $3.50.[308] As Burleson indicated in his final report, while the initial schedule of charges "must be regarded as a war measure," the revised schedule "rests on commercial considerations" and "a sound commercial principle" and now "must be regarded as one item in the unified and standardized system of telephone charges."[309]

On December 13, 1918, after persistent lobbying by Vail, Burleson also authorized a 20% increase in long-distance rates effective January 21, 1919, which augmented AT&T's revenues by roughly $10 million.[310] He justified the rate increase in part by the abnormal economic conditions brought about by the war.[311] In addition, "[t]he purpose of the new schedule of telephone toll rates is to standardize the long distance service throughout the country and to establish uniform charges."[312] Burleson lamented the interconnectivity problems that variations in rates had caused and cast standardization as a program to include all Americans in the telecommunications system: "Under the toll rates now established, the toll service is an intercommunity, interstate, and interregional service, available to all Americans, at all times, everywhere."[313]

The order raising long-distance rates also called for a fifty percent reduction between the hours of 8:30 p.m. and midnight as well as a seventy-five percent reduction in rates between midnight and 4:30 a.m.[314] The final report indicated that night rates were the application of a business principle

---

307. U.S. POST OFFICE, OFFICE OF POSTMASTER GEN., BULL. NO. 15, MODIFIED SERVICE CONNECTION CHARGES, ORDER NO. 2352 (Nov. 18, 1918), *reprinted in* U.S. POST OFFICE, *supra* note 235, at 68, 68.

308. *Id.* at 69.

309. U.S. POST OFFICE, OFFICE OF POSTMASTER GEN., *supra* note 264, at 5, 13.

310. U.S. POST OFFICE, OFFICE OF POSTMASTER GEN., BULL. NO. 22, TOLL RATE SCHEDULE, ORDER NO. 2495 (Dec. 13, 1918), *reprinted in* U.S. POST OFFICE, *supra* note 235, at 75, 77; DANIELIAN, *supra* note 17, at 256.

311. U.S. POST OFFICE, OFFICE OF POSTMASTER GEN., *supra* note 264, at 5, 14.

312. OFFICE OF INFO., POST OFFICE DEP'T (Feb. 6, 1919), ASB, Box 22.

313. *Id.* One of the other service changes was the addition of person-to-person calling that precluded charges when the particular person being called could not be located. *Id.*

314. U.S. POST OFFICE, OFFICE OF POSTMASTER GEN., *supra* note 264, at 78. This followed a similar order imposing reduced telegraphy rates for "night messages." U.S. POST OFFICE, OFFICE OF POSTMASTER GEN., BULL. NO. 17, TELEGRAPH RATES ON "NIGHT MESSAGES", ORDER NO. 2354 (Nov. 18, 1918), *reprinted in* U.S. POST OFFICE, *supra* note 235, at 70, 70. The night message would be transmitted to a receiving station which would transcribe the message and then place the message in the regular mail. Burleson explained the importance of this service: "A very great increase of traffic between distant points is expected to result from this low rate. A letter may take four or five days with no alternative but the payment of one dollar. This gives the alternative of one-half dollar service, and brings together the distant parts of the country about three days closer together." OFFICE OF INFO., POST OFFICE DEP'T (Nov. 19, 1918), ASB, Box 22.

developed in the context of the telegraph that was both "simple and of universal recognition"[315]:

> It recognizes the fact that the facilities of any industry must provide for carrying the maximum business load, and that if this load can be distributed over the 24 hours of each day, a larger amount of business can be carried on with relatively less investment than if the business done is crowded into relatively few hours of each day.[316]

In short, Burleson recognized that telecommunications networks must be sized according to the peaks in traffic and engaged in an early example of traffic shaping through peak-load pricing in an attempt to increase the efficiency of the network and to reduce the cost of service.

On March 19, 1919, Burleson also approved an increase in local rates recommended by the operating board.[317] Together these rate increases totaled roughly $50 million.[318] Unsurprisingly, these rate increases were quite unpopular with consumers. In April, Burleson sought to address criticisms by releasing data showing that the proposed rate increase was more modest than the overall rate of wartime inflation:

> The increase of 20% in telegraph rates should be considered in comparison with the 100% increase in other prices, and it is less than that found necessary to add to the railroad freight rates and is no greater than has been made generally in other public utility rates, in order to obviate financial collapse.[319]

Burleson further explained that despite being required to carry unprofitable business, e.g., government communications, the Post Office had not been able to reduce the price paid for the materials involved.[320] The problem also included the expansion of the system into rural underserved areas. As Burleson explained:

> The extension of the telegraph service into fields that are less profitable than are the great business centers, and the handling of Government business claimed by both companies to be at a loss of 50% of the operating cost, are a charge upon the gross revenues which is escaped by a company which avoids the rendering of this necessary public service.[321]

These rate increases engendered public anger, as one of the primary rationales for government ownership had been that the absence of desire for

---

315. U.S. POST OFFICE, OFFICE OF POSTMASTER GEN., *supra* note 264, at 5, 14.

316. *Id.* at 14.

317. Letter from Albert Burleson, Postmaster Gen., to U. N. Bethell, Chairman, U.S. Telegraph and Telephone Administration's Operating Board (Mar. 19, 1919), *quoted in* DANIELIAN, *supra* note 17, at 258.

318. DANIELIAN, *supra* note 17, at 260.

319. OFFICE OF INFO., POST OFFICE DEP'T 1 (Apr. 12, 1919), ASB, Box 23.

320. *Id.* at 2.

321. *Id.* at 1.

profit would lead to lower rates.[322] Burleson's decision to increase local rates also antagonized state regulatory agencies, which successfully obtained injunctions against $16 million of the rate increases in ten states across the country.[323] The Supreme Court overturned these injunctions in *Dakota Central Telephone Co. v. South Dakota*,[324] which was argued on May 5–6, 1919, and decided on June 2, 1919.[325] Writing for an 8–1 majority, Justice White held that the war power was complete and sufficient to uphold Congress's decision to take over a public utility.[326] Moreover, the state police power did not require that the judiciary carve out realms of state prerogative.[327] In other words, Congress's authority over the telecommunications system under the War Power was complete and included the ability to set rates for intrastate services.[328]

Despite the rate increases, the telephone system still operated at a substantial loss, which under the terms of the agreement the government had to make good.[329] As such, the government owed AT&T a deficiency payment of $13 million, although AT&T forgave $4 million of it "to facilitate prompt and economical settlement."[330] The Treasury allocated an additional $4 million to compensate the independents.[331]

C. *Returning the Wires*

On December 13, 1918, Chairman John Moon (D-Tenn.) of the House Committee on the Post Office and Post Roads introduced a proposal directing the Postmaster General "to negotiate contracts for the purchase of any or all telephone lines . . . subject to the approval of Congress."[332] The House Committee on the Post Office and Post Roads reviewed the bill in January 1919 in a series of hearings.[333] On January 29, the Committee issued a report entitled "Extension of Government Control of Telegraph and Telephones."[334] While the report indicated that "many of the committee desired a longer time

---

322. COHEN, *supra* note 15, at 38.

323. DANIELIAN, *supra* note 17, at 260–62.

324. 250 U.S. 163 (1919).

325. *Id.* at 163.

326. *Id.* at 183. Justice Brandeis dissented without opinion. *Id.* at 188.

327. *Id.* at 185–87.

328. *Id.* at 187.

329. DANIELIAN, *supra* note 17, at 268.

330. *Id.* (citation omitted).

331. *See id.* at 268–69 (noting that the bill to the Treasury totaled just over $13 million after totaling the $9 million paid to AT&T and the payments due to telegraph and independent telephone companies).

332. H.R.J. Res. 368, 65th Cong. (1918), *reprinted in Government Control of the Telegraph and Telephone Systems: Hearings on H.R.J. Res. 368 Before the H. Comm. on the Post Office and Post Roads*, 65th Cong., pt. 1, at 3 (1919).

333. *Government Control of the Telegraph and Telephone Systems: Hearings on H.R.J. Res. 368 Before the H. Comm. on the Post Office and Post Roads*, 65th Cong., pts. I–III (1919).

334. H.R. REP. NO. 65-1012 (3d Sess. 1919).

for extension of Government control of telegraph and telephone lines, the majority are of the opinion that the lines should be returned to the owners on December 31, 1919."[335]

On May 19, Vail and the President of the United States Independent Telephone Association sent a letter to Congress requesting the return of their telephone properties.[336] The Senate Committee on Interstate Commerce conducted hearings on May 29,[337] while the House Committee on Interstate and Foreign Commerce conducted hearings on May 30–31 and June 4–5.[338]

By this time, government ownership had lost many of its key constituencies. The Supreme Court had just handed down its decision upholding the local rate increase on June 2, much to the dismay of the state regulatory commissions and consumers.[339] Labor tensions were brewing that would culminate in the second telegraph operators' strike on June 11, and the unions were frustrated by the fact that government ownership appeared to hurt their bargaining position.[340] The armistice had been in place for over half a year, and the negotiations that would culminate in the June 28 signing of the Treaty of Versailles were approaching their conclusion.[341] The change in the political winds is well illustrated by the titles of the hearings. Unlike the January hearings, which referred to "Extension of Government Control," the hearings of late May and early June spoke of the "Relinquishment" and the "Return" of the telephone system.[342]

After the war, the public emergency rationale for government ownership of the wires ceased. Public hostility to the rate increases and labor strife had dogged the period of government control.[343] Government operation continued to be dogged by accusations of widespread censorship.[344] In this context, Congress held hearings before the House Committee on Interstate and Foreign Commerce on whether to return the wires in June 1919.[345] The

---

335. *Id.* at 5.

336. Letter from T. N. Vail, President, Am. Tel. & Tel. Co., & G. W. Robinson, President, U.S. Indep. Tel. Ass'n, to Congress (May 19, 1919), *reprinted in* DANIELIAN, *supra* note 17, at 266.

337. *Relinquishment of Government Control of Telephone and Telegraph Lines: Hearings Before the S. Comm. on Interstate Commerce*, 66th Cong. (1919).

338. *Return of the Wire Systems Hearings*, *supra* note 202.

339. N. Pac. Ry. Co. v. North Dakota, 250 U.S. 135, 151–52 (1919).

340. *See* STEPHEN H. NORWOOD, LABOR'S FLAMING YOUTH: TELEPHONE OPERATORS AND WORKER MILITANCY, 1827–1923, at 200–01, 207–08 (1990) (describing a strike called by the Commercial Telegraphers' Union of America during a large-scale strike of telephone operators and noting that government ownership benefited management).

341. *See* RODNEY P. CARLISLE, WORLD WAR I 289–90 (2007) (chronicling the negotiation of the Treaty of Versailles from the armistice to the signing of the treaty).

342. *Relinquishment of Government Control of Telephone and Telegraph Lines: Hearings Before the S. Comm. on Interstate Commerce*, 66th Cong. (1919); *Return of the Wire Systems: Hearings Before the H. Comm. on Interstate and Foreign Commerce on H.R. 421*, 66th Cong. (1919).

343. DANIELIAN, *supra* note 17, at 267; NORWOOD, *supra* note 340, at 157–58.

344. Douglas, *supra* note 94, at 283.

345. *Return of the Wire Systems Hearings*, *supra* note 202; DANIELIAN, *supra* note 10, at 267.

companies launched a full-fledged lobbying effort to convince Congress to relinquish control.[346] Burleson fought in vain to hold back the tide. He wanted the wires to stay under government control, although "at lower rates."[347]

Representative John J. Esch (R-Wis.) presided over the hearings, and they began with consideration of a written statement by Burleson.[348] Burleson argued that private control would not solve the basic problem that the government had faced during the war: increased costs of materials and labor. He wrote, "The extraordinary increased cost of operation and maintenance which has been fastened on [the telephone and telegraph operators] as a result of the war will continue for some time after control passes from the Government."[349] Burleson remained adamant that "the various systems should be coordinated as to operation."[350] Monopoly was not necessary to maximize efficiency, but consolidation was necessary to avoid "wasteful competition and the economic loss occasioned by duplication of plant and force."[351] In order to facilitate this consolidation, Burleson recommended that Congress enact a law allowing any telecommunications company to "purchase the property of any telegraph or telephone company, or any part thereof, or consolidate with any other telegraph or telephone company, or pool its traffic and facilities with any other telegraph or telephone company," subject to the approval of the ICC.[352] Burleson recognized that if government ownership was going to end, then the best that he could hope for was that Congress would encourage consolidation of the industry via relaxation of the antitrust regulatory scheme.

State regulators did not necessarily share Burleson's belief in consolidation. For instance, Carl D. Jackson, chairman of the Railroad Commission of Wisconsin, which oversaw telecommunications in the state, testified that Congress simply needed to return the wires to private control, rather than encouraging consolidation via additional legislation.[353] Jackson explained that the problem of duplication, and attendant waste, was nonexistent in his state because Wisconsin prohibited duplication and required interconnection.[354] Although Wisconsin had approximately one thousand independent operators and diverse rates, costs were low and the

346. DANIELIAN, *supra* note 17, at 266–68.

347. *Id.* at 269–70.

348. *Return of the Wire Systems Hearings*, *supra* note 210, at 5–9 (statement of A.S. Burleson, Postmaster General).

349. *Id.* at 6.

350. *Id.*

351. *Id.*

352. *Id.*

353. *Id.* at 10, 12 (statement of Carl D. Jackson, Chairman, Railroad Commission of Wisconsin).

354. *Id.* at 10.

technology in use was adequate to cover present needs.[355] He disagreed with Burleson's belief in uniform rates, stating that urban service cost more than rural service due to the increased need for central stations in dense areas.[356] Jackson explained that in a city like Milwaukee, density actually drove up costs because of the expenses related to operating central stations.[357] If Wisconsin moved to a uniform rate, then it could unfairly hurt rural customers.[358] Jackson declared, "Such a thing as a uniform postage-stamp rate for telephone service throughout the United States is unthinkable."[359] And, on the general question of return of the wires, Jackson was clearly in favor of the resumption of private control.[360]

Also at issue in the hearings was why the rates had gone up during the war. There were lingering suspicions that nationalization, in and of itself, had driven up costs.[361] Nathan Kingsbury, Vice President of AT&T, took a different approach and testified that costs had skyrocketed during the war years for three reasons.[362] First, wages had increased approximately 50% during the war years, AT&T's payroll rising from $110 million in 1915 to $175 million in 1919.[363] Second, the price of copper had increased by 100% in 1917 over the cost in 1914, although it had dropped some in the most recent years.[364] The cost of lead-covered cable increased 45%, the cost of in-house manufactured goods increased 25%, and the cost of purchased manufactured goods increased 75%.[365] Third, the cost of capital had also increased. Prewar, capital was relatively cheap and AT&T's credit was good; during the war, interest rates for AT&T had risen 2%, with capital only being available at close to 7%.[366] Kingsbury explained that this was a particular hardship on telephone companies because they had constant need for new capital. He stated, "You have got to keep on building all the time every day."[367] Every new phone required, on average, a $150 capital investment, and AT&T installed 168,000 new phones in the first quarter of

---

355. *Id.* at 10–11.

356. *Id.* at 12.

357. *Id.*

358. *Id.*

359. *Id.*

360. *Id.* at 11 ("We certainly are against any further control of the companies by the Federal Government. We see no benefit to be derived from it, and we ask that they be returned to private ownership . . . .").

361. *See id.* at 45 (question from Rep. Alben Barkley to Nathan Kingsbury, Vice President, AT&T) (asking whether increases in costs were larger under government control than they would have been under private control).

362. *Id.* at 15.

363. *Id.* at 16.

364. *Id.*

365. *Id.*

366. *Id.* at 16–17.

367. *Id.* at 16.

1919.[368] Considering the entire situation, Kingsbury explained, "Up to the time we went to war in 1917 it had been possible to get along without large increases in revenue" via increases in efficiency.[369] Now, AT&T needed rate increases to keep up with rising costs.[370]

Considering the effect of government control during the war, Kingsbury stated that Burleson had not interfered in the operations of AT&T.[371] Nationalization caused a small loss of morale and increased pressure on wages.[372] Recasting the rationale for Burleson's connection charge, Kingsbury stated that rather than it being a subsidy for rural development, "the principal purpose" of the charge was to depress demand during the war years given the material shortages.[373] Kingsbury's explanation of why the government lost money during the takeover was simply "[b]ecause it could not get the rates up as soon as the expenses went up."[374] Pressed on this point by Congressman Edward L. Hamilton (R-Mich.), Kingsbury reiterated that the failure of the rates to adjust quickly to increased costs during the war was "the whole story" of why the government lost money.[375]

F.B. MacKinnon, vice president of the U.S. Independent Telephone Association, also believed that the wires should be returned, and he agreed with Kingsbury on the need for remedial legislation.[376] MacKinnon, like Kingsbury, blamed any deterioration of service on war shortages.[377] Specifically, he blamed the "general labor conditions, scarcity of labor, and inability to hold operators."[378]

Joseph P. Hayes, National President of the Association of Western Union Employees, also encouraged Congress to return the wires.[379] The unions resented that they had lost ground on wages during Burleson's tenure, and they sought to return the wires to private control so that they could better

---

368. *Id.*

369. *Id.* at 17.

370. *Id.* at 17–18.

371. *Id.* at 21.

372. *Id.* at 23, 36.

373. *Id.* at 19.

374. *Id.* at 46.

375. *Id.* The entire dialogue between Kingsbury and Hamilton is worth reproducing:

> Mr. HAMILTON. In brief, what are the prime reasons why the Government has lost money on the operation of the lines?
>
> Mr. KINGSBURY. Because it could not get the rates up as soon as the expenses went up.
>
> Mr. HAMILTON. That is the whole story?
>
> Mr. KINGSBURY. That is the whole story practically, and that is going to be the reason. If these properties are turned back to us without some legislation, we will not be able to get those rates up in time to prevent very serious loss and financial embarrassment on the part of these companies, which is going right down to our stockholders.

Id.

376. *Id.* at 112.

377. *Id.* at 120.

378. *Id.*

379. *Id.* at 73–74.

pressure the companies for wage increases.[380] Hayes reported that his members were eager to end government control, stating, "I find that our people, the workers, as a whole throughout the country, unlike the railroad workers, are very much committed to private control."[381]

John C. Koons, First Assistant Postmaster General, represented the Administration at the end of the hearings.[382] Koons stated that he was not opposed to the return of the wires, and he recognized the need for some remedial legislation.[383] Koons sought to deflect blame for the perceived deterioration in service during the takeover. He stated that service had suffered because the telephone companies "had released thousands of people for service in the Army. Hundreds of their very best men were in the Army."[384] Seeking to counter Kingsbury's explanation of the connection charge, Koons said that the primary rationale was to take the burden of installations off of those that did not change lines.[385] Koons said that some businesses moved frequently, and thus steady subscribers subsidized those moves.[386] He explained, "So we fixed an installation charge in order to take the burden off of the subscribers who did not make changes and ought not to be made to bear the cost, and also in order to reduce the demand for extensions and installations. We had to do it."[387]

Rounding out the hearing was the statement of J.A. Pratt, representing the United Telephone Company of the State of Wisconsin and the Wisconsin State Telephone Association.[388] Pratt decried the takeover, and was blunt about its effect, "I do not believe anything in the last 12 years of the history of regulation in the State of Wisconsin has done more to weaken in the minds of the people of that State the theory of regulation than the acts of the Post Office Department."[389] For Pratt, the takeover had eroded citizen confidence not just in publicly owned telecommunications, but also in regulation generally.

---

380. *See id.* at 74–75 (stating that Sunday pay had been reduced from time-and-a-half basis to straight-time basis, that maximum pay increases had been reduced from 15% to 10%, and that the union "[felt] that the company ha[d] shown an inclination to deal with us . . . squarely and fairly").

381. *Id.* at 74.

382. *Id.* at 183.

383. *Id.* at 183–84.

384. *Id.* at 199.

385. *Id.* at 209.

386. *Id.*

387. *Id.* Koons's explanation is not mutually exclusive with Kingsbury's. Both could be correct: the connection charge could have been both a deterrent to new service requests and a reduction of the subsidy for businesses that frequently changed lines. In the hearings, however, the political question was what would be the public rationale for the charge. Kingsbury was willing to bluntly state that it was to deter new requests, while Koons wanted to portray it as a fair allocation of costs. When Kingsbury was pressed on why, if material were so short, AT&T did not just cease installations, he responded, "Telephone companies can not go out of business," meaning that they could not afford to go out of the business of new installations. *Id.* at 20.

388. *Id.* at 219.

389. *Id.*

Considering the hearings as a whole, the voices for retaining government control were few.[390] AT&T, the independent operators, the state regulators, and the operators' unions all wanted the wires returned to private control. AT&T wanted to return to its business plan, which included continued expansion and gradual rate increases approved by state commissions.[391] Federal control chafed the independent operators, and they too, like AT&T, wanted to get back to their prewar business plans.[392] The state regulators had been stripped of their power to set rates while the wires were under government control by the just announced *Dakota Central* decision on June 2, 1919.[393] Not surprisingly, they yearned for the prewar system in which they exercised influence by setting state-by-state rates.[394] The unions, having been stymied by Burleson and the issues surrounding striking against the government during the war, also sought return to private control.[395] All that stood on the other side was a diffuse set of ideas about the inefficiencies of a competitive and private decentralized system. Burleson, overwhelmed by the negative impression of his tenure, could not hold onto the wires.

On June 4, the Senate Committee issued a report entitled "Return of Telephone, Telegraph, and Cable Lines" that proposed setting the outer limit of the return of the wires at sixty days after the bill's enactment.[396] On June 16, the House Committee issued a report entitled "To Repeal the Telephone and Telegraph Act" that would require the restoration of the wires at midnight on the last day of the calendar month the bill was signed into law.[397] Both chambers adopted the House's language on June 27,[398] and the President signed the legislation into law on July 11.[399] Per the terms of the statute, the telephone system left government control at midnight on

390. Considering what would have happened to AT&T had the government not nationalized the wire, Kingsbury stated, "If the companies had continued to be managed by their owners, we would have been diligently at work during all that time on these rate matters." *Id.* at 152.

391. *See id.* at 18 (statement of Nathan Kingsbury, Vice President, AT&T) (discussing AT&T's plans for expansion and rate increases prior to government take over).

392. *See, e.g.*, *id.* at 10 (stating that independent operators in Wisconsin see "nothing relating to Government control since the Government has taken possession of the telephone companies which has appealed to the patrons of the telephone companies").

393. Dakota Cent. Tel. Co. v. South Dakota, 250 U.S. 163 (1919).

394. Charles M. Elmquist, President and General Solicitor of the National Association of Railway and Utility Commissioners, testified during the congressional hearing on returning the wires that state commissions set rates in forty-five states at that time. *Return of the Wire Sytems Hearings*, *supra* note 202, at 163.

395. *See id.* at 74 (statement of Joseph P. Hayes, National President, Association of Western Union Employees) ("I find that our people, the workers, as a whole throughout the country, unlike the railroad workers, are very much committed to private control.").

396. S. REP. NO. 66-4, at 1 (1919).

397. H.R. REP. NO. 66-45, at 2 (1919).

398. 58 CONG. REC. 1906–07, 1924–25 (1919).

399. Act of July 11, 1919, ch. 10, 41 Stat. 157.

July 31.[400] The act provided that the rate increases that Burleson had approved would continue in force for up to four months.[401]

In the end, a number of factors brought the nation's experiment with a publicly owned telephone system to an end. As an initial matter, assurances that the initial proposal was only temporary placed a natural limit on the prospects for extending the period of government ownership. Indeed, Burleson continued to be dogged by accusations of trying to make the arrangement permanent.[402] The rate hikes that Burleson had authorized were intensely unpopular. Representative Aswell, the sponsor of the original legislation, said, "I owe it to my people and to Congress to apologize for my resolution if government control means increase in rates."[403] Burleson's decisions had alienated key constituencies, such as the labor unions and the state regulatory authorities, not to mention the consumers that had hoped for less expensive service.[404] And both legislators and the public had the strong sense that the network had been poorly run by the government.[405]

After signing the order returning the wires to private control,[406] Burleson wrote a personal letter of thanks to Theodore Vail, then Chairman of the Board of Directors of AT&T.[407] Reflecting on this period, Burleson wrote "to express [his] heartfelt appreciation" for Vail's assistance.[408] Burleson praised Vail's unselfishness, and hoped that the future of the wire service would involve "the same successful control and direction" which it had received under Vail's administration.[409] Wishing Vail "many years of health and happiness," Burleson signed off, "[y]our sincere friend," and an exceptional experiment in American telecommunications came to an end.[410]

AT&T emerged from this period in decent shape. Rates were raised and standardized.[411] The state regulatory commissions had been prevented from

---

400. *Id.*

401. *Id.*

402. *See* 58 CONG. REC. 1347 (1919) (reporting debate over the government's authority to control the wire systems and change rates).

403. *The First Step*, TELEPHONY, Apr. 19, 1919, at 11.

404. *E.g.*, *Burleson Rapped on All Sides*, FOURTH EST., May 10, 1919, at 15, 15–16; Kenneth N. Bickers, *Transformations in the Governance of the American Telecommunications Industry*, *in* GOVERNANCE OF THE AMERICAN ECONOMY 77, 90 (John L. Campbell et al. eds., 1991); JOHN, *supra* note 17, at 403; MAY, *supra* note 17, at 50.

405. *See, e.g.*, 58 CONG. REC. 1347 (1919) (decrying the inefficiency of the wire systems under government control); *Representative Aswell Apologizes*, 43 AM. ECONOMIST 235, 235 (1919) (attacking the government's "experiment[ing] at the expense of the public" by seizing the wires and reporting widespread opposition to the same).

406. OFFICE OF INFO., POST OFFICE DEP'T, ORDER 3380 (July 30, 1919), ASB, Box 24.

407. Letter from Albert Sydney Burleson to Theodore N. Vail, AT&T (July 30, 1919), ASB, Box 24.

408. *Id.*

409. *Id.*

410. *Id.*

411. BROOKS, *supra* note 16, at 158–59.

blocking national rate increases.[412] New universal service charges were added that subsidized development, and AT&T was permitted to retain its long-desired service connection charge despite the fact that it was justified as a war measure.[413] The unpopularity of the rate hikes and the labor troubles blunted calls for nationalization.[414] The Republican Party platform of 1920 was able to crow that we "took from the incompetent Democratic administration the administration of the telegraph and telephone lines of the country and returned them to private ownership."[415] Proposals to revive government ownership of the telephone system would continue to appear throughout the 1920s, but none were able to garner any substantial support.[416] Indeed, the nation's unhappy experience with government control of the telephone system is widely regarded as the death knell for calls for government ownership of telecommunications.[417]

III. Implications of the Government Takeover

The government takeover yields new insight into several key questions of telecommunications policy. First, it provides a new view of the reasons that the telephone network collapsed into a monopoly. Second, it provides a new perspective on the origins of universal service. Third, it adds a new twist to the development of state–federal relations. Fourth, it sheds new light on the proper scope of government intervention by identifying characteristics that are well and poorly suited to governmental control.

A. The Reemergence of Monopoly

One of the historical puzzles concerning the early telephone industry is how AT&T was able to reestablish its monopoly. By 1907, AT&T's market share had dropped below 50%.[418] And yet, the Bell System's market share had reached 80% by 1934.[419] Commentators typically attribute the

---

412. Bickers, *supra* note 404, at 90–91.

413. BROOKS, *supra* note 16, at 157–58.

414. *See id.* at 158 (discussing opposition to nationalization); COHEN, *supra* note 15, at 38 (noting opposition to rate increases brought about by nationalization); Bickers, *supra* note 17, at 152–54 (explaining why labor troubles in part made lowering rates infeasible and detailing objections to these rate increases).

415. Republican Platform of 1920, *in* NATIONAL PARTY PLATFORMS 1840–1972, *supra* note 41, at 232.

416. Cox & Byrnes, *supra* note 15, at 30 n.39.

417. JOHN, *supra* note 17, at 405 ("The failure of 'postalization' legitimated the ownership and operation of the telephone, telegraph, and cable by private corporations that would become a hallmark of managerial capitalism and a defining feature of the twentieth-century American political economy. Never again in the twentieth century would government ownership of the telephone and the telegraph occupy so prominent a place on the national political agenda."); MAY, *supra* note 17, at 54 (calling "the most lasting effect" of the takeover "was to discredit the principle of state socialism").

418. COHEN, *supra* note 15, at 27.

419. BROCK, *supra* note 15, at 177.

reestablishment of monopoly to one of two causes. First, they suggest that the presence of scale economies or network economic effects rendered telephony a natural monopoly.[420] Second, they argue that the reemergence of monopoly is the result of the unwillingness of antitrust authorities to curb Theodore Vail's ambitions.[421] A close look at the history of the government takeover and the data that it generated reveals that neither factor was decisive. Instead, industry consolidation was directly influenced by deliberate government policy.

1. Natural Monopoly.—One possible explanation for the reemergence of Bell dominance is that the telephone network is a natural monopoly. Indeed, many distinguished observers regard this as uncontroversial.[422] Natural monopoly is believed to be the result of the supply-side scale economies associated with constantly declining costs or the demand-side scale economies associated with network economic effects. The circumstances surrounding the takeover make clear that neither provides a convincing explanation.

a. Scale Economies.—The most frequently cited explanation for the reemergence of monopoly is the economies of scale associated with high fixed costs.[423] The presence of unexhausted economies of scale causes unit costs to decline as volume increases.[424] When average costs decline, the firm with the largest volume can underprice its rivals, which causes it to take even more share of the market. If the economies of scale remain unexhausted, markets that begin as competitive will collapse into natural monopolies.

A revisionist history has emerged pointing out that telephone service was not a declining cost industry. In particular, when switching was performed manually by operators sitting at a switchboard, it did not scale.[425] The deployment of mechanical switches would eventually change this limitation, but AT&T did not begin deploying mechanical switches until the

---

420. Weiman & Levin, *supra* note 142, at 104.

421. *See* BROOKS, *supra* note 16, at 143–45 (discussing Vail's view of regulation and the government's initial nonregulation of telephone rates).

422. *See, e.g.*, FAULHABER, *supra* note 6, at 107 ("Indeed, until the late 1960s few questioned that the telephone industry was a natural monopoly."); HUBER ET AL., *supra* note 6, at 86 ("Is the telephone industry (or any part of it) a natural monopoly? Until the 1960s, the answer was generally presumed to be yes, from end to end."); 3B PHILLIP E. AREEDA & HERBERT HOVENKAMP, ANTITRUST LAW ¶ 787c, at 366 (3d ed. 2008) ("Until the 1960s or 1970s long distance telephone connections between local exchanges in the United States were considered as much a natural monopoly as the local exchanges themselves.").

423. *Cf.* BROOKS, *supra* note 16, at 133–34 (outlining how AT&T's economies of scale advantage allowed it to buy up competitors and work towards a monopoly); JEFFREY CHURCH & ROGER WARE, INDUSTRIAL ORGANIZATION 55 (2000) (identifying long-run fixed costs as a source of economies of scale).

424. CHURCH & WARE, *supra* note 423, at 54.

425. Milton Mueller, *The Switchboard Problem: Scale, Signaling, and Organization in Manual Telephone Switching, 1877–1897*, 30 TECH. & CULTURE 534, 559 (1989).

1920s.[426] Both AT&T and its competitors found a need to request rate increases as their operations grew.[427]

Interestingly, contemporary observers recognized the absence of scale economies in the telephone industry.[428] Indeed, Nathan Kingsbury made precisely this point in his testimony regarding the return of the wire systems. Kingsbury stated that installation cost more in the cities than in the rural areas because of the need for additional central switching stations.[429] Each central station could only take about 10,000 lines, and as a city grew, AT&T was forced to install new central stations, with new trunk lines between those stations.[430] Trunk lines were exceedingly expensive.[431] Moreover, as the central stations grew in a city, the company was required to employ additional inter-operator connectors.[432] While automatic switches could overcome some of these issues, the costs of installing new trunk lines between new central stations could not be ameliorated.[433] Summing up this situation, Kingsbury stated, "The profit per unit decreases as the number of units increases."[434]

Interestingly, the takeover provided data in which the diseconomies of scale were apparent. In 1916, the government collected data on the operations of telephone systems,[435] apparently to assist it with the management of these companies.[436] Plotting cost per telephone against the

426. *E.g.*, H.R. DOC. NO. 76-340, at 261 (1939); A HISTORY OF ENGINEERING AND SCIENCE IN THE BELL SYSTEM: THE EARLY YEARS (1875–1925), at 552–53, 611–12 (M.D. Fagen ed., 1975); ROBERT J. CHAPUIS, 100 YEARS OF TELEPHONE SWITCHING (1878–1978), at 249 (1982); Joan Nix & David Gabel, *The Introduction of Automatic Switching into the Bell System: Market Versus Institutional Influences*, 30 J. ECON. ISSUES 737, 738 (1996).

427. MUELLER, *supra* note 10, at 36–37; *see also* Weiman & Levin, *supra* note 142, at 104 (discussing AT&T's growth and the increasing costs that ensued).

428. *E.g.*, J. MAURICE CLARK, STUDIES IN THE ECONOMICS OF OVERHEAD COSTS 321 (1923) ("Telephone companies . . . show no signs of economy with increased size, but rather the opposite.").

429. MUELLER, *supra* note 10, at 41.

430. Kingsbury said, "[A]s the city grows it is necessary to install a larger and larger number of central offices because a girl's arm is just so long, and as the manual switchboards are constructed she can only reach about 10,000 stations with her arm." *1919 House Hearings*, *supra* note 338, at 31 (statement of Nathan Kingsbury, Vice President, AT&T).

431. *Id.*

432. *Id.* at 32.

433. *See Return of the Wire Systems Hearings*, *supra* note 202, at 35 (describing technology that eliminated the 10,000-line limit but still required trunk lines).

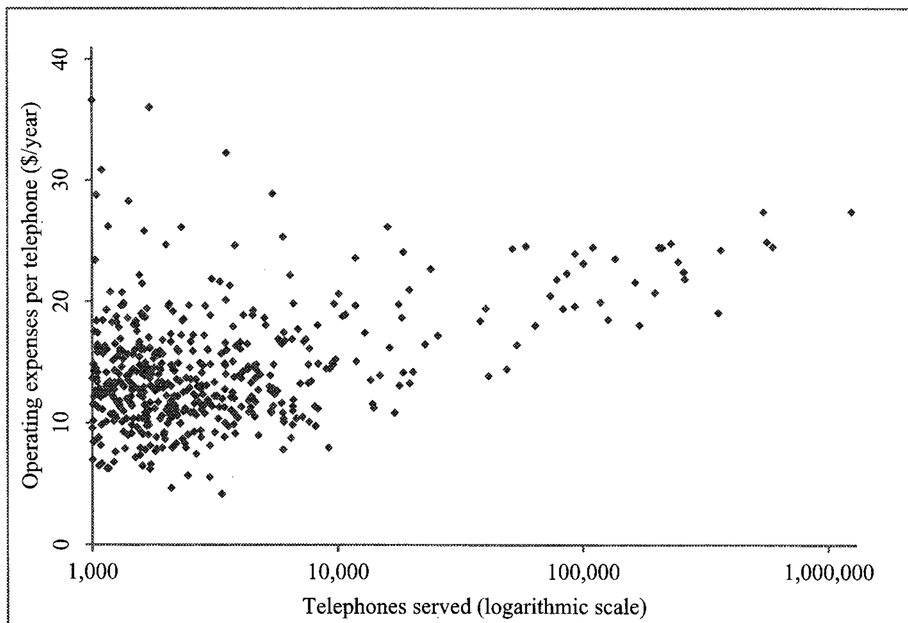
434. *Id.* at 30. "The difficulty is that the larger the number of units you serve in the telephone business, under conditions that exist requiring a larger and larger investment per unit, a larger and larger operating cost per unit, that larger investment and larger cost goes up so fast that the larger the number of units you serve, the more it costs per unit to serve them." *Id.*

435. Memorandum from Bureau of Statistics, Interstate Commerce Comm'n on Tel. Cos. & Tel. Cos. Reporting to the Interstate Commerce Comm'n for the Calendar Year 1916 (Aug. 20, 1918) *available at* http://archive.org/details/TelephoneAndTelegraphCompaniesReportingToTheIccFor1916.

436. Douglas Galbi, *Early U.S. Telephone Industry Data*, GALBITHINK.ORG, http://www.galbithink.org/telcos/early-telephone-data.htm.

number of subscribers reveals a clear upward trend, suggesting the presence of diseconomies of scale rather than scale economies.

Figure 1. Cost Scaling Among U.S. Telephone Companies in 1916



Source: Douglas Galbi, *Dis-Economies in Communications Networks*, PURPLE MOTES (Apr. 25, 2010), http://purplemotes.net/2010/04/25/economies-and-dis-economies-in-communications-networks/.

Telephone companies facing increasing costs found themselves constantly having to ask municipal regulators to approve rate increases.[437] This was particularly difficult for the independents who offered the benefits of cheaper rates as the principal reason for being allowed to enter.[438]

b. Network Economic Effects.—Dominant positions are also often attributed to another economic concept known as network economic effects. Network economic effects exist when the value of a network increases with the number of subscribers.[439] To use a classic example, consumers during the 1980s who were choosing between the two leading videocassette recorder (VCR) formats (Sony Betamax and VHS) did not really care about their

437. MUELLER, *supra* note 10, at 66.

438. *Id.* at 37.

439. DANIEL F. SPULBER & CHRISTOPHER S. YOO, NETWORKS IN TELECOMMUNICATIONS 4 (2009).

technical capabilities.[440] What determined the value is which format the majority of other VCR owners would adopt.[441] The telephone system is regarded as a paradigmatic example of a network that exhibits network economic effects.[442] As AT&T noted in its 1901 Annual Report:

> That the system be complete and of the greatest utility, it is necessary that as many persons as possible should be connected to it as to be able to talk or be talked to by telephone. . . . [The user's] advantage as a telephone subscriber is largely measured by the number of persons with whom he may be put in communication.[443]

AT&T similarly observed in its 1908 Annual Report, "A telephone—without a connection at the other end of the line—is . . . one of the most useless things in the world. Its value depends on the connection with the other telephone—and increases with the number of connections."[444]

Network economic effects can give large companies a competitive advantage. The fact that larger networks are more valuable provides strong incentives for new customers to opt for the larger network.[445] This in turn makes the largest network still larger, further reinforcing its competitive advantage.[446] This advantage can come from having more local subscribers.[447] The market leader could ensure that it alone enjoyed those advantages simply by refusing to interconnect with the other network.[448]

440. *See* Hiroshi Ohashi, *The Role of Network Effects in the US VCR Market, 1978–1986*, 12 J. ECON. & MGMT. STRATEGY 447, 449 (2003) (explaining that improvements in Beta's product quality were not enough to overcome network effects).

441. *Id.* at 448.

442. *See* Christopher S. Yoo, *Network Neutrality, Consumers, and Innovation*, 2008 U. CHI. LEGAL F. 179, 223 n.139 (collecting authorities).

443. Ithiel de Sola Pool et al., *Foresight and Hindsight: The Case of the Telephone*, *in* THE SOCIAL IMPACT OF THE TELEPHONE 127, 131 (Ithiel de Sola Pool ed., 1977) (ellipsis and alteration in original) (quoting AT&T CO., ANNUAL REPORT OF THE DIRECTORS OF AMERICAN TELEPHONE & TELEGRAPH COMPANY TO THE STOCKHOLDERS FOR THE YEAR ENDING DECEMBER 31, 1901, at 6 (1902)).

444. AT&T CO., ANNUAL REPORT OF THE DIRECTORS OF AMERICAN TELEPHONE & TELEGRAPH COMPANY TO THE STOCKHOLDERS FOR THE YEAR ENDING DECEMBER 31, 1908, at 21 (1909).

445. Howard A. Shelanski & J. Gregory Sidak, *Antitrust Divestiture in Network Industries*, 68 U. CHI. L. REV. 1, 8 (2001).

446. *Id.* at 8–9.

447. *See* Mark A. Lemley, *Antitrust and the Internet Standardization Problem*, 28 CONN. L. REV. 1041, 1046 n.19 (1996) (recognizing network effects in phone service and noting that phone companies were originally successful as local-only providers); MUELLER, *supra* note 10, at 72–73 (noting that demand for long-distance service was initially very low).

448. Shelanski & Sidak, *supra* note 445, at 8; *see also* BROOKS, *supra* note 16, at 114 (describing how Bell's policy against interconnection gave it a competitive advantage over independents). *But see* LIPARTITO, *supra* note 14, at 250 n.4 ("The notion that Bell's refusal to interconnect was a potent competitive weapon is an article of faith in telephone literature.").

Alternatively, others have suggested that greater reach made possible by AT&T's key long-distance patents made its network more desirable.[449]

Looking at only the national numbers, arguments that AT&T's overall size gave it a strategic advantage are hard to reconcile with the fact that in 1907, the independents controlled more subscribers than did AT&T.[450] The independents could thus nullify whatever advantages AT&T enjoyed simply by interconnecting with one another.[451] That said, national numbers are somewhat misleading in that subscribers during the World War I era made almost exclusively local calls.[452] What mattered, then, was the percentage of customers that any particular company controlled locally, not nationally. Although AT&T continued to enjoy a strong position in the northeast and mid-Atlantic states, the independents were the market leaders in the Midwest.[453] The independents were also stronger in small towns and rural communities that AT&T had neglected.[454] In those areas, local network economic effects would have favored the independents, not AT&T.[455] The fact that AT&T also lost market share in markets which it entered first further cuts against network economic effects as a source of competitive advantage.[456]

Moreover, it is recognized that customer heterogeneity can ameliorate network economic effects.[457] If subscribers place a higher value on a small subset of people, what matters is not the total number of people who subscribe, but rather whether the people most important to that subscriber join the network.[458] When this is the case, different groups can segregate into

---

449. BROCK, *supra* note 15, at 119; FAULHABER, *supra* note 14, at 3; John V. Langdale, *The Growth of Long-Distance Telephony in the Bell System: 1875–1907*, 4 J. HIST. GEOGRAPHY 145, 155 (1978).

450. *See* Langdale, *supra* note 449, at 152 (providing a breakdown of 1907 phone ownership).

451. *See* Roger G. Noll & Bruce M. Owen, *The Anticompetitive Uses of Regulation:* United States v. AT&T, *in* THE ANTITRUST REVOLUTION 290, 292 (John E. Kwoka, Jr. & Lawrence J. White eds., 1989) (noting that the independents were able to effectively compete by interconnecting).

452. *See* Robert MacDougall, *The People's Telephone: The Political Culture of Independent Telephony*, 1 BUS. & ECON. HIST. ONLINE 1, 13 (2003), www.thebhc.org/publications/BEHonline/2003/MacDougall.pdf (remarking that even as late as 1930, less than half of 1% of all telephone calls crossed state lines).

453. Jeffrey E. Cohen, *The Telephone Problem and the Road to Telephone Regulation in the United States, 1876–1917*, 3 J. POL. HIST. 42, 49 (1991).

454. *Id.* at 48.

455. MUELLER, *supra* note 10, at 59, 62.

456. Cohen, *supra* note 453, at 49.

457. *See* Christopher S. Yoo, *When Antitrust Met Facebook*, 19 GEO. MASON L. REV. 1147, 1151–53 (2012) (noting that increased heterogeneity can work to counterbalance certain network effects); SPULBER & YOO, *supra* note 439, at 140 ("[N]etwork externalities may be substantially mitigated if user preferences are nonuniform.").

458. *See* Bob Briscoe et al., *Metcalfe's Law Is Wrong: Communications Networks Increase in Value as They Add Members—But by How Much? The Devil Is in the Details*, IEEE SPECTRUM, July 2006, at 34, 37 (pointing out that different connections within a network have different values).

different networks without any significant loss in value.[459] What little loss remains can be substantially ameliorated if gateways exist between the networks.[460]

As Milton Mueller has pointed out, this is precisely what occurred in the early telephone network.[461] "Classes and neighborhoods divided themselves into user communities."[462] When they needed to call the other network, they had easy access to bridge technologies, such as payphones or free phones maintained by drugstores and saloons to attract business.[463] Perhaps the best evidence that network economic effects did not give AT&T any advantage in local markets is the fact that the independents expressed little interest in interconnecting with AT&T.[464] Clearly, the independents did not see their inability to reach a larger number of customers as a competitive disadvantage.

All of these considerations undercut the suggestion that AT&T was able to use network economic effects in local telephone markets to restore its dominance. But what about long distance? As an initial matter, interstate long-distance calling represented a trivially small fraction of overall telephone revenues.[465] As one customer noted in 1909, truly long-distance telephoning was "of little commercial or social importance."[466]

---

459. *See* S.J. Liebowitz & Stephen E. Margolis, *Are Network Externalities a New Source of Market Failure?*, *in* 17 RESEARCH IN LAW AND ECONOMICS 1, 18–19 (Richard O. Zerbe, Jr. & William Kovacic eds., 1995) (describing two scenarios where a network increases in value not only by increasing the sheer number of users, but also by increasing the desirability of communication between those users).

460. Yoo, *supra* note 457, at 1153–54.

461. MUELLER, *supra* note 10, at 20–29.

462. *Id.* at 85; *see also* MacDougall, *supra* note 452, at 13 (describing the choice between AT&T's extensive network and an independent's intensive network was often one that divided along class lines).

463. MUELLER, *supra* note 10, at 82, 85; *see also* BROCK, *supra* note 15, at 110 (observing that there was often not much to be gained in the first place by connecting telephone networks with each other).

464. MUELLER, *supra* note 10, at 10, 51, 78–79; Gabel, *supra* note 138, at 353–54.

465. *See* MUELLER, *supra* note 10, at 72–73 & n.50 (noting that "[t]he demand for telephone connections between points over 200 miles apart was still restricted to a tiny minority of users" and citing an AT&T report from 1900 that 98% of calls placed from cities and 95% of calls placed from small towns were to points within 50 miles); MacDougall, *supra* note 462, at 13 n.46 (citing a 1905 statement by an Independent that 98% of all long-distance calls were placed to points within a one hundred-mile radius). For later statements to the same effect, see *Smith v. Ill. Bell Tel. Co.*, 282 U.S. 133, 147 (1930) (reporting that interstate calls constituted 0.5% of all telephone traffic); *Hearings on S. 6 Before the Comm'n. on Commc'ns of the H. Comm. on Interstate Commerce*, 71st Cong. 1565, 1585–86 (1930) (statement of Joseph B. Eastman, Comm'r, Interstate Commerce Comm'n) (reporting that interstate traffic represented 0.47% of all exchange calls and 0.46% of total exchange revenue and that if exchange and toll calls were combined, intrastate traffic represented 1.36% of all calls and 9.9% of revenue).

466. Gansey R. Johnson, *Telephone Combination: Would It Serve a Good Purpose?*, TELEPHONY, Jan. 2, 1909, at 5, 7.

What little long distance traffic existed tended to be regional.[467] In regional long distance, AT&T enjoyed no technological advantage, as both the independents and the Bell System simply connected adjacent exchanges.[468] Indeed, AT&T President Frederick Fish was forced to concede in 1903 that the company held "no controlling patents on long distance telephone apparatus or systems" and that "long distance lines of some commercial value [could] be constructed and operated by anyone."[469]

Moreover, AT&T had focused most of its attention on connecting distant points.[470] The skeletal pattern that resulted made AT&T weaker with respect to short-haul long distance.[471] The independents' focus on intensive coverage of smaller areas put them in a stronger position.[472] In the words of the president of one independent, his company "has the near long distance points, the Bell [has] the far-off."[473] Between the two, it was the near-long-distance points that mattered more.[474]

It is thus hard to see how either scale economies or network economic effects could have been the means through which AT&T reestablished its monopoly. The answer must lie elsewhere.

2. The Supposedly Lax Enforcement of the Antitrust Laws.—Other commentators attribute the reemergence of monopoly to a failure of antitrust enforcement in two ways. First, some assert that government intervention occurred after AT&T had reestablished its monopoly position and that the authorities should have forced AT&T to divest its newly acquired properties.[475] Second, some argue that antitrust authorities implemented the Kingsbury Commitment in a way that permitted the Bell System to continue to merge to monopoly.[476] Specifically, the Commitment did not prevent AT&T from acquiring independent local telephone companies with which they did not directly compete.[477] Many scholars have claimed that the

467. *See* MUELLER, *supra* note 10, 72–73 (noting that "[n]o more than 5 percent of all telephone calls were to points more than fifty miles away" and that "the real source of competitive advantage was comprehensive coverage of a particular region corresponding to the interest of the majority of telephone users").

468. *Id.*

469. *Id.* at 72 n.49.

470. *See id.* at 73 (noting that prior to 1894 Bell pursued the long-distance market "to the exclusion of most others" and that "[t]he new emphasis on intensive toll line development within the licensee companies' territories was actually a sharp departure from the old Bell vision").

471. *Id.* at 73, 90.

472. *See* MacDougall, *supra* note 462, at 11–12 (arguing that AT&T "regarded its long lines as a major competitive weapon" but that "middle distance connections became a key competitive weapon for the independents").

473. FREDERICK S. DICKSON, TELEPHONE INVESTMENTS AND OTHERS 41 (1905).

474. *Id.*

475. *See supra* note 162 and accompanying text.

476. *See supra* notes 163–64 and accompanying text.

477. *See supra* note 163 and accompanying text.

antitrust authorities allowed mergers of competitive companies so long as the transaction involved a swap of lines.[478]

Claims that the Kingsbury Commitment arrived too late are belied by the fact that the independents still controlled 45% of the national market at the time of the settlement,[479] leaving AT&T's share well below the threshold needed to constitute a monopoly.[480] Moreover, as noted above, regional share mattered more than national share, and in many regions, independents still enjoyed majority positions.[481] Had the Kingsbury Commitment simply stabilized the industry structure that existed in 1913, it would not have inevitably collapsed back into monopoly.

---

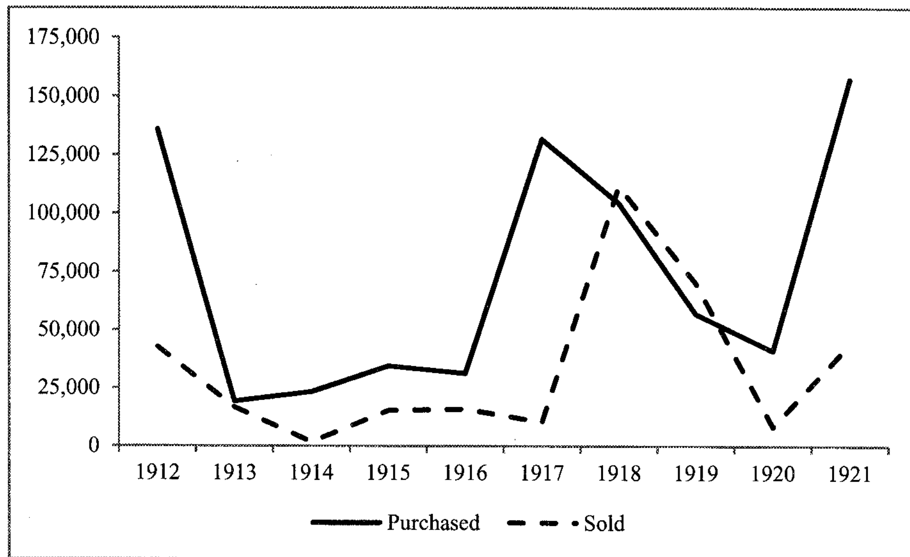
478. *See supra* note 164 and accompanying text.

479. U.S. DEP'T OF COMMERCE, HISTORICAL STATISTICS OF THE UNITED STATES: COLONIAL TIMES TO 1970 (pt. 2) 783 (Bicentennial ed. 1975); *see also* COHEN, *supra* note 15, at 27 (stating that AT&T's market share "dipped below 50 percent"); MUELLER, *supra* note 10, at 133 (noting that at the time of the agreement dual service remained in 13% of all communities with exchanges in the United States); Krishna P. Jayaker & Harmeet Sawhney, *Universal Service: Beyond Established Practice to Possibility Space*, 28 TELECOMM. POL'Y 339 (2004) (noting that "significant market share remained with the independents until 1921, when the Willis-Graham Act again permitted the Bell System to acquire non-affiliated companies") ; Steve G. Parsons & James Bixby, *Universal Services in the United States: A Focus on Mobile Communications*, 62 FED. COMM. L.J. 119, 125 (arguing that independents maintained significant market share until 1921).

480. *See* United States v. Aluminum Co. of Am., 148 F.2d 416, 424 (2d Cir. 1945) (expressing doubts that 60% to 64% market share would be enough to constitute a monopoly in the aluminum market).

481. *See supra* notes 453–54, 472 and accompanying text.

Figure 2. Telephone Lines Purchased and Sold by the Bell System, 1912–1921



Source: FED. COMMC'NS COMM'N, REPORT ON CONTROL OF TELEPHONE COMMUNICATIONS: CONTROL OF INDEPENDENT TELEPHONE COMPANIES 42 tbl.v (1937).

The suggestion that the Kingsbury Commitment did nothing to slow AT&T's acquisition strategy is also belied by the facts. As Figure 2 shows, the three years following the imposition of the Kingsbury Commitment saw a sharp drop in the number of lines acquired by the Bell System. Had AT&T been simply allowed to swap lines with its competitors, its acquisition numbers would have remained high and been counterbalanced by an equal number of lines acquired by the independents. Yet this is not the pattern observed following 1913. The fact that the number of lines acquired by the Bell System dropped in 1913, 1914, 1915, and 1916 suggests that the Kingsbury Commitment was effective in slowing AT&T's acquisition policy for at least three years.

3. Government-Mandated Consolidation as a Missing Consideration.— The biggest factor missing from the explanation of the reestablishment of the Bell System is the government takeover during World War I. As noted above, industry consolidation was one of Burleson's central policies during the takeover. Consequently, acquisitions spiked in 1918, only to drop off again in 1919 and 1920 after the return of the wires until the Willis-Graham Act completely abrogated antitrust review of telecommunications mergers in

1921.[482] From this explanation, monopoly was not the sole result of the antitrust authorities' refusal to curb the corporate ambitions of Theodore Vail. Instead of being asleep at the switch, the government was one of the primary drivers of the return to monopoly.

There is one aspect of the data that does not completely fit this story: the upsurge in Bell acquisitions in 1917.[483] The impetus for this change came not from Bell, but rather from the independents. Having finished the build out of the areas that Bell had ignored, independents found that further revenue growth required going head-to-head with Bell.[484] Intensive competition is much more expensive and less profitable than an extensive race for the market.[485] Faced with the prospects of vigorous competition, many independents began to explore reaching some form of accommodation.

In 1915, the independents appointed a committee to explore consolidating into the Bell System.[486] It was this committee that proposed complying with the spirit, if not the letter, of the Kingsbury Commitment by "permit[ting] the acquisition by the Bell System of Independent properties by means of a division of territory, so long as in such a division the Bell System should not acquire more property or territory than it relinquished."[487] This solution "would conform to the probable spirit" of the Kingsbury Commitment "by continuing the prohibition on the expansion of the Bell System at the expense of competition" while still allowing a division of territory.[488] The Justice Department effectively accepted this modification in 1917.[489]

The critical support for the reemergence of mergers in 1917 thus came from the independents, who also provided the impetus for the Willis-Graham Act's abolition for all antitrust scrutiny of telecommunications mergers.[490]

---

482. *See supra* Figure 2.

483. *See* Bickers, *supra* note 17, at 126–27 (remarking on the increase in consolidations of telephone companies during this time).

484. *See* MUELLER, *supra* note 10, at 55–60 (discussing phases of expansion employed by independents).

485. Christopher S. Yoo, *Product Life Cycle Theory and the Maturation of the Internet*, 104 NW. U. L. REV. 641, 646–47, 666–67 (2010).

486. Bickers, *supra* note 17, at 123–24.

487. Letter from H.D. Critchfield, Sales Dep't, Automatic Electric Co., to F.H. Woods 2 (May 18, 1915), *quoted in* Bickers, *supra* note 17, at 125.

488. *Id.* (emphasis omitted) (internal quotation marks omitted).

489. Bickers, *supra* note 17, at 126; CHARLES A. PLEASANCE, THE SPIRIT OF INDEPENDENT TELEPHONY 86 (1989); *see also* H.R. Doc. No. 340, 76th Cong., 1st Sess. (1939) (noting that the Kingsbury Commitment was generally understood to repent acquisition of competing telephone companies until after January 1918, when it became generally understood that it was not a violation of The Kingsbury Commitment for the Bell System to acquire competing telephone stations, if at the same time the Bell System sold an equal or comparable number of Bell-owned stations to an independent).

490. 61 CONG. REC. 1983 (1921) (remarks of Rep. Winslow) ("The bill was brought to the attention of the committee by . . . the so-called independent telephone companies of the United States. . . . [T]hey have represented to the committee . . . that if the opportunity to sell or

Consolidation enjoyed the support of key consumer groups who needed both services and did not want to pay for two subscriptions.[491] This fact makes it difficult to lay this development at the feet of Theodore Vail. The fact that both competitors and consumers endorsed the change makes it hard to characterize it as government capitulation to a monopolist, as antitrust enforcement authorities find it difficult to oppose mergers when neither competitors nor consumers object.

This was particularly so in the case of telephony, which as noted above is largely intrastate in character.[492] Because federal jurisdiction extended only to interstate matters, mergers between two local telephone companies were arguably beyond the Justice Department's jurisdictional reach.[493] That is why on September 7, 1914, when authorizing AT&T's acquisition of a competing local telephone company in Spokane, Washington, the Justice Department announced that the Kingsbury Commitment was not meant to prevent communities from eliminating dual service if they so chose.[494] Similarly, on December 7, 1914 (less than a year after the issuance of the Kingsbury Commitment), the Report of the Attorney General noted that the settlement "does not mean that where there are two telephone systems in a city or town there never can be a consolidation into a single system."[495] On the contrary, it "leaves local communities generally free to have one telephone system, if they desire," so long as the resulting consolidated company maintained all previous long-distance interconnections.[496] If sufficient local political support existed for the merger, the Justice Department would be hard pressed to oppose it.[497]

In any event, all of these dynamics were soon rendered moot by the government takeover, as Burleson condoned and accelerated the process of industry consolidation.[498] The strength of the Postmaster General's convictions rendered the views of both AT&T and the independents irrelevant. Even if it was not the only factor, Burleson's relentless support

---

consolidate is not afforded to them they are liable to go through the condition of bankruptcy . . . ."); *see also* Bickers, *supra* note 17, at 127; Gabel, *supra* note 138, at 353 ("[T]he independents joined Bell in seeking passage of the Willis-Graham Act of 1921, which permitted the merger or consolidation of competing telephone companies."); Robinson, *supra* note 14, at 8 ("[T]he independents joined AT&T in supporting a lifting of restrictions on AT&T acquisitions.").

491. COHEN, *supra* note 15, at 33; MUELLER, *supra* note 10, at 140–44; Gabel, *supra* note 138, at 348; Weiman & Levin, *supra* note 142, at 122, 124.

492. *See supra* notes 465–73 and accompanying text.

493. STONE, *supra* note 3, at 193.

494. United States v. AT&T Co., No. 6082 (D. Or. Sept. 7, 1914) (order modifying decree), *reprinted in* DECREES AND JUDGMENTS IN FEDERAL ANTI-TRUST CASES, JULY 2, 1890–JANUARY 1, 1918, at 497, 497–99 (Roger Shale ed., 1918).

495. ATT'Y GEN. ANN. REP. 14 (1914).

496. *Id.*

497. MUELLER, *supra* note 10, at 133–34.

498. STONE, *supra* note 3, at 199; VIETOR, *supra* note 16, at 172–73.

for industry consolidation has been largely overlooked as one of the reasons for AT&T's return to dominance.[499]

B. The Origins of Universal Service

A vibrant debate has emerged over the origins of universal service. The conventional wisdom argues that universal service was the result of AT&T's commitment to broad geographic coverage, exemplified by AT&T's endorsement of "'One System,' 'One Policy,' 'Universal Service.'"[500] Critics of this position suggest that the concept of universal service is more modern, being the product of AT&T's attempt to justify the continuation of its monopoly when facing the emergence of competition during the 1960s and 1970s.[501] Richard John has offered an intriguing third interpretation, arguing that universal service represents Theodore Vail's attempt to emulate the postal system, influenced by his experience running the Railway Mail System in between his stints at AT&T.[502]

There can be no question that AT&T endorsed the idea of providing telephone service to all Americans. For example, its 1907 Annual Report intoned that the "'universality'" that was the Bell System's strength "carries with it . . . the obligation to occupy and develop the whole field," including semi-urban and rural areas as well as urban areas.[503] Its 1910 Annual Report affirmed the company's "belie[f] that some sort of connection with the telephone system should be within the reach of all" and that the telephone network would ultimately become a medium for all electronic

---

499. Only a handful of works acknowledge Burleson's role, and those that do offer no more than a few words. *See* H.R. Doc. No. 340, 76th Cong., (1939) (devoting a single sentence to the government's conscious policy of "caus[ing] the coordination and consolidation of competing systems wherever possible"); FISCHER, *supra* note 15, at 50 (devoting a single sentence to how "the wartime experience of coordination between AT&T and the independents accelerated the unification of the industry"); HUBER ET AL., *supra* note 14, § 4.4.1, at 354 ("During World War I, from 1918 to 1919, the Postmaster General took over operation of the telephone industry and . . . directed the competing local systems to consolidate into a single national network."); MUELLER, *supra* note 10, at 133 (making a passing reference to "World War I-induced centralization"); VIETOR, *supra* note 16, at 172–73 (devoting a single sentence to how government control "ran squarely up against the Justice Department's prohibition on consolidation of competing exchanges," as Burleson "sought to eliminate competition and integrate operations wherever possible"); Peters, *supra* note 15, at 257 (noting Burleson's view that "government operation and control of the telephone system 'would undoubtedly cause the coordination and consolidation of competing systems wherever possible'").

500. AT&T CO., ANNUAL REPORT OF THE DIRECTORS OF AT&T COMPANY TO THE STOCKHOLDERS FOR THE YEAR ENDING DECEMBER 31, 1909, at 18 (1910); *see* Pool et al., *supra* note 444, at 131 ("[T]he goal of universality, which became one of the watchwords of the Bell system, was there from the beginning.").

501. MUELLER, *supra* note 10, at 150–52; *see also* ROBERT W. CRANDALL & LEONARD WAVERMAN, WHO PAYS FOR UNIVERSAL SERVICE? 6–8 (2000) (explaining that although the phrase was first used by Theodore Vail, the company "revived the notion of universal service" with a "new definition" in response to competition during the 1960s and 1970s).

502. JOHN, *supra* note 17, at 388; Richard R. John, *Theodore N. Vail and the Civic Origins of Universal Service*, 28 BUS. & ECON. HIST. 71, 76–79 (1999).

503. AT&T, *supra* note 146, at 28.

communications "from every one in every place to every one in every other place, a system as universal and as extensive as the highway system of the country which extends from every man's door to every other man's door."[504]

Critics such as Milton Mueller have dismissed these statements as nothing more than an "oratorical jab."[505] It was not until the independents forced AT&T's hand that the company began to live up to the promise of its earlier rhetoric.[506] Instead, they regard universal service as an invention of the late 1960s and 1970s to justify the continuation of the Bell monopoly.[507] The primary mechanism was to use the process of separations to allocate a higher proportion of network elements used both for local and long-distance service, such as the loop and the switch, to long-distance rates.[508] The effect was to overcharge for long-distance service in order to cross subsidize local service.[509] This process did not begin until the 1950s and did not reach full stride until the late 1960s and 1970s.[510]

While it is true that long distance–local cross subsidies did not emerge until later in the history, it was only one of several cross subsidies built into telephone rates. Another key cross subsidy takes advantage of the fact that the higher density makes providing service less costly in urban areas than rural areas.[511] Using rate averaging to impose a uniform price effectively permits rates paid by urban users to cross subsidize those paid by rural users.[512]

Although these scholars are correct that the long distance–local cross subsidy did not emerge until the 1960s and 1970s, the urban–rural cross subsidy associated with rate averaging was well established in the postal service before World War I.[513] The takeover allowed this postal concept to become a staple of telephone policy as well.[514] In the words of one commentator:

> During this period of government ownership, the decision was made to set standard long-distance rates throughout the country, based on average costs. In other words, subscribers calling from large cities would pay above costs in order to provide a subsidy to those in rural areas. So, early in the century cross-subsidization began, embraced by the industry, which rarely question the premise behind the

504. *Id.* at 23.
505. MUELLER, *supra* note 10, at 100.
506. *Id.* at 101–03.
507. *Id.* at 151–52.
508. *Id.* at 151–55.
509. *Id.* at 159.
510. *Id.* at 160–61.
511. Thierer, *supra* note 16, at 277.
512. *Id.* at 276–77.
513. JOHN, *supra* note 17, at 379; John, *supra* note 502, at 75–76.
514. 52 CONG. REC. 849 (1915) (statement of Rep. David Lewis); POSTMASTER GEN., *supra* note 26, at 10; JOHN, *supra* note 17, at 379, 387.

arrangement that the ability to communicate with subsidized subscribers was of value to the subsidizing subscribers.[515]

Following the war, the state public utility commissions would follow the federal government's example when setting local rates.[516] By the 1920s, statewide rate averaging had become a standard feature of the regulatory landscape.[517]

So John's suggestion that the postal model of universal service provided an early influence on universal service policy in the telephone industry appears to be well taken, but with a somewhat different twist. As noted above, John believed that the mechanism through which these concepts were incorporated into telecommunications policy was Vail's experience with the Railway Mail Service.[518] While John is correct about the influence of the postal system on telephone rates, Vail's experiences were reinforced by a much more direct mechanism: Rate averaging was imposed on the telephone system by the Postmaster General himself.

C. *Federal–State Relations*

The government takeover during World War I also had a profound influence on federal–state relations. Burleson's order to raise and standardize national rates was met with immediate resistance from state regulatory agencies who sought to enjoin the rate increases.[519] Injunctions were granted in ten states across the country, from Florida to Pennsylvania to South Dakota.[520] AT&T's general counsel, N.T. Guernsey worked with Post Office Solicitor William Lamar to get a test case quickly to the Supreme Court.[521] In March, the South Dakota Supreme Court enjoined the Dakota Central Telephone Company from raising rates.[522] Guernsey convinced Lamar to fight the injunction, and Lamar obtained approval from the Attorney General to challenge the ruling.[523] As Danielian describes it, "Thus the United States Attorney General's office, the Post Office Department, and the Bell System were mobilized, hand in hand, to defend the Postmaster General's order for increased Bell telephone rates."[524]

515. 1 LEONARD S. HYMAN ET AL., THE NEW TELECOMMUNICATIONS INDUSTRY: EVOLUTION AND ORGANIZATION 81 (1987), *quoted in* Thierer, *supra* note 16, at 276.

516. Thierer, *supra* note 16, at 277.

517. VIETOR, *supra* note 16, at 173–74; Warren G. Lavey, *The Public Policies that Changed the Telephone Industries into Regulated Monopolies*, 39 FED. COMM. L.J. 171, 188–89 (1987).

518. *See supra* note 502 and accompanying text.

519. DANIELIAN, *supra* note 17, at 260.

520. *Id.* at 262.

521. *Id.*

522. *Id.*

523. *Id.* at 263.

524. *Id.*

The case of *Dakota Central Telephone Co. v. South Dakota* was argued on May 5–6, 1919, and decided on June 2, 1919.[525] As noted earlier, the Supreme Court upheld the rate increases by a vote of 8–1.[526] The Court depicted the issue as similar to the one recently resolved in *Northern Pacific Railway Co. v. North Dakota*,[527] where the Court upheld the federal government's ability to set intrastate railroad rates under Congress's war power.[528] In the case at bar, Justice White saw the same principle at work: Congress could lawfully take over a public utility under its war power,[529] and state police power did not create protected enclaves of state prerogative.[530] Justice White wrote:

> Conceding that it was within the power of Congress, subject to constitutional limitations, to transplant the state power as to intrastate rates into a sphere where it, Congress, had complete control over telephone lines because it had taken possession of them and was operating them as a governmental agency, it must follow that, in such sphere there would be nothing upon which the state power could be exerted except upon the power of the United States, that is, its authority to fix rates for the services which it was rendering through its governmental agencies.[531]

Accordingly, the Court overruled the injunction and Burleson's rate hikes were upheld.[532] The victory ended up being somewhat pyrrhic. Frustration with the rate increases approved by the Court led the states to lend their support to returning the wires to private control as quickly as possible.[533] Moreover, the Supreme Court would subsequently curb federal power in *Smith v. Illinois Bell Telephone Co.*[534] by denying the ICC the authority to affect intrastate rates.[535] Instead, the states would bear that responsibility after the cost of any assets used for both interstate and intrastate service was apportioned between the two services.[536] This movement culminated with the inclusion of a provision in the Communications Act of 1934 disavowing any FCC jurisdiction over

525. 250 U.S. 163 (1919).

526. *See supra* notes 326–28 and accompanying text.

527. 250 U.S. 135 (1919).

528. *Id.* at 151–52.

529. *Dakota Cent. Tel.*, 205 U.S. at 183–84.

530. *Id.* at 187.

531. *Id.*

532. As May notes, the Court's decisions occurred after "the experiments were marked for extinction," and thus the Court may have felt less need to become involved with the issues. MAY, *supra* note 17, at 57.

533. Bickers, *supra* note 17, at 154–55.

534. 282 U.S. 133 (1930).

535. *Id.* at 159–60.

536. *Id.* at 148–49.

"charges, classifications, practices, services, facilities, or regulations for or in connection with intrastate communication service."[537]

A comparison with how similar issues were resolved in railroading illustrates the strength and influence of state opposition to federal power mobilized by Burleson's ham-fisted policies.[538] When confronted with a similar issue with respect to railroads in *The Shreveport Rate Case*,[539] the Court held that the ICC had jurisdiction over interstate rates as well as "all matters having such a close and substantial relation to interstate traffic that the control is essential or appropriate" to the security and efficiency of interstate service.[540] The Court further concluded, "The fact that carriers are instruments of intrastate commerce, as well as of interstate commerce, does not derogate from the complete and paramount authority of Congress over the latter."[541] In short, "[w]herever the interstate and intrastate transactions of carriers are so related that the government of the one involves the control of the other, it is Congress, and not the State, that is entitled to prescribe the final and dominant rule."[542] The Court reaffirmed "the principle that Congress in the exercise of its paramount power may prevent the common instrumentalities of interstate and intrastate commercial intercourse from being used in their intrastate operations to the injury of interstate commerce" even though "intrastate transactions of interstate carriers may thereby be controlled."[543] Any other conclusion would contradict the principle of federal supremacy.[544]

The Court reiterated this conclusion after the war in *Railroad Commission v. Chicago, Burlington & Quincy Railroad*,[545] in which it held that "[e]ffective control of [interstate traffic] must embrace some control over [intrastate traffic] in view of the blending of both in actual operation. The same rails and the same cars carry both. The same men conduct them. Commerce is a unit and does not regard state lines . . . ."[546] When interstate and intrastate commerce "are so mingled together that the supreme authority,

537. 47 U.S.C. § 152(b) (2006).

538. *See* HUBER ET AL., *supra* note 14, § 3.2.3–.5, at 216–18 (noting that states wanted federal authority to be sharply limited).

539. 234 U.S. 342 (1914).

540. *Id.* at 351.

541. *Id.*

542. *Id.* at 351–52.

543. *Id.* at 353.

544. *Accord* Minnesota Rate Cases, 230 U.S. 352 (1913). This case held:

> [T]he full control by Congress of the subjects committed to its regulation is not to be denied or thwarted by the commingling of interstate and intrastate operations [and that] the execution by Congress of its constitutional power to regulate interstate commerce is not limited by the act that intrastate transactions may have become so interwoven therewith that the effective government of the former incidentally controls the latter.

Id. at 399.

545. 257 U.S. 563 (1922).

546. *Id.* at 588.

the Nation, cannot exercise complete effective control over interstate commerce without incidental regulation of intrastate commerce, such incidental regulation is not an invasion of state authority or a violation of the proviso."[547] The Supremacy Clause gives the federal government the power to "impose any reasonable condition on a State's use of interstate carriers for intrastate commerce it deems necessary or desirable."[548]

Had the Court applied these same principles to telephony, it would have found that interstate and intrastate phone calls were just as intertwined as rail traffic and upheld federal jurisdiction over local rates. And yet in *Smith*, the Court mandated that all property used for both interstate and intrastate calling be separated into interstate and intrastate portions, calling this separation "essential to the appropriate recognition of the competent governmental authority in each field of regulation."[549] The Court recognized that "the difficulty in making an exact apportionment of the property is apparent," but nonetheless optimistically concluded that "reasonable measures" would be sufficient.[550] As noted above, Congress codified this understanding by enacting section 2(b) of the Communications Act of 1934, which courts and commentators have recognized was specifically enacted to prevent the extension of *Shreveport*-type rules to telephony.[551]

It is likely that the Post Office's dismal record running the telephone system and its willingness to brush aside the interests of state regulators rendered policy makers less inclined to condone strong federal jurisdiction over telephony. Moreover, the fact that the Supreme Court's *Dakota Central* decision upholding federal authority over local telephony was upheld as a war measure may have left the Justices feeling constrained to come to the opposite conclusion after the exigency had passed.

D. *The Limits of Government*

Perhaps the biggest question is that having taken over the telephone system, why did the government give it back? Burleson openly harbored ambitions of making government ownership permanent, as did leaders in the U.S. military.[552] Doing so would also have brought U.S. policy into conformity with the rest of the world.

---

547. *Id.*

548. *Id.* at 590.

549. 282 U.S. 133, 148 (1930); *accord id.* at 149 ("The proper regulation of rates can be had only by maintaining the limits of state and federal jurisdiction . . . .").

550. *Id.* at 150; *accord* HUBER ET AL., *supra* note 14, § 3.3.3, at 223–24.

551. La. Pub. Serv. Comm'n v. FCC, 476 U.S. 355, 372 (1986); Computer & Commc'ns Indus. Ass'n v. FCC, 693 F.2d 198, 216 n.99 (D.C. Cir. 1982); N.C. Util. Comm'n v. FCC, 537 F.2d 787, 793 n.6 (4th Cir. 1976); Matthew S. Bewig, *Federalism and Telecommunications: On the Right Wavelength?*, 59 GEO. WASH. L. REV. 1190, 1194–95 (1991); Richard McKenna, *Preemption Under The Communications Act*, 37 FED. COMM. L.J. 1, 13–14 (1985).

552. *See supra* note 197 and accompanying text.

One challenge that government ownership struggles to surmount is the problem of raising risk capital. Unlike private enterprises, which can issue equity, governments must finance any capital improvements through taxes, additional revenue from operations, or debt.[553] Thus, because telephone systems in Europe were primarily developed by government actors, capital could not be raised by issuing equity to shareholders willing to assume entrepreneurial risk, and the systems faced political challenges that private enterprises did not need to confront.[554]

In addition, government-owned enterprises tend to avoid risk rather than maximize economic welfare. Government operators in Europe "used their market power to protect themselves from risk rather than to maximize profits."[555] These governments had invested heavily in telegraph systems, and the development of the telephone posed significant risks to telegraph revenues: "All the public agencies attempted to protect their telegraph services from telephone competition, even when they controlled both telegraph and telephone."[556] Even when European governments did not grant an exclusive monopoly to government agencies, these very same agencies "used similar tactics to those of private companies to extend their power."[557] Without the threat of antitrust, there was nothing to restrain unfair practices on the part of the government agencies.

In the United States, development was much more rapid owing to the competitive environment. Competition encouraged innovation and experimentation. For instance, "[t]he existence of several companies allowed various beliefs as to the elasticity of demand to be tested and prevented slow growth through a mistaken belief that the demand was inelastic."[558] But the pace and pattern of development in the United States was not merely influenced by the *absence* of government ownership, as discussed above. U.S. development was also influenced by the *presence* of state and local regulation, the possibility of antitrust enforcement, and the possibility of nationalization.[559]

Each of these three factors influenced development of the U.S. network, such that it cannot be accurately said that the development of telecommunications in Europe and the United States diverged because of the presence of one factor, e.g., ideological or political. It is more accurate to

---

553. *See* Johannes M. Bauer, *Regulation and State Ownership: Conflicts and Complementarities in EU Telecommunications*, 76 ANNALS PUB. & COOPERATIVE ECON. 151, 155 (2005) (explaining that government-owned enterprises have fewer financing options than their private counterparts).

554. *See* BROCK, *supra* note 15, at 146–47 (discussing the differences in development of long-distance telegraphs in the United States and Europe).

555. *Id.* at 145.

556. *Id.* at 146.

557. *Id.* at 145.

558. *Id.* at 144.

559. *See* John, *supra* note 17, at 410 ("Politics always mattered.").

say that U.S. development at this time was influenced by the example of European regulation, and accordingly began to move towards a more similar result.[560] In other words, there was a transatlantic dialogue about the proper way to develop a national telephone network.

But perhaps the most important reason cited during the 1919 hearings on whether the government should return the telephone system to private control was the government's inability to control costs.[561] Both of these concerns should serve as cautionary tales to contemporary advocates of networks operated by governments. Indeed, the government's struggles to run the telephone network during World War I are part of a larger tradition identifying circumstances under which common carriage regulation is most likely to work well. It is best suited to industries such as water and natural gas, in which technology is static, market shares are stable, and the fact that a network has already been built out reduces the emphasis on investment incentives.

All of these considerations should give modern proponents of government ownership of telecommunications networks considerable pause. Indeed, the most salient examples appear to confirm these lessons. On a more optimistic note, this episode also provides reassurance about how justifications based on national emergencies need not necessarily be enduring.

Conclusion

The brief, one-year government takeover of the U.S. telephone system during World War I is rarely analyzed at any length by commentators. When it is discussed, it is often dismissed as an ad hoc event with few implications. A closer inspection of the history and dynamics of this episode in history yields a host of answers to a number of ongoing academic disputes. It reveals that the reconsolidation of the telephone industry during the early twentieth century was the result of conscious government policy as well as the consequences of economic features of the market. As such, it provides another example where government actors rather than corporate magnates were movers in curtailing competition.[562]

The history also sheds new light on the origins of universal service, showing that it was the result of direct application of the ratemaking principles developed for the postal system. In the process, it underscores the

---

560. *See generally*, DANIEL T. RODGERS, ATLANTIC CROSSINGS: SOCIAL POLITICS IN A PROGRESSIVE AGE (1998) (documenting cross-Atlantic influences on Progressive reformers).

561. *See Return of the Wire Systems*, *supra* note 202, at 33–36 (discussing the high costs associated with telephone systems).

562. *See* FRED S. MCCHESNEY, MONEY FOR NOTHING: POLITICIANS, RENT EXTRACTION, AND POLITICAL EXTORTION 17–19 (1997) (arguing that regulatory bodies often deliberately create economic rents in order to have something to redistribute).

legacy of rate averaging as a source of universal service funding that long antedates the more recent imposition of long distance–local cross subsidies.

It illustrates how the invocation of emergency powers represents a two-edged sword. On the one hand, the exigencies of war make actions easier to justify. On the other hand, courts that have upheld a governmental action on the basis of that exigency may later prove reluctant to uphold similar actions taken under more normal circumstances.

Perhaps most importantly, the episode sheds new light on the circumstances under which governmental operation of a communications network is likely to succeed. The experience suggests that such an arrangement works best when the technology is relatively stable, the risks are well defined, providing service does not require substantial new investments, and the political coalition supporting government operation has realistic expectations. These insights can provide considerable guidance to policy makers considering reversing the trend toward privatization and returning to patterns of government ownership.

Remapping the Path Forward: Toward a Systemic View of Forensic Science Reform and Oversight

Jennifer E. Laurin*

The 2009 Report of the National Academy of Sciences (NAS) on the state of forensic science in the American criminal justice system has fundamentally altered the landscape for scientific evidence in the criminal process, and is now setting the terms for the future of forensic science reform and practice. But the accomplishments of the Report must not obscure the vast terrain that remains untouched by the path of reform that it charts. This Article aims to illuminate a critical and currently neglected feature of that territory: namely, the manner in which police and prosecutors, as upstream users of forensic science, select priorities, initiate investigations, collect and submit evidence, choose investigative techniques, and charge and plead cases in ways that have critical and systematic, though poorly understood, influences on the accuracy of forensic analysis and the integrity of its application in criminal cases. By broadening our understanding of how forensic science is created and used in criminal cases—by adopting a systemic perspective—the Article points to a raft of yet unaddressed issues concerning the meaning of scientific integrity and reliability in the context of investigative decisions that are by and large committed to the discretion of decidedly unscientific actors. Critically, the Article demonstrates that systemic dynamics affecting upstream use of forensic science might well undermine the reliability-enhancing goals of the reforms advocated by the National Academy Report. As the NAS Report begins to set the agenda for active conversations around legislative and executive action to reform forensic science, it is critical to consider these questions. Moreover, the Article suggests that the embrace of science as a unique evidentiary contributor within the criminal justice system problematizes some of the bedrock assumptions of American criminal procedure that have, to date, prevented more robust doctrinal intervention in the investigative stages and decisions that the Article explores.

Introduction 1052
I. The NAS Report's View 1058
A. A Brief History of a Path Forward 1058
1. *A History of Neglect* 1059
2. *The "Independence" Question* 1064

---

* Assistant Professor, The University of Texas School of Law. For generosity in providing helpful feedback over the course of this Article's development I thank Erica Beecher-Monas, Simon Cole, Brandon Garrett, Steve Saloom, and participants in the University of Texas School of Law Drawing Board Lunch, the AALS Workshop on Women Rethinking Equality, and the University of Texas School of Law-Instituto Tecnológico Autónomo de México Symposium. For indispensable research assistance I am grateful to Brad Estes, Alex Holmes, Sam Rettew, and especially Ryan Goldstein.

B. The NAS Report's (More and Less Implicit) Premises 1066
1. *The Science–Justice Link* 1068
2. *Funds and Firewalls* 1072
II. What's Missing 1075
A. Upstream Discretion 1076
1. *Evidence Collection.* 1076
2. *Obtaining and Using Forensic Analysis* 1087
3. *Enter NAS* 1100
III. Another Path Forward: Systemic Integration of Forensic Science ... 1106
A. The NAS Report: A Fuller Reform Agenda 1107
1. *Research, Standard Setting, and Training* 1107
2. *Independence Reconsidered* 1111
3. *Oversight* 1112
B. The Criminal Justice Conversation Must Account for Forensic Science and Upstream Dynamics—Particularly Assuming a World of More Early Science 1114
Conclusion 1118

Introduction

Over the course of the last half century, science has moved from the periphery to a place of prominence in the investigation and prosecution of crime. Analysis of physical evidence, particularly with recent scientific and technological advances in the arena of DNA, has been embraced as advancing the fundamental epistemic goals of the criminal justice system by enhancing society's ability to connect the guilty with their misdeeds and, even more powerfully, enabling exculpation of the innocent. As the reliability of more traditional investigative tools such as eyewitness identifications and confessions has been increasingly scrutinized, the comparative accuracy of scientific evidence has been hailed.[1] Yet at the same time, news headlines continually reveal laboratory- or analyst-level breakdowns, and many of the hundreds of exonerations seen in recent decades—through DNA testing or otherwise—have exposed error or outright

1. *See, e.g.*, Williams v. Illinois, 132 S. Ct. 2221, 2227 (2012) (commenting that other forms of proof are less reliable than DNA evidence); Davis v. Mississippi, 394 U.S. 721, 727 (1969) ("[F]ingerprinting is an inherently more reliable and effective crime-solving tool than eyewitness identifications or confessions and is not subject to such abuses as the improper line-up and the 'third degree.'"); Kenworthey Bilz, *Self-Incrimination Doctrine Is Dead; Long Live Self-Incrimination Doctrine: Confessions, Scientific Evidence, and the Anxieties of the Liberal State*, 30 CARDOZO L. REV. 807, 834 (2008) (arguing that "both confessions and eyewitness identifications suffer from a high risk of failing to provide inculpatory evidence on a guilty suspect"); President George W. Bush, State of the Union Address (Feb. 2, 2005), *available at* http://www.msnbc.msn.com/id/6902913/ns/politics-state_of_the_union/t/full-text-state-union-speech (calling for the expanded use of DNA evidence to prevent wrongful convictions).

fraud committed under the guise of "scientific" opinion.[2] A significant body of critical academic commentary on the forensic science field strongly suggests structural rather than individual causes of these ills: the surprisingly thin research base for many forensic methodologies;[3] systematic compromises to the quality of crime laboratory output due to under-resourcing and the undue influence of police and prosecutorial agendas on scientific analysis;[4] poor access to, and use of, expert resources by defense counsel;[5] and lax scrutiny of scientific evidence by courts.[6]

---

2. *See Browse Cases*, NAT'L REGISTRY EXONERATIONS, http://www.law.umich.edu/special/exoneration/Pages/browse.aspx (showing 1,060 total exonerations since 1989 and 239 in cases involving false or misleading science as of January 28, 2013). DNA exonerations that reveal the error of previous expert opinion provide the most dramatic demonstration of erroneous science. *See* BRANDON L. GARRETT, CONVICTING THE INNOCENT: WHEN CRIMINAL PROSECUTIONS GO WRONG 89–90 (2011) (reporting that 61% of trials in DNA exonerations featured invalid conclusions drawn from the evidence); Michael J. Saks & Jonathan J. Koehler, *The Coming Paradigm Shift in Forensic Identification Science*, 309 SCIENCE 892, 892–95 (2005) (arguing that a model for scientifically sound identification science, changing legal admissibility standards for expert testimony, high error rates across forensic science, and the discovery of wrongful convictions are all driving a paradigm shift in forensic identification science); Jonathan Saltzman & Mac Daniel, *Man Freed in 1997 Shooting of Officer*, BOS. GLOBE, Jan. 24, 2004, http://www.boston.com/news/local/articles/2004/01/24/man_freed_in_1997_shooting_of_officer/?page=full (reporting on a case of a fingerprint "match" disproved through DNA exoneration). But equally important, albeit far more challenging and contested, questions have been raised in cases of scientific opinion that has evolved over time to a point of rejecting the theory or application on which the conviction was premised. *See, e.g.*, *Ex parte* Henderson, 384 S.W.3d 833, 833–34 (Tex. Crim. App. 2012) (affirming a grant of habeas corpus relief for a death row inmate based on "new developments in the science of biomechanics" causing scientific experts at the initial trial to change conclusions regarding the cause of death).

3. Jennifer L. Mnookin et al., *The Need for a Research Culture in the Forensic Sciences*, 58 UCLA L. REV. 725, 744–60 (2011) (extolling the need for a research culture in the forensic sciences and arguing that a research culture would increase the accuracy and reliability of forensic science); Michael J. Saks & David L. Faigman, *Failed Forensics: How Forensic Science Lost Its Way and How It Might Yet Find It*, 4 ANN. REV. L. & SOC. SCI. 149, 150–68 (2008) (describing scientific and technological failures of "nonscience forensic sciences," the role of courts in propping up flawed disciplines, and proposals to make "these nonscientific forensic sciences . . . scientific").

4. Paul C. Giannelli, *The Abuse of Scientific Evidence in Criminal Cases: The Need for Independent Crime Laboratories*, 4 VA. J. SOC. POL'Y & L. 439, 441 (1997) (noting that while scientific evidence is superior to other types of evidence, there have been a number of abuses of scientific evidence, "including perjury by expert witnesses, faked laboratory reports, and testimony based on unproven techniques," and that too many experts have a police–prosecution bias); Randolph N. Jonakait, *Forensic Science: The Need for Regulation*, 4 HARV. J.L. & TECH. 109, 191 (1991) (concluding that "[a]ll available information indicates that forensic science laboratories perform poorly," and arguing for increased regulation of crime laboratories).

5. Peter J. Neufeld & Neville Colman, *When Science Takes the Witness Stand*, SCI. AM., May 1990, at 46, 52–53 (noting that defendants rarely have adequate resources to challenge scientific evidence and calling for independent oversight of forensic methods).

6. *See, e.g.*, ERICA BEECHER-MONAS, EVALUATING SCIENTIFIC EVIDENCE: AN INTERDISCIPLINARY FRAMEWORK FOR INTELLECTUAL DUE PROCESS 122 (2007) ("The struggle over what counts as science is in dire straits when it comes to capital sentencing proceedings . . . ."); DAVID H. KAYE, THE DOUBLE HELIX AND THE LAW OF EVIDENCE 244–59 (2010) (describing how the adversarial process affected DNA science); Keith A. Findley, *Innocents at Risk: Adversary Imbalance, Forensic Science, and the Search for Truth*, 38 SETON HALL L. REV. 893, 895–96 (2008) (arguing that "cognitive biases, institutional pressures, and systemic choices (including everything from

Thus, forensic evidence is both special and mundane. It is special in its potential to identify and exclude with a degree of reliability that sets it apart from more traditional forms of proof in criminal investigations (eyewitness identification, confessions, informants, and the like). But it is also, like all evidence produced by humans in the crucible of the criminal justice system, susceptible to error, bias, manipulation, rationing, and other dynamics that compromise its reliability both in theory and in practice.

For at least three decades, academic observers (largely legal scholars, joined by a handful of social scientists and a smattering of commentators within the tiny community of academic forensic science) were nearly alone in grappling with this vexing duality. But with the release of the 2009 Report of the National Academy of Sciences (NAS), *Strengthening Forensic Science in the United States: A Path Forward* (the NAS Report), these critiques have been nudged from the margins into the policy mainstream. In three hundred pages, the NAS Report criticized the absence of validation for virtually every forensic methodology;[7] pointed to widespread deficiencies in funding, training, and standard setting in forensic science;[8] and laid further blame at the feet of courts for "continu[ing] to rely on forensic evidence without fully understanding and addressing the limitations of different forensic science disciplines."[9] Importantly, however, the thrust of the NAS Report was not ultimately pessimistic, but rather, as the title implies, forward-looking: its prestigious authors clearly viewed the future of criminal justice as bound up with the future of forensic science.[10] Toward that end, the NAS Report proffered thirteen recommendations for comprehensive reform of the forensic science field, which, in sum, call for broader training and standardization of laboratory work, an ambitious program for expanding research and education directed at improving forensic science, and most controversially, institutional independence of laboratories from law enforcement institutions and the formation of a new federal agency, the National Institute of Forensic Science (NIFS), charged with funding and agenda setting in the forensic sciences.[11]

The NAS Report has been widely heralded as a watershed, and its analysis and recommendations look to be setting the terms of academic and

---

police training to judicial rules of evidence and procedure) combine to enforce a type of tunnel vision, which makes it very difficult for a wrongly accused, and ultimately wrongly convicted, person to be vindicated").

7. COMM. ON IDENTIFYING THE NEEDS OF THE FORENSIC SCI. CMTY. ET AL., NAT'L RESEARCH COUNCIL OF THE NAT'L ACADS., STRENGTHENING FORENSIC SCIENCE IN THE UNITED STATES: A PATH FORWARD 127–82 (2009) [hereinafter NAS REPORT].

8. *Id.* at 183–239.

9. *Id.* at 85–110.

10. *See id.* at 4–5 (discussing the impact of advances in forensic science on criminal justice).

11. *See id.* at 19–33 (summarizing all thirteen recommendations).

policy debates concerning forensic science for the foreseeable future.[12] But academic and policy agendas tethered to the NAS Report will be deficient in a critical respect. Like the overwhelming majority of the scholarship and criticism that so heavily influenced it, the light shined by the Report is focused almost exclusively on the primary site of forensic science *production*—the laboratory—as the relevant site of reform. But this ignores a critical set of dynamics affecting forensic science: namely, the manner in which upstream *users* of forensic science—police and prosecutors, to be precise—will select priorities, initiate investigations, collect and submit evidence, choose investigative techniques, and charge and plead cases in ways that have critical and systematic, though poorly understood, influences on the accuracy of forensic analysis and the integrity of its application in criminal cases.

These dynamics have featured (albeit often below the surface) in many known and even more unknown cases in which forensic science has failed to live up to, or even frustrated, the truth-facilitating function it is deployed to serve. Consider, for example, the fact that both cases in which the Supreme Court has taken on the question of postconviction access to (putatively exonerative) DNA testing feature challenges not to *shoddy* science but rather to *incomplete* science: both William Osborne and Hank Skinner have argued that prosecutors and police in their original investigations opted not to test available and potentially exculpatory evidence and relied in convicting them on less than the best available scientific evidence.[13] Consider, similarly, the

---

12. *See, e.g.*, Mnookin et al., *supra* note 3, at 734 (outlining consensus among scholars and practitioners embracing the essence of the NAS Report recommendations). As this Article was being prepared for press, the Department of Justice and National Institute of Standards and Technology announced the launch of a National Commission on Forensic Science, thus bringing to fruition at least some version of the NAS Report's central proposal for national forensic science oversight. *See* Press Release, Nat'l Inst. of Standards & Tech., Department of Justice and National Institute of Standards and Technology Announce Launch of National Commission on Forensic Science (Feb. 15, 2013), *available at* http://www.nist.gov/oles/doj-nist-forensic-science021513.cfm. Two bills have been introduced in Congress to implement the Report's recommendations. An early bill introduced by Senator Leahy of Vermont has recently been joined by legislation directing $299 million to forensic science research and standards development. *See* Forensic Science and Standards Act of 2012, H.R. 6106, 112th Cong. (2012) [hereinafter Forensic Science Act], *available at* http://democrats.science.house.gov/sites/democrats.science.house.gov/files/documents/H.R.%206106%20-%20Forensic%20Science%20and%20Standards%20Act%20of%202012.pdf (proposing to establish a national forensic science research program); Criminal Justice and Forensic Science Reform Act of 2011, S. 132, 112th Cong. (2011) [hereinafter Criminal Justice Act], *available at* http://www.opencongress.org/bill/112-s132/show (proposing the establishment of an Office of Forensic Science and a national Forensic Science Board). Action is also afoot at the Executive level. COMM. ON SCI., NAT'L SCI. & TECH. COUNCIL, CHARTER OF THE SUBCOMMITTEE ON FORENSIC SCIENCE 1–2 [hereinafter CHARTER OF THE SUBCOMMITTEE ON FORENSIC SCIENCE], *available at* http://www.whitehouse.gov/sites/default/files/sofs_charter_2012_signed.pdf (explaining the goal of the subcommittee as "to assess the practical challenges of implementing . . . recommendations [in the NAS Report] and [to] advise the White House on how best to achieve the goals outlined in the . . . report"); *see also infra* Part III.

13. *See* Brief for the Respondent at 8–9 & n.3, Dist. Attorney's Office for the Third Judicial Dist. v. Osborne, 129 S. Ct. 2308 (2009) (No. 08-6) (discussing the State's decision to not conduct

fact that in a recent study of the first 250 DNA exonerations in the United States, analysis of investigative documents and trial transcripts revealed that at least 34 defendants were initially tried for their crimes despite the known contemporaneous existence of arguably exculpatory forensic evidence.[14]

The reform agenda of the NAS Report has little to say about the critical questions raised by these cases, which center not on laboratory-based practices, but rather on the exercise of upstream discretion by other law enforcement actors. These actors exercise a range of discretion in selecting, submitting, and utilizing scientific evidence in criminal cases, and they do so within professional, organizational, and legal contexts that create particular incentives and, at times, pathologies in regard to these tasks. Critically, not only do these dynamics at times play a dispositive role in determining the impact of science in an investigation—as illustrated by the Osborne and Skinner examples above—but they also, perhaps more commonly, play a contributory role by reacting and adjusting, perhaps unexpectedly and perhaps perversely, to the work of laboratory actors. Accounting for these dynamics thus requires a broader view—a systemic view—than is afforded by a laboratory-centric lens.

Injection of a systemic view is the aim of this Article. It proceeds as follows. Part I describes the NAS Report and identifies two premises that gird its view of forensic science and contribution to the field: (1) that more good science as early as possible in the life cycle of criminal investigations will further the goal of enhancing the substantive quality of criminal justice; and (2) that greater monetary resources and more independence for laboratories are necessary and sufficient conditions to achieve premise (1). Part I further aims to situate the NAS Report's adoption of these (implicit) premises in a particular historical, intellectual, and political context to suggest that the relative narrowness of the path forward that it charted was, though understandable and valuable in its own right, far from inevitable. Part II problematizes the premises identified in Part I by illuminating upstream dynamics driven not by laboratory-based practitioners, but rather

---

RFLP testing and to rely upon microscopic examination of hairs—an analysis which is no longer accepted as a valid basis for identification standing alone); Appellant's Opening Brief at 31–33, Skinner v. State, No. AP-76,675 (Tex. Crim. App. Feb. 2, 2012), 2012 WL 591289, *31–33 (discussing facts); *see also* Skinner v. Switzer, 131 S. Ct. 1289, 1298 (2011) (holding that a convicted state prisoner seeking DNA testing of evidence may assert that claim in a civil rights action); *Osborne*, 129 S. Ct. at 2323 (declining to recognize a freestanding constitutional right of access to DNA testing).

14. This is based on a conservative count of the publicly available data from Professor Brandon Garrett's study. *See* BRANDON L. GARRETT, CONVICTING THE INNOCENT: WHEN CRIMINAL PROSECUTIONS GO WRONG app. (2011) [hereinafter GARRETT, APPENDIX TO CONVICTING THE INNOCENT], *available at* http://www.law.virginia.edu/pdf/faculty/garrett/convicting_the_innocent/garrett_forensics_appendix.pdf (providing a table of characteristics of forensic evidence used at trial from a list of 250 later-exonerated convicts); *see also* GARRETT, *supra* note 2, 6–11 (discussing the reasons behind the wrongful convictions of the first 250 people to be exonerated by DNA evidence). Interestingly, in only three of those cases were the exculpatory results *not* disclosed to defense counsel. *See* GARRETT, APPENDIX TO CONVICTING THE INNOCENT, *supra*.

by police and prosecutors.[15] Focusing on evidence gathering, the decision to obtain forensic testing, and the investigative response to forensic analysis, the discussion aims to demonstrate that, currently, decisions committed almost entirely to the discretion of police and prosecutors are marked both by underutilization—with surprisingly little physical evidence being collected, analyzed, and relied upon in investigations—and qualitatively suboptimal utilization—marked by mutually reinforcing dynamics of late-in-time testing, less-than-thorough follow-up investigation in response, and systematic discounting of exculpatory science. Of course, these pathologies do not infect every case, and even where they do, the impact varies tremendously. But to the extent their causes are structural and systemic—as Part II argues—there is reason for concern both that laboratory-centric reform will not alone meaningfully enhance the quality of forensic science in the criminal justice system, and that dynamics of usage might frustrate or be frustrated by some of the NAS Report's recommendations. The NAS Report's recommendations to preserve and even expand the (properly circumscribed) use of non-DNA forensic techniques of individualization through comparison and to create "independent" crime laboratories are held out, somewhat speculatively, as examples of reform proposals that might well spur unintended and undesirable consequences when upstream actors are included in the mix. Ultimately, though, many of the questions raised in Part II are empirical and not answerable on the current record of knowledge. Hence, expanded and refocused research is a major area of reform embraced by this Article.

While the primary goal of the Article is diagnostic, Part III provisionally outlines some proposals for widening the NAS Report's path forward by adopting a more systemic vision for forensic science oversight and reform. In part, the aim is to capitalize on the current momentum for reform[16] by offering a series of possible complements to the existing agenda outlined by the NAS Report: the Report's research and standard-setting proposals should encompass police and prosecution practices such as evidence gathering, testing decisions, and disclosure regimes; the Report's "independence" recommendation should be reflected upon in light of the concerns raised in Part II; and policy makers should prioritize state-level oversight as a complement, or alternative, to proposed national oversight of the field. But Part III reaches beyond the NAS Report blueprint as well. An additional part of the forward-looking agenda is to address a pervasive inattention to forensic science within criminal procedure scholarship. This field has recently featured innovative work emphasizing the complex interactions among criminal justice actors, the institutions in which they operate, and the legal strictures that guide them, prompting something of an upstream turn in

15. For reasons explained below, the analysis is confined to state and local, nonfederal actors, although it will be generalizable to a certain extent. *See infra* Part II.

16. *See supra* note 12.

the literature.[17] As Part II reveals, these dynamics are quite relevant in considering the role of scientific evidence in criminal cases, and yet criminal procedure scholars, in contrast to evidence scholars, have not for the most part given sustained attention to the forensic science field.[18] Thus, Part III concludes with a preliminary sketch of how the goals of enhancing forensic science integrity challenge some of the foundations of criminal procedure doctrine in ways not previously considered, particularly by problematizing the enormous degree of discretion granted to police and prosecutors in pursuing investigations and building cases.

I. The NAS Report's View

A. *A Brief History of a Path Forward*

The NAS Report was the fruit of an unprecedented congressional charge to the National Academy of Sciences to conduct a comprehensive examination of the entire field of forensic science across all disciplines.[19] Its undertaking represented a shift from an attitude of near-total relative governmental neglect with regard to the forensic sciences, despite the steadily increasing centrality of scientific evidence to the criminal justice system over several decades.[20] It also, relatedly, represented a concerted challenge (the magnitude of which remains to be seen) to a tradition of top-to-bottom law enforcement control of the forensic sciences. This Part examines these dynamics in turn.

---

17. Work exemplifying this trend includes Darryl K. Brown, *The Decline of Defense Counsel and the Rise of Accuracy in Criminal Adjudication*, 93 CALIF. L. REV. 1585 (2005), Gerard E. Lynch, *Our Administrative System of Criminal Justice*, 66 FORDHAM L. REV. 2117 (1998), Daniel Richman, *Prosecutors and Their Agents, Agents and Their Prosecutors*, 103 COLUM. L. REV. 749 (2003), and William J. Stuntz, *The Uneasy Relationship Between Criminal Procedure and Criminal Justice*, 107 YALE L.J. 1 (1997).

18. *See, e.g.*, Kent Roach, *Wrongful Convictions and Criminal Procedure*, 42 BRANDEIS L.J. 349, 363–64 (2004) (observing that "rigid compartmentalization" in legal training segregates the field of evidence law, where forensic science naturally sits, and typically holds questions concerning criminal procedure and criminal justice at bay); Andrea Roth, *Book Review*, 62 J. LEGAL EDUC. 377, 383–85 (2012) (reviewing WILLIAM J. STUNTZ, THE COLLAPSE OF AMERICAN CRIMINAL JUSTICE (2011)) (observing that one of the leading criminal procedure and criminal justice theorists of our time, the late William Stuntz, largely ignored scientific evidence in his work).

19. *See* H.R. REP. NO. 109-272, at 121 (2005) (Conf. Rep.) (charging the National Academy of Sciences "to conduct a study . . . as described in the Senate report"); S. REP. NO. 109-88, at 46 (2005) (listing eight charges).

20. *See, e.g.*, RICHARD SAFERSTEIN, CRIMINALISTICS: AN INTRODUCTION TO FORENSIC SCIENCE 6–8 (7th ed. 2001) (reviewing history and recounting factors); Edward J. Imwinkelried, *A New Era in the Evolution of Scientific Evidence—A Primer on Evaluating the Weight of Scientific Evidence*, 23 WM. & MARY L. REV. 261, 261–62 (1981) (reviewing history); Joseph L. Peterson & Anna S. Leggett, *The Evolution of Forensic Science: Progress Amid the Pitfalls*, 36 STETSON L. REV. 621, 623–40 (2007) (same).

1. A History of Neglect.—President Lyndon Johnson's Commission on Law Enforcement and Administration of Justice made the development of scientific and technological capacity in the criminal justice system a federal priority, and hence the 1970s saw the first federal grant money to develop scientific capacity for state and local law enforcement agencies.[21] But expanding the production of and demand for scientific evidence only further pressed on the existing lack of institutionalized commitment to funding or overseeing the quality of scientific evidence within law enforcement organizations.[22] Oversight of crime laboratories eventually developed, but only from within the field, primarily via the wholly voluntary accreditation program developed by the American Society of Crime Lab Directors Laboratory Accreditation Board (ASCLD/LAB).[23] Many forensic subfields eventually promulgated standards of practice through Scientific Working Groups (SWGs) comprised of practitioners of specific forensic science techniques.[24] But protocols developed through those efforts were typically little more than nonbinding guidelines, rarely institutionalized as policies within laboratories and often rooted in little more than the accumulated

---

21. *See* PRESIDENT'S COMM'N ON LAW ENFORCEMENT & ADMIN. OF JUSTICE, THE CHALLENGE OF CRIME IN A FREE SOCIETY 270–71 (1967) (recommending greater research and capacity building in forensic science).

22. *See, e.g.*, NAT'L ADVISORY COMM'N ON CRIMINAL JUSTICE STANDARDS & GOALS, REPORT ON POLICE 304–05 (1973) (observing that "[t]oo many police crime laboratories have been set up on budgets that preclude the recruitment of qualified, professional personnel," and that "[t]oo often the laboratory is not considered a primary budget item and is one of the first units to suffer when budgets are trimmed"); Peterson & Leggett, *supra* note 20, at 623–25 (noting that "[w]hile the growth [of crime-laboratory services] was necessary, it was unregulated and without clear guidance from, or adherence to, national standards. Thus, . . . some of the underlying problems of quality assurance and minimum scientific standards simply multiplied.").

23. *See About ASCLD/LAB*, AM. SOC'Y OF CRIME LABORATORY DIRECTORS/LABORATORY ACCREDITATION BOARD, http://www.ascld-lab.org/about_us/aboutoverview.html ("ASCLD/LAB has been accrediting crime laboratories since 1982 and currently accredits most of the federal, state and local crime laboratories in the United States plus forensic laboratories in six countries outside of the United States."). Even today only nine states require that crime laboratories be accredited, although approximately 83% of publicly funded laboratories have pursued accreditation nonetheless. MATTHEW R. DUROSE ET AL., BUREAU OF JUSTICE STATISTICS, U.S. DEP'T OF JUSTICE, NCJ 238252, CENSUS OF PUBLICLY FUNDED FORENSIC CRIME LABORATORIES, 2009, at 7 (2012), *available at* http://www.bjs.gov/content/pub/pdf/cpffcl09.pdf; NAT'L CONFERENCE OF STATE LEGISLATURES, DNA LAWS DATABASE tbl.6 (2010), *available at* http://www.ncsl.org/portals/1/Documents/cj/Table6AccredLaboratories.pdf; *see also* Donald Kennedy, *Forensic Science: Oxymoron?*, 302 SCIENCE 1625, 1625 (2003) (asserting that "despite repeated calls for accreditation and oversight, many government crime labs continue to lack either one"). While ASCLD/LAB has filled a much-needed oversight gap and occasionally placed itself on the forefront of investigating scientific fraud, it has nevertheless been widely criticized as far too insular and uncritical of the forensic science field to perform meaningful oversight. *See* Ryan M. Goldstein, Note, *Improving Forensic Science Through State Oversight*, 90 TEXAS L. REV. 225, 238–39 & n.114 (2011) (reviewing some of the criticisms).

24. *See* Mnookin et al., *supra* note 3, at 771–72 (describing how SWGs are formed and what they ought to look like); *see also About Scientific Working Groups*, SWGFAST, http://www.swgfast.org/AboutSWGs.htm (last modified Feb. 2011) ("Since the early 1990s, American and International forensic science laboratories and practitioners have collaborated in Scientific Working Groups . . . to improve discipline practices and build consensus standards.").

wisdom of practitioners.[25] Indeed, this reality reflected the fact that forensic science methodologies were themselves an incredible grab bag from the standpoint of their theoretical or experimental basis and reliability of results. Some, such as blood analysis, possessed a set of theories and techniques rooted in nonforensic scientific research; others, such as fingerprint comparison and a range of other comparison techniques such as hair microscopy or firearms analysis, were guided by an experientially generated set of protocols, developed and passed down wholly in the context of law enforcement applications.[26]

Development of forensic DNA applications in the 1980s was thus a watershed not simply from the standpoint of what the technology itself offered, but more systemically because of the scientific paradigm that it injected into the field.[27] It is well beyond the scope of this Article (and a feat already accomplished by others) to offer a detailed account of the path from DNA's vanguard appearance in American courtrooms in the mid-1980s to its transformation of criminal investigations and prosecutions and ascendance to the status of "gold standard" for scientific evidence.[28] But that journey, along a trail marked by a series of pitched legal, scientific, and political

---

25. *See, e.g.*, GARRETT, *supra* note 2, at 105–06 (describing the scientifically invalid recommendation of the Scientific Working Group on Shoeprint and Tire Tread Evidence that the shoe print comparison method allows "definite conclusion of identity"); Harry T. Edwards, *The National Academy of Sciences Report on Forensic Sciences: What It Means for the Bench and Bar*, 51 JURIMETRICS 1, 11 (2010) (asserting that SWGs were of "questionable value" due, among other reasons, to lack of enforcement, lack of empirical measurement of effectiveness, and excessive vagueness); Mnookin et al., *supra* note 3, at 772–73 (describing how SWGs "despite the scientific label in the name . . . have a rather tenuous relationship with research science," and concluding that while "practitioner-led SWGs may often reach appropriate, thoughtful, and perhaps even research-based conclusions, . . . they also risk being guided by and influenced by populist practitioner pressures"); *see also* State v. Dominguez, No. 01-10-00428-CR, 2011 WL 3207766, at *5–8 (Tex. App.—Houston [1st Dist.] July 28, 2011, pet. ref'd.) (not designated for publication) (recounting testimony in a capital case of the founding director of the Scientific Working Group on Dog Scent and Orthogonal Detector Guidelines, who identified departures from dog scent lineup best practices but contended that the procedure in the case was still reliable).

26. *See, e.g.*, Joseph L. Peterson & Penelope N. Markham, *Crime Laboratory Proficiency Testing Results, 1978–1991, I: Identification and Classification of Physical Evidence*, 40 J. FORENSIC SCI. 994, 994, 1007 (1995) (demonstrating wide variability in accuracy, shown through proficiency testing, among classification-oriented forensic techniques, and summarizing earlier studies with consistent results); Joseph L. Peterson & Penelope N. Markham, *Crime Laboratory Proficiency Testing Results, 1978–1991, II: Resolving Questions of Common Origin*, 40 J. FORENSIC SCI. 1009, 1027–28 (1995) (same).

27. *See, e.g.*, Erin Murphy, *What 'Strengthening Forensic Science' Today Means for Tomorrow: DNA Exceptionalism and the 2009 NAS Report*, 9 LAW, PROBABILITY & RISK 7, 9–15 (2010) (summarizing the history and contrasting the scrutiny of DNA to the inattention vis-à-vis other forensic sciences).

28. *See, e.g.*, Peterson & Leggett, *supra* note 20, at 654 (recounting the history and attainment of the "gold standard" label); Erin Murphy, *The New Forensics: Criminal Justice, False Certainty, and the Second Generation of Scientific Evidence*, 95 CALIF. L. REV. 721, 731–34 (2007) (discussing the history of forensic DNA and its role in upending traditional techniques of criminal investigation).

battles that have come to be known as "the DNA wars,"[29] was marked by three features of particular relevance to the present account.

First, in contrast to all forensic science techniques that preceded DNA analysis—what Erin Murphy helpfully categorizes as "first-generation" forensic science—DNA came to the criminal justice system as a tool developed by research scientists and used in a range of nonforensic applications—features that did not guarantee reliability, but stood in contrast to the law enforcement and otherwise nonscience origins of many precursor techniques.[30] Second, forensic DNA analysis retained much of the technical and theoretical standardization that characterized its home base of research science when it entered the criminal justice system, and also was met by a level of judicial, political, and ultimately regulatory scrutiny that immediately and significantly outpaced that which had attached to first-generation forensic science.[31] The DNA Identification Act of 1994 (which established, among other things, a national system of DNA databases and a national DNA Advisory Board) as well as regular and substantial congressional funding of the field are only two of many examples of the exceptional level of formal, external regulation of forensic DNA.[32]

But these features prompted a third dynamic that presaged a theme that resurfaced in the 2009 NAS Report: that the hotly contested process of erecting a legal infrastructure for oversight of forensic DNA was closely bound up with questions of the proper demarcation of the law–science divide in the forensic sciences more broadly. While a community of researchers and defense advocates aimed to situate forensic DNA and oversight of it squarely within the province of the "pure" scientific field that gave birth to it, another powerful bloc, led by the FBI, contended that *forensic* DNA was a unique field amenable to regulation only from within, i.e., from within the

---

29. *See, e.g.*, JAY D. ARONSON, GENETIC WITNESS: SCIENCE, LAW, AND CONTROVERSY IN THE MAKING OF DNA PROFILING 146–72 (2007) (detailing the series of battles known as "the DNA wars"); Simon A. Cole, *Fingerprint Identification and the Criminal Justice System: Historical Lessons for the DNA Debate*, *in* DNA AND THE CRIMINAL JUSTICE SYSTEM: THE TECHNOLOGY OF JUSTICE 63, 63–90 (David Lazer ed., 2004) (detailing lessons learned from the DNA debate).

30. Murphy, *supra* note 28, at 726–31.

31. *See, e.g.*, Murphy, *supra* note 27, at 9–11 (summarizing the history and noting the implementation of regulation and standardization in forensic science).

32. DNA Identification Act of 1994, Pub. L. No. 103-322, tit. XXI, subtit. C, 108 Stat. 2065 (codified as amended at 42 U.S.C. §§ 14131–14134 (2006)); *see also* DNA Analysis Backlog Elimination Act of 2000, Pub. L. No. 106-546, 114 Stat. 2726 (codified as amended in scattered sections of 10, 18, 28 & 42 U.S.C.) (allowing states to carry out DNA analyses for use in CODIS and to collect and analyze DNA samples); Crime Identification Technology Act of 1998, Pub. L. No. 105-251, tit. I, 112 Stat. 1870 (codified as amended at 42 U.S.C. § 14601 (2006)) (authorizing the federal government to provide states with information in order to establish an integrated approach to develop information and identification technologies and systems to combat crime). The federal Coverdell grant program is a rare example of a more broadly available funding source for forensic science. Science, State, Justice, Commerce, and Related Agencies Appropriations Act, 2006, Pub. L. No. 109-108, tit. I, 119 Stat. 2290, 2302 (allocating funds for Paul Coverdell Forensic Science Improvement Grants).

law-enforcement-based forensic science community.[33] A tug and pull between these perspectives shaped the earliest congressional foray into DNA oversight: the establishment of the national system of DNA databases and of a DNA Advisory Board comprised in part of independent scientists appointed with the approval of the FBI Director.[34] It also shaped early efforts by the National Academy's National Research Council (the same body responsible for the 2009 NAS Report) to develop standards for DNA typing—efforts that prompted the drawing of battle lines between the "forensic DNA as 'pure' science" and "forensic DNA as 'forensic' science" camps, and that ultimately by many accounts featured significant technical and regulatory concessions to the second camp.[35] The point for present purposes is not a scientific one but a political one: efforts at external, governmental involvement in the field of forensic science has historically been bound up with (and by) questions of law enforcement discretion and territorialism. Indeed, efforts to launch the sort of broad-based review eventually undertaken by the NAS in 2009 were stymied for years by infighting among law enforcement and defense leaders, forensic practitioners, and policy makers about whether a congressionally authorized assessment of the field should be undertaken by the National Institute of Justice (NIJ), an arm of the Department of Justice (DOJ), or by a scientific authority like the National Academy.[36]

In any event, governmental interest in, and funding and regulation of, DNA did not spill over to other forensic science techniques, despite the fact that these non-DNA applications dominated (as they continue to dominate) the caseloads of crime laboratories.[37] Indeed, the greatest impact of forensic DNA was undoubtedly to highlight the *absence* of oversight with regard to most other forensic techniques, as well as the enormous and increasingly glaring gap between the promise of scientific reliability and the increasingly exposed reality of forensic practice.[38] Reports of laboratory breakdowns

---

33. *See, e.g.*, ARONSON, *supra* note 29, at 147 (arguing that the FBI "sought to maintain the apparent boundary between scientific and legal issues").

34. *Id.* at 152–53.

35. *See id.* at 153–72 (discussing the NRC reports and the controversy surrounding population genetics); *see also* Richard Lempert, *After the DNA Wars: Skirmishing with NRC II*, 37 JURIMETRICS 439, 451–53, 465–66 & nn.63–64 (1997) (discussing the history of the reports and outlining their differing recommendations regarding calculation of match probability); William C. Thompson, *Accepting Lower Standards: The National Research Council's Second Report on Forensic DNA Evidence*, 37 JURIMETRICS 405, 406–07 (1997) (characterizing the second report as "inadequate and less rigorous" than the first).

36. Murphy, *supra* note 27, at 13–14 & n.50.

37. *See* DUROSE ET AL., *supra* note 23, at 4 tbl.5 (reporting that in 2009, 66% of requests for analysis to laboratories were for nonbiological evidence); NAS REPORT, *supra* note 7, at 41 (stating that DNA analysis comprises only about 10% of laboratory caseloads).

38. *See* ARONSON, *supra* note 29, at 3 ("As the history of DNA profiling demonstrates, [forensic] technologies have limitations that only become apparent when they are applied in practice and are challenged by people who have a vested interest in pointing out their shortcomings."); Michael J. Saks & Jonathan J. Koehler, *What DNA "Fingerprinting" Can Teach the Law About the*

from Houston to North Carolina to Detroit pointed to a litany of similar structural roots: a history of competition with other police divisions for limited resources; failure to hire, train, and retain qualified analysts; and caseload pressures that exacerbated other organizational deficiencies to further cause slipshod work and enhance analysts' vulnerability to pressure from police and prosecutors.[39] Indeed, these were circumstances that were compromising the integrity of results even in the new "gold standard."[40]

But if similar themes were emerging from the accelerating reports of scandal and mismanagement, the particular institutional contexts from which these themes were distilled varied enormously. Just as forensic science encompasses a wide array of disciplines, so too do the institutional settings in which forensic science is produced exhibit significant variation. Organizations approximating the CSI-fueled public imagination of a freestanding crime laboratory, replete with microscopes and staffed by dedicated forensic analysts, are where much of forensic science is produced; but a significant number of forensic science tasks are actually traditionally and still conducted within police organizations in nonlaboratory settings—including, depending upon the particular jurisdiction, crime scene work and fingerprint and other comparative and interpretative analyses (e.g., blood spatter, photographic analysis, handwriting comparison).[41] Among the former category of crime laboratories, the vast majority are under the control of a law enforcement entity at some level of government, some are publicly run yet institutionally independent, and a significant amount of forensic analysis is done as well by private, for-profit labs.[42] Among public laboratories, jurisdiction might encompass an entire state, a particular region, a specific local jurisdiction, or as in the case of the FBI laboratory, a hybrid arrangement with a primary customer (federal law enforcement) as well as a

---

Rest of Forensic Science, 13 CARDOZO L. REV. 361, 372 (1991) (noting that the buzz surrounding DNA evidence "should not obscure the fact that most forensic sciences, including DNA typing, rely on assumptions that have not yet been verified by empirical testing").

39. *See generally* MICHAEL R. BROMWICH, FINAL REPORT OF THE INDEPENDENT INVESTIGATOR FOR THE HOUSTON POLICE DEPARTMENT CRIME LABORATORY AND PROPERTY ROOM (2007), *available at* http://www.hpdlabinvestigation.org/reports/070613report.pdf (describing the Houston crime lab's failures in DNA, serology, and other disciplines); Mandy Locke et al., *Scathing SBI Audit Says 230 Cases Tainted by Shoddy Investigations*, NEWS & OBSERVER, Aug. 27, 2010, http://www.newsobserver.com/2010/08/19/635632/scathing-sbi-audit-says-230-cases.html (describing widespread failures at the North Carolina state crime lab); Ben Schmitt & Joe Swickard, *Troubled Detroit Police Crime Lab Shuttered: State Police Audit Results 'Appalling,' Wayne County Prosecutor Declares*, DETROIT FREE PRESS, Sept. 26, 2008, http://truthinjustice.org/detroit-lab.htm (discussing ballistics analysis errors and widespread lab failures).

40. *See, e.g.*, BROMWICH, *supra* note 39, at 116–50 (detailing numerous problems with the Houston crime lab's DNA testing).

41. *See* NAS REPORT, *supra* note 7, at 38–39 (discussing how some forensic disciplines are conducted by scientists while others are conducted by nonscientists); Randall A. Childs et al., *Survey of Forensic Science Providers*, 1 FORENSIC SCI. POL'Y & MGMT. 49, 54–56 & tbl.9 (2009) (examining forensic functions performed by city police and sheriff departments).

42. *See* NAS REPORT, *supra* note 7, at 57–64 (discussing various configurations of crime laboratories).

mandate to perform casework on request from other jurisdictions.[43] This tremendous fragmentation in the field creates both an oversight vacuum and an oversight challenge, particularly from the standpoint of centralized (read, federal governmental) authorities.[44]

In sum, Congress's call to the NAS in 2006 was an important and long-overdue measure to address a near-total failure within policy circles to address what the accompanying Senate Report described as an "absence of data"—or what might be more accurately described as a preliminary diagnosis of failure—with respect to the national infrastructure for the production of forensic science.[45] Moreover, it presented the NAS Report drafters with a task of immense factual and political complexity. In ultimately committing the project to the National Academy of Sciences, widely viewed as the government's preeminent scientific authority, Congress for its part appeared to call for a striking and deliberate turn toward recapturing the *scientific* bona fides of the field. But the history of DNA oversight was no doubt in the minds of many participants in the review process. And, notwithstanding a general consensus that the forensic sciences required, at least, greater funding and infrastructural improvement, by undertaking a broad-based review of forensic disciplines that had long been wholly the province of law enforcement, the NAS's effort was likely to engender even more controversy.[46] In any event, Congress's precommitment that the NAS Report would be an endeavor generated quite truly by the "scientific" "academy" *and not* law enforcement or even the broader criminal justice community did not relieve the Report's drafters of the heavy burden of history and criminal justice politics.[47]

2. The "Independence" Question.—The notion of scientific and law enforcement entanglement, lurking as it did somewhat below the surface of

---

43. *See id.*

44. *See id.* at 77 (discussing the challenges of fragmentation).

45. S. REP. NO. 109-88, at 46 (2005) ("The results of these studies are indicative of a larger problem within the forensic science and legal community: the absence of data. While a great deal of analysis exists of the requirements in the discipline of DNA, there exists little to no analysis of the remaining needs of the community outside of the area of DNA.").

46. *Cf.* Murphy, *supra* note 27, at 16 (arguing that the availability of DNA typing helped foster a political push for the NAS Report).

47. *See* NAS REPORT, *supra* note 7, at 13–14 (acknowledging the delicate political balance between federal and state programs within the forensic science community); Paul C. Giannelli, Daubert *and Forensic Science: The Pitfalls of Law Enforcement Control of Scientific Research*, 2011 U. ILL. L. REV. 53, 88 & nn.238–40 (describing evidence of political battles between the DOJ/NIJ and the National Academy of Sciences over the authority to conduct forensic science research); Solomon Moore, *Science Found Wanting in Nation's Crime Labs*, N.Y. TIMES, Feb. 4, 2009, http://www.nytimes.com/2009/02/05/us/05forensics.html?_r=1&pagewanted=all (noting the storm brewing amongst federal law enforcement agencies in anticipation of the NAS Report); *see also* William J. Stuntz, *The Pathological Politics of Criminal Law*, 100 MICH. L. REV. 505, 523–57 (2001) (detailing the institutional dynamics and political incentives affecting the balance of power between prosecutors and legislatures).

the quiet seas of governmental (non)involvement in forensic science, loomed large in the increasingly pitched academic criticisms of the forensic science field—including those advanced by prominent scholars whose perspectives would emerge prominently in the NAS Report. An overriding concern in the literature was the lack of institutional independence among the overwhelming majority of crime laboratories. Organizationally, crime laboratories are by and large under the administrative control of police and prosecutorial agencies.[48] But critics have noted that the entanglement of science and law enforcement (or, to some, dominance of the latter over the former) is more foundational as well.

The attention and funds of the federal government brought to bear on forensic science—DNA and non-DNA methodologies alike—have emanated largely from the priorities and professional infrastructure of law enforcement, and, putting aside DNA testing and toxicology analysis, were divorced from both the theoretical underpinnings and the professional culture of the sciences, a feature that critics bemoaned as suppressing scientific rigor or, worse, infecting science-based claims in the field with intolerable bias.[49] Overwhelmingly, the empirical bases for the expert claims made by forensic practitioners—assertions, for example, of "matches" between fingerprints or between bullets and weapons or between shed hairs—lay in the accumulated observation of examiners rather than in a scientific theory developed and tested in a neutral laboratory setting.[50] To the extent that data exists to substantiate the scientific claims made by forensic experts, government agencies in possession of that data have been reluctant to open it to scrutiny, complicating meaningful efforts at verification and undermining the scientific hallmarks of collaboration and replicability.[51] And to the extent that the federal government sponsors independent research in forensic science, priorities are generally set and funds disseminated by the Department of Justice's research arm, the National Institute of Justice, whose grant administration has been heavily criticized for its lack of independence from the DOJ's own law enforcement agenda.[52] The SWGs that had taken

---

48. *See, e.g.*, Giannelli, *supra* note 4, at 469–76 (proposing a solution to the problems stemming from crime laboratories' lack of institutional independence from police departments).

49. *See* Giannelli, *supra* note 47, at 88–90 (advocating for more rigorous research conducted by scientists independent from law enforcement); Mnookin et al., *supra* note 3, at 765–67 (describing how situating forensic laboratories within law enforcement agencies created a "divide between research values and forensic practice").

50. *See, e.g.*, Mnookin et al., *supra* note 3, at 737–40 (stating that there is insufficient scientific research to empirically ascertain the accuracy of various forensic techniques).

51. *See* Roger Koppl, *How to Improve Forensic Science*, 20 EUR. J.L. & ECON. 255, 257 (2005) (discussing several features of the organization of forensic science that lower the quality of scientists' work); Saks & Faigman, *supra* note 3, at 152–53 (characterizing the approach of law enforcement toward novel forensic technologies as "bureaucratic" and empirically unsound).

52. *See, e.g.*, Craig M. Cooley & Gabriel S. Oberfield, *Increasing Forensic Evidence's Reliability and Minimizing Wrongful Convictions: Applying* Daubert *Isn't the Only Problem*, 43

the lead in developing what standards of practice did exist for non-DNA methodologies were funded by the Department of Justice and administered under the auspices of the FBI, NIJ, and Drug Enforcement Agency.[53]

Thus, a majority (perhaps even consensus)[54] view of academic observers was that the forensic science field was grossly underperforming in relation to its claims of scientific validity and reliability, and that a crucial causal factor in that underperformance was the field's lack of structural, professional, indeed epistemological independence from its law enforcement consumers. These scholarly views would come to be enormously influential in the NAS Report's conclusions, as many of the leading proponents of this view either served on or testified before the blue ribbon panel charged with creating the NAS Report.[55] But if professional identity has played a formative role in the creation of knowledge within the forensic sciences, so too has it shaped critique of the field. Significantly (and unsurprisingly), the overwhelming majority of legal academic commentary in regard to forensic science has emerged from the realm of evidence scholarship, a fact that has focused critiques on the nature of forensic evidence itself and courtroom regulation of it. By contrast (and perhaps more surprisingly), criminal procedure scholars, whose gaze more commonly extends outside the courtroom to the work of police and prosecutors, have given little sustained attention to forensic science.[56] However valuable the academic contributions in this area have been—all the more so for standing as a rare example of dramatic scholarly synergy with real-world reform—the disciplinary orientation of the dominant view has left unilluminated existing, and pressing, concerns that emerge when the forensic sciences are viewed as one facet of the broader criminal justice system.

B. *The NAS Report's (More and Less Implicit) Premises*

Armed with an important but challenging mandate, the NAS formed a blue ribbon Committee on Identifying the Needs of the Forensic Science Community to conduct the study. Its congressional charge specified that the Committee was to be "independent" and broadly representative of interests

---

TULSA L. REV. 285, 375 (2007) (describing allegations of NIJ ineffectiveness in grant oversight); Giannelli, *supra* note 47, at 88 (detailing reports of NIJ bias).

53. *See* Mnookin et al., *supra* note 3, at 771–73 (noting that many forensic practice standards were developed by SWGs that are funded by the DOJ and operate under the auspices of the FBI laboratory); *Scientific Working Groups*, NAT'L INST. JUST., http://www.nij.gov/topics/forensics/lab-operations/standards/scientific-working-groups.htm (last modified July 25, 2012) (listing funding sources for various SWGs).

54. *See generally* Mnookin et al., *supra* note 3 (presenting these critiques as the consensus view).

55. *See* NAS REPORT, *supra* note 7, at v, xi–xii (listing evidence scholar Margaret Berger as a member of the Committee, and noting testimony by Professors Giannelli, Kaye, Mnookin, Risinger, and Saks).

56. *See supra* note 18 and accompanying text.

within the "forensics community"[57]: the sixteen-member Committee included leading research scientists, forensic scientists, a criminal defense attorney, a former federal prosecutor, legal academics, and a federal judge.[58] No non-forensic-practitioner police officers (or sitting prosecutors), however, were among its ranks—a fact not lost on those constituencies.[59] Over the course of two years, the Committee heard hundreds of hours of testimony from stakeholders in the field—practicing forensic scientists, social scientists, academics, prosecutors and defense attorneys, and federal, state, and local law enforcement officials[60]—and "engaged in independent research."[61]

The factual findings of the Committee contained in its three-hundred-page report largely confirm the most pessimistic accounts of the forensic sciences that have been circulating in academic and, to some extent, professional quarters.[62] Underfunding, unvalidated methodologies, and untrained and unregulated analysts are described as, if not the norm, a disturbingly pervasive reality among the nearly 400 public crime laboratories in the United States.[63] Techniques long relied on by law enforcement and accepted by courts—pattern identification analysis ranging from fingerprints to shoe prints to ballistics, hair and fiber analysis, and questioned document analysis, among other disciplines—are called out as never having been systematically and scientifically validated.[64] The Report's thirteen recommendations for reform are no less sweeping. It calls for large-scale reform of crime laboratory practices, from development of standardized protocols for conducting and documenting particular forensic tests to development of a national code of ethics and greater proficiency oversight.[65] It outlines an ambitious agenda for expanding research and education directed at improving laboratory practice at its foundation, calling for

57. S. REP. NO. 109-88, at 46 (2005) (directing the formation of "an independent Forensic Science Committee" to include "members . . . representing operational crime laboratories, medical examiners, and coroners; legal experts; and other scientists as determined appropriate").

58. NAS REPORT, *supra* note 7, app. A at 287–302 (listing members and biographies).

59. *See id.* (chronicling the biographical information of each committee member); *see also Strengthening Forensic Science in the United States: Hearing Before the S. Comm. on the Judiciary*, 111th Cong. 15 (2009) (statement of Barry Matson, Deputy Director, Alabama District Attorneys Association) (responding to the NAS Report and noting that "[t]he absence of prosecutors on the National Academy of Sciences Committee on Forensic Sciences has not been lost on those of us serving every day in the trenches of America's courtrooms"); Meredith Mays, *Forensic Science Reform Continues*, POLICE CHIEF (Nov. 2009), http://www.policechiefmagazine.org/magazine/index.cfm?fuseaction=display&article_id=1938&issue_id=112009 (noting the objection of the International Association of Chiefs of Police that "the report was developed without input from law enforcement practitioners and recommend[ing] their input be sought").

60. NAS REPORT, *supra* note 7, app. B at 303–14 (listing committee meeting agendas).

61. *Id.* at 2.

62. *See supra* notes 3–6 and accompanying text (discussing deficiencies in the forensic sciences).

63. *See* DUROSE ET AL., *supra* note 23, at 1–10 (describing significant backlogs and gaps in accreditation, funding, and training at federal, state, and municipal crime labs).

64. NAS REPORT, *supra* note 7, at 127–82.

65. *Id.* at 22–26.

independent research into the "accuracy, reliability, and validity in the forensic science disciplines" as well as "observer bias" on the part of analysts;[66] development of graduate education programs in the forensic sciences;[67] encouragement of continuing legal education in forensic science for law students, practitioners, and judges;[68] and enhancement of two specific applications of forensic science: fingerprint and AFIS[69] and homeland security.[70] Finally, two specific proposals are advanced as foundational to all of the above: establishing a new federal agency, the National Institute of Forensic Science, charged with funding and agenda setting in the forensic sciences, including carrying out the Report's laboratory and research recommendations, and restructuring crime laboratories to make them "independent" from law enforcement organizations.[71]

While the recommendations themselves are detailed and the factual findings animating them clear, the path forward sketched by the Report is premised on several far more implicit assumptions concerning the relationship of forensic science to the criminal justice system more generally. Two are key for present purposes: (1) that more good science will enhance the quality of criminal justice; and (2) that a better-resourced and more independent forensic community will achieve premise (1). Part II will demonstrate that a fuller accounting for the use of forensic science in the criminal justice system complicates the coherence of these premises. First, however, it is worth reflecting on the presence and significance of these assumptions embedded within the NAS Report.

1. The Science–Justice Link.—The NAS Report paints a discomforting picture of the current state of forensic science.[72] Against this bleak backdrop, it is striking that the tenor and prescriptive thrust of the NAS Report is a commitment to promote the *expansion* of forensic science and a belief that more forensic science, so long as it is *good*, will advance the ends of justice in the criminal law.[73] The view is of a field failing to live up to its potential, and the solution for this failure is to address the multiple facets of "underresourc[ing]" in forensic science, not to reconsider the primacy placed on scientific evidence in criminal investigations and prosecutions.[74] In this

66. *Id.* at 22, 24.

67. *Id.* at 27–28.

68. *Id.* at 28.

69. *Id.* at 31–32.

70. *Id.* at 32–33.

71. *Id.* at 19–21.

72. *See supra* notes 57–71 and accompanying text.

73. *See* NAS REPORT, *supra* note 7, at 20 (lauding the benefits that will obtain from strengthening forensic science).

74. *See id.* at 14–15 ("Being underresourced also means that the tools of forensic science . . . are not as strong as they could be, thus hindering the ability of the forensic science disciplines to excel at informing investigations, providing strong evidence, and avoiding errors in important ways.").

regard, the NAS Report echoes concerns expressed both by law enforcement and defense-minded communities that "the investigative capabilities of forensic science are not being realized,"[75] inhibiting both detection and correction in criminal cases.

The NAS Report's premise that maximizing science will in turn maximize the reliability of outcomes in the criminal justice system might be problematized from at least three angles.[76] First, while the notion that forensic science holds the dual promise of convicting the guilty and freeing the innocent has become something of a requisite incantation in the literature, it is not without important critics. Scholars of science, including several who have been engaged in the forensic science debates, have problematized faith in science, particularly as a tool of inclusion rather than exclusion in the criminal justice system. Sheila Jasanoff has cautioned that the always imperfect process of scientific production is bound to be systematically less reliable when wielded in the service of catching and convicting criminals (a goal already girded by motivational and cognitive biases) rather than exonerating them.[77] More pointedly, Professor Jasanoff has observed, "The issue . . . is not whether DNA profiling can in theory provide unambiguous proofs of identity, but whether society is capable of generating DNA

75. Kevin J. Strom & Matthew J. Hickman, *Unanalyzed Evidence in Law-Enforcement Agencies: A National Examination of Forensic Processing in Police Departments*, 9 CRIMINOLOGY & PUB. POL'Y 381, 381, 391–93 (2010); *see also* CAL. TASK FORCE ON FORENSIC SERVS., FORCE REPORT 68 (2003), *available at* oag.ca.gov/sites/all/files/pdfs/publications/bfs_bookmarks.pdf ("Unfortunately, the rules by which investigators currently prioritize cases and evidence for examination by resource-constrained laboratories mitigate against the use of [new forensic tools to identify suspects] for cases that are not the very most serious or highest profile."); JEREMIAH GOULKA ET AL., RAND, TOWARD A COMPARISON OF DNA PROFILING AND DATABASES IN THE UNITED STATES AND ENGLAND 1 (2010), *available at* http://www.rand.org/content/dam/rand/pubs/technical_reports/2010/RAND_TR918.pdf (noting the U.S. law enforcement community's envy of the United Kingdom's broader and more accelerated use of forensic DNA in investigations); KEVIN STROM ET AL., THE 2007 SURVEY OF LAW ENFORCEMENT FORENSIC EVIDENCE PROCESSING: FINAL REPORT 4-4 (2009), *available at* https://www.ncjrs.gov/pdffiles1/nij/grants/228415.pdf ("[S]ome U.S. law enforcement agencies continue to have only a limited understanding of the full benefits of forensic evidence and a mindset that forensic evidence is beneficial mainly for prosecuting crimes, not for developing new leads in investigations."); Nancy Petro, *Early DNA Testing Could Prevent Nightmare of Wrongful Charges*, WRONGFUL CONVICTIONS BLOG (Aug. 20, 2012), http://wrongfulconvictionsblog.org/2012/08/20/early-dna-testing-could-prevent-nightmare-of-wrongful-charges/ (suggesting that DNA testing should be utilized as quickly as possible to protect wrongfully charged innocents).

76. This Article puts to the side concerns about the privacy and other civil liberties consequences of expanding law enforcement use of scientific evidence and the question of whether losses on those scores are outweighed by or even commensurate to gains in accuracy of criminal justice outcomes, taking up only an internal critique of the link between scientific production and criminal justice accuracy. For one recent example of a comprehensive treatment of the civil liberties debate, see SHELDON KRIMSKY & TANIA SIMONCELLI, GENETIC JUSTICE: DNA DATA BANKS, CRIMINAL INVESTIGATIONS, AND CIVIL LIBERTIES 225–320 (2011) (exploring privacy, civil liberties, and civil rights concerns from expanded exploitation of DNA).

77. Sheila Jasanoff, *Just Evidence: The Limits of Science in the Legal Process*, 34 J.L. MED. & ETHICS 328, 337 (2006).

evidence that is free from bias and error."[78] Simon Cole and Jennifer Mnookin, among others, have made similar points in regard to fingerprint comparison.[79] All three scholars likely had in mind the issue of whether a forensic analyst could validly and reliably make comparisons among, as it were, profiles or prints, as an issue distinct from whether profiles or prints were in theory individualizable. But the same question could be asked of whether other members of society—say investigators or prosecutors—inject bias or error into the process of obtaining or using scientific evidence. Indeed, Professors Cole and Lynch have made this point in challenging the widely presumed objectivity and reliability of DNA and other "second-generation" forensic technologies that are now trusted by law enforcement (and the NAS Report) to not only confirm but also develop suspects in the first instance, noting the crucial role of police and prosecutors in deploying science in a case.[80] Thus, the notion that more good science leads fairly inexorably to more good criminal justice outcomes disregards or at least downplays the *dynamic* features of the relationships among scientific production, scientific usage, and the attainment of epistemological goals in the criminal justice system.

Second, the NAS Report's confidence in the future of forensic science is methodologically agnostic and thereby rejects any notion that DNA and other second-generation forensics would largely supplant, or at least significantly marginalize, their more traditional cousins. Rather, traditional techniques must, but can, learn from DNA—and the government should support this aim through funding for validation research and laboratory resources.[81] While conceding that the "scientific foundation" of these disciplines "is currently limited" and that some might have limited prospects for validation to support claims of individualization that are currently made in courtrooms, the Report contends that they "might have the capacity (or the potential) to provide probative information to advance a criminal investigation."[82] The view that some forensic methods might not offer sufficiently reliable or discrete information for a jury's assessment, but can be reliably wielded by law

---

78. *Id.* at 331.

79. *See, e.g.*, Simon A. Cole, *Grandfathering Evidence: Fingerprint Admissibility Rulings from* Jennings *to* Llera Plaza *and Back Again*, 41 AM. CRIM. L. REV. 1189, 1197–99 (2004); Jennifer L. Mnookin, *The Courts, the NAS, and the Future of Forensic Science*, 75 BROOK. L. REV. 1209, 1225–26 (2010).

80. *See* Simon A. Cole & Michael Lynch, *The Social and Legal Construction of Suspects*, 2 ANN. REV. L. & SOC. SCI. 39, 56 (2006) (concluding that new data-mining technologies used to develop and prosecute suspects "reproduce many of the racial and other forms of discrimination that characterize discretionary criminal justice practices"); Murphy, *supra* note 28, at 725 ("[T]he very characteristics that instill such confidence in the second generation—their technical complexity, reliance on databasing, and breadth of application—in fact aggravate the conditions that ultimately caused widespread failures in the first generation.").

81. *See* NAS REPORT, *supra* note 7, at 133 ("The probative power of these [traditional] methods can be high .... [and] can be improved by strengthening the methods' scientific foundations and practice, as has occurred with forensic DNA analysis.").

82. *Id.* at 127.

enforcement to generate or exclude suspects, is plausible but far from intuitively correct.[83] More to the point, it effectively commits significant discretion to investigators and prosecutors to select among scientific tools to aid in investigations.

A third and final point to be made concerning the NAS Report's presumed tie between science and reliable criminal justice outcomes is that it is almost entirely focused on a single producer–user path—namely, from the laboratory to the courtroom. To be sure, the Report understands that police (and to some extent, prosecutors) will utilize the results of forensic analysis to make decisions in the course of criminal cases long before arrest, much less trial. Nevertheless, the overriding emphasis of the Report's recommendations concerning both research and standard setting—recommendations that form the heart of its reliability-enhancing project—are proposed with an eye to trial-oriented decisions such as the admissibility of evidence and the form and integrity of testimony presented to judges and juries.[84] Indeed, it is against the backdrop of a view that "the courts have been utterly ineffective in addressing" the problem of invalid or unverified science that the Report concludes that the forensic science profession itself must shore up its inputs to adjudication.[85]

Of course, the overwhelming majority of charged criminal cases never see the light of trial (and its attendant admissibility hearings), since some 95% terminate in a guilty plea.[86] More to the point, the overwhelming majority of physical evidence never sees the fluorescent light of a laboratory, and much of what a laboratory *could* produce is never actually generated for use in criminal cases.[87] The reasons for this aspect of the "underutilization" question are complex, and to be sure are at least partly related to laboratory-centered circumstances—in particular, as the NAS Report noted, work backlogs that might discourage utilization.[88] But, as the next Part discusses in detail, an array of decisions made outside a laboratory by police and prosecutors deciding what evidence to collect, what evidence to submit for testing, and how (if at all) to act upon analysis that a laboratory can or does perform are undoubtedly factors to be considered in the equation. The fact that, for all of these reasons, police and prosecutors are themselves playing

---

83. *Cf.* Emily Hammond Meazell, *Super Deference, The Science Obsession, and Judicial Review as Translation of Agency Science*, 109 MICH. L. REV. 733, 734–35, 753–54 (2011) (considering whether, in light of scientific uncertainty, courts ought to defer to scientific findings of administrative agencies).

84. NAS REPORT, *supra* note 7, at 127.

85. *See id.* at 53 (adding that "judicial review, *by itself*, is not the answer" (emphasis added)).

86. HINDELANG CRIMINAL JUSTICE RESEARCH CTR., UNIV. AT ALBANY, SOURCEBOOK OF CRIMINAL JUSTICE STATISTICS ONLINE tbl.5.46.2002 [hereinafter SOURCEBOOK OF CRIMINAL JUSTICE STATISTICS ONLINE], *available at* http://www.albany.edu/sourcebook/pdf/t5462002.pdf. Plea rates are lower, though still quite high, for rapes, murders, and other violent crimes. *Id.*

87. *See infra* Part II.

88. NAS REPORT, *supra* note 7, at 37.

critical and frequently dispositive roles in "adjudicating" the significance of scientific evidence in a particular case[89] is a reality that the NAS Report does not engage.

2. Funds and Firewalls.—If the NAS Report hitched the wagon of criminal justice to the star of forensic science, it understood resource constraints and professional infrastructure as the primary, and interconnected, barriers on the path forward. As the previous section discussed, the Committee identified at the outset that the greatest concern raised by its investigation was "underresourcing" in the forensic sciences: inadequate resources devoted to foundational research to validate forensic methodologies and inadequate laboratory resources to support consistent, reliable, and expeditious work.[90] At the same time, the Committee concluded that a necessary condition for harnessing the justice-advancing potential of the forensic science field is the severance of its administrative and, critically, professional ties to the law enforcement institutions that utilize its product.[91] This view is reflected in many of the Report's recommendations, but perhaps most clearly in the controversial proposals to remove public forensic laboratories "from the administrative control of law enforcement," and to establish a new national oversight body housed outside the Department of Justice.[92] The goal for the field should be *professionalization*: "[F]ull adoption of [the] scientific culture," which is viewed as entailing commitments that are distinct from the law (and criminal law in particular), and which in turn give scientific evidence a special claim to validity within legal culture.[93] The NAS Report calls for NIFS to superintend the development of all the hallmarks of professional solidification[94]: educational pathways,[95] technical standards of practice,[96] ethical norms,[97] and centralized

89. *See* Rachel E. Barkow, *Institutional Design and the Policing of Prosecutors: Lessons from Administrative Law*, 61 STAN. L. REV. 869, 871 (2009) (discussing the adjudicative role that prosecutors play in light of plea rates); Lynch, *supra* note 17, at 2149 ("[P]rosecutors, in their discretionary charging and plea bargaining decisions, are acting largely as administrative, quasi-judicial decision-makers . . . .").

90. NAS REPORT, *supra* note 7, at 77.

91. *See id.* at 23 ("The best science is conducted in a scientific setting as opposed to a law enforcement setting.").

92. *Id.* at 24, 80–81; *see also id.* at 18 ("In sum, the committee concluded that advancing *science* in the forensic science enterprise is not likely to be achieved within the confines of DOJ.").

93. *Id.* at 111, 125; *see also id.* at 111 (quoting Sir Isaac Newton's description of the scientific method).

94. *See generally, e.g.*, PROFESSIONALIZATION (Howard M. Vollmer & Donald L. Mills eds., 1966) (chronicling the processes and consequences of professionalization across a number of different occupations); Richard H. Hall, *Professionalization and Bureaucratization*, 33 AM. SOC. REV. 92, 93 (1968) (discussing "structural and attitudinal" attributes of professionalism).

95. NAS REPORT, *supra* note 7, at 218–21.

96. *Id.* at 194 ("Standards and best practices create a professional environment that allows organizations and professions to create quality systems, policies, and procedures and maintain autonomy from vested interest groups.").

certification and accreditation of practitioners.[98] Research and educational programs are to reflect greater connectivity with the broader scientific community, and the Committee repeatedly draws on examples from other technical and scientific professions as models for a professional infrastructure for forensic science.[99]

But critically, in the view of the NAS Report, a direct relationship existed among the three aims of enhancing the quality of criminal justice, addressing underutilization, and achieving forensic science independence and professionalization. Professionalization would contribute directly to reliability of outcomes by addressing issues of bias, both motivational and cognitive, in forensic analysis. Motivationally, the Committee credited the concerns from academia and the defense community (substantiated at least anecdotally by the forensic science scandals of the last decade) that forensic analysts view themselves as aligned with the interests of law enforcement, and that "[b]ecause forensic scientists often are driven in their work by a need to answer a particular question related to the issues of a particular case, they sometimes face pressure to sacrifice appropriate methodology for the sake of expediency."[100] Cognitively, the Committee was moved by the burgeoning social science literature—as well as the spectacular failure of fingerprint analysis in the FBI's investigation of Brandon Mayfield in connection with the 2005 Madrid subway bombing—demonstrating that information often made available to forensic analysts by the law enforcement officials with whom they work (about the identities of suspects, or about the general strength of the state's case, for example) can unconsciously and unintentionally lead to false conclusions systematically skewed in the state's favor.[101]

Also, however, professionalization is viewed as making a more indirect contribution to reliability by addressing the problem of underresourcing. In a world of limited budgets, the Committee concluded that crime laboratory administrators were systematically disadvantaged in competing against other priorities in a police organization to obtain the budgetary capacity necessary to generate reliable analysis.[102] More broadly, resources devoted to the field

97. *Id.* at 212.

98. *Id.* at 25.

99. *Id.* at 194, 195, 208, 212 (referring to examples from "technical professions" including clinical laboratory research, medicine, and engineering).

100. *Id.* at 23–24; *see, e.g.*, Chris Swecker & Michael Wolf, An Independent Review of the SBI Forensic Laboratory 25 (2010) (finding in a review of the North Carolina State Crime Laboratory that 230 instances of negative or inconclusive laboratory test results were never disclosed to police or prosecutors, but that there were no instances of positive or confirmative results that were not disclosed), *available at* http://www.ncids.com/forensic/sbi/Swecker_Report.pdf.

101. *See* NAS REPORT, *supra* note 7, at 123–24 (citing work by, among others, Itiel Dror, and discussing the Mayfield case).

102. *See id.* at 183–84 (observing that in a system where the laboratory administrator reports to the head of the law enforcement agency, "significant concerns related to the independence of the laboratory and its budget" are raised); *cf.* Jan S. Bashinski & Joseph L. Peterson, *Forensic Sciences,*

and funneled through law enforcement institutions like the DOJ (and the NIJ) would not be devoted to the pursuit of scientific truth (wherever it may lie), but to self-serving confirmation of the already-developed infrastructure for forensic evidence.[103]

In any event, administrative segregation is viewed as a key feature of professionalism and the synergy with resourcing and general reliability concerns, and there is no suggestion that any tension exists among those three facets of the Report's view of the field.[104] Unsurprisingly, these features of the NAS Report's reform agenda garnered the most negative response: while forensic science practitioners, police, and prosecutors largely welcomed calls for better resourcing and expanded research in the field, they opposed calls for outside regulation as well as for the institutional independence of laboratories.[105] Also unsurprisingly, given the academic consensus prior to the Report, none of the opposition to independence or national oversight emanating from the forensic science and law enforcement communities has enjoyed a serious defense from outside law enforcement circles.[106] But politics and self-interest aside, there may well be a more

---

in LOCAL GOVERNMENT POLICE MANAGEMENT 488, 503–04 (William A. Geller ed., 3d ed. 1991) (arguing that laboratories located within law enforcement agencies should be situated "as high in the [police] organization as possible" and cautioning that placing laboratories in lower echelons "may condemn [them] to a reduced level of budgetary support and may keep [them] from developing effective and necessary communications and rapport with investigative units").

103. NAS REPORT, *supra* note 7, at 80.

104. *See id.* at 23 ("The best science is conducted in a scientific setting as opposed to a law enforcement setting.").

105. *See* NAT'L DIST. ATT'YS ASS'N, RESOLUTION IN SUPPORT OF EFFORTS TO STRENGTHEN FORENSIC SCIENCE IN THE UNITED STATES (2010), *available at* http://www.ndaa.org/pdf/NDAA_strengthen_forensic_science_resolution_4_10.pdf (supporting calls for more research and greater funding but opposing any requirement that crime laboratories be independent of law enforcement); NAT'L DIST. ATT'YS ASS'N, TURNING THE INVESTIGATION ON THE SCIENCE OF FORENSICS (2011), *available at* http://www.theiai.org/current_affairs/20111207_TurningInvestigationOnScienceOfForensics.pdf (noting areas of agreement with and objection to the Report's recommendations); Jonathan J. Koehler, *Forensic Science Reform in the 21st Century: A Major Conference, A Blockbuster Report and Reasons To Be Pessimistic*, 9 LAW, PROBABILITY & RISK 1, 3–6 (2010) (detailing objections to the NAS Report from the FBI, forensic science organizations, and Congress); *The ASCLD/LAB Board Reacts to the NAS Study*, ASCLD/LAB NEWSL., Mar. 26, 2009, at 3, 3, *available at* http://www.ascld-lab.org/communications/newsletters/2009_march_newsletter.pdf (reporting the view of Association of Crime Lab Directors largely embracing the Report's recommendations but rejecting the independence and NIFS proposals); Meredith Mays, *Forensic Science Reform Continues*, POLICE CHIEF (Nov. 2009), http://www.policechiefmagazine.org/magazine/index.cfm?fuseaction=display&article_id=1938&issue_id=112009 (describing objections of International Association of Chiefs of Police to the NAS Report); Joseph Polski, *Forensic Science: A Critical Concern for Police Chiefs*, POLICE CHIEF (Sept. 2009), http://www.policechiefmagazine.org/magazine/index.cfm?fuseaction=display_arch&article_id=1887&issue_id=92009 (same).

106. *See, e.g.*, D. Michael Risinger, *The NAS/NRC Report on Forensic Science: A Path Forward Fraught with Pitfalls*, 2010 UTAH L. REV. 225, 225–26, 236–39 [hereinafter Risinger, *Path Forward*] (arguing that "the [NAS] Report has now made it untenable to treat criticisms as simply the cavils of uninformed academics with nothing better to do" though calling the NIFS and independence proposals politically infeasible); Koehler, *supra* note 105, at 2 (noting that the "blockbuster" Report is institutionally significant in its call for laboratory independence); Mnookin

complicated relationship among the three aims of professionalism, resourcing, and reliability than what the NAS Report presumes—or so the next Part will contend.

II. What's Missing

This Part tells a different sort of story about forensic science, one that does not take place primarily in the laboratory, or even in the courtroom, but rather at crime scenes, in squad rooms, and in prosecutors' offices in the course of criminal investigations. It aims to sketch the numerous ways in which law enforcement actors exercise discretion in regard to the production and use of forensic science that is (like much of criminal investigation and pretrial prosecutorial work) largely unexamined and unregulated. The picture that emerges problematizes both premises of the NAS Report by adding dynamics indispensable to forensic science production and usage that will sometimes prevent laboratory-level reform from positively affecting the quantity and quality of scientific evidence in the criminal justice system, and might well be exacerbated by some of the NAS Report's reform proposals.

It is well to acknowledge that many of the ground-level concerns discussed below are well-known to criminal justice practitioners and to the law enforcement community in particular. Especially with regard to crime scene work and evidence collection more generally, management and research-priority setting within the police profession (at least at the upper echelons) has acknowledged the importance of and made some strides in previously neglected arenas.[107] However the academic literature, particularly within the United States and certainly in the realm of legal scholarship, has not by and large reflected this priority, and in particular has done little to connect the (largely police-driven) dynamics of evidence gathering to the (laboratory-centered) forensic science production process. Both facts may have influenced the NAS Report's excision of these issues from its focus: on the one hand, there exists a perception that questions concerning evidence gathering are well in hand in the law enforcement field; on the other, academic commentators influential to the Report have largely treated such questions as peripheral.

A caveat is necessary at the outset. By necessity, the account in Part II traffics significantly in generalization. To do otherwise would be an undertaking both enormous, accounting as it must for the massive and diverse set of actors, organizations, and practices represented by the

---

et al., *supra* note 3, at 761 (supporting the Report's recommendation of laboratory independence). *But see generally* D. Michael Risinger, *The NAS/NRC Report on Forensic Science: A Glass Nine-Tenths Full (This Is About the Other Tenth)*, 50 JURIMETRICS 21 (2009) (criticizing the Report for not going far enough in prescribing standards of laboratory practice by merely calling for further research on the subject).

107. *See, e.g.*, Roberta Julian et al., *"Get It Right the First Time": Critical Issues at the Crime Scene*, 24 CURRENT ISSUES IN CRIM. JUST. 25, 26 (2012) (discussing factors contributing to "crime scene processing [coming] to be recognized as a critical stage in the forensic process").

country's more than 15,000 law enforcement agencies and 2,300 prosecutors' offices,[108] and ultimately impossible, stymied as it is by a relative paucity of empirical data about the practices that are most relevant to the account.[109] This Article responds to that challenge with something of a dance between the general and the particular, the theoretical and the anecdotal, the speculative and the data driven. I make one significant qualifying cut, confining the analysis (broad brush though it is) to nonfederal criminal law; the role of the FBI crime laboratory, the unique structure and role of U.S. Attorney's Offices vis-à-vis federal (as well as state) investigators, and the distinctive (and, when compared to the states on aggregate, far smaller) caseload on the federal side (including perhaps most significantly a relative dearth of violent crimes and property crimes) render the story on the federal level sufficiently distinctive to be laid to the side. The account that follows attempts to marshal what empirical evidence is available with respect to the criminal justice system and forensic science usage to offer a plausible account of what might be happening on the ground; indeed, one independent aim here is to call scholarly attention to the social science literature that does exist in this area, which though limited is also underutilized in legal academic accounts of the field. To be sure, this is an approach that runs the grave risk of saying both too much (in overgenerality) and too little (in the selectivity of its focus). Nevertheless, it is a starting point for questioning the completeness of the NAS Report's agenda, and for beginning to forge ahead down an admittedly more complex, but hopefully more fruitful, path forward.

A. Upstream Discretion

1. Evidence Collection.—It is perhaps too obvious to state that the quality and reliability of forensic science is entirely dependent upon the

---

108. *See* BRIAN A. REAVES, BUREAU OF JUSTICE STATISTICS, U.S. DEP'T OF JUSTICE, NCJ 212749, CENSUS OF STATE AND LOCAL LAW ENFORCEMENT AGENCIES, 2004, at 1 (2007), *available at* http://bjs.ojp.usdoj.gov/content/pub/pdf/csllea04.pdf (reporting 12,766 local police departments and 3,067 sheriffs' offices); BRIAN A. REAVES, BUREAU OF JUSTICE STATISTICS, U.S. DEP'T OF JUSTICE, NCJ 238250, FEDERAL LAW ENFORCEMENT OFFICERS, 2008, at 1 (2012) (reporting that twenty-four different federal agencies employ 96% of the 120,000 federal law enforcement agents), *available at* http://bjs.ojp.usdoj.gov/content/pub/pdf/fleo08.pdf; STEVEN W. PERRY & DUREN BANKS, BUREAU OF JUSTICE STATISTICS, U.S. DEP'T OF JUSTICE, NCJ 234211, PROSECUTORS IN STATE COURTS, 2007—STATISTICAL TABLES 2 (2011), *available at* http://bjs.ojp.usdoj.gov/content/pub/pdf/psc07st.pdf (reporting 2,330 state prosecutors' offices); EXEC. OFFICE FOR U.S. ATT'Y, U.S. DEP'T OF JUSTICE, UNITED STATES ATTORNEYS' ANNUAL STATISTICAL REPORT: FISCAL YEAR 2010, MESSAGE FROM THE DIRECTOR (2010), *available at* http://www.justice.gov/usao/reading_room/reports/asr2010/10statrpt.pdf (describing the United States Attorney system as consisting of 94 headquarters offices and 138 staffed branch offices nationwide).

109. BRIAN FORST, IMPROVING POLICE EFFECTIVENESS AND TRANSPARENCY: NATIONAL INFORMATION NEEDS ON LAW ENFORCEMENT 2, 6 (2008) (encouraging BJS and NIJ to do more data collection on police management and "use of technology" in order to raise conviction rates).

quality and reliability of the processes by which the analyzed evidence is collected.[110]

By "collection," I refer to a range of activities that occur in relation to spatial or physical locations—primarily, a geographic location where a crime occurred, the body of a victim (as when a sexual assault kit is collected), or the body of a suspect (as when reference samples are collected for comparison to evidence found on a victim or at a crime scene). In addition to evidence gathering, other activities such as transportation (typically first to a police storage facility to await possible transmittal to a laboratory) and documentation also must occur in this stage.[111] The stakes are high. If evidence is not identified and gathered, or if it is collected, transported, and stored in a deficient manner,[112] items that could have established an element of a crime, implicated a perpetrator, or exculpated a suspect, could be destroyed or lost—or worse, could generate inaccurate results.[113] Somewhat less obviously and far less glamorously, if the steps of evidence collection are not documented to show what, how, and when physical evidence was collected, exploitation of that evidence could be compromised (by, for example, precluding admissibility under chain-of-custody rules), and the ability for downstream actors to evaluate the integrity of the evidence and the investigation that uncovered it will be compromised.[114]

---

110. *See* Bashinski & Peterson, *supra* note 102, at 493 ("The effectiveness of a forensic operation rests on the ability of the police department's evidence recovery system to recognize, preserve, document, and retrieve relevant physical evidence."); VERNON J. GEBERTH, PRACTICAL HOMICIDE INVESTIGATION: TACTICS, PROCEDURES, AND FORENSIC TECHNIQUES 167 (3d ed. 1996) (recognizing that the proper collection and preservation of physical evidence is an essential step in homicide investigations); HENRY C. LEE ET AL., HENRY LEE'S CRIME SCENE HANDBOOK 1 (2001) (contending that processing crime scene evidence is the most crucial step in a forensic case); MIKE REDMAYNE, EXPERT EVIDENCE AND CRIMINAL JUSTICE 23 (2001) (noting that the "most fundamental" issue affecting the use of forensic science is that "[t]race evidence cannot be used unless the police are aware of its existence and usefulness, and know how to collect and preserve it"); JOHN K. ROMAN ET AL., THE DNA FIELD EXPERIMENT: COST-EFFECTIVENESS ANALYSIS OF THE USE OF DNA IN THE INVESTIGATION OF HIGH-VOLUME CRIMES 11 (2008) ("Biological specimens must be properly collected, stored, and submitted to the crime lab to get a sample that can be analyzed.").

111. *See, e.g.*, TECHNICAL WORKING GRP. ON CRIME SCENE INVESTIGATION, CRIME SCENE INVESTIGATION: A GUIDE FOR LAW ENFORCEMENT 17–32 (2000) (directing officers to document each stage of crime scene response and evidence collection).

112. Typically, though not always, physical evidence is gathered from a crime scene and transported to a nonlaboratory police storage facility to await a decision concerning submission for testing. JOSEPH PETERSON & IRA SOMMERS, THE ROLE AND IMPACT OF FORENSIC EVIDENCE IN THE CRIMINAL JUSTICE PROCESS 22 (2010), *available at* https://www.ncjrs.gov/pdffiles1/nij/grants/231977.pdf.

113. *See* Bashinski & Peterson, *supra* note 102, at 497 ("For example, potential evidence that could exclude a suspect may be overlooked at the scene if the crime scene examiner has prematurely focused on a particular theory of reconstruction. Or physical evidence capable of answering a critical investigative question may never be analyzed if an investigator is unaware of its potential value.").

114. *See, e.g.*, *id.* at 492–93 (discussing basic principles of forensic science including standards, controls, and evidence-handling operations).

Evidence collection practices and priorities unquestionably suppress forensic analysis in at least some cases in which it might be thought to aid in detection or arrest, though as discussed below the prevalence and magnitude of this effect is difficult to estimate.[115] The consequences of deficient evidence collection are borne by identifiable, putatively innocent suspects as well. O.J. Simpson's highly publicized criminal trial popularized such concerns, when his attorneys successfully neutralized evidence of a DNA match between Simpson's blood and that found at the scene of his wife's (and her companion's) murder by arguing that police and laboratory sloppiness and corruption, rather than Simpson's presence at the crime scene, explained the inculpatory evidence.[116] In New York's notorious Central Park Jogger case, failure to properly handle evidence likely led to the identification of hairs that were transferred at the police station and not the crime scene.[117] One of the Supreme Court's recent forays into actual innocence claims concerned plausible allegations of flawed blood spatter analysis as a result of poor evidence handling and transport.[118] The recent exoneration of Texan Michael Morton, convicted in 1987 of murdering his wife, featured a near miss in this regard: a bandana, initially overlooked by investigators because it was located beyond the confines of what they adjudged the crime scene, was recovered by an enterprising relative of the victim and became critical in identifying the actual murderer and exonerating Morton nearly twenty-five years after his conviction.[119] There is fair reason to think that Morton's (tragically delayed) near miss is not an isolated occurrence.[120]

---

115. *See, e.g.*, Erin Mulvaney, *Mislabeled Rape Kit from 2003 Linked to Other Assaults*, HOUS. CHRON., Oct. 10, 2012, http://www.chron.com/news/houston-texas/article/2003-rape-kit-untested-for-9-years-due-to-3937748.php (describing a nine-year time lag in identifying a sexual assault suspect due to police mislabeling of the rape kit); *see also infra* notes 137–40 and accompanying text.

116. *See, e.g.*, ARONSON, *supra* note 29, at 192–93 (describing a defense strategy of arguing that contamination occurred in crime scene processing).

117. TASK FORCE ON WRONGFUL CONVICTIONS, N.Y. STATE BAR ASS'N, FINAL REPORT OF THE NEW YORK STATE BAR ASSOCIATION'S TASK FORCE ON WRONGFUL CONVICTIONS 94 (2009) [hereinafter N.Y. STATE BAR ASS'N, FINAL REPORT], *available at* http://www.nysba.org/Content/ContentFolders/TaskForceonWrongfulConvictions/FinalWrongfulConvictionsReport.pdf.

118. *See* House v. Bell, 547 U.S. 518, 543–46 (2006) (describing an allegation of shoddy evidence storage and transport undermining blood spatter evidence in a capital case). These risks at the collection stage are only exacerbated in the DNA era—not simply because of the potential enormity of a missed opportunity to subject evidence to DNA testing, but also because, ironically, the sensitivity of the most current techniques of DNA analysis increases the potential for contaminated evidence to yield profiles of individuals who are not legitimate suspects in a crime. *See* Murphy, *supra* note 28, at 724 (asserting that new forms of forensic evidence can actually exacerbate the conditions that lead to wrongful convictions); *see also* Dist. Attorney's Office for the Third Judicial Dist. v. Osborne, 557 U.S. 52, 80–81 (2009) (Alito, J., concurring) (noting that DNA testing is unique both in its promise of certainty and in its sensitivity to contamination).

119. Pamela Colloff, *The Innocent Man, Part One*, TEX. MONTHLY, Nov. 2012, http://www.texasmonthly.com/story/innocent-man-part-one.

120. *See Ex parte* Miles, 359 S.W.3d 647, 651–52, 653 n.2 (Tex. Crim. App. 2012) (recounting a detective's failure to search an area in the bushes where an eyewitness stated the shooter in

This is in part because the degree of selectivity that is occurring at this early stage is more dramatic than might commonly be appreciated. The best (albeit limited) empirical data that exists indicates that, across the board, significantly less physical evidence is collected in most cases than is available; that the rates of collection vary widely across categories of crime; and that this gap between collection potential and actuality has not meaningfully diminished even as forensic science has become a more central feature of criminal cases. A recent study sponsored by the National Institute of Justice revealed (across four urban jurisdictions) that some physical evidence was collected in nearly all homicide investigations initiated during the study period, but in only 30% of assault investigations and only 20% of burglary investigations.[121] Even in rape cases, which to most minds probably stand apart as consistently offering physical evidence that is likely to be of probative value, only 64% featured crime scene evidence collection.[122] To some extent, these percentages reflect variation among crimes in terms of manner and area of commission that will affect the existence of physical evidence: biological evidence is less likely to be left behind in a nonviolent property crime than a sexual assault, while fingerprints are at least as, if not more, likely to be left in the former than the latter.[123] But "natural" variation is not the whole story, as shown by a number of innovative studies that have actually *checked* the evidence collection work of crime scene responders[124] or have compared evidence collection rates in a given jurisdiction with and

---

question had hidden); KAYE, *supra* note 6, at 258 ("Of course, the laboratory is not going to come back with a report that states that suspect I is the driver if its STR testing plainly excludes the suspect as the source of material swabbed from the driver's airbag, but what might additional swabs from both airbags show?"); Robert B. Bates, *Curing Investigative Tunnel Vision*, POLICE CHIEF, Jan. 1987, at 41, 41 (discussing the investigative impact of overly selective collection of fingerprint evidence); Jonathan Schuppe & William Kleinknecht, *Evidence of a Crisis*, STAR LEDGER, Jan. 30, 2006, http://www.nj.com/starledger/speacialprojects/index.ssf?/news/ledger/stories/013006_essex murder_main.html (describing a beleaguered Essex County Crime Scene Unit and a Newark homicide investigation in which no physical evidence was recovered). But of course, where documentation practices are weak, it is more difficult to identify the extent of the problem. *See, e.g.*, TEX. FORENSIC SCI. COMM'N, REPORT OF THE TEXAS FORENSIC SCIENCE COMMISSION: WILLINGHAM/WILLIS INVESTIGATION 20–21 (2011), *available at* http://www.fsc.state.tx.us/documents/FINAL.pdf (rejecting the allegation that potentially probative debris from fire scenes were discarded without examination in the investigation of two arson cases, but finding that the collection and examination of the debris were completely undocumented). Critically, documentation is likely to be the weakest in the cases that will naturally receive the least scrutiny: relatively minor offenses carrying relatively low sentences that are likely to lead to early negotiated guilty pleas.

121. PETERSON & SOMMERS, *supra* note 112, at 46, 63, 77.

122. *Id.* at 95. In fact, biological evidence, quintessentially associated with sexual assault investigations, was collected in only 54% of the investigations studied in Peterson's research. *Id.*

123. *See, e.g.*, *id.* at 21 (noting that the level of interaction between the offender and the victim or scene affects what evidence is available).

124. *See generally* BRIAN PARKER & JOSEPH PETERSON, PHYSICAL EVIDENCE UTILIZATION IN THE ADMINISTRATION OF CRIMINAL JUSTICE (1972) (detailing their study of the physical evidence gathered from crime scenes and input into crime laboratories).

without the intervention of training or standardized collection procedures.[125] Human selectivity and error are undoubtedly major drivers.

What factors animate selectivity? Most generally, institutional and individual resource constraints are clearly the overarching driver. Not all crimes can possibly be met with a full investigative response along any dimension, forensic evidence gathering and analysis being only one. As with other investigative resource-allocation decisions in police departments, evidence collection activities are largely determined by the seriousness of a reported offense.[126] Across the board, departments will devote the greatest resources of personnel, equipment, and time to homicides, sexual assault, and other violent crimes against persons—less so to property crimes, and even less so to run-of-the-mill drug crimes, despite their significantly greater numerical share of caseloads.[127] Initial response to the least serious offenses will be the most standardized and least elaborate: collection of obviously relevant objects and perhaps fingerprint processing in property cases, but likely no search for and collection of, say, less apparent biological material for DNA analysis (despite technological advances permitting analysis of invisible, trace quantities of DNA in such cases).[128]

Having a picture in mind of *who* is doing this work is critical to make this account of institutional and individual resource constraints more particular. First, and consistent with the NAS Report's express cabining and

---

125. JOHN ROMAN ET AL., POST-CONVICTION DNA TESTING AND WRONGFUL CONVICTION 4 (2012) (noting that convictions of sexual assault are more likely to involve a determinate DNA profile than convictions of nonsexual assault because of standardized evidence collection in rape cases); *see also* Rebecca Campbell et al., *The Effectiveness of Sexual Assault Nurse Examiner (SANE) Programs: A Review of Psychological, Medical, Legal, and Community Outcomes*, 6 TRAUMA, VIOLENCE & ABUSE 313, 315 (2005) (discussing the effect of SANE programs on rates of evidence recovery in sexual assault cases); Donald Johnson et al., *Use of Forensic Science in Investigating Crimes of Sexual Violence: Contrasting Its Theoretical Potential with Empirical Realities*, 18 VIOLENCE AGAINST WOMEN 193, 194 (2012) (noting the role that DNA analysis has played in linking sexual offenders to their victims); *cf.* HER MAJESTY'S INSPECTORATE OF CONSTABULARY, UNDER THE MICROSCOPE—REFOCUSED: A REVISIT TO THE INVESTIGATIVE USE OF DNA AND FINGERPRINTS 2–3, 8 (2002) [hereinafter UNDER THE MICROSCOPE], *available at* http://www.hmic.gov.uk/media/under-the-microscope-20020601.pdf (reporting that an increase in crime scene personnel in the United Kingdom yielded a substantial increase in DNA and fingerprint recovery).

126. *See* JOSEPH L. PETERSON ET AL., FORENSIC EVIDENCE AND THE POLICE: THE EFFECTS OF SCIENTIFIC EVIDENCE ON CRIMINAL INVESTIGATIONS 97–98 (1984) (concluding from statistics on cities' evidence collection that police investigators "will usually go to greater lengths collecting information to attempt to solve personal crimes than they will for property crimes").

127. *See* SOURCEBOOK OF CRIMINAL JUSTICE STATISTICS ONLINE tbl.4.6.2010, *available at* http://www.albany.edu/sourcebook/pdf/t462010.pdf (showing that more property and drug crimes than violent crimes were charged in the United States from 2001 to 2010).

128. Bashinski & Peterson, *supra* note 102, at 494; *see also* ROMAN ET AL., *supra* note 110, at 148 ("Forensic labs should provide data to police describing the attributes of evidence collected that are associated with a higher probability of suspect identification . . . ."); Anita Hassan, *Investigators Using 'Touch DNA' to Solve Property Crimes*, HOUS. CHRON., Mar. 12, 2012, http://www.chron.com/news/houston-texas/article/DNA-is-solving-property-crimes-3397341.php (discussing the rise of microscopic DNA sampling in combating property crime).

exclusion of evidence collection from its inquiry,[129] evidence collection is only occasionally, and at most partly, a laboratory function. Only about half of the nation's crime laboratories even engage in crime scene response, and fewer than half of state laboratories do so (with a greater percentage of municipal and county laboratories offering such services).[130] Even where crime laboratories do play a role in evidence collection, they are typically in any given case only one of multiple responsible parties.[131]

The primary engines of evidence collection occupy the opposite end of the specialization spectrum, as they are typically patrol officers[132]—the most junior, least trained, and most overtasked personnel in the police hierarchy.[133]

---

129. NAS REPORT, *supra* note 7, at 183.

130. *See* DUROSE ET AL., *supra* note 23, at 3 (stating that 52% of the nation's crime labs engaged in crime scene response activities and that both county labs (62%) and municipal labs (71%) were more likely than state labs (44%) to be directly involved in crime scene investigations in 2009); *see also* BARRY A.J. FISHER ET AL., INTRODUCTION TO CRIMINALISTICS: THE FOUNDATION OF FORENSIC SCIENCE 5 (2009) (explaining that the seriousness of the crime determines whether first responders or more specialized personnel collect evidence at the crime scene).

131. FRANK HORVATH & ROBERT T. MEESIG, A NATIONAL SURVEY OF POLICE POLICIES AND PRACTICES REGARDING THE CRIMINAL INVESTIGATION PROCESS: TWENTY-FIVE YEARS AFTER RAND 76 (2001), *available at* https://www.ncjrs.gov/pdffiles1/nij/grants/202902.pdf ("[I]n most agencies evidence-related duties are not assigned predominantly to any one type of individual or position. Rather, they are more likely to be shared among patrol officers . . . investigators . . . and evidence technicians . . . ."); KEITH INMAN & NORAH RUDIN, PRINCIPLES AND PRACTICE OF CRIMINALISTICS: THE PROFESSION OF FORENSIC SCIENCE 62 (2001) ("The individual making decisions about what evidence to collect and the person given the responsibility to collect it vary widely between jurisdictions, so it is difficult to generalize."); ROMAN ET AL., *supra* note 110, at 23–24 (describing the range of evidence collection responsibilities among five studied jurisdictions).

132. *See* HORVATH & MEESIG, *supra* note 131, at 6 (describing the investigative role of patrol officers); FRED E. INBAU ET AL., SCIENTIFIC POLICE INVESTIGATION 36 (1972) (same); PAUL L. KIRK & LOWELL W. BRADFORD, THE CRIME LABORATORY: ORGANIZATION AND OPERATION 104–05 (1965) (same); TECHNICAL WORKING GRP. ON CRIME SCENE INVESTIGATION, *supra* note 111, at 14–16 (same).

133. *See* William H. Bieck et al., *The Patrol Function*, *in* LOCAL GOVERNMENT POLICE MANAGEMENT, *supra* note 102, at 59, 59 (describing the patrol officer as a "master generalist" who deals with a wide variety of calls from the public). With regard to training, requirements vary by state and department. *See generally* BRIAN A. REAVES, BUREAU OF JUSTICE STATISTICS, U.S. DEP'T OF JUSTICE, NCJ 222987, STATE AND LOCAL LAW ENFORCEMENT TRAINING ACADEMIES, 2006 (2009), *available at* http://bjs.ojp.usdoj.gov/content/pub/pdf/slleta06.pdf (describing state and local law enforcement academies in terms of their personnel, expenditures, facilities, curricula, and trainees); Nancy Marion, *Police Academy Training: Are We Teaching Recruits What They Need to Know?*, 21 POLICING: AN INT'L J. POLICE STRATEGIES & MGMT. 54, 56–57 (1998) (listing the minimum hours of basic training that most police cadets receive at training academies (400 hours) and describing the attendance policy at one state-accredited police training academy in a university setting). In Texas, for example, a course in crime scene investigation is required only for "Intermediate Peace Officer" certification. TEX. COMM'N ON LAW ENFORCEMENT OFFICER STANDARDS & EDUC., RULES HANDBOOK 91 (2011), *available at* http://www.tcleose.state.tx.us/publications/publications_gen/Rules%20Handbook_101011.pdf. Basic certification requires a 44-hour course in "Criminal Investigation," which includes one unit in "Protection of and Crime Scene Search." *See* TEX. COMM'N ON LAW ENFORCEMENT OFFICER STANDARDS & EDUC., THE BASIC COURSE—REGULAR OFFICERS (2008), *available at* http://www.tcleose.state.tx.us/publications/

Certainly, patrol officers are commonly—in small and large jurisdictions, and across offense categories—first responders to a reported crime, meaning that they will perform critical foundational tasks such as establishing a crime scene, the spatial parameters of which will guide the search for evidence, and securing that scene to prevent human and environmental contamination.[134] In small departments that lack the personnel to differentiate between patrol and investigative functions, or in most departments in the case of a less serious crime, these same first responders will likely personally perform (or at least coordinate) the full task of evidence collection.[135] Where a department's size permits role differentiation or the reported crime is more serious—homicides, sexual assaults, serious assaults, or violent property crimes—a trained investigator, such as a detective, will typically direct or personally conduct evidence collection after the patrol officer's initial preliminary crime scene response.[136] Some law enforcement organizations, though perhaps surprisingly few, have specially trained evidence technicians or crime scene responders that are institutionally located within patrol or other units (including crime laboratories);[137] in the smallest departments, they will likely simply be police officers who primarily perform other functions (either patrol or investigative) but have some additional training or experience in evidence collection.[138] In the most serious crimes, particularly homicides, there may

---

publications_gen/history_of_the_bpoc_course.pdf (listing required courses for the "618-Hour Basic Peace Officer Course," including a 44-hour Criminal Investigation Course).

134. GEBERTH, *supra* note 110, at 43 ("In almost all instances, the first officer to arrive at any homicide crime scene is the uniformed patrol officer.").

135. *Id.*

136. *See* JOHN E. ECK, SOLVING CRIMES: THE INVESTIGATION OF BURGLARY AND ROBBERY 259–63 (1983) (explaining that patrol officers are typically the first to arrive on scene and suggesting that specialist technicians should be called to routine crime scenes only under special circumstances); HORVATH & MEESIG, *supra* note 131, at 6, 34–35, 41 (reporting that most preliminary investigation is conducted by patrol officers and observing that the involvement of specialized investigators in investigative tasks is surprisingly limited); Frank Horvath & Robert Meesig, *The Criminal Investigation Process and the Role of Forensic Evidence: A Review of Empirical Findings*, 41 J. FORENSIC SCI. 963, 966 (1996) (explaining that a patrol officer's performance of evidence collection has a "significant impact on whether the case will . . . be assigned for follow-up investigation" and noting that if an officer fails "to recognize or to collect potentially valuable evidence, particularly from a suspect, the case outcome is likely to be adversely affected").

137. JAN M. CHAIKEN, THE CRIMINAL INVESTIGATION PROCESS VOLUME II: SURVEY OF MUNICIPAL AND COUNTY POLICE DEPARTMENTS 25 (1975) (reporting that more than 80% of departments have evidence technicians); HORVATH & MEESIG, *supra* note 131, at 75–76 (reporting that only 45% of agencies employ evidence technicians); *see also* SAFERSTEIN, *supra* note 20, at 16 (noting that in most police forces "a patrol officer or detective is charged with the responsibility of collecting the evidence," and that the officer's effectiveness "will be dependent on the extent of his or her training and working relationship with the laboratory").

138. *See* Childs et al., *supra* note 41, at 49 (reporting that 60% of police and sheriffs' departments have one or more employees who work directly on forensic services); ROMAN ET AL., *supra* note 110, at 70 (describing case processing for burglaries in Los Angeles as involving first a responding officer, then a CSI technician, then a burglary detective); HORVATH & MEESIG, *supra* note 131, at 76 (stating that in most agencies evidence-related duties are likely to be shared among patrol officers, investigators, and evidence technicians). Strikingly, in Horvath & Meesig's survey

be policies in place whereby one or more evidence technicians (perhaps but not necessarily assigned to a crime laboratory) respond to the scene and assist in evidence collection—though this is more likely a decision largely within the discretion of an investigator, or even (to the extent she retains responsibility for the crime scene) a patrol officer.[139]

For patrol officers, crime scene response is typically competing with calls for other crimes, ticket quotas, routine patrol responsibilities, and any number of other tasks that make up a patrol officer's diverse portfolio of daily responsibilities.[140] With limited time, patrol officers are undoubtedly triaging their work on any given call; there is good reason to suspect that meticulous performance of tasks associated with evidence collection will be less attractive and receive short shrift to "real" police work such as interviewing witnesses.[141] Even the limited regime of evidence collection in the run-of-the-mill case is likely to be performed quickly—and perhaps shoddily.

In serious cases, where more time and human capital are devoted to evidence collection, selectivity is driven by more deliberate exercises of discretion. Evidence collection is shaped by and contributes to a process of case construction, as officials, particularly police, make judgments in selecting and prioritizing the collection of some subset of available material. Processing a crime scene, for example, requires attention not only to the technical requirements of evidence collection, but also to far more subjective judgments about what areas are likely to yield evidence of a crime.[142] In one case, a researcher observed a homicide detective issue "specific directions" to crime scene unit personnel to fingerprint "everything possible in the bedroom where the victim was found," perform "[n]o fingerprint examination in the second bedroom or of the deceased's car," and take prints in the kitchen and lounge only "near where blood marks are found or where objects that are likely to have been handled by the offender are located."[143]

---

of law enforcement organizations, 12% of responding agencies indicated that evidence technicians were *not* required to have any specialized training. *See id.* (reporting that 88% of the agencies surveyed required specialized training).

139. *See generally* PETERSON & SOMMERS, *supra* note 112.

140. ROMAN ET AL., *supra* note 110, at 106, 148 (describing pressure that patrol experienced to quickly process crime scenes and become available for incoming calls).

141. *See, e.g.*, PETERSON ET AL., *supra* note 126, at 46 ("Patrol officers seldom rope off a crime scene or ban other police personnel from the scene except in the most extraordinary situations. Most officers . . . are more interested in interviewing witnesses and completing their preliminary report so that they may resume patrol activities.").

142. *See, e.g.*, MARTIN INNES, INVESTIGATING MURDER: DETECTIVE WORK AND THE POLICE RESPONSE TO CRIMINAL HOMICIDE 157 (2003) (describing the process of evidence collection as both methodical and improvised); LEMOYNE SNYDER, HOMICIDE INVESTIGATION: PRACTICAL INFORMATION FOR CORONERS, POLICE OFFICERS, AND OTHER INVESTIGATORS 21 (2d ed., 3d prtg. 1973) (encouraging investigators to employ a "rigid routine" in evidence collection modified to fit the particular crime scene).

143. INNES, *supra* note 142, at 209.

David Kaye has shared a similar account from the perspective of a crime scene specialist called to an automobile crash:

> Detective Thompson requested that I swab and impound both airbags in vehicle I. . . . I noted red stains . . . on both the driver and the front passenger airbags. . . . I swabbed a representative spot of the bloodstains on the driver's airbag. . . . I cut and collected the driver's airbag. . . . I then swabbed a representative spot on the front passenger airbag. . . . At the direction of Detective Thompson no further action was taken.[144]

Of course, particularly where these judgment calls are made by personnel who will remain involved in an investigation—a circumstance more typical in serious cases where a detective usually responds to a crime scene—they are likely to determine not only the available physical evidence, but also to shape investigators' understanding of the relevance of that evidence (and, perhaps, the irrelevance of evidence not selected in the initial canvass) to the case.[145] Such assessments will be better or worse depending on the crime and the personnel involved, of course, which only reinforces the importance of documenting the choices made early on.

Aside from the quite real but largely informal constraints created by organizational and individual capacity, there are also more formal constraints on law enforcement discretion in regard to the collection of physical evidence. From the standpoint of legal doctrine, these are important but relatively limited. In the main, police need not collect or save any quantum of evidence with any particular degree of competence, so long as they work in good faith.[146] The most significant direct constraints flow from the Fourth Amendment: there is no general "crime scene" exception to the Fourth Amendment's requirement of "reasonable" searches and seizures, and thus police may not enter premises and gather evidence absent an imminent threat to public safety, consent to entry, or the obtaining of a warrant.[147] But where the space at issue is the site where a crime has occurred—the moment of

144. KAYE, *supra* note 6, at 258.

145. *See* LEE ET AL., *supra* note 110, at 1–2 (characterizing crime scene investigation as the first stage in crime "reconstruction").

146. *See* Arizona v. Youngblood, 488 U.S. 51, 58 (1988) (rejecting the contention that due process was violated by the loss of potentially favorable evidence, in favor of the standard requiring a showing of official bad faith to make out the constitutional destruction-of-evidence claim); California v. Trombetta, 467 U.S. 479, 488–89 (1984) (rejecting the argument that failure to preserve breath test samples for later analysis violated due process unless their exculpatory value was apparent and comparable evidence was unavailable to the defendant).

147. As the Supreme Court has noted:

> [P]olice may make warrantless entries onto premises if they reasonably believe a person is in need of immediate aid and may make prompt warrantless searches of a homicide scene for possible other victims or a killer on the premises, but we rejected any general "murder scene exception" as "inconsistent with the Fourth and Fourteenth Amendments—. . . the warrantless search of [an] apartment was not constitutionally permissible simply because a homicide had recently occurred there."

Flippo v. West Virginia, 528 U.S. 11, 14 (1999) (citations omitted).

initial crime scene response—none of those requirements presents a formidable or time-consuming barrier. A warrant is readily obtainable where an officer has received a report (and thus has probable cause to believe) that a crime has occurred at a given location;[148] even better, consent to search can often be obtained; [149] and in all events, the current trend in Fourth Amendment doctrine is to lower the remedial risk to police who proceed without a warrant, at least in arguably close cases.[150]

Of course, these rules only directly constrain the initial decision to gather evidence, and say little about the particular manner in which that evidence will be gathered. But undoubtedly there are trickle-down effects. If a warrant must be obtained to search a location or a person, that increases the likelihood that more senior law enforcement personnel, or perhaps even a district attorney, will be involved, which might in turn enhance the expertise and care that is brought to the task. The warrant-application process also permits a moment of reflection (and documentation) concerning the value, known or hypothesized, of evidence to an investigation—an event that might serve some disciplining function, particularly later in an investigation as theories of a case start, for better or for worse, to coalesce.[151] On the other hand, the prospect of delay might compromise evidence (although exigency doctrines permit warrantless searches and seizures to prevent imminent destruction),[152] or it might present just enough hassle to disincentivize the request to collect it in the first place,[153] particularly in less serious cases or where police are not otherwise convinced of the value of the effort.

---

148. *See* Yale Kamisar, *Does (Did) (Should) the Exclusionary Rule Rest on a "Principled Basis" Rather than an "Empirical Proposition"?*, 16 CREIGHTON L. REV. 565, 569–70 (1983) (describing search warrants as easy to obtain).

149. *See* Mark Hutchins, *Crime Scene Searches*, PROSECUTOR, Nov./Dec. 1999, at 25, 26 (observing that "[w]arrants are seldom necessary for making the initial entry" to a crime scene because of the consent and exigent circumstances doctrines).

150. *See* Jennifer E. Laurin, Messerschmidt *and Convergence in Action: A Reply to Comments on* Trawling for *Herring*, 112 COLUM. L. REV. SIDEBAR 119, 123 (2012) (arguing that the Supreme Court has recently given police "'ample' space to err" when deciding liability for departures from constitutional requirements).

151. *Cf.* Chip Heath et al., *Cognitive Repairs: How Organizational Practices Can Compensate for Individual Shortcomings*, 20 RES. ORGANIZATIONAL BEHAV. 1, 3 (1998) (describing the concept of cognitive repairs, organizational practices that people use to correct initial errors in judgment).

152. *See, e.g.*, Kentucky v. King, 131 S. Ct. 1849, 1858 (2011) (permitting warrantless entry to prevent the destruction of evidence); Schmerber v. California, 384 U.S. 757, 770 (1966) (permitting warrantless collection of blood for purposes of toxicology analysis because "the delay necessary to obtain a warrant, under the circumstances, threatened the 'destruction of evidence'"); Kaliku v. United States, 994 A.2d 765, 780 (D.C. 2010) (holding warrantless collection of penile swab from an assault suspect justified under the exigent circumstances exception); State v. Dupree, 462 S.E.2d 279, 282–83 (S.C. 1995) (permitting warrantless search of a suspect's mouth given the risk that drugs were inside and would be swallowed).

153. *See* RICHARD VAN DUIZEND ET AL., THE SEARCH WARRANT PROCESS: PRECONCEPTIONS, PERCEPTIONS, PRACTICES 19 (1985) (quoting a detective as saying: "Actually, there are a lot of warrants that are not sought because of the hassle. . . . I don't think you can forgo a case because of the hassle of a search warrant, but you can . . . work some other method."); David A. Sklansky, *Quasi-Affirmative Rights in Constitutional Criminal Procedure*, 88 VA. L. REV. 1229, 1250–51

With regard to the scope and manner of evidence collection, regulation comes largely from departmental policies, which are minimal and varied, particularly in smaller jurisdictions. Evidentiary chain of custody as a prerequisite to admissibility of course plays some role in spurring policies (or at least standardized practices) governing the collection, packaging, documentation, and transfer of evidence.[154] Nevertheless, observers of the field have remarked that such policies are often lacking, and that evidence collection, documentation, and storage are widely viewed as low-priority matters within departments.[155]

Significantly, much of the work in promulgating and encouraging good practices among police responsible for evidence collection appears to fall to crime laboratories. This is in part because, thanks to CODIS regulations, they face their own legal requirements for documenting and certifying the integrity of evidence they receive.[156] But despite the fact that the vast majority of laboratories are under the control of law enforcement, relationships between crime laboratories and their police "customers" vary in terms of closeness and functionality. Indeed, notwithstanding the NAS Report's overriding concern about coziness in these relations, the dominant theme sounding from criminal justice researchers as well as practitioners is one of a need for *greater* coordination than what currently exists between laboratory-based personnel and law enforcement, particularly with regard to evidence-collection tasks.[157]

(2002) (observing that "getting a warrant still takes hours or days" and that "formal procedures for seeking a telephonic warrant are often cumbersome" and that as a result police simply forego them); William J. Stuntz, *Warrants and Fourth Amendment Remedies*, 77 VA. L. REV. 881, 891–92 (1991) (observing that the warrant requirement certainly deters some searches).

154. *See* TECHNICAL WORKING GRP. ON CRIME SCENE INVESTIGATION, *supra* note 111, at 37, 41, 43.

155. John E. Eck & Gerald L. Williams, *Criminal Investigations*, *in* LOCAL GOVERNMENT POLICE MANAGEMENT, *supra* note 102, at 131, 144; GEBERTH, *supra* note 110, at 805 ("Often agencies do not put enough emphasis on [the crime scene investigative] phase of the investigation . . . ."); PETERSON ET AL., *supra* note 126, at 46 (noting that patrol officers are "rather blase" about going to great investigative lengths and prefer to "complet[e] their preliminary report so that they may resume patrol activities"); Joseph T. Latta & William P. Kiley, *Property and Evidence Control—The Hidden (and Ticking) Time Bomb*, CALEA UPDATE MAG., June 2007, *available at* http://www.calea.org/calea-update-magazine/issue-94/property-and-evidence-control-hidden-and-ticking-time-bomb (discussing widespread problems with evidence storage and disposition stemming from the perception that property room control is low priority); *see also* INNES, *supra* note 142, at 208 ("The early stages of an investigation are frequently chaotic and confused, and . . . problems were often experienced in exerting control over the scene in the earliest stages. It was frequently the case that the continuity of records . . . was not maintained . . . .").

156. Julie Samuels et al., *Collecting DNA from Arrestees: Implementation Lessons*, NIJ J., June 2012, at 18, 22, *available at* http://www.nij.gov/journals/270/arrestee-dna.htm.

157. *See, e.g.*, Bashinski & Peterson, *supra* note 102, at 495 (advocating for coordination); ROMAN ET AL., *supra* note 110, at 51, 73, 148 (describing how collaboration is critical and describing the breakdown in evidence collection in some jurisdictions because of poor coordination); SAFERSTEIN, *supra* note 20, at 16 (advising coordination but noting the lack thereof in many agencies).

2. *Obtaining and Using Forensic Analysis.*—In its last four terms, the Supreme Court has decided two cases concerning constitutional rights of access to DNA testing: *District Attorney's Office for the Third Judicial District v. Osborne*[158] and *Skinner v. Switzer*.[159] Both cases directly raised the question of whether a defendant could challenge the state's opposition to providing postconviction access to physical evidence for the purpose of DNA testing.[160] But both also illustrated the consequences of decisions made by police and prosecutors that affected the availability and significance of forensic evidence in the cases. In Osborne's case, DNA analysis on semen recovered from the rape victim included Osborne in a relatively large group of potential donors, but only after a putative accomplice fingered Osborne and the victim gave an equivocal identification; the state opted not to pursue more discriminating DNA analysis.[161] In Skinner's, a significant amount of evidence recovered from the crime scene, including blood-stained clothing, was never analyzed, despite the fact that fingerprints found on the murder weapon did not match Skinner's.[162] In both cases, the scientific evidence that *was* offered was ambiguous—inculpatory but far from definitive[163]—making other forms of evidence crucial (in particular eyewitness accounts obtained by police before the forensic evidence was available) and raising the question of whether further analysis might not have yielded more probative, and perhaps exculpatory, results. In other words, police and prosecutorial roles in the production and usage of forensic evidence funnels may well have been material to the reliability of the scientific evidence offered at trial.

As it turns out, these cases are far from aberrational. Once gathered, evidence typically makes its way to a storage location (often though not always within a police department) to await further decisions about its fate. Will it be submitted to a crime laboratory for analysis? If so, what will that analysis entail? And if that analysis is obtained, what will be the investigative response? Each of these decisions is far less automatic, far more variable, and far more subject to the discretion of upstream, nonlaboratory actors than is commonly understood.[164]

As to the first question, forensic evidence is funneled through the criminal justice system in a manner that results in surprisingly stark drop-offs from rates of collection to rates of submission to analysts and rates of analysis actually performed. Recent attention to rape kit backlogs has

158. 557 U.S. 52 (2009).

159. 131 S. Ct. 1289 (2011).

160. *Id.* at 1293; *Osborne*, 557 U.S. at 61.

161. *See* Osborne v. State, 110 P.3d 986, 989–90 (Alaska Ct. App. 2005).

162. Skinner v. State, 956 S.W.2d 532, 536 (Tex. Crim. App. 1997).

163. *See* Appellant's Opening Brief at 31–33, Skinner v. State, No. AP-76,675 (Tex. Crim. App. Feb. 2, 2012), 2012 WL 591289, *31–33 (discussing the evidence proffered at trial).

164. *See* PETERSON & SOMMERS, *supra* note 112, at 122 ("[I]t is clear that criminal justice officials external to the laboratory screen much of the forensic evidence and have a major influence on evidence examination priorities and practices.").

exposed one manifestation of this phenomenon: evidence collected from the bodies of sexual assault victims, sometimes in cases yet unsolved, has been found in a number of jurisdictions to be filling the storage lockers of police departments, never having been submitted for laboratory analysis.[165] What limited empirical data exists suggests that the phenomenon of collected but unexploited potential forensic evidence is the rule rather than the exception. In an NIJ-sponsored study of crime laboratories in Indiana and Los Angeles, Joseph Peterson found significant drop-offs between evidence collection and evidence analysis across (though variable among) crimes.[166] Thus, evidence was collected in 30% of assault investigations, but submitted in only 12% and analyzed in only 9%.[167] Even in homicide investigations, which nearly always entailed the collection of physical evidence, only 89% of cases saw that evidence submitted to laboratories, and only 81% of cases had the submitted evidence analyzed.[168] Sexual assaults featured even more dramatic funneling, with evidence collected in 64% of cases, submitted in only 32%, and analyzed in only half the cases in which it was submitted.[169] Peterson's results were consistent with the forensic evidence funnel revealed in other similar studies.[170] Variation exists among types of forensic analysis as well. Some types of forensic testing are practically essential to establishing the elements of certain criminal offenses; for example, in drug offenses, where the chemical composition of a substance must be determined, evidence will invariably be submitted for forensic toxicological analysis (though whether it is analyzed before the case is disposed of is another story, as discussed below). Across a wider variety of offenses, fingerprints are the most consistently submitted and analyzed type of evidence for identification purposes—more commonly, and more consistently, than biological evidence

---

165. *See, e.g.*, SARAH TOFTE, TESTING JUSTICE: THE RAPE KIT BACKLOG IN LOS ANGELES CITY AND COUNTY 3 (2009) (discussing the backlog in Los Angeles County); Brandi Grissom, *Thousands of Rape Kits Sit Untested for Decades, but Change Would Be Costly*, N.Y. TIMES, Jan. 27, 2011, http://www.nytimes.com/2011/01/28/us/28ttkits.html?pagewanted=all&_r=0 (discussing a similar situation in Texas); Kari Lyderson, *Law Came Too Late for Some Rape Victims*, N.Y. TIMES, July 8, 2010, http://www.nytimes.com/2010/07/09/us/09cnckits.html?_r=1&adxnnl=1&adxnnlx=1352823755-gqfzAuFZ5XVukBgIXdUxfw (discussing similar situations in Illinois and other jurisdictions, including Detroit, Houston, Los Angeles, San Antonio, and San Diego); Megan Twohey, *Dozens of Rape Kits Not Submitted for Testing by Chicago Suburban Police Departments*, CHI. TRIB., June 14, 2009, http://www.chicagotribune.com/news/local/chi-rape-kits-14-jun14,0,3619454.story?page=1 (discussing a similar situation in Chicago).

166. *See generally* PETERSON & SOMMERS, *supra* note 112.

167. *Id.* at 46.

168. *Id.* at 77–78.

169. Johnson et al., *supra* note 125, at 206.

170. *See* STROM ET AL., *supra* note 75, at xi (reporting that in a 2007 survey of national law enforcement organizations, among unsolved homicide, rape, and property crime cases, 14%, 18%, and 23%, respectively and on average, featured unsubmitted forensic evidence); Travis C. Pratt et al., *This Isn't CSI: Estimating the National Backlog of Forensic DNA Cases and the Barriers Associated with Case Processing*, 17 CRIM. JUST. POL'Y REV. 32, 36–37 (2006) (discussing their findings that over 200,000 U.S. rape and homicide cases contain biological evidence that has not been sent for testing).

susceptible to DNA analysis, which even when collected is frequently never submitted.[171]

Moreover, once evidence is submitted to a laboratory, a series of decisions must be made about the type and timing of testing that will occur. To some extent, these are decisions that are typically and properly the subject of laboratory protocols. Biological evidence, for example, will typically be "screened" to determine whether it contains substances that might be susceptible to further testing, and laboratories typically do this testing first and automatically after a request for testing is made.[172] Some evidence might be susceptible to multiple types of analysis (for example, an item that potentially bears both blood and latent fingerprints), and scientific protocols will dictate the order of testing to preserve the various samples. But typically, there are also choices to be made that fall within the investigative wheelhouse. Of course, in some (perhaps many) cases, police will approach evidence submission decisions in a fairly routine fashion. In the mine-run of drug cases, for example, evidence suspected of being or containing an illicit substance will undoubtedly be submitted for toxicology analysis.[173] In more complex cases, however, there are likely to be multiple items of evidence, some or all of which could be subjected to a variety of different tests. A shoe, for example, might be examined for trace evidence (such as hairs), for the presence of biological material (such as blood), and for comparison to a shoe print left at a crime scene. If initial testing identified a stain on the shoe as blood, there might be a variety of different types of analysis available to try to associate that blood with an individual: for example, a small or degraded piece of evidence might be susceptible to DNA testing with certain techniques, such as mini-STR analysis (permitting the testing of much smaller samples but with less discriminating results than the now-standard method of STR analysis) or mtDNA analysis (more widely available for small samples than mini-STR but far less discriminating), but the choice will often entail trade-offs with respect to how narrow a field of inclusion will be drawn.[174]

---

171. *See* PETERSON & SOMMERS, *supra* note 112, at 46, 63 (revealing the gap between submission and analysis of fingerprints versus biological evidence in assault and burglary cases); Johnson et al., *supra* note 125, at 206 (demonstrating the same gap in rape cases); Strom & Hickman, *supra* note 75, at 391–93 (showing the number and types of cases containing unsubmitted forensic evidence).

172. *See* NATHAN JAMES, CONG. RESEARCH SERV., R41800, DNA TESTING IN CRIMINAL JUSTICE: BACKGROUND, CURRENT LAW, GRANTS, AND ISSUES 9 (2012) (explaining that one factor in the forensic casework backlog is the "time-consuming" process of biological evidence screening before further testing can begin).

173. *See* PETERSON ET AL., *supra* note 126, at 59 (noting that in "[c]ases of suspected drug possession . . . the identification of the substance is one of the crucial items of information required to prove the crime").

174. *See, e.g.*, Erin Murphy, *The Art in the Science of DNA: A Layperson's Guide to the Subjectivity Inherent in DNA Typing*, 58 EMORY L.J. 489, 493–94 (2008) (discussing a variety of available DNA-analysis techniques); *Mini-STR Testing*, DNA DIAGNOSTICS CENTER, http://www.forensicdnacenter.com/dna-ministr.html (championing mini-STR testing over standard

There is a further, arguably crucial, issue of timing. Notwithstanding significant popular and academic attention to the recent ability of police to use DNA and (to a lesser extent) fingerprint databases to generate suspects through "cold search[es]" early in criminal cases,[175] it remains true that the vast majority of forensic evidence is generated and used not early but quite late in a case. Available empirical evidence demonstrates that while physical evidence might well be submitted for analysis prior to arrest, in the overwhelming majority of cases testing typically is not performed until after a suspect is in custody.[176] Indeed, it is quite common for forensic analysis not to be performed until well after charges are filed, and frequently not until shortly before a trial occurs in a case.[177] This is in large part—maybe in a but-for sense entirely—a reality of the press of caseloads on most crime laboratories, particularly in (high-volume) drug-analysis requests[178] and any request for (time-intensive) DNA analysis.[179] Significant wait times typically accompany requests for analysis in the ordinary course, often exceeding thirty days in these higher volume request categories.[180] For reasons of tactics or legal constraint (discussed in further detail below), decisions to

---

STR testing when DNA evidence is limited in quality or quantity); *see also* Dist. Attorney's Office for the Third Judicial Dist. v. Osborne, 557 U.S. 52, 62 (2009) (acknowledging that advances in DNA testing culminating in STR technology have made possible tissue matching with a suspect to a near certainty).

175. *See, e.g.*, Cole & Lynch, *supra* note 80, at 44–45 (attributing an increased principal usage of "[t]race evidence recovered from crime scenes" to late twentieth-century "databases indexed according to biometric information (fingerprints and DNA profiles)").

176. *See* Johnson et al., *supra* note 125, at 210 (reporting that while physical evidence was a predictor of arrest it was analyzed prior to arrest in only 1.6% of cases); *see also* PETERSON & SOMMERS, *supra* note 112, at 123 (same).

177. *See* PETERSON & SOMMERS, *supra* note 112, at 22 ("Evidence sometimes remains in the property room for brief or extended periods of time while the investigation is proceeding and sometimes until suspects are identified, standards are being sought, or a decision is being made whether to pursue or terminate the investigation.").

178. Although relatively few requests for DNA analysis are made, each case is extremely time and resource intensive, and is competing against an ever-increasing stream of work flowing from state and federal laws requiring the routine collection and analysis of DNA from individuals who enter the criminal justice system. *See* DUROSE ET AL., *supra* note 23, at 4 & tbl.5 (estimating that the total number of requests for forensic services received exceeded the total number completed by publicly funded forensic crime labs); TEX. FORENSIC SCI. COMM'N, STAKEHOLDER ROUNDTABLE REPORT 12 (2012), *available at* http://www.fsc.state.tx.us/documents/StakeholderRoundtableReport-June62012.pdf (identifying stakeholder dissatisfaction with turnaround times among the "key issues and challenges" to quality and timeliness of forensic services); Michelle Villarreal, *Department of Public Safety Crime Labs Limits DNA, Drug Testing*, CORPUS CHRISTI CALLER-TIMES (Sept. 1, 2012, 5:25 PM), http://www.caller.com/news/2012/sep/01/department-of-public-safety-crime-labs-limits/ (reporting a Department of Public Safety crime lab policy change limiting the DNA testing technicians will perform in response to their increased workload).

179. *See* DUROSE ET AL., *supra* note 23, at 4 (estimating that "60% of the estimated 887,000" forensic biology requests were backlogged in 2009 due to increases in collection of DNA samples); TEX. FORENSIC SCI. COMM'N, *supra* note 178, at 10 (questioning how a forensic examiner could be qualified to conduct independent casework but unqualified for certification in DNA testing); Villarreal, *supra* note 178 (reporting a recent limitation of the DNA testing technicians will perform to manage increased workload).

180. DUROSE ET AL., *supra* note 23, at 3–4.

arrest, and frequently even to charge, simply cannot await the completion of testing, at least in cases in which the low threshold of probable cause can be met with other evidence. Of course, not every case *is* handled in the ordinary course. Laboratories might have formal testing priority policies based on, say, some combination of court deadlines and offense type, giving priority to cases set for trial or other hearings—to take a typical example.[181] In many laboratories, however, the approach is more improvisational: evidence is accepted (as submitted at the discretion of law enforcement), and testing and timing decisions are made on the basis of some combination of formal policy, informal practice, and negotiation between analysts and case investigators or prosecutors.[182] There is likely to be at least some degree of communication, and perhaps even collaboration, among laboratory personnel and law enforcement "customers" in navigating these resource considerations; indeed, as with evidence collection, practitioners and criminal justice researchers view the facilitation of ongoing communication as a need and a goal for the field.[183] Often, it is a matter of the squeaky wheel getting the grease—meaning, again, that the ball is often in the court of police and prosecutors.[184]

Finally, there is the question of investigative and prosecutorial response to scientific evidence that is obtained. Of course, forensic evidence will not by its own force solve a crime; the thing never speaks for itself.[185] Many forensic science techniques can only, at the top of their games, prove the nature of an object or instrumentality of crime: chemical drug analysis establishing a powder as cocaine, for example, or firearms examination linking a spent bullet cartridge to a weapon. And even those forensic methodologies that purport to identify linkages between physical evidence and human sources—DNA analysis, fingerprint comparison, and hair microscopy, for example—are only one piece of an investigative puzzle. Thus, DNA shed at a crime scene only points to a suspect if there is no lawful explanation for its presence. And, as in Osborne's and Skinner's case, even forensic evidence putatively capable of "identifying" its source might

---

181. *See* MOLLY E. GRISWOLD & GERARD B. MURPHY, IT'S MORE COMPLEX THAN YOU THINK: A CHIEF'S GUIDE TO DNA 16–20 (2010), *available at* http://policeforum.org/library/dna-forensics/DNA-Book10.10.pdf (describing prioritization approaches); PETERSON ET AL., *supra* note 126, at 69, 76–77 (describing formal laboratory policies dictating submission and analysis).

182. *See* NAS REPORT, *supra* note 7, at 61 (reiterating the warning from the Los Angeles Sheriff's Department Crime Laboratory Director that labs "triage" cases, prioritizing violent crimes and cases handled by persistent investigators who pester the lab frequently).

183. GRISWOLD & MURPHY, *supra* note 181, at 23–25.

184. NAS REPORT, *supra* note 7, at 61; *see also* Justin Fenton, *Criminals' DNA Ignored*, BALT. SUN, Sept. 27, 2008, http://articles.baltimoresun.com/2008-09-27/news/0809270007_1_dna-crime-lab-found-on-evidence (reporting that police told crime lab technicians not to follow-up on DNA found at crime scenes in at least six open homicide and sexual assault cases and three closed burglary cases).

185. *See* Cole & Lynch, *supra* note 80, at 55 (stating that forensic "evidence must have a material relation to the crime that renders it suspicious" to have probative value).

(often) do so to a degree of discrimination that requires further linking evidence to convince a fact finder (meaning, prior to trial, an investigator, a magistrate, or a prosecutor) that an identified suspect is probably guilty. But importantly, to the extent that forensic evidence is typically available only *after* an arrest, this is frequently the province of the prosecutor, who will generally have a number of options other than the one that typically receives the most focus—taking the case to trial. Rather, she will be faced with questions such as whether to charge a suspect or instead wait for forensic analysis or pursue additional lines of testing, as well as whether the availability (or not) of forensic evidence will affect negotiations to resolve the case by plea.

Perhaps even more so than was the case with evidence collection, police enjoy enormous discretion in connection with evidence submission and requests for testing. There are relatively few formal constraints. Most commonly, accepted practice or (rarely) formal policies will simply delegate to a lead investigator the decision of whether, when, and with what direction to submit evidence for forensic testing.[186] She is most likely to be highly resource constrained, and critically, in all but the most serious or specialized instances, carrying far more cases than she could ever investigate thoroughly.[187] So, the name of the game is typically doing the minimum required to close a case by arrest.[188] To be sure, there are legal constraints to the extent that the substantive penal law forms the backdrop to any investigation: to effect an arrest, probable cause is required, which in turn requires reference to the substantive law underlying a given offense.[189] Some forensic analysis is effectively required in order to establish an element of an offense: drug analysis in narcotics offenses, for example, or testing for the presence of semen in many sexual assault cases (particularly if a victim is

---

186. *See* GRISWOLD & MURPHY, *supra* note 181, at 17–18 (describing various approaches to prioritization). Indeed, the exceptions prove the rule. One of the most widely remarked and controversial consequences of the above-described revelations of unsubmitted, untested sexual assault kits has been the adoption in some police departments and crime laboratories of mandatory testing policies for all evidence collected in rape investigations. *Compare* Megan Twohey, *Illinois to Test Every Rape Kit*, CHI. TRIB., July 6, 2010, http://articles.chicagotribune.com/2010-07-06/news/ct-met-rape-kit-law-20100706_1_untested-kits-crime-lab-dna-backlog (describing Illinois legislation requiring testing of all rape kits backed by victims' groups), *with* JOSEPH PETERSON ET AL., SEXUAL ASSAULT KIT BACKLOG STUDY v (2012), *available at* https://www.ncjrs.gov/pdffiles1/nij/grants/238500.pdf (reporting that law enforcement was generally skeptical of mandatory testing).

187. *See, e.g.*, Villarreal, *supra* note 178 (reporting that Texas crime lab requests related to drug and alcohol crimes increased nearly 500% within six years and describing a new policy allowing prosecutors to select which misdemeanor-related lab requests are analyzed).

188. *See, e.g.*, United States v. Jacobsen, 466 U.S. 109, 122–23 (1984) (finding a field test by federal agents that "merely discloses whether or not a particular substance is cocaine" does not violate legitimate privacy interests); United States v. Place, 462 U.S. 696, 707 (1983) (finding that a sniff inspection of a traveler's luggage by a trained narcotics-detection dog does not violate privacy interests).

189. U.S. CONST. amend. IV.

unavailable to testify).[190] But in many cases, scientific evidence is just one potential form of proof, and even if forensic evidence would be desirable or even necessary for discharging the state's burden at trial, the low threshold of probable cause to arrest or charge can typically be met without it.[191]

In the context of forensic science, these pressures against the marginal utility of additional investigative effort are only reinforced, at least as a general matter, by the prevailing culture of policing. Students of the field continue to depict it as a highly bureaucratic enterprise, generally resistant to innovation both at the management level and at the level of individual police actors.[192] Challenges to policing from "outsiders" is typically viewed with the greatest skepticism; the contributions of those who are not "team players" are unlikely to be embraced, at least where they do not conform to existing police priorities.[193] And so, researchers have noted that to the extent forensic science has been embraced, it is in service of police-defined objectives (determined in traditionally police-led ways), not as generative of lines of inquiry or tactics for pursuing them in a given case.[194]

---

190. PETERSON & SOMMERS, *supra* note 112, at 22–23.

191. For example, in drug cases, police will typically have a variety of "field tests" at their disposal, permitting them to make presumptive drug identifications that are sufficient as a legal matter to establish probable cause to arrest and charge—albeit perhaps not admissible in court to prove the chemical composition of a drug. *See, e.g.*, Florida v. Harris, 133 S. Ct. 1050, 1056 (2013) (holding that an alert by a dog trained in narcotics detection may establish probable cause to search a vehicle based on the ordinary "totality-of-the-circumstances" standard regardless of whether documentation supports the reliability of the dog); People v. Swamp, 646 N.E.2d 774, 775–76 (N.Y. 1995) (holding that a positive result of a field test for identifying cocaine was sufficient to authorize an indictment without formal laboratory results); Villarreal, *supra* note 178 (emphasizing the importance of crime scene investigation given tighter regulations on forensic testing); *see also* Marc G. Kurzman & Dwight Fullerton, *Drug Identification*, *in* SCIENTIFIC AND EXPERT EVIDENCE 521, 523–54 (Edward J. Imwinkelried ed., 2d ed. 1981) (discussing the threshold of proof required to sustain a conviction when no scientific proof is available).

192. *See generally* JEROME H. SKOLNICK & DAVID H. BAYLEY, COMMUNITY POLICING: ISSUES AND PRACTICES AROUND THE WORLD 49–65 (1988) (discussing challenges to community policing stemming from the "attitudes, internal divisions, belief systems, traditions, [and] values" of police departments).

193. *See id.* at 50 (discussing the causes of the "solidarity" concept in police culture).

194. *See* Cole & Lynch, *supra* note 80, at 50 (discussing empirical data that many law enforcement agencies still do not use DNA testing as an investigatory tool and noting their reluctance to submit DNA evidence in "suspectless" cases); Horvath & Meesig, *supra* note 136, at 966 (stating that "unless a suspect has been identified, the crime scene evidence is typically not analyzed"); David A. Schroeder & Michael D. White, *Exploring the Use of DNA Evidence in Homicide Investigations: Implications for Detective Work and Case Clearance*, 12 POLICE Q. 319, 337 (2009) ("[D]etectives consistently indicated that they would use DNA evidence when needed; they just did not need it that often."); *cf.* Edward R. Maguire, *Structural Change in Large Municipal Police Organizations During the Community Policing Era*, 14 JUST. Q. 547, 569–70, 572 (1997) (describing the failure of community policing policies to create structural changes within police organizations and suggesting that structural changes only occur when they have symbolic value and are nondisruptive to day-to-day activities); James J. Willis et al., *COMPSTAT and Bureaucracy: A Case Study of Challenges and Opportunities for Change*, 21 JUST. Q. 463, 485 (2004) (recounting that commanders used data generated through COMPSTAT merely to fulfill "traditional 'crime control' function[s]" rather than to identify and respond to particular crime problems as COMPSTAT promised).

Ironically, these dynamics are likely to be particularly pronounced with respect to the types of forensic methodologies that society (and the NAS Report drafters) might most want police to embrace: those with the strongest footing in science rather than policing, such as DNA, for example, rather than fingerprint analysis and other techniques rooted in police training and culture.[195] Studies suggest that this remains a prevailing attitude by demonstrating that DNA analysis tends to be requested by police far more in cases with very low clearance rates; it is a tool of last, rather than first, resort.[196] It is also entirely consistent with the notable enthusiasm in the law enforcement community for expanded use of early DNA analysis to permit exploitation of the CODIS suspect databases in property crimes investigations.[197] These are categories of cases that currently have extremely low clearance rates because there are typically no witnesses and little other evidence available for police to pursue along traditional lines: clearing property crimes is difficult, and many reported incidents never even pass the initial department screen of being worth the bother of opening a file.[198] At the same time, though, there is concern that the type of investigative follow-up that should occur after obtaining a "hit" in such cases is frequently absent: no gathering of elimination samples to establish that a recovered DNA profile was not the result of lawful access, much less investigation to provide evidence confirming the plausibility of the database-generated suspect.[199]

Of course, there is at least one institutional check on police discretion in all criminal cases—the prosecutor, who has the authority to accept or decline a case for prosecution, and must (as a matter of both legal and ethical constraint) assess anew whether probable cause supports charging the case. Prosecutors' offices can, and sometimes do, undertake efforts to make police more or less sensitive to postarrest dynamics, with the aim of improving the quality of cases submitted, decreasing declination rates and plea bargaining, and freeing up prosecutorial resources.[200] But there is good reason to suspect

---

195. Joseph L. Peterson et al., *The Uses and Effects of Forensic Science in the Adjudication of Felony Cases*, 32 J. FORENSIC SCI. 1730, 1733–34 (1987).

196. Schroeder & White, *supra* note 194, at 337.

197. *See* ROMAN ET AL., *supra* note 110, at 13–18 (discussing the components of CODIS and their investigatory uses).

198. *See, e.g.*, William K. Rashbaum & Joseph Goldstein, *DNA Match Tying Protest to 2004 Killing Is Doubted*, N.Y. TIMES, July 11, 2012, http://www.nytimes.com/2012/07/12/nyregion/suspected-dna-link-to-2004-killing-was-the-result-of-a-lab-error.html?_r=0 (noting that DNA evidence is particularly useful in property crime cases in which the police have no suspect).

199. *See, e.g.*, *id.* (discussing an unsolved murder case where laboratory error likely caused a false positive); *Texas Crime Labs Ill-Equipped to Handle Coming Volume of Touch-DNA Cases*, GRITS FOR BREAKFAST (July 15, 2012, 9:42 AM), http://gritsforbreakfast.blogspot.com/2012/07/texas-crime-labs-ill-equipped-to-handle.html (expressing concern that crime labs will be unable to keep up with growth in demand for their services).

200. *See, e.g.*, Marc L. Miller & Ronald F. Wright, *The Black Box*, 94 IOWA L. REV. 125, 148–53 (2008) (discussing factors prosecutors use when deciding where to allocate their limited resources); Ronald Wright & Marc Miller, *The Screening/Bargaining Tradeoff*, 55 STAN. L. REV.

that where such procedures do not exist as a matter of office policy or culture, prosecutors have limited power or inclination to prompt greater police effort. Despite the vaunted duty to see "that justice shall be done,"[201] the overriding reality of most prosecutors' work is to see that dockets are managed. Prosecutors have their own press of caseloads, only a small number of which will go to trial—at which time forensic evidence might well be viewed as critical whenever it is available, if only to deliver on juror expectations.[202] Scrutiny of a police case file is typically light as a general matter, and perhaps even more so in regard to scientific evidence.[203] The limited available studies of the effect of forensic evidence on charging decisions show that prosecutors screen, if at all, simply for the possibility of forensic analysis, not for its existence or content.[204] And so, again, the incentive in most cases will be to pursue scientific evidence late if at all, and only as a corroborator; otherwise, forensics are far more hassle than they are worth.[205]

Consider what is likely to be lost when forensic analysis occurs late in an investigation—postarrest or postcharge. To the extent that scientific evidence is susceptible to databasing (chiefly DNA, fingerprint, and firearms), delayed analysis means a missed opportunity to identify suspects or links to other crimes, or to save investigative resources in generating those leads. To the extent that scientific evidence has substantial *dis*confirming ability—i.e., to exclude suspects identified through other means—delay in analysis risks a prolonged investigation into an innocent individual.[206] Late

---

29, 30–35 (2002) (advocating enhanced "screening" by prosecutors to cut down on the number of negotiated plea bargains that take place).

201. Berger v. United States, 295 U.S. 78, 88 (1935).

202. *See, e.g.*, Donald E. Shelton, *Juror Expectations for Scientific Evidence in Criminal Cases: Perceptions and Reality About the "CSI Effect" Myth*, 27 T.M. COOLEY L. REV. 1, 23 (2010) (finding that jurors have high expectations that they will be presented with scientific evidence); Tom R. Tyler, *Viewing* CSI *and the Threshold of Guilt: Managing Truth and Justice in Reality and Fiction*, 115 YALE L.J. 1050, 1083 (2006) (asserting that *CSI* might raise juror standards and noting that jurors may view scientific evidence as overly conclusive).

203. *See, e.g.*, Daniel Givelber, *Meaningless Acquittals, Meaningful Convictions: Do We Reliably Acquit the Innocent?*, 49 RUTGERS L. REV. 1317, 1362 (1997) ("Unless the police report on its face reveals an inconsistency or barrier to conviction, the prosecutor accepts the general conclusion of the police without making an independent investigation or evaluation of the evidence." (quoting LLOYD L. WEINREB, DENIAL OF JUSTICE: CRIMINAL PROCESS IN THE UNITED STATES 58 (1977)).

204. Peterson et al., *supra* note 195, at 1752.

205. *See, e.g.*, Joshua K. Marquis & Velva M. Walter, *CSI Effect—Does It Really Exist?*, TALKING JUST. (Oct. 16, 2007, 3:50 PM), http://web.archive.org/web/20080821125119/http://communities.justicetalking.org/blogs/day17/archive/2007/10/16/csi-effect-does-it-really-exist.aspx (noting that forensic crime shows have prompted expensive tests even when the defendant was "caught in the act" by police and witnesses).

206. *See, e.g.*, Brief of Amici Curiae New York County District Attorney's Office & the New York City Office of the Chief Medical Examiner in Support of Respondent at 10, Williams v. Illinois, 132 S. Ct. 2221 (2012) (No. 10-8505) (reporting that "over a recent twelve-month period, nearly one in ten suspect profiles tested by the OCME for the Manhattan DA's Office resulted in an exoneration").

analysis also compromises the ability of both prosecutors and defense attorneys to fully vet the strength of the overall case and of the scientific evidence in particular, potentially distorting charging decisions, plea negotiations, and of course, trial preparations.[207]

There are also more subtle, and perhaps more pervasive (though admittedly not fully understood) concerns, stemming from the risk that systematic skews in how police and prosecutors process information and make decisions could thwart or even pervert the objective value of scientific evidence. A warehouse of research in behavioral psychology suggests that once individuals form particular beliefs, they pursue and process information in a manner that aims to verify that belief.[208] Trained professionals—police and prosecutors among them—are hardly immune to such cognitive biases; indeed, the dynamic of "tunnel vision" in criminal investigations has long been acknowledged.[209] Most relevant, perhaps, is the category of "confirmation bias" that leads individuals systematically to both credit and pursue information that is consistent with prior belief, and to discredit and ignore

---

207. *See* Daniel Givelber, *Lost Innocence: Speculation and Data About the Acquitted*, 42 AM. CRIM. L. REV. 1167, 1180–81 (2004) (describing how prosecutors and defense attorneys rely only on information, including forensic evidence analysis, that is readily available when making pretrial decisions).

208. Raymond S. Nickerson, *Confirmation Bias: A Ubiquitous Phenomenon in Many Guises*, 2 REV. GEN. PSYCHOL. 175, 176 (1998).

209. *See, e.g.*, Karl Ask et al., *The 'Elasticity' of Criminal Evidence: A Moderator of Investigator Bias*, 22 APPLIED COGNITIVE PSYCHOL. 1245, 1246 (2008) (discussing the problem of confirmation bias in criminal investigations); Karl Ask & Pär Anders Granhag, *Motivational Bias in Criminal Investigators' Judgments of Witness Reliability*, 37 J. APPLIED SOC. PSYCHOL. 561, 579–80 (2007) (finding disconfirmation bias regarding witness statements in a study group of Swedish investigators); Barbara O'Brien, *Prime Suspect: An Examination of Factors That Aggravate and Counteract Confirmation Bias in Criminal Investigations*, 15 PSYCHOL., PUB. POL'Y & L. 315, 328–29 (2009) (demonstrating that in a study group of college students the act of naming a suspect produced confirmation-bias effects, including believing or valuing available evidence to the extent that it confirmed suspect guilt); Barbara O'Brien, *A Recipe for Bias: An Empirical Look at the Interplay Between Institutional Incentives and Bounded Rationality in Prosecutorial Decision Making*, 74 MO. L. REV. 999, 1033 (2009) (noting that prosecutors, like other professionals, are not immune from cognitive bias); *see also* Bates, *supra* note 120, at 41 (noting that "tunnel vision" can result in the loss of evidence); D. Kim Rossmo, *Failures in Criminal Investigation*, POLICE CHIEF (Oct. 2009), http://www.policechiefmagazine.org/magazine/index.cfm?fuseaction=display_arch&article_id=1922&issue_id=102009 (giving an example of "tunnel vision" leading to a wrongful conviction). Indeed, studies have demonstrated that professionals with specialized expertise in a relevant area can exhibit *exacerbated* cognitive limitations—as when police trained in interrogation demonstrated *inferior* capacity to judge truth and deception in suspects as compared to untrained college student subjects. *See, e.g.*, Ask & Granhag, *supra*, at 579–80 (identifying confirmation bias in controlled study of trained investigators); Saul M. Kassin & Christina T. Fong, *"I'm Innocent!": Effects of Training on Judgments of Truth and Deception in the Interrogation Room*, 23 LAW & HUM. BEHAV. 499, 511–12 (1999) (finding that trained investigators were more likely than lay people to incorrectly "detect" deception in suspects); Christian A. Meissner & Saul M. Kassin, *"He's Guilty!": Investigator Bias in Judgments of Truth and Deception*, 26 LAW & HUM. BEHAV. 469, 478–79 (2002) (same); Eric Rassin, *Blindness to Alternative Scenarios in Evidence Evaluation*, 7 J. INVESTIGATIVE PSYCHOL. & OFFENDER PROFILING 153, 162 (2010) (finding that trained investigators, in contrast with laypeople, exhibited blindness toward alternative suspects).

information that is inconsistent.[210] Critically, the later scientific evidence is available in a criminal investigation, the more strongly police, as well as prosecutors, are likely to be intellectually committed to a particular understanding of the case; indeed, where scientific analysis of evidence is not done until after charges are filed, the commitment is strongly to a particular suspect.[211]

Cognitive bias of this sort is likely to have particularly perverse effects with respect to precisely the types of forensic evidence that, from a reliability-enhancing perspective, we should be most concerned about: exculpatory science, and science that is *less* than the "gold standard." On the former count, confirmation bias and tunnel vision have been widely accepted as causes of erroneous disregard, rejection, or recharacterization of exculpatory evidence by both police and prosecutors,[212] and the anecdotal evidence is that the force of science does not render forensic evidence immune to this pressure. Far from it, as the cases of thirty-four DNA exonerees convicted in spite of the presence of known exculpatory evidence demonstrate.[213] In Jeffrey Deskovic's case, DNA evidence obtained post-charge, which established he was not the source of semen found in the classmate he was accused of raping and murdering, was disregarded in the face of a (now known to be) false confession.[214] The late scientific evidence was reconciled with prior investigative commitment by adopting a baseless theory about the young victim's sexual history.[215] In Neil Miller's case, pre-DNA blood group testing that excluded him as a source of the semen found on the bed where he was accused of raping a woman was disregarded in

---

210. *See* Karl Ask & Pär Anders Granhag, *Motivational Sources of Confirmation Bias in Criminal Investigations: The Need for Cognitive Closure*, 2 J. INVESTIGATIVE PSYCH. & OFFENDER PROFILING 43, 47 (2005) (proposing that time pressure, norms of decisiveness, and early commitment to suspects impel need for "cognitive closure" by investigators); O'Brien, *supra* note 209, at 316 (discussing "confirmation bias").

211. Alafair S. Burke, *Improving Prosecutorial Decision Making: Some Lessons of Cognitive Science*, 47 WM. & MARY L. REV. 1587, 1605–06 (2006) (suggesting that because prosecutors have and presumably discharge an ethical duty to charge a defendant only if they are reasonably certain of that defendant's guilt, "[i]f additional evidence arises, selective information processing comes into play"); *see also* ECK, *supra* note 136, at 156–57 (elucidating the shift in police behavior in an investigation once a prime suspect is identified, at which point the focus becomes proving the prime suspect's involvement in the crime); Keith A. Findley & Michael S. Scott, *The Multiple Dimensions of Tunnel Vision in Criminal Cases*, 2006 WIS. L. REV. 291, 308 ("When what a person expects to see is the result of the person's own generation of hypotheses . . . , the personal investment in those hypotheses will reinforce the tendency to perceive or overvalue confirming information and to miss or irrationally undervalue disconfirming information.").

212. *See, e.g.*, GARRETT, *supra* note 2, at 266–68 (recounting how "cognitive bias" changed the conclusions arrived at by a group of experts); *see also infra* notes 214–19 and accompanying text.

213. *See generally* GARRETT, APPENDIX TO CONVICTING THE INNOCENT, *supra* note 14 (describing the invalid evidence presented in these thirty-four cases).

214. *See* LESLIE CROCKER SNYDER ET AL., REPORT ON THE CONVICTION OF JEFFREY DESKOVIC 1–3 (2007), *available at* http://www.westchesterda.net/Jeffrey%Deskovic%20Comm%20Rpt.pdf (summarizing the facts of the investigation which culminated in Deskovic's arrest).

215. *Id.* at 27–29.

favor of the victim's identification, and information from the crime laboratory that semen from strikingly similar rape cases came from an individual with the same rare blood group type was ignored by police.[216] Prosecutors argued that the source must have been a boyfriend, whose own blood was never collected or tested; DNA testing more than nine years after Miller's conviction revealed that the source was not a boyfriend but the true perpetrator, a serial rapist.[217] In James Edwards's murder trial, jurors were told that blood evidence collected at the crime scene that was inconsistent with the victim and Edwards was left by a store employee—but again, no testing was ever conducted to verify that theory, which turned out to be false.[218] The list goes on.[219]

Confirmation bias can also, perversely, lead investigators or prosecutors to embrace less probative or less reliable evidence than what objective scientific evaluation would counsel. Consider, for example, the Massachusetts case of Edmond Burke, in which despite DNA testing that excluded him as the source of saliva found on a rape victim, Burke was arrested and prosecuted on the strength of forensic odontology examination that matched his teeth with the bite mark that caused the saliva to be deposited.[220] The controversial use of canine scent lineups often reflects this dynamic as well, as dog scent experts have a curious habit of intervening late in cases in which other forensic evidence has proved unsuccessful in linking identified suspects to crimes.[221] The concern is obvious: that the heat of pursuit (of a suspect or a theory) blinds investigators or prosecutors to the inconvenient but more fair-minded conclusion that the best available science counsels a new investigative path.

None of this should be taken to suggest, though, that the introduction of scientific evidence into an investigation at an earlier stage necessarily

---

216. Brandon L. Garrett & Peter J. Neufeld, *Invalid Forensic Science Testimony and Wrongful Convictions*, 95 VA. L. REV. 1, 82 (2009) (summarizing the Miller case).

217. *Id.*; GARRETT, APPENDIX TO CONVICTING THE INNOCENT, *supra* note 14, at 26.

218. Ruth Fuller & Dan Hinkel, *Convict Cleared in '94 Waukegan Slaying*, CHI. TRIB., May 30, 2012, http://articles.chicagotribune.com/2012-05-30/news/ct-met-edwards-charges-dropped-20120530_1_dna-evidence-murder-case-reckling-case.

219. *See, e.g.*, *Ex parte* Brandley, 781 S.W.2d 886, 890 (Tex. Crim. App. 1989) (berating the lead investigator of a rape-murder for refusing to entertain the possibility of Brandley's innocence and to follow leads pointing towards three janitors who saw the victim just before the attack); N.Y. STATE BAR ASS'N, FINAL REPORT, *supra* note 117, at 91–96 (describing multiple cases including those of Scott Fappiano, Anthony Faison, Hector Gonzalez, Charles Shepard, and the Central Park jogger defendants, in which known inconsistencies between evidence and theory of guilt existed, and additional potentially exculpatory forensic testing was not conducted); Garrett & Neufeld, *supra* note 216, at 81–83 (describing cases in which police and prosecutors failed to conduct elimination or comparison testing).

220. Burke v. Town of Walpole, 405 F.3d 66, 81–85 (1st Cir. 2005).

221. *See* Winfrey v. State, 291 S.W.3d 68, 69–74 (Tex. App.—Eastland 2009, pet. granted) (describing how multiple laboratory-based methods of analysis failed to connect crime scene evidence and individuals who police had developed as suspects through informants, leading to multiple dog scent lineups, each following interviews with informants allegedly reporting inculpatory information), *rev'd by* 323 S.W.3d 875 (Tex. Crim. App. 2010).

obviates concerns about the cognitive biases of investigators and prosecutors. Consider the textbook "early science" circumstance: development of a suspect following collection of evidence from a crime scene, immediate laboratory submission, and testing that permits database searching for the source of the evidence—the owner of a resulting DNA profile or a detected latent fingerprint. The Brandon Mayfield case illustrated in rather notorious fashion that such early science is far from infallible—and indeed that it might trigger investigative responses that are less truth seeking and skeptical than they should be, or might have been *absent* scientific evidence. In Mayfield's case, multiple FBI fingerprint examiners erred in "matching" Mayfield's prints to those detected on evidence recovered from the scene of the 2005 Madrid subway bombing, and Mayfield, an Oregon attorney who happened to be Muslim and somewhat active in the local Islamic community, was wrongly arrested for the crime.[222] While the laboratory-based error garnered the most attention (and critical response), another feature of the investigation is relevant for present consideration: after obtaining the results of the fingerprint comparison, FBI agents obtained a warrant for Mayfield's arrest based in part on allegations that he was likely to flee, given that he had no passport yet had previously traveled to Madrid illegally.[223] In other words, what turned out to be an exculpatory fact—that government agencies had no record of Mayfield ever having *been* in Spain—was resolved against Mayfield on the strength of the fingerprint "match."[224]

Cases like this illustrate the risk that forensic science's "apparent credibility[] leav[es] the process of detection, evidence gathering and investigation hidden. The canopy of science obscures the primitive analytic tools that persist."[225] In Mayfield's case, and in most other known instances of questionable investigative paths following early suspect development through databases, we might have expected the high-profile and violent nature of the crimes at issue to trigger the greatest investment of investigative resources, as well as the most vigorous testing by defense counsel. The push to expand the tools of early suspect development through the tools of forensic science, including in property offenses and other high-volume crime, would seem to intensify these risks. Consider the recent exoneration of Dwayne Jackson in Las Vegas, convicted following a guilty plea to robbery,

---

222. OFFICE OF THE INSPECTOR GEN., U.S. DEP'T. OF JUSTICE, A REVIEW OF THE FBI'S HANDLING OF THE BRANDON MAYFIELD CASE 1–2 (2006), *available at* http://www.justice.gov/oig/special/s0601/final.pdf.

223. *Id.* at 2, 58.

224. *See id.* at 252 ("The only factual underpinning for th[e] inference [that Mayfield had false travel documents] was the existence of a fingerprint believed to be Mayfield's on the plastic bag [of bomb detonators].").

225. P.K. Manning, *Technology's Ways: Information Technology, Crime Analysis and the Rationalization of Policing*, 1 CRIM. JUST. 83, 84 (2001); *see also* Cole & Lynch, *supra* note 80, at 48–49 (discussing cold hits in DNA and fingerprint analysis); William C. Thompson, *Tarnish on the 'Gold Standard': Understanding Recent Problems in Forensic DNA Testing*, CHAMPION, Jan./Feb. 2006, at 10, 13–14 (describing cold hits that resulted from cross contamination).

following an investigation in which the only evidence against him was a CODIS hit.[226] Jackson was innocent, as police discovered years later when crime scene samples originally matched to Jackson yielded another CODIS hit to a different individual; the source of the error was a laboratory sample mix-up.[227] In a resource-restricted environment, scientific evidence may well cut short investigation that is still called for, and with far more appealing plea offers coming in these cases, we should expect less-than-full adversarial testing of that work, as Dwayne Jackson's case tragically illustrates.[228]

3. *Enter NAS.*—Leading law enforcement researchers surveying the current state of affairs with respect to forensic science usage have concluded, "[P]hysical evidence is not collected in most cases investigated by police; when it is collected, much of it is not scientifically analyzed; and when it is analyzed, it is used not to promote investigative efficiency, but rather to bolster prosecutorial proceedings."[229] The previous two Parts explain why that state of affairs exists and why it is likely to persist.

The NAS Report itself has little to offer in response to the concerns raised by these observations. Simply creating the conditions for more and better quality science to be produced by laboratories will not address critical features of underproduction, underutilization, and qualitatively suboptimal exploitation of forensic science in the criminal justice system. To be sure, there are systemic resource constraints that, if addressed at the laboratory level, are likely at least partly to improve the state of affairs with respect to suppressed rates of evidence submission and analysis usage; undoubtedly, this behavior is at least in part a resigned response to laboratory backlogs.[230] But the causes are likely more endogenous, and more entrenched, as well.[231] Affirmative rather than trickle-down intervention appears to be called for.

---

226. Jackie Valley, *Metro Reviewing DNA Cases After Error Led to Wrongful Conviction*, LAS VEGAS SUN, July 7, 2011, http://www.lasvegassun.com/news/2011/jul/07/dna-lab-switch-led-wrongful-conviction-man-who-ser (discussing the wrongful conviction of Dwayne Jackson after the laboratory switched samples and wrongly matched him to crime scene evidence); Lawrence Mower & Doug McMurdo, *Las Vegas Police Reveal DNA Error Put Wrong Man in Prison*, LAS VEGAS REV.-J. (July 7, 2011, 10:22 AM), http://www.lvrj.com/news/dna-related-error-led-to-wrongful-conviction-in-2001-case-125160484.html (same).

227. Mower & McMurdo, *supra* note 226; Valley, *supra* note 226.

228. *See supra* notes 196–99 and accompanying text (describing problems with respect to DNA in property crimes).

229. Horvath & Meesig, *supra* note 136, at 965.

230. NAS REPORT, *supra* note 7, at 37.

231. For whatever self-reporting is worth in this context, it is instructive that recent surveys of forensic science "customers" reveal laboratory wait times to be only one of many reasons cited for nonsubmission of evidence. *See* Pratt et al., *supra* note 170, at 39–41 (providing numerous reasons for nonsubmission of evidence including lack of funding, expectation of a guilty plea, and no identifiable suspect). There is also good reason to doubt that laboratory backlogs are likely to substantially diminish, at least in the near future; indeed, despite significant dedicated federal funding prior to and since the NAS Report, they have not yet shrunk. *See* MARK NELSON, NAT'L INST. OF JUSTICE, MAKING SENSE OF DNA BACKLOGS, 2010—MYTHS VS. REALITY 3 (2011), *available at* https://www.ncjrs.gov/pdffiles1/nij/232197.pdf (presenting graphs showing steady

Moreover, the concern is not just that the reform agenda of the NAS Report might not address critical concerns, but that it might in some instances unwittingly confound them. Two features of the NAS Report's proposals are illustrative. Consider first the NAS Report's commitment to putting more scientific evidence into the hands of investigators early in investigations. This is perhaps most directly reflected in its call for expanded interoperability of Automated Fingerprint Identification Systems (AFIS) and its conclusion that a range of identification-orientated forensic disciplines (such as hair microscopy, handwriting analysis, or forensic odontology) might be well suited for investigative use although they may never be validated sufficiently to support the claims of individualization currently made in courtroom testimony.[232] Investigations might, for example, be aided by the ability of these "second-best" techniques to exclude suspects or narrow an otherwise wide field of suspects (to a manageable number if not one), especially where there is not opportunity (for lack of appropriate evidence) or resources (for lack of time or laboratory capacity) for preferred techniques such as DNA analysis, and especially if training of law enforcement actors on the scientific limits of such techniques is expanded.[233] These may well be right-headed proposals on balance. Particularly with regard to AFIS, it is likely that the expanded ability to generate suspects based upon fingerprints recovered from crime scenes—a capacity currently limited by a surprising array of proprietary barriers to interoperability of multiple jurisdictions' computerized fingerprint databases[234]—would permit police to pursue the investigation of cases (property crimes in particular) that are currently closed without resolution for lack of evidence. Moreover, it might well be that a range of forensic disciplines provide, on balance, more reliable inclusions of suspects than other more traditional forms of evidence (think, especially, eyewitness evidence)[235]—and certainly might permit highly reliable *exclusion*.

But the preceding discussion suggests that the NAS Report's ideal vision for how expanded emphasis on and access to these "second-best"

---

increase in productivity and steady increase in submissions leading to persistent increase in backlogs). Expanded collection of DNA from convicted and, increasingly, arrested individuals—a practice that may well dramatically increase if the Supreme Court ratifies the authority of police to collect DNA from arrestees in the case raising that issue this Term—has a major impact on laboratory resources in this regard. *See* Maryland v. King, 133 S. Ct. 1, 2 (2012) (discussing the likelihood of the Court granting certiorari in a case raising the issue of DNA collection); *see also* DUROSE ET AL., *supra* note 23, at 4 (reporting that 75% of DNA analysis requests were for arrestee and offender samples).

232. NAS REPORT, *supra* note 7, at 127.

233. *See id.* at 234–37 (proposing expansion of "education" for users as well as producers of forensic science).

234. *Id.* at 272–75 (discussing technical challenges in working towards AFIS interoperability).

235. So, for example, it may be that hair examiners can say with a higher degree of confidence that two hairs are consistent with being from the same source, than an eyewitness can say that a face seen in a lineup is in fact the face she saw at a crime scene; the latter would be more discriminating if true, but it might be highly likely to be false.

techniques will enhance investigations is simply unlikely to come to materialize—at least absent concerted intervention (beyond "education")[236] in the conditions and culture of most investigative environments. Cases like Edmond Burke's illustrate the possibility that the limited capacity of investigators to fair-mindedly evaluate competing scientific evidentiary conclusions in a case diminishes the prospect for scientific evidence to be put to the best possible use in a criminal case, and creates a risk that second-best science will lock investigators into error.[237] Thus, an early "inclusion," however nonexclusive, might trigger a level of *commitment* by investigators that is not easily revised—even by putatively more valid or reliable scientific evidence.[238] Second-best science might also be particularly likely to interact in undesirable ways with traditional, nonscientific investigative techniques. The motivations and pressures that cause investigators to have a "restricted" view of the value of forensic evidence as primarily a tool of confirmation rather than investigation, and to typically view the best investigative path as the one representing the shortest distance to case closure, create the risk that forensic evidence of this sort will be seen as most usefully deployed to enhance tried-and-true investigative methods—suspect interviews and interrogations among them.[239] This could well be an effective strategy from the standpoint of closing cases, as scientific evidence has been established as an important factor in obtaining confessions.[240] From the standpoint of reliability concerns, however, this should give us pause; scientific evidence has the capacity to convince both the guilty and the innocent that confession is in their interests,[241] and confessions in turn diminish the capacity to independently scrutinize the reliability of the forensic evidence that prompts them.[242]

Consider also the NAS Report's much-debated call for laboratory independence. Critically, apart from the merits of the recommendation, its very meaning has been the subject of significant debate. While strong proponents of crime laboratories operating wholly outside of law enforcement control have claimed the NAS Report adopted their viewpoint,

236. *See supra* note 233.

237. *See supra* notes 208–28 and accompanying text.

238. *See* Ask & Granhag, *supra* note 210, at 47 (discussing the role of "commitment" in triggering effects of cognitive bias).

239. Horvath & Meesig, *supra* note 136, at 966.

240. *See* RICHARD A. LEO, POLICE INTERROGATION AND AMERICAN JUSTICE 143–47 (2008) (discussing interrogation techniques involving scientific or allegedly scientific evidence that law enforcement officers use in order to elicit confessions).

241. *Id.* at 147; *see also* TRUE STORIES OF FALSE CONFESSIONS 193–202 (Steven A. Drizin & Rob Warden eds., 2009) (describing false confessions in the Central Park jogger case).

242. *See, e.g.*, GARRETT, *supra* note 2, at 25–26 (discussing the case where a man named Douglas Warney made a false confession); Dan Simon, *The Limited Diagnosticity of Criminal Trials*, 64 VAND. L. REV. 143, 181 (2011) (explaining that when one item in the chain of evidence is erroneous the result can be an "escalation of error" because "evidence items are not truly independent of one another").

others—in particular, groups representing law enforcement interests—have argued that a more modest accommodation, "administrative independence" within a law enforcement organization, is consistent with the Report's aims.[243] Supporters of the NAS Report's agenda have been largely dismissive of this minimalist take on "independence," viewing its advocates (somewhat understandably) as tainted by the very dynamics of control that motivated the recommendation in the first place.[244]

Politics aside, a systemic view of forensic science oversight suggests that there are very real concerns raised about the design of an independent laboratory system that should be considered in connection with reform. How would fully independent crime laboratories assist, if at all, with evidence collection and crime scene response? To the extent a role is envisioned, would independence delay response or undermine working relationships between law enforcement officers and their "independent" colleagues? Will evidence collection instead fall within the domain of law enforcement, and do resources exist to manage that transition?[245] In the course of an investigation, would fully independent laboratories mean more formalized submission and prioritization decisions for testing? Would discretion in regard to submission and prioritization be committed to investigators or to analysts? These are critical questions to be answered in connection with the independence proposal.

Of course, as Part I aimed to show, there were fair reasons why the NAS Report itself did not squarely take on these concerns, and thereby assume a more systemic (and correspondingly disruptive) posture toward the role of forensic science in the criminal justice system. Nevertheless, these concerns cannot fall out of the reform agenda that is solidifying in the Report's

243. *Compare* Paul C. Giannelli, *Independent Crime Laboratories: The Problem of Motivational and Cognitive Bias*, 2010 UTAH L. REV. 247, 262–66 (supporting complete independence but proposing more limited alternatives that might achieve those aims), *with* Kenneth E. Melson, *Embracing the Path Forward: The Journey to Justice Continues*, 36 NEW ENG. J. ON CRIM. & CIV. CONFINEMENT 197, 217 (2010) ("Advocacy groups critical of forensic science have latched on to the complete removal of crime laboratories from law enforcement. . . . Several organizations oppose the removal of crime laboratories from law enforcement agencies, but support different degrees of autonomy within the parent law enforcement agencies.").

244. *See* Giannelli, *supra* note 47, at 89–90 (concluding that the NAS Report's recommendations are warranted and that "[t]he government has not only failed to conduct the needed research, it has thwarted any effort to do so").

245. PETERSON ET AL., *supra* note 126, at 76 (listing five basic considerations in prioritizing evidence); ROMAN ET AL., *supra* note 110, at 9, 349–50 (discussing consequences of delayed crime laboratory response to crime scenes, poor working relationships among police and crime scene responders, and the need to tailor individual training protocols to particular institutional cultures); STATE OF N.Y. OFFICE OF THE INSPECTOR GEN., REPORT OF INVESTIGATION OF THE MONROE COUNTY PUBLIC SAFETY LABORATORY 30–32 (2012), *available at* http://www.ig.ny.gov/pdfs/MonroeCountyLaboratoryReport.pdf (discussing significant laboratory management failures); Horvath & Meesig, *supra* note 136, at 966–67 (urging that "improvements in the mutual exchange of information among investigators and others involved in the collection, analysis, and use of physical evidence would enhance the value" of that evidence more than reforming the identity or institutional role of particular actors in collection and analysis).

aftermath, and that is likely to dictate the terms of both policy response and academic inquiry going forward. In this regard, there is value in some comparative reflections. Consider the recent government-commissioned inquiry into the state of pediatric forensic pathology in Ontario, Canada, prompted by a number of high-profile wrongful convictions and other events indicating serious deficiencies in forensic pathology practice. The Goudge Report, as the massive four-volume document is known, examined and issued 169 recommendations concerning the forensic pathology field, including proposals largely analogous to those of the NAS Report, touching on training, education, professional protocols, and oversight of practitioners.[246] But while "much of the focus must be on forensic pathologists and the issues surrounding their training, education, accreditation, oversight, and accountability," the Report includes "specific recommendations . . . designed to ensure that [police, prosecutors, and defense attorneys] will be as effective as possible in th[e] task" of "protecting the public against the introduction of flawed or misunderstood pediatric forensic pathology into the system."[247] In recognition of the fact "that other participants in the criminal justice system have important roles to play in protecting the public against the introduction of flawed or misunderstood pediatric forensic pathology into the system," the Report's recommendations also target police and prosecutors (among others), and encourage changes to practices in the substance and timing of information flow among pathologists, coroners, and police, increased training for police with regard to pediatric forensic death investigation and general confirmation bias in investigation, and special considerations for prosecutors in regard to case preparation and disclosure issues.[248]

Consider also the British example. In recent years, the British government has produced at least two reports on scientific evidence in criminal investigations that have examined the roles of producers and users as integrated functions in the field of forensic science.[249] This integration extends as well to national oversight of forensic science, which has long been a feature of the British system, most recently through the roles of the Forensic Science Regulator and the Forensic Science Advisory Counsel.[250] In addition to the promulgation of standards for forensic science practice at

246. STEPHEN T. GOUDGE, INQUIRY INTO PEDIATRIC FORENSIC PATHOLOGY IN ONTARIO 437–57 (2008), *available at* http://www.attorneygeneral.jus.gov.on.ca/inquiries/goudge/report/index.html.

247. *Id.* at 46, 437; *see also id.* at 437–69 (devoting an entire chapter of recommendations to this concept).

248. *Id.* at 437–57.

249. *See* UNDER THE MICROSCOPE, *supra* note 125, at vii (explaining that the purpose of the 2002 report is to examine the police response to its first iteration, released in 2000).

250. *See Forensic Science Regulator*, HOME OFFICE, http://www.homeoffice.gov.uk/agencies-public-bodies/fsr (describing the function of the independent Forensic Science Regulator located within the British Home Office).

the laboratory level, the Forensic Science Advisory Council has also undertaken the task of promulgating "End-User Requirements" for forensic science services, directed to the roles of police, prosecutors, judges, and others in utilizing scientific evidence.[251] Matters addressed by the End-User Specialist Group include, for example, codes of conduct relating to crime scene canvassing, the appropriate roles for evidence gatherers, and proper interactions between police and forensic pathologists.[252]

This is not, of course, to suggest that the Canadian or British models of forensic science oversight ought to be, or even could be, borrowed by the United States, particularly given the important differences between these countries and the United States in terms of criminal justice administration as a general matter. Both Canada and the United Kingdom feature far more centralized bureaucratic control over policing and criminal adjudication in general and forensic science in particular, and both have many fewer provincial and local police and prosecutorial organizations.[253] Further, in the United Kingdom, policing has, as a general matter, been at the vanguard of a broader trend toward data-based accountability for government services for the last three decades,[254] which has provided incentive and opportunity for creating and funding the capacity to scrutinize and regulate law enforcement use of forensic science.[255] Nevertheless, the Canadian and British examples show that the instinct to view oversight through a more systemic lens is not quite so counterintuitive as it would seem when set against the dominant American paradigm.

---

251. *See* Forensic Sci. Regulator, Terms of Reference for the End User Specialist Group (Mar. 21, 2011), *available at* http://www.homeoffice.gov.uk/publications/agencies-public-bodies/fsr/end-user-sg-terms-of-ref?view=Binary.

252. *See* End-User Specialist Grp., Minutes of Meeting (Mar. 2, 2010), *available at* http:// www.homeoffice.gov.uk/publications/agencies-public-bodies/fsr/end-user-sg-02032010mins?view=Binary (discussing a regulatory structure and code of practice for forensic science).

253. *See* Kent Roach & M.L. Friedland, *Borderline Justice: Policing in the Two Niagaras*, 23 AM. J. CRIM. L. 241, 255–60 (1996) (differentiating the local police structures in Canada and the United States); Kent Roach, *Forensic Science and Miscarriages of Justice: Some Lessons from Comparative Experience*, 50 JURIMETRICS 67, 75 (2009) (detailing the more centralized structure of the criminal justice systems in Canada and the United Kingdom and its impact on the regulation of forensic science in both nations); *cf.* Matthew C. Waxman, *Police and National Security: American Local Law Enforcement and Counterterrorism After 9/11*, 3 J. NAT'L SECURITY L. & POL'Y 377, 387–89 (2009) (contrasting American and British institutional relationships concerning policing and the resulting challenges in information sharing and oversight).

254. *See* ROBIN WILLIAMS & PAUL JOHNSON, GENETIC POLICING: THE USE OF DNA IN CRIMINAL INVESTIGATIONS 102–05 (2008) (discussing the trend).

255. *See id.* (analyzing the application of economic efficiency-based principles to the forensic science context and recommendations deriving from the "New Public Management" era in the United Kingdom). This was largely the outgrowth of a Thatcher-era embrace of New Public Management and related strategies of broad privatization. *See generally* Paul Roberts, *What Price a Free Market in Forensic Science Services?: The Organization and Regulation of Science in the Criminal Process*, 36 BRIT. J. CRIMINOLOGY 37 (1996) (examining the tension between free market reforms and the need for effective regulation).

III. Another Path Forward: Systemic Integration of Forensic Science

The primary goal of this Article is diagnostic: it aims to reveal deficiencies in forensic science usage that are systemic, driven by forces outside the laboratory, and unlikely to be addressed—indeed, perhaps exacerbated—by the prevailing approach to reform. Developing remedies for those ills will be at least as complex and multifaceted as the dynamics that drive them, and is thus an ongoing project for future work (my own and, hopefully, that of others). But one must strike while the iron is hot. The NAS Report has generated not only academic foment but also a relatively energetic level of response from policy corners. The Report's call for national oversight looks to have been heeded in at least some form, as the Attorney General recently announced the formation of a thirty-member National Commission on Forensic Science charged with "recommend[ing] strategies for enhancing quality assurance in forensic science units."[256] Meanwhile, the last three years have seen an array of presidentially created working groups discussing the merits of the NAS Report proposals; white papers from that process, which are in turn likely to shape pending legislation, are soon to issue.[257] Legislatively, two bills have been introduced in Congress, each largely tracking the NAS Report's proposed reform agenda with regard to theoretical and applied research and expanded standard setting in the forensic sciences, as well as the creation of national oversight capacity; neither specifically incorporates or addresses the laboratory-independence recommendation.[258] To the extent the NAS Report is driving these processes—and it undoubtedly is—now is a critical time to raise a dissenting, or at least qualified concurring, voice in relation to the trajectory of the path forward.

Accordingly, this Part outlines two categories of reforms that might widen the NAS Report's path forward to address systemic concerns. The first category takes the NAS Report's proposals on their own terms and suggests ways to supplement them in critical respects that account for the concerns raised in Part II: the Report's research and standard-setting agenda should encompass police and prosecution practices such as evidence gathering, testing decisions, and disclosure regimes; the Report's "independence" recommendation should be reflected upon in light of further research and the concerns raised in Part II; and policy makers should prioritize enhancing state-level oversight as a complement, or alternative, to national oversight of the field. If the proposals are not surprising or groundbreaking, they do have the advantage of feasibility, plausibly fitting

---

256. *See* Notice of Establishment of the National Commission on Forensic Science and Solicitation of Applications for Commission Membership, 78 Fed. Reg. 12,355 (Feb. 22, 2013).

257. *See, e.g.*, CHARTER OF THE SUBCOMMITTEE ON FORENSIC SCIENCE, *supra* note 12 (describing progress of the subcommittee in creating draft reports in response to the NAS recommendations and forthcoming white paper).

258. *See supra* note 12.

into the reform landscape that is already taking shape in the aftermath of the NAS Report; happily, some have the advantage of potentially being *more* feasible than the Report's proposed reforms. The second category of proposals is more ambitious as well as more provisional, suggesting that the goals of enhancing forensic science use and integrity challenge foundational themes of existing criminal procedure doctrine.

A. *The NAS Report: A Fuller Reform Agenda*

1. Research, Standard Setting, and Training.—The lion's share of the NAS Report's recommendations aim to enhance our confidence in the forensic sciences by broadening and deepening the knowledge base concerning both the scientific foundations of its disciplines as well as the technical aspects of forensic science practice. A corollary aim is for this research to inform the development of more formal, detailed, and binding standards of practice for analysts.[259] So too should police and prosecutorial practices vis-à-vis forensic science be put under the proverbial microscope.

Much like the non-DNA forensic disciplines in the years prior to the NAS Report, the subject of police and prosecutorial practices with respect to scientific evidence has been relatively neglected by U.S. researchers and, critically, by federal research grants.[260] An extremely small community of researchers has been the recipient of such funds over the past four decades, and while its work, particularly that of Joseph Peterson, has been critical in opening the black box of forensic science at the investigative stage, the work has occurred in only a small number of jurisdictions, confounding generalization.[261] Perhaps most disappointingly, outside the limited context of expanding use of DNA analysis from homicides and sexual assaults to the field of property crimes, there has been little follow-up research to translate observed patterns of utilization into best (or at least better) practices.[262]

---

259. *See supra* Part I.

260. *See Fiscal Year 2011 Awards*, NAT'L INST. JUST., http://www.nij.gov/funding/awards/2011-table.htm (last modified June 20, 2012) (showing all NIJ grant awards, including those related to forensic science); *Fiscal Year 2010 Awards*, NAT'L INST. JUST., http://www.nij.gov/funding/awards/2010-table.htm (last modified Jan. 12, 2011) (same); *see also* Analysis of 2010–2011 Forensic Science Award Abstracts (on file with author) (showing that among 678 total grants related to Forensic Science, 363 had available abstracts to review for applicability to nonlaboratory work, and only 10.7% had any such component, almost exclusively concerning crime scene work); *see also supra* note 32 and accompanying text (describing limited pot of non-DNA-specific funding for forensic science).

261. For only a sample of Professor Peterson's field-defining work, see *supra* notes 20, 26, 102, 112, 124, 126, 186, and 195.

262. A recent NIJ-sponsored study on the use of DNA in property crimes helpfully combined quantitative and qualitative assessment in this regard, ROMAN ET AL., *supra* note 110, at 139–53, and followed up on an earlier assessment of forensic evidence utilization in two urban areas, *see* Tom McEwen, The Role and Impact of Forensic Evidence in the Criminal Justice System, Final Report 2–3 (Dec. 13, 2010), *available at* https://www.ncjrs.gov/pdffiles1/nij/grants/236474.pdf (reporting results of NIJ grant-funded research in San Diego and Denver).

Particularly glaring in its absence from the field is a substantial body of qualitative, observational studies illuminating why and how it is that the data on significant filtering of physical and scientific evidence, reported by Peterson and a small number of other researchers, came to pass. Also relatively neglected has been the issue of cognitive bias, an area susceptible to study,[263] and which the NAS Report singled out as a priority on the laboratory side of the equation.

With respect to standard setting and incorporation of best practices into policy guidance concerning forensic science, law enforcement users lag even farther behind the laboratories that are the subject of the NAS Report. Reviews of forensic science practices that have followed the NAS Report have recognized this deficiency,[264] but their calls for reform have not enjoyed the prominence of the NAS's recommendations, and have not to date influenced draft-implementing legislation. By contrast, a number of jurisdictions have in recent years been convinced to adopt (sometimes voluntarily but often following legislative mandate) detailed policies informed by the best-available science concerning eyewitness identification and interrogations—two critical tools of investigation, to be sure.[265] That similar energy has not been seen with regard to forensic evidence undoubtedly owes in large part to the comparative lack of research in that arena, though there may also be a sense that practices with respect to forensic science in the investigative or other pretrial contexts defy standardization.

As a blanket premise, this is misguided. Indeed, as with SWGs and other voluntary, nonbinding standard-setting bodies in the forensic science field, there are currently examples of voluntary efforts internal to the police and prosecution fields to propose best practices bearing on forensic evidence, though adoption has been limited.[266] Evidence collection, for example, is a field of practices that are susceptible to far more standardization than is currently seen, particularly with regard to role differentiation and

---

263. To the author's knowledge, the most illuminating research in this area with respect to police has been done in Sweden; little has been done in the United States. *See supra* note 209.

264. *See* N.Y. STATE BAR ASS'N, FINAL REPORT, *supra* note 117, at 98–99 (pointing out the discrepancies in standards between crime labs and law enforcement personnel performing forensic disciplines that are not regulated by the New York Commission on Forensic Science).

265. *See, e.g.*, *Eyewitness Identification*, INNOCENCE PROJECT, http://www.innocenceproject.org/fix/Eyewitness-Identification.php (describing jurisdictions that have adopted requirements concerning law enforcement identification procedures); *False Confessions & Mandatory Recording of Confessions*, INNOCENCE PROJECT, http://www.innocenceproject.org/fix/False-Confessions.php (discussing jurisdictions that have adopted requirements concerning law enforcement interrogation procedures); *see also* GARRETT, *supra* note 2, at 270 (discussing reform momentum).

266. *See, e.g.*, AM. BAR ASS'N, ABA STANDARDS FOR CRIMINAL JUSTICE: DNA EVIDENCE 1 (3d ed. 2007); LISA R. KREEGER & DANIELLE M. WEISS, AM. PROSECUTORS RES. INST., DNA EVIDENCE POLICY CONSIDERATIONS FOR THE PROSECUTOR 3–5 (2004), *available at* http://www.ndaa.org/pdf/dna_evidence_policy_considerations_2004.pdf (proposing case assessment team model for DNA cases); COMM'N ON ACCREDITATION FOR LAW ENFORCEMENT AGENCIES, LAW ENFORCEMENT AGENCY STANDARDS chs. 41, 42, 83 (1998) (containing standards relating to forensic evidence and investigations).

coordination among patrol, investigative, and scientific personnel. But even more investigative tasks could be treated less improvisationally than they are in the status quo. One could imagine more default standards for what evidence will be submitted prior to arrest in particular cases—similar to the sorts of investigative "checklists" that have been proposed as mechanisms to improve documentation and discovery in criminal cases.[267] Similarly, one could imagine the development of standards of practice concerning the type of forensic analysis that will, in the ordinary course, be conducted in particular types of cases, taking evaluation of both resource and reliability questions out of the hands of individual investigators in individual case contexts and committing law enforcement to standards of evidentiary evaluation that are rooted in the scientific objectivity they aim to exploit in their investigations. Perhaps most simple in theory though elusive in current practice, standards of documentation of evidence gathering by police are crucial to later efforts to reconstruct and deconstruct those sometimes-formative decisions made about investigative paths taken in a case.[268]

More structural reforms might be entertained as well. Some have proposed that laboratories be organized to assign "independent case managers" with responsibility for interfacing with law enforcement over submission and testing decisions, thereby segregating that assessment function from analysis.[269] Analogously, Rachel Barkow and others have proposed more segregation within prosecutors' offices to separate investigative from adjudicative decision making.[270] One could well imagine a more modest version of Professor Barkow's proposal taking hold with respect to decisions concerning scientific evidence, either within police departments or prosecutors' offices: a dedicated supervisor (or committee) could oversee decisions regarding forensic evidence submission and follow-up response to testing, thereby countering some of the effects of time pressures, habit, and cognitive bias that might limit an individual investigator's or prosecutor's consideration of the question. Whereas

267. *See* COMM'N ON ACCREDITATION FOR LAW ENFORCEMENT AGENCIES, LAW ENFORCEMENT AGENCY STANDARDS ch. 42 (mandating checklists for investigators); *New Perspectives on* Brady *and Other Disclosure Obligations: Report of the Working Groups on Best Practices*, 31 CARDOZO L. REV. 1961, 1974 (2010) (presenting the proceedings of the working groups endorsing use of checklists by prosecutors and police).

268. *See supra* notes 116, 154–55 and accompanying text (discussing current deficiencies in documentation of evidence gathering, and proposals by the Technical Working Group on Crime Scene Investigation, CALEA, and outside observers to formalize documentation requirements).

269. *See* O'Brien, *supra* note 209, at 1045–46 (discussing structural debiasing options and their limitations); D. Michael Risinger et al., *The* Daubert/Kumho *Implications of Observer Effects in Forensic Science: Hidden Problems of Expectation and Suggestion*, 90 CALIF. L. REV. 1, 46–47 (2002) (suggesting the creation of an evidence control unit that is staffed by "the most highly trained and highly respected personnel in the laboratory . . . responsible not only for coordinating work among examiners in different specialties, but also for being the sole contact point between the entity requesting the test and the laboratory").

270. Barkow, *supra* note 89, at 869–70.

Professor Barkow's proposal draws inspiration from administrative law,[271] the instinct here is instead to incorporate some of the lessons of science into the investigative and prosecutorial realm, at least where those realms are drawing on scientific products to do their work.[272]

Finally, just as training and education are viewed in the NAS Report as a critical piece of the reform puzzle, so too with respect to police and prosecutors. Calls from criminal justice researchers to ensure that law enforcement officials fully understand "the value of forensic evidence for *investigative* purposes,"[273] rather than just adjudicative advantage, should be heeded as part of the forensic science reform agenda. So too, though, should users of forensic science be trained in the risks of overreliance on science, including the risks of erroneous "cold hits" from databases and cognitive bias.[274]

All of these proposals posit a "fix" to some of the pathologies that emerge from the portrayal in Part II. If, accordingly, they offer the promise of better, more reliable outcomes in the criminal justice system, one could flippantly say that their advantages are self-evident. But of course, even if such a claim were backed by the sort of empirical validation that has yet to occur, certainty regarding the feasibility of implementation and magnitude of benefit will ultimately be elusive. In the meantime, the demands of these proposals in terms of behavioral change, cultural shift, and resource allocation mean that they are unlikely to be attractive candidates for adoption. That said, all of them are of a piece with analogous changes that have been seen among isolated police and prosecutor organizations, often in the wake of some public breakdown highlighting the risks that Part II describes.[275] Indeed, an advantage of the sort of sustained attention to these

271. *Id.* at 896 ("In the case of agencies, the law mandates structural separation within the agency itself or aggressive judicial review of the record to ensure unbiased decision making. . . . [A] corrective modeled along the lines of the APA's separation requirement would be feasible and desirable in the case of federal prosecutors' offices.").

272. *Cf.* Ask et al., *supra* note 209, at 1246 (noting that while good scientific practice dictates that "the person administering a procedure or recording an observation to the fullest possible extent be blind to the hypothesis underlying the study, to avoid any influence from preconceptions and preferences with regard to the results[,] this stringent standard is not feasible in criminal investigations," creating a challenge for investigators who must "discount the knowledge of the case that is irrelevant to the assessment of reliability").

273. Strom & Hickman, *supra* note 75, at 398 (emphasis added); *see also* HORVATH & MEESIG, *supra* note 131, at 112.

274. *See* GOUDGE, *supra* note 246, at 447–48 (recommending that Ontario police "be trained to be vigilant against confirmation bias in their investigative work . . . . This training is best accomplished through increased professionalism, an enhanced awareness of the risks of confirmation bias, the promotion of an evidence-based culture, and complete transparency regarding what is communicated between the police and the forensic pathologist.").

275. *See, e.g.*, Memorandum from David W. Ogden, Deputy Att'y Gen., U.S. Dep't of Justice, to Dep't Prosecutors (Jan. 4, 2010), *available at* http://www.justice.gov/dag/dag-memo.pdf (announcing mandatory discovery policies in federal prosecutors' offices in wake of *Brady* scandals including Ted Stevens's prosecution); Minjae Park, *Davis to Law Enforcement: Report Rape Kit Backlog*, TEX. TRIB. (May 30, 2012) http://www.texastribune.org/texas-legislature/82nd-legislative-

matters that this Article calls for is that such responses might be informed by research concerning best practices rather than ad hoc political priorities.

2. *Independence Reconsidered.*—In the view of many observers, prior to the NAS Report, independence was the *sine qua non* of forensic science reform.[276] As a political matter, it is unlikely that any strong form of this proposal will be implemented via any federally sponsored agenda. But some modified structural reforms aimed at the ills of bias and professional control that the Report sought to remedy may well emerge.[277] Moreover, the ideal of independence may well continue to be touted and could take hold as a matter of state-level reform. Hence, it is well to consider the lessons of this Article's systemic view for the independence debate.

As already previewed in Part II, the account herein suggests a number of risks from independence, at least in its strongest form. At the same time, it suggests that laboratory independence might not solve some of the bias concerns by its own force that, while not addressed by the NAS Report, nevertheless bedevil the integration of scientific evidence into criminal investigations.

The twin goals of building professionalism in forensic science and creating organizational firewalls against undue investigative influence over scientific analysis should not and need not obscure some of the advantages of collaboration between forensic scientists and law enforcement personnel. One of the more important contributions of a move toward organizational reform of the sort that the NAS Report calls for could and should be to identify those areas where collaboration is most advantageous, as well as those areas that, from the standpoint of reliability and integrity of investigative outcomes, are best committed to laboratory discretion and best committed to law enforcement discretion. Where tasks like evidence collection, evidence submission, and testing priority should be the subject of reasoned inquiry into best practices, not ad hoc determination. The goal in this context should be similar to what Dan Richman has argued for in relation to other dynamics of criminal investigations: "to promote teamwork" and to enable "each player [to] orient[] to his distinct institution and professional

---

session/wendy-davis-author-rape-kit-law-clarifies-intent (describing legislation enacted to mandate testing of rape kits).

276. *See, e.g.*, Giannelli, *supra* note 243, at 247–48 & nn.4–9 (citing a number of journal articles and other sources asserting that laboratories associated with law enforcement agencies suffer from inherent biases).

277. *See* Risinger, *Path Forward*, *supra* note 106, at 239 ("[A]ny hope of congressional action to coerce or encourage the establishment of independence of forensic labs from law enforcement control is also dead on arrival."); Jamie Downs, Speech at Texas Criminal Justice Integrity Unit–Texas Forensic Science Commission, Joint Conference on Forensic Science (June 4, 2012), *available at* https://www.dropbox.com/sh/z0adceaub177maq/kNqQCO_LWy/TCJIU-FSC%202012%20Forensic%20Science%20Seminar%20Video/FSS2012-auditorium-2.rm (reporting that working groups are unlikely to agree on independence recommendations).

culture" such that "interaction presents less a risk of capture than an opportunity for both productive collaboration and mutual monitoring."[278]

In other words, laboratory independence is an institutional arrangement that requires thoughtful calibration, which in turn will likely require more information than we currently possess about best practices in the division of responsibilities and collaboration between and among investigators, prosecutors, and crime laboratory personnel. Here, then, is another area where empirical work should be funded and pursued. Researchers might, for instance, investigate both quantitative and qualitative measures of how successfully law enforcement works with crime laboratories (for example, both crime-scene-response times from a crime-laboratory-centered evidence collection team, and investigators' narrative accounts of how effectively they work with such personnel). Of particular value might be comparisons that could be generated in those jurisdictions that use a variety of forensic service providers (for example, splitting evidence submission between a departmental crime laboratory and, when resources demand it, an independent private laboratory). Further research opportunities are presented by individual jurisdictions that take up the NAS Report's call to alter the organizational or administrative structures of their crime laboratories to enhance "independence." Such decision points offer natural experiments to observe and measure the effects of such changes.

3. *Oversight.*—In proposing the formation of NIFS to oversee forensic science at the national level, the NAS Report ran head-on into a long-entrenched resistance from the law enforcement and forensic science communities alike to outside regulation of the field.[279] Perhaps the only proposal that would garner more opposition would be oversight of law enforcement and prosecutors themselves. And yet, to a certain extent, the logical takeaway from Part II is that the actions of those critical players in the production and use of forensic science cannot be immune from scrutiny if the goal of enhancing the contributions of scientific evidence is to be taken seriously.

From the standpoint of the systemic view advanced by this Article, the proposed NIFS (and its apparent, at least partial embodiment in the recently announced National Commission on Forensic Science)[280] suffers from two deficiencies. The first is substantive. Oversight in the form of research, standard setting, and training that focuses on laboratory-level concerns and excludes questions of evidence procurement and usage will miss, and may confound, dynamics that are critical in ensuring that scientific evidence makes positive contributions to the criminal justice system. While it is far too soon to render a verdict on the National Commission on Forensic

278. Richman, *supra* note 17, at 813.

279. *See supra* Part I.

280. *See supra* note 256 and accompanying text.

Science, the announced scope of its mission—"enhancing quality assurance in forensic science units"—suggests the possibility that it might embody the very narrowest conception of what counts as forensic science oversight, taking the NAS Report's scope as a ceiling rather than a floor.[281]

The second and related deficiency is structural. In service of the NAS Report's goals of standardization and uniformity, *national* oversight sacrifices geographic and professional footholds in the thousands of state and local jurisdictions where its standard setting will most often play out. Indeed, as the Report acknowledges, "[o]versight of the forensic science community and medical examiner system will sweep broadly into areas of criminal investigation and prosecution";[282] yet, by design, national oversight lacks a certain context sensitivity necessary for understanding much less overseeing or reforming the work of law enforcement. Consider evidence collection and preservation, for example, which is an important issue for forensic science oversight, but which is closely tied to issues of institutional organization, local practices, and even state law.[283]

Policy makers should therefore consider supplementing national oversight with state-level institutes of forensic science (SIFSs). SIFSs could be administered through NIFS, or could be created through a more robust version of the existing Coverdell grant program, which conditions federal funds for state forensic science programs on the designation of a state agency responsible for receiving and investigating complaints stemming from the conduct of crime laboratories.[284] Apart from the Coverdell mandate, a number of states have already created forensic science commissions, which might well be appropriate, existing entities to serve as SIFSs. Models range from bodies with narrow mandates to accredit or otherwise create standards for laboratory practice, to entities having investigative functions as well, with authority to investigate complaints of forensic science negligence or

---

281. Notice of Establishment of the National Commission on Forensic Science and Solicitation of Applications for Commission Membership, 78 Fed. Reg. 12,355, 12,356 (Feb. 22, 2013) (detailing responsibilities that track NAS Report's professionalization goals and also including a focus on "the intersection of forensic science and the courtroom").

282. NAS REPORT, *supra* note 7, at 17.

283. *See, e.g.*, Erica Solange Deray, Note, *The Double-Helix Double-Edged Sword: Comparing DNA Retention Policies of the United States and the United Kingdom*, 44 VAND. J. TRANSNAT'L L. 745, 755–58 (2011) (summarizing the history of DNA collection in America with a focus on state-level differences); Jeffrey E. Nicoson, *A Case for Certiorari: Whether Federal Courts Should Consider State Law When Admitting State-Collected Electronic Surveillance Evidence*, 46 U. LOUISVILLE L. REV. 335, 337–38 (2007) (exploring differences between state wiretapping laws).

284. *See* 42 U.S.C. § 3797k(4) (2006) (requiring that laboratories receiving federal grants create mechanisms for external independent investigations); *Oversight of the Department of Justice's Forensic Grant Programs: Hearing Before the S. Comm. on the Judiciary*, 110th Cong. 2–3 (2008) (statement of Glenn A. Fine, Inspector General, U.S. Department of Justice); *Oversight of the Justice For All Act: Hearing Before the S. Comm. on the Judiciary*, 110th Cong. 172–74 (2008) (statement of Peter Neufeld on behalf of The Innocence Project); *see also* Garrett & Neufeld, *supra* note 216, at 94 (discussing the neglected state of Coverdell oversight mandate).

misconduct on their own motion or through the receipt of complaints.[285] While none of these entities possesses a mandate to engage in oversight of police and prosecutors, those with broader investigative functions have done important work in assessing the role of forensic evidence in the construction and investigation of a case, in some instances exposing deficiencies in the manner in which police and prosecutors have responded to scientific evidence.[286] These state-level entities are well positioned to develop recommendations for areas of need or opportunity for research, or even for best practices in regard to the types of issues raised by Part II. Of course, proximity also carries the risk of undue influence, and it may well be that a SIFS entity is, though more knowledgeable about local context, far less politically inclined to push for change.[287] One way to address this concern is to conceive of SIFSs as advisory bodies for a national oversight entity, responsible for supplying local data that could inform proposed (and financially incentivized) standards emanating from a national entity.

B. The Criminal Justice Conversation Must Account for Forensic Science and Upstream Dynamics—Particularly Assuming a World of More Early Science

An implicit though important lesson of Part II is that criminal procedure doctrine currently has little role to play in the dynamics that this Article aims to highlight. True, the Fourth Amendment formally serves as a pervasive regulatory backdrop for the work of law enforcement in the investigative stage of a case, particularly where gathering and exploiting *stuff* is concerned—that is to say, in relation to search and seizure—as is fundamentally the case in relation to all physical evidence in a case. But for reasons described above—in particular, because of diminished warrant strictures and weakened remedies—Fourth Amendment doctrine serves as a fairly light constraint and an even weaker affirmative mandate in regard to law enforcement activity around forensic evidence.[288] Other areas of crim-

---

285. *See generally* Goldstein, *supra* note 23 (discussing models).

286. *See generally, e.g.*, TEX. FORENSIC SCI. COMM'N, *supra* note 120 (discussing an expert report that found that the conclusions made by Fire Marshall Vasquez in the Willingham arson case were unscientific, and recommending a greater role for prosecutors and defense attorneys, as well as judges, in acting as gatekeepers to ensure that forensic expert testimony should not be allowed unless conclusions made by those experts are reliable).

287. *Cf.* William C. Thompson & Rachel Dioso-Villa, *Turning a Blind Eye to Misleading Scientific Testimony: Failure of Procedural Safeguards in a Capital Case*, 18 ALB. L.J. SCI. & TECH. 151, 189–91 (2008) (discussing the reluctance of a Virginia state board to investigate a fundamental evidentiary issue in a capital murder case by determining that the issue could be passed on to the appellate courts).

288. An exception to this rule is the barrier that the Fourth Amendment may create to the suspicion-less collection of DNA specimens for purposes of populating the CODIS databases. The Supreme Court appears likely to take up the question of whether the Fourth Amendment bars such collection and retention on a showing of probable cause to arrest this Term. *See* Maryland v. King, 133 S. Ct. 1, 2 (2012) (granting a stay of a lower court judgment against the state of Maryland in

inal procedure doctrine are even less relevant: while the Fifth, Sixth, and Fourteenth Amendments speak (haltingly) to police work in relation to eyewitness identification and suspect interrogations,[289] notions of due process are essentially viewed as irrelevant to scientific evidence absent evidence of fabrication or framing in the course of an investigation.[290] *Brady* doctrine entitles the defense, as a feature of due process, to favorable information within the control of the state, but this has to date been limited to known rather than potentially exculpatory evidence and continues to be tethered solely to trial.[291]

This is reflective of more foundational commitments revealed in the prevailing majorities (and occasionally, vocal dissents) in United States Supreme Court cases over the last half century of evolving constitutional criminal procedure. American criminal procedure doctrine is fairly preoccupied with policing a declared line between the inquisitorial and accusatorial features of our criminal justice system—the former attending the investigative and charge-screening work that police and prosecutors do in constructing a case, and the latter attending the trials in which that evidence is tested.[292] Closely related, of course, is a tradition of substantial discretion enjoyed by police and prosecutors in their core pretrial tasks (investigation and charging, respectively), which has led courts to exercise minimal oversight in regard to decision making by those individuals.[293] Indeed, what oversight exists is always carefully and expressly calibrated so as not to disrupt these lines: warrant doctrine cabins police and prosecutorial

---

litigation over arrestee DNA collection, based on likelihood that certiorari would be granted and judgment reversed).

289. *See, e.g.*, Colorado v. Connelly, 479 U.S. 157, 163–64 (1986) (discussing the Fourteenth Amendment's protections against certain types of police interrogatory techniques); Manson v. Brathwaite, 432 U.S. 98, 110–14 (1977) (analyzing the Fourteenth Amendment's requirements for admissibility of identification testimony); United States v. Wade, 388 U.S. 218, 221–22, 224–25 (1967) (analyzing whether the conduct of a lineup and subsequent courtroom identification violated a defendant's Fifth or Sixth Amendment rights).

290. *See supra* note 146.

291. *See* Moore v. Illinois, 408 U.S. 786, 794–95 (1972) (finding "no constitutional requirement that the prosecution make a complete and detailed accounting to the defense of all police investigatory work on a case" and therefore holding that the prosecution was not required to disclose nonexculpatory statements of witnesses).

292. McNeil v. Wisconsin, 501 U.S. 171, 181 n.2 (1991) ("Our system of justice is, and has always been, an inquisitorial one at the investigatory stage . . . .").

293. *See, e.g.*, Town of Castle Rock v. Gonzalez, 545 U.S. 748, 761 (2005) (discussing the "deep-rooted nature of law-enforcement discretion" in the American criminal justice system); Arizona v. Youngblood, 488 U.S. 51, 58 (1988) (discussing the considerable deference afforded police judgments concerning the significance of evidence as the basis for rejecting a proposed constitutional duty "to preserve all material that might be of conceivable evidentiary significance in a particular prosecution"); Wayte v. United States, 470 U.S. 598, 607 (1985) (explaining reasons why the "Government retains 'broad discretion' as to whom to prosecute"); Brinegar v. United States, 338 U.S. 160, 176 (1949) (expressing concern that law enforcement not be "unduly hamper[ed]" by too stringent a probable cause standard).

discretion with magistrate review, but only so much;[294] *Brady* doctrine has been doggedly constrained;[295] and, at least until recently, the Court has roundly resisted the notion that the plea bargaining which displaces trial practice in the overwhelming majority of criminal cases should enjoy something like the scrutiny that attends courtroom proceedings.[296]

Embracing the centrality of scientific evidence in the criminal justice system and understanding that evidence (as the NAS Report does) as a distinctive evidentiary product created under conditions of scientific rather than forensic legitimacy challenges this settled understanding, at least with respect to the role of forensic evidence in the inquisitorial, investigative stages of criminal proceedings.[297] Others precede me in making this sort of observation in the context of the accusatorial stages of criminal adjudication, and in doing important work to think through how modes of scientific inquiry might appropriately reshape the adversarial modes of legal inquiry with respect to forensic science.[298] The concerns raised herein push on a different front, however. The question is whether quite so much of what transpires in the investigative, inquisitorial stages of our criminal justice system is appropriately committed to the unfettered discretion of police and prosecutors, at least where forensic evidence is concerned. It is fair to ask why some standards to promote attitudes and practices consistent with scientific values—including, in particular, the value of *inquiry*—should not be imposed upon those investigative stages and decisions that rely upon scientific evidence. Critically, whereas police and prosecutors surely do possess comparative competence in making the mine-run of evaluations and decisions in the development, evaluation, and funneling of cases prior to

---

294. *See, e.g.*, Hudson v. Michigan, 547 U.S. 586, 601–02 (2006) (concluding that a violation of the "knock-and-announce rule" did not require suppression of evidence); Franks v. Delaware, 438 U.S. 154, 155–56 (1978) (holding that the Fourth Amendment requires a hearing be held on a defendant's request when the defendant makes a substantial preliminary showing that a false statement was made in a warrant affidavit and that allegedly false statement was necessary to the finding of probable cause).

295. Connick v. Thompson, 131 S. Ct. 1350, 1369 (2011) (Scalia, J., concurring) (characterizing the claim of right to potentially exculpatory forensic evidence as lying on the "frontier" of *Brady* doctrine); United States v. Ruiz, 536 U.S. 622, 633 (2002) (rejecting an extension of *Brady* to the pretrial plea stage); United States v. Bagley, 473 U.S. 667, 674–75 (1985) (emphasizing that *Brady* supplements rather than supplants adversarialism).

296. *See* Lafler v. Cooper, 132 S. Ct. 1376, 1397 (2012) (Scalia, J., dissenting) ("In the United States, we have plea bargaining a-plenty, but until today it has been regarded as a necessary evil.").

297. *See* BRIAN FORST, ERRORS OF JUSTICE: NATURE, SOURCES, AND REMEDIES 41 (2004) (arguing that scientific evidence problematizes the adversarial–inquisitorial line); Susan Haack, *Irreconcilable Differences?: The Troubled Marriage of Science and Law*, 72 LAW & CONTEMP. PROBS. 1, 12–13 (2009) (contrasting the core business of science and that of a legal system); *see also* Dist. Attorney's Office for the Third Judicial Dist. v. Osborne, 557 U.S. 52, 74 (2009) ("DNA evidence will undoubtedly lead to changes in the criminal justice system. It has done so already.").

298. *See generally, e.g.*, Giannelli, *supra* note 47 (arguing that there is more than adequate support for the NAS Report's conclusions that meaningful reform requires an independent agency); Mnookin et al., *supra* note 3 (discussing the need for a research culture in the forensic sciences); Murphy, *supra* note 28 (challenging the new orthodoxy of forensic science).

trial, any such claim with regard to forensic evidence is far less clear—all the more so in a post-NAS Report world that acknowledges a professionally and ethically distinct field attending the production of such evidence.

What might be the consequences of pressing on the inquisitorial–accusatorial line that has tended to demarcate the appropriate boundaries of judicial oversight? An adequate answer depends on much fuller examination than can be undertaken in the confines of this Article, but some provisional thoughts are in order. Fourth Amendment doctrine might take a more demanding view of the warrant requirement where forensic evidence is concerned. Imagine, for example, a requirement that exculpatory facts concerning forensic evidence be detailed in an arrest warrant if they would be credited by a reasonable officer—a requirement that goes beyond current doctrine. As a matter of incentives, such a requirement might force more reflection on the part of investigators, as well as naturally bring additional perspectives into the process, since more demanding warrant requirements are likely to encourage supervisory or prosecutorial review.[299] The right of defense access to favorable evidence would also be the subject of reconsideration under these different assumptions. The notion that information concerning forensic evidence should not be subject to disclosure at the earliest possible stage—i.e., promptly after it is available to the state—seems a tactically driven, uncomfortable fit with the goals of greater exploitation of science in the criminal justice system, all the more so in a world in which forensic analysts are understood as independent scientific contributors to the adjudicative process.[300] More innovatively, we might imagine expanded opportunities for defendants to affirmatively test the science-supported premises of an investigation, including (as a small number of states have instituted) a right to compel testing of evidence *not* developed by the state.[301]

---

299. *See* Richman, *supra* note 17, at 782 ("The more technically demanding a warrant process is, the more a prosecutor can use her statutory role to scrutinize an agency's investigation."); *see also* Findley & Scott, *supra* note 211, at 384 (discussing that the best way to overcome the issue of tunnel vision is to preserve all notes and evidence for a third-party advisory investigator to analyze); O'Brien, *supra* note 209, 327–28 (finding that forcing individuals to articulate why a guilt hypothesis might be wrong mitigated confirmation bias, but that forcing individuals to generate multiple hypotheses had no such effect); *cf.* Heath et al., *supra* note 151, 5–22 (providing examples of a number of organizational practices that may effectively repair the cognitive shortcomings of individuals).

300. Indeed, a small number of states have adopted discovery reforms reflecting a robust concurrence with this view. *See* Paul C. Giannelli, *Forensic Science*, 33 J.L. MED. & ETHICS 535, 539–40 (2005) (arguing that full pretrial discovery disclosure should be required in criminal cases).

301. *See* 725 ILL. COMP. STAT. ANN. 5/116-5 (West 2008) (permitting a defendant in any case where DNA may be relevant to the defense investigation or at trial to move the court for an order requiring the state police to conduct certain genetic tests or to make certain comparisons or searches within the database); GA. CODE ANN. § 24-4-63 (2010) (providing similar rights); AM. BAR ASS'N, *supra* note 266, at 140 (providing similar rights through Standard 16-8.3); N.Y. STATE BAR ASS'N, FINAL REPORT, *supra* note 117, at 99 (arguing for provision of similar rights); *see also* Lynch, *supra* note 17, at 2117 (arguing that the pretrial phase of a case might bear more adversarial process).

For the time being, the Supreme Court has registered its resistance to rethinking any fundamental features of criminal procedure doctrine in light of the influence of forensic science, although its posture has been more "wait and see" than "nevermore."[302] Thinking outside existing criminal procedure boxes is therefore more than a proverbial academic exercise. As the Supreme Court watches and state legislatures and courts pursue more innovative responses to the pressures of forensic science on our settled understandings of competence, deference, and oversight in criminal investigation and adjudication, it becomes all the more important to develop a principled approach to reconciling these accommodations within existing criminal procedure theory and doctrine.

Conclusion

The NAS Report has fundamentally altered the landscape for forensic science in the criminal justice system. This is to be celebrated. But the accomplishments of the Report must not obscure the vast terrain that remains untouched by the path of reform that it charts. This Article has aimed to illuminate one important aspect of that currently neglected territory: namely, the manner in which upstream *users* of forensic science—police and prosecutors—will select priorities, initiate investigations, collect and submit evidence, choose investigative techniques, and charge and plead cases in ways that have critical and systematic, though poorly understood, influences on the accuracy of forensic analysis and the integrity of its application in criminal cases. By broadening our understanding of how forensic science is created and used in criminal cases—by adopting a systemic perspective—we begin to see a raft of yet unaddressed issues concerning the meaning of scientific integrity and reliability in the context of investigative decisions that are by and large committed to the discretion of decidedly unscientific actors. Moreover, we see that decisions with respect to oversight of one corner of the system—the laboratory, in the case of the NAS Report—cannot be made in isolation, lest responses from other corners render that oversight ineffective or counterproductive. Undoubtedly, this Article has raised at least as many questions as it has answered. This is for the best. The account here only scratches the surface of the sorts of systemic concerns that we might reflect upon, and that hopefully the active reform conversations will take up, as we commit our criminal justice system to more and more institutionally entrenched forensic science.

302. *See* Dist. Attorney's Office for the Third Judicial Dist. v. Osborne, 557 U.S. 52, 74–75 (2009) (pointing to "active[]" efforts by state governments to manage the "challenges" and "opportunities" to the criminal justice system posed by DNA as reason to act cautiously in constitutionalizing a right of access to evidence).

Book Reviews

What Do We Talk About When We Talk About the Constitution?

AMERICA'S UNWRITTEN CONSTITUTION: THE PRECEDENTS AND PRINCIPLES WE LIVE BY. By Akhil Reed Amar. New York, New York: Basic Books, 2012. 615 pages. $29.99.

FRAMED: AMERICA'S 51 CONSTITUTIONS AND THE CRISIS OF GOVERNANCE. By Sanford Levinson. New York, New York: Oxford University Press, 2012. 448 pages. $29.95.

Reviewed by Akhil Reed Amar* & Sanford Levinson**

I. Dear Akhil,

It is certainly not surprising that *America's Unwritten Constitution* is remarkably stimulating, informative, and challenging. You are surely correct that one cannot possibly understand the American constitutional system simply by reading the text of the Constitution (or, for that matter, reading decisions of the judiciary ostensibly "interpreting" the text). Instead, one must not only look at long-established American practices but also at social movements and transcendent moments in American history—the Gettysburg Address and Martin Luther King's "Dream" speech are two that you emphasize[1]—that have provided the rationales for how we understand those practices (and, on occasion, become willing to transform them). Your Constitution is necessarily a "living Constitution," for the American people, as active agents of their own constitutional destinies, are constantly debating one another about what constitutes its deep meanings; they constantly create new movements, which in turn generate new political leaders committed to particular understandings. This is one way of understanding not only the civil rights movement that is so important to both of us, but also the Tea Party, which cannot be understood without paying careful attention to its

---

* Sterling Professor of Law, Yale Law School.

** W. St. John Garwood and W. St. John Garwood, Jr. Centennial Chair in Law, University of Texas Law School; Professor of Government, University of Texas.

Readers should not believe that Akhil and I are in the habit of footnoting our correspondence with each other! Thus I am grateful to the editors of the *Texas Law Review* for their sometimes remarkable diligence—as with tracking down the potential negative effects of taking aspirin—in verifying my often off-the-cuff assertions or allusions (e.g., the House of Atreus). I'm sure I speak for Akhil as well.

1. AKHIL REED AMAR, AMERICA'S UNWRITTEN CONSTITUTION: THE PRECEDENTS AND PRINCIPLES WE LIVE BY 247 (2012).

narratives of the Constitution and calls both for fidelity to its ostensible norms and for amendments, such as repeal of the Seventeenth Amendment,[2] that would return the Constitution to its intended—and they argue better—embrace of a far stronger form of federalism than we in fact have today.

It is also not surprising that we continue to have some quite fundamental disagreements, whatever our personal closeness. Each of us recognizes in his respective acknowledgments the importance of our relationship over what is now more than a quarter century,[3] which includes for the last decade our service as co-editors of a casebook in constitutional law, *Processes of Constitutional Decisionmaking*.[4] But that does not mean, of course, that we have become clones of one another. We met initially when you came to Austin for a symposium on Philip Bobbitt's then recently published *Constitutional Fate*,[5] and we bonded during the course of what turned out to be (at least) a two-hour visit to the monument to Confederate war dead in front of the Texas State Capitol. As noted in our casebook, that monument presents what might be described as the "standard" Southern view of the War:

> DIED FOR STATE RIGHTS GUARANTEED UNDER THE CONSTITUTION.
>
> THE PEOPLE OF THE SOUTH, ANIMATED BY THE SPIRIT OF 1776, TO PRESERVE THEIR RIGHTS, WITHDREW FROM THE FEDERAL COMPACT IN 1861. THE NORTH RESORTED TO COERCION. THE SOUTH, AGAINST OVERWHELMING NUMBERS AND RESOURCES, FOUGHT UNTIL EXHAUSTED.[6]

We debated at length whether this is a "possible" interpretation of the War in relation to the Constitution, which is a very different question from whether it is the "best" interpretation. Your view, I think it is safe to say, is that this does not rise to the level of a "possible" interpretation—that it would deserve an "F" if submitted on a final examination. My view was that it is, for better or worse, a possible view, because the 1787 Constitution, correctly interpreted, is ambiguous (or, in the language of the 1980s, when we first met, "indeterminate"). In the interim, neither of us has changed our fundamental view.

Thus I was startled (though I should not have been surprised) to see your declaration that "the original Constitution emphatically denied state

---

2. *See* Matt Bai, *Tea Party's Push on Senate Election Exposes Limits*, N.Y. TIMES, June 1, 2010, http://www.nytimes.com/2010/06/02/us/politics/02bai.html (discussing Tea Party movement members in several states calling for repeal of the Amendment).

3. AMAR, *supra* note 1, at 597; SANFORD LEVINSON, FRAMED: AMERICA'S 51 CONSTITUTIONS AND THE CRISIS OF GOVERNANCE 399 (2012).

4. PAUL BREST ET AL., PROCESSES OF CONSTITUTIONAL DECISIONMAKING (2006).

5. PHILIP BOBBITT, CONSTITUTIONAL FATE: THEORY OF THE CONSTITUTION (1982).

6. BREST ET AL., *supra* note 4, at 219.

authority to unilaterally secede."[7] As many times as I have read the Constitution, I quite literally don't see this "emphatic[] deni[al]." At most, I see an aspiration, set out in the Preamble, which both you and I admire greatly, that the new Constitution would indeed be the instrument creating a "Perfect Union." For me, the Preamble is the equivalent of the hope announced during a traditional wedding ceremony—admittedly in the form of a promise—that this union will last forever, until death separates the two partners. What this means, however, in both written and unwritten American legal culture, is only that one *hopes* that the marriage will generate a truly lasting commitment because of mutual appreciation of what marriage brings; however, in the not-unlikely circumstance that that won't actually happen, then divorce, amicable or otherwise, is a possibility. Everyone, whether those at the altar or the onlooking audience, knows this.

Lincoln, unequivocally one of your heroes, offered a mocking response to those defending the legitimacy of secession: "In their view, the Union, as a family relation, would not be anything like a regular marriage at all, but only as a sort of free-love arrangement [laughter] to be maintained on what that sect calls passionate attraction."[8] Unlike Lincoln's audience, I don't view this as a laughing matter. After all, consider a sentence that appeared in the penultimate paragraph of the penultimate draft of the Declaration of Independence, where Thomas Jefferson noted the British misconduct served to generate "the last stab to agonizing affection," so that "manly spirit bids us to renounce forever these unfeeling brethren."[9] The affection necessary to maintain political unity, especially in a federal political system that can be understood only against the background of *dissensus* and potential for "disaffection," can never be taken entirely for granted. Such dissensus had made a truly unified government—or even the more truly centralized system that James Madison in fact yearned for in Philadelphia and was, to his chagrin, systematically denied by his fellow delegates—impossible. We *were* a "house divided," and it was the perhaps dubious premise of the 1787 Constitution that it was building a structure that would enable such a house to stand instead of inevitably imitating the House of Atreus by generating a tragic and bloody carnage.[10]

---

7. AMAR, *supra* note 1, at 85.

8. ABRAHAM LINCOLN, *Speech at Indianapolis, Indiana* (Feb. 11, 1861), *in* ABRAHAM LINCOLN: SPEECHES AND WRITINGS 1859–1865, at 201, 202 (Don E. Fehrenbacher ed., 1989).

9. DRAFT OF THE DECLARATION OF INDEPENDENCE para. 34 (U.S. 1776), *available at* http://www.monticello.org/site/jefferson/transcript-declaration-independence-rough-draft.

10. The Greek myth of the cursed family follows the descendants of Atreus, who served King Thyestes his murdered sons for dinner. Atreus's son Agamemnon continues the slaughter when he sacrifices his daughter Iphigenia to further the Trojan War. Agamemnon's wife Clytemnestra avenges Iphigenia by murdering Agamemnon with the help of her lover Aegisthus, who is Thyestes's surviving son. Agamemnon's son Orestes concludes the legend when he kills Clytemnestra and Aegisthus. Helene P. Foley, *Introduction* to AESCHYLUS, THE ORESTIA, at vi, vii (Peter Meineck trans., Hackette Pub. Co. 1998).

You have written with true eloquence, both in this book and in your earlier "biography" of the Constitution,[11] of the ravages generated by the "rotten compromise," in the words of Israeli philosopher Avishai Margalit, that entrenched slavery and made us, in the title of Don Fehrenbacher's last work, a "slaveholding republic."[12] We'll never know, of course, whether Romeo and Juliet, had they lived, could have successfully surmounted the bitter divisions between the Capulets and Montagues. We do know that this proved unsuccessful, as an empirical matter, with regard to the "free North" and the slaveholding South. And both sides claimed constitutional support, including, inevitably, the support of America's unwritten norms. One might well perceive one of those norms as a willingness to engage in bisectional compromise—think of the Missouri Compromise or the Compromise of 1850. Both, not at all coincidentally, involved the issue triggered by what I suspect both of us view as the most important constitutional event of 1803, the Louisiana Purchase (far more important than *Marbury*)[13]: Would slaveowners be welcome in the vast new territories that constituted the reality of American expansionism (a topic, I should note, that does not garner extensive discussion in your book)? There is a reason that Justice Catron, in his concurrence in *Dred Scott*[14] (reprinted, I think uniquely, in our casebook) could claim that the basic premise of "EQUALITY" (as he spelled it) guaranteed that slaveowners would have the same ability to take their legal property, as defined by state law, into the territories operated by a fiduciary Congress in the equal interest of all citizens.[15]

Perhaps the greatest difference between us is that I basically accept in a way that you do not William Lloyd Garrison's view of the 1787 Constitution as a "Covenant with Death and an Agreement with Hell."[16] Pacts with the Devil make their own claim to "honorable" lawyers of the Marshallian persuasion; he insisted, after all, on separating the role of the "jurist" from that of the "moralist," a distinction repeated, of course, by his successor Roger Taney in *Dred Scott*.[17] Garrison notably burnt the Constitution and, indeed, suggested that there should be "No Union with Slaveowners."[18] Imagine for a moment that he was actually successful in generating a secessionist movement within New England, so that the six New England

---

11. AKHIL REED AMAR, AMERICA'S CONSTITUTION: A BIOGRAPHY (2005).

12. LEVINSON, *supra* note 3, at 43–44 (quoting AVISHAI MARGALIT, ON COMPROMISES AND ROTTEN COMPROMISES (2010)); *see also* DON E. FEHRENBACHER, THE SLAVEHOLDING REPUBLIC: AN ACCOUNT OF THE UNITED STATES GOVERNMENT'S RELATIONS TO SLAVERY (2002).

13. Marbury v. Madison, 5 U.S. 137 (1803).

14. Dred Scott v. Sandford, 60 U.S. 393 (1857).

15. *Dred Scott*, 60 U.S. at 529 (Catron, J., concurring); BREST ET AL., *supra* note 4, at 248–49.

16. Resolution Adopted by the Massachusetts Anti-Slavery Society (Jan. 27, 1843) (quoted in *The "Covenant with Death,"* LIBERATOR, Mar. 13, 1863, at 1, col. 3).

17. The classic formulation of this distinction appears in Marshall's opinion in *The Antelope*, 23 U.S. 66, 121 (1825): "Whatever might be the answer of a moralist to this question, a jurist must search for its legal solution . . . ."

18. BREST ET AL., *supra* note 4, at 253.

states tried to secede on antislavery grounds. Would you have supported James K. Polk or Franklin Pierce if they sent troops to prevent secession and preserve the Union and, if so, why? Is your opposition to secession within the United States simply positivistic—i.e., it is, you believe, the unequivocal command of the Constitution that any faithful constitutionalist must adhere to—or does it rest on a stronger moral principle? One such principle, of course, is antislavery, and one might support the War as basically a "humanitarian intervention" by which the slaveowners' regimes would be transformed. But if that is the basis for one's support, then the operative principle is surely not "secessionist movements are always to be opposed, wherever and whenever they occur."

Both of us agree that the Constitution was far better after the slaughter of War and the addition of the "Reconstruction Amendments." One of your most striking chapters is the defense of the Fourteenth Amendment—and the Military Reconstruction necessary to realize its ratification—as Congress simply acting under its power to guarantee to the states a "Republican Form of Government."[19] Your chapter is eloquent and incisive. I certainly find it a "possible" view of the Guarantee Clause,[20] even as one must recognize that such a strong reading of the Guarantee Clause by Congress is quite literally unique, both unprecedented prior to 1866 and left untapped since then as a defense for subsequent congressional actions regarding state governments. I continue to accept Bruce Ackerman's view of the Amendment as a basically extra-constitutional addition, the extra-constitutionality being necessary because of the truly egregious Article V and its setting of basically insurmountable hurdles to those striving for fundamental change.[21]

But even if the Constitution is decidedly better after the Reconstruction Amendments, whatever their provenance, for me it is not sufficiently better to warrant the love you so movingly display for the Constitution. And, as you well know, what has generated my deep alienation from the Constitution has almost literally nothing to do with the Reconstruction Amendments or, for that matter, the doctrinal Constitution as enunciated by the judiciary and others, but instead, the fundamental institutional structures established in 1787 and left remarkably unchanged since then. What most disappoints me about your new book is your confidence that clever lawyering can provide adequate "workarounds" to what you so obviously believe yourself to be real problems. At one level, you might be right: Should the general public be convinced that the unamended Constitution is taking us over a cliff, as I sometimes believe to be the case, it might well accept the determination of a clever lawyer that we can avoid that fate.

The problem, though, is that a dreadful part of our "unwritten Constitution" is that it should be treated as a basically sacred document and

19. AMAR, *supra* note 1, at ch. 2.

20. U.S. CONST. art. 4, § 4.

21. 2 BRUCE ACKERMAN, WE THE PEOPLE: TRANSFORMATIONS, 99–119 (1998).

thus subjected to little, if any, radical criticism. I don't think this can be traced in any uncomplicated way to 1787 itself, save for Madison's hope, expressed in *Federalist No. 49*, that the Constitution be treated as an object of reverence and "venerat[ed]" rather than coldly analyzed.[22] A century ago, when Woodrow Wilson and Theodore Roosevelt ran against each other (and William Howard Taft) for the presidency, they both were more than willing to offer fundamental critiques of constitutional practices and norms.[23] My own view is that one of the consequences of World War II was an unwarranted faith in the Constitution as the unquestioned symbol of what we were fighting for. Even though (should it be because of?) both of the candidates in 2012 were graduates of the Harvard Law School, we received not one moment of serious discussion of the adequacy of the Constitution from either of them. Of course, Yale Law School-educated Bill Clinton was no better. Since neither of us is truly complacent about the American polity, we might wonder whether this contemporary silence has anything to do with the way constitutional law is taught at our leading law schools.

Along with reading your undoubtedly lively responses to my arguments above, I hope you write as well of what changes you would like to see in what might be termed the pedagogy of the Constitution. Both of us, in recent years, have been teaching *undergraduates* as well as fledgling lawyers. Both of us, I have no doubt, hope to reach a wide audience of our fellow citizens. But I am curious which, as between law schools and undergraduate classes, you expect to be the actual venue for reading and confronting your always imaginative takes on the Constitution. If, as I suspect, the answer is the latter, then how would you change the culture of the legal academy to make it more receptive to your distinctive way of approaching the Constitution?

In friendship and fondness,

Sandy

II. Dear Sandy,

It is great to be back in conversation with you—this time, about your new book and mine. By highlighting the classic constitutional question of secession in your opening missive to me, you have chosen a great place to begin our epistolary exchange—namely, at the beginning. Not just at the beginning of our quarter-century friendship, which did indeed start with an intense and extended debate over the Confederacy's legal theory, but also at the beginning of our Constitution's text and history. For (as you know) I believe that the secession issue is powerfully illuminated by the Constitution's opening sentence (A.K.A. the Preamble) and also by the epic

---

22. *See* THE FEDERALIST NO. 49, at 311 (James Madison) (Clinton Rossiter ed., 1961).

23. Sanford Levinson, *Our Imbecilic Constitution*, CAMPAIGN STOPS, N.Y. TIMES (MAY 28, 2012, 8:36 PM), http://campaignstops.blogs.nytimes.com/2012/05/28/our-imbecilic-constitution/.

yearlong continental conversation in 1787–1788 that accompanied the Constitution's ratification.

Let me first summarize my legal reasoning, and then step back and discuss some of the implications of this reasoning—what this reasoning says about the similarities and the differences between your approach and mine, in our two new books and elsewhere.

The terse text that we call the Constitution begins, in its opening sentence, by calling itself a "Constitution."[24] Not a "league," not a "confederacy" or a "confederation," nor a treaty based on states that retain their full "sovereignty" and "independence"—highly significant legal keywords that were all prominently featured in the opening passages of a predecessor document, the 1781 "Articles of Confederation,"[25] that the 1787–1788 document aimed to wholly displace. In lieu of the old 1781 compact/league/treaty/confederation, what was instead being offered up to the American people in 1787–1788 was made clear by the new document's opening words: a "Constitution" obviously modeled on the extant *state constitutions*. The literally primary purpose of this 1787–1788 document was to form a "more perfect" union—one that would safeguard "common defence" and thereby preserve "the Blessings of Liberty."[26]

The terse text's penultimate section, Article VI, explains quite clearly what the legal status of "this Constitution" would be, once ratified.[27] The text was to be—and has legally always remained—"the supreme Law of the Land."[28] The Articles of Confederation had never described themselves as "law," much less as *supreme* law. Nowhere was the old "Congress" under the Articles described as a "legislature" or a "lawmaker," and this old body in fact was more of an international assembly/war council on the model of today's NATO Council and UN Security Council. Under the Articles of Confederation, state officials were not obliged by oath to treat the Articles themselves or congressional edicts pursuant to the Articles as supreme domestic law trumping the contrary commands of state legislatures (which were in fact described by the Articles as "legislatures" and whose commands were thus seen by that 1781 document as true *laws*).[29]

In emphatic and unambiguous contrast, the 1787–1788 Constitution makes clear that all state officials are indeed oath bound to follow the Constitution as supreme law,[30] and further makes clear that nothing that a state does unilaterally—nothing in any future state constitution or state

---

24. U.S. CONST. pmbl.

25. ARTICLES OF CONFEDERATION of 1781, arts. I, II, III, IV.

26. U.S. CONST. pmbl.

27. U.S. CONST. art. VI, para. 2.

28. U.S. CONST. art. VI, para. 2.

29. *See* AMAR, *supra* note 11, at 40–41, 64–65, 301 (describing the relationship between state legislatures and federal law under the Articles).

30. U.S. CONST. art. VI, para. 3.

statute—can change the hierarchical status of the U.S. Constitution and federal statutes and treaties pursuant to that Constitution as the supreme law of the land.[31] Period. No ifs, ands, or buts. Disgruntled individuals are always free to leave, but *the land* remains as part of America, and the *supreme law* that governs that *land* is the U.S. Constitution "*notwithstanding*" any "*Contrary*" unilateral state action.[32] Unilateral state secession is simply not provided for, and indeed is emphatically ruled out. (If a state did actually retain a right to unilaterally exit, surely we would expect to see all sorts of rules about how that could happen—for example, how a departing state would need to shoulder its fair share of the pre-existing national debt and guarantee peace with neighboring union states. But we don't see any of that stuff in the text because, to repeat, unilateral state secession is not allowed by the Constitution, which makes itself supreme regardless of what individuals within an individual state might say or do.) As I sometimes say to my kids, "What part of 'No!' did you not understand?" Sandy, it's just that simple, textually.

But of course there is far, far—far!—more legal evidence than this, and I spent some thirty pages in the opening chapter of my 2005 book (*America's Constitution: A Biography*—the predecessor to my latest volume, *America's Unwritten Constitution*) laying out this evidence.[33] Just a few highlights. Article III explicitly says that anyone who wages war against the Union commits "Treason," even if that individual is supported by, and supportive of, his anti-Union state government.[34] When Antifederalist Luther Martin explicitly objected to precisely this result at Philadelphia, he was pointedly outvoted—and in the ratification conversation he brought the issue prominently to the attention of his fellow Americans, who once again outvoted him by ratifying the clear antisecession rules of Article III and the document as a whole.[35] Article VII made clear that no state could be bound by the new Constitution unless it chose to sign on: Precisely because each state was indeed sovereign and independent prior to 1787, no state could bind any other.[36] But in obvious and unmistakable contrast, Article V made clear

31. U.S. CONST. art. VI, para. 2.

32. U.S. CONST. art. VI, para. 2.

33. *See* AMAR, *supra* note 11, at 5–39 (deducing from a variety of textual differences that the Articles allowed for unilateral secession as part of each State's sovereignty and that the Constitution disavowed this as a failure of the Articles).

34. U.S. CONST. art. III, § 3; *see also* AMAR, *supra* note 11, at 242–45 (discussing the Treason Clause and the debates surrounding its ratification).

35. AMAR, *supra* note 11, at 242.

36. *See* U.S. CONST. art. VII ("The Ratification of the Conventions of nine States, shall be sufficient for the Establishment of this Constitution between States so ratifying the Same."); *see also* AMAR, *supra* note 11, at 34–35 ("Article VII made it clear that the people of [a non-ratifying] state were a distinct sovereign entity free to vote down the new Constitution and ignore it."). *But cf.* Donald S. Lutz, *Constitutional Bricolage?: A Commentary on Akhil Reed Amar's* America's Constitution: A Biography, 57 SYRACUSE L. REV. 311, 314 (2007) ("[Amar's] use of Article VII is imaginative and probably correct, but this and other pieces of evidence are trotted out repeatedly in

that once a state joined the new union, it could indeed be bound by future constitutional amendments agreed to by the conventions of enough other states, even if its own state convention voted against these future amendments.[37] The clear structural logic here, reinforcing text and history, is that a state would no longer be fully sovereign once it chose to join the new Constitution. And without full state sovereignty, secession as a legal, constitutional right simply has no leg to stand on.[38]

Now turn briefly to consider the epic continental conversation that accompanied the Constitution's ratification. At Philadelphia, James Madison clearly said that while a sovereign state might secede from a treaty/league/confederation if the sovereign state deemed the treaty to have been breached or to threaten its sovereign existence, no such secession was permitted by nonsovereign subparts within a true legal "Constitution"—e.g., cities or counties within a state, or states within the proposed federal Constitution.[39] Madison here invoked the famed British legal authority Blackstone, whose bestselling treatise had clearly explained that the 1707 union of Scotland and England forbade unilateral secession of either part.[40] In *Federalist No. 5*, Publius made clear that this indivisible 1707 union was indeed *the precise model* for what was now being proposed four score years later in America.[41] Indeed, Publius explicitly linked the words of the Preamble about the formation of a "more perfect union" with the language

---

different combinations willy nilly, and some . . . leave this reader less willing to buy other pieces of evidence at face value.").

37. *See* U.S. CONST. art. V (requiring only two-thirds of the States to call a constitutional convention and only three-fourths of the states' approval to bind all the states under a constitutional amendment).

38. Two sidenotes: First, the natural right to revolt in case of true tyranny might remain as a right above all law, but this right has nothing to do with whether there happens to exist majoritarian electoral support within the boundaries of any particular state. In any event, Jeff Davis and company were never able to point to anything truly tyrannical justifying their treason. These rebels purported to secede before Lincoln had even taken office; Lincoln in 1860 had scrupulously played by the rules in winning election; so had congressional Republicans; courts remained open to hear valid legal disputes; and anti-Lincoln forces had full freedom of speech—even as they denied that freedom to others, and indeed made a mockery of the natural rights of Southern slaves and violated many legal rights of pro-Lincoln citizens in the South, citizens who were legally entitled to all the protections of the U.S. Constitution, including the right to have their land governed by that Constitution rather than by those aiming to overturn binding law and valid continental elections by force of arms.

Second, even though a state may not unilaterally secede, the Union itself might lawfully decide to dissolve via various legal federal procedures provided for by the text itself—constitutional amendments, federal statutes, federal treaties, even federal presidential elections. But all of these procedures would involve democratic decision making by the Union as a whole and not unilateral decision making merely by one geographical part of the Union called a "state."

39. AMAR, *supra* note 11, at 31.

40. *See id.* at 30–32, 36 (summarizing Blackstone's argument and crediting it as the source for Madison's own "breached treaty defense").

41. THE FEDERALIST NO. 5 (John Jay), *supra* note 22, at 44–45, 48; AMAR, *supra* note 11, at 36.

that had accompanied the 1707 unification of the British isle.[42] In *Federalist No. 11*, Publius explicitly described the Constitution's proposed union as "strict and indissoluble," and indeed the entire opening section of the *Federalist Papers*—encompassing Numbers 2 through 9—was premised on the need for an indivisible American union on the 1707 English–Scotch model.[43] Only if internal land borders were demilitarized could Americans prevent states from warring against each other and keep Europe from playing divide and conquer in the New World. Unilateral secession was wholly inconsistent with the main structural argument for union sketched out in these essays.

And this argument was raised not just in these essays, but in the ratification conversations themselves, where leading Federalists repeatedly said that the new union would not permit unilateral secession.[44] As I explain in my book, "[James] Wilson contrasted traditional 'confederacies' that historically 'have all fallen to pieces' with the proposed Constitution, in which 'the bonds of our union' would be 'indissolubly strong.'"[45] Wilson himself had emigrated from Scotland, as had North Carolina's Federalist Governor Samuel Johnston, who could not have been clearer in thought or expression: "[T]he Constitution must be the supreme law of our land; otherwise, it would be in the power of any one state to counteract the other states, and withdraw itself from the Union."[46]

In every state, Antifederalists got the message and loudly warned their audience that if the Constitution were ratified, their state would lose its sovereignty and be unable to exit unilaterally.[47] Again and again and again the Antifederalists said this. *Yet never in this entire year long continental conversation—not once!—did any prominent Federalist say that each state would indeed have the right to leave if subsequently dissatisfied.* Sandy, if secession was permitted, why didn't the Federalists say so? Surely a money-back guarantee/right to return the purchase no questions asked would have been a great selling point if true.

But it was not true, legally. No state ratifying convention explicitly purported to reserve a right of unilateral secession in the course of approving the federal Constitution.[48] And when the unilateral secession issue arose at

---

42. *See* THE FEDERALIST NO. 5 (John Jay), *supra* note 22, at 44–45 (praising this element of the unification); AMAR, *supra* note 11, at 36 (linking Jay's praise of this language to the Preamble).

43. AMAR, *supra* note 11, at 36.

44. *See id.* at 34–37 (recounting the Federalists' comments relevant to unilateral secession made during the ratification debates).

45. *Id.* at 36–37.

46. *Id.* at 37.

47. *Id.* at 38; *see also id.* at 35 (outlining a number of sovereignty-based criticisms of the Constitution leveled by Antifederalists).

48. *See* Kenneth M. Stampp, *The Concept of a Perpetual Union*, 65 J. AM. HIST. 5, 20 (1978) ("No state convention made the right of secession the subject of extended inquiry . . . ."). *But see* AMAR, *supra* note 11, at 38 n.84 ("[The New York] vote goes unmentioned by the great historian

the New York ratifying convention, Federalists explicitly rejected the idea, and did so at the risk of losing the ultimate ratification vote.[49] The final vote—with no secession reservation—was a nailbiter, 30 to 27.[50] Had New York voted no, it is doubtful that the fledgling Constitution would have actually succeeded without the deep-water ports of New York City and control over West Point and the mighty Hudson River, the geographic keys to the continent. In short, when ratification hung in the balance, the Federalists emphatically rejected the secession idea—at the risk of losing everything. Alexander Hamilton proclaimed that the plain language of Article VI "stands in the way" of any subsequent right of unilateral secession.[51] John Jay pronounced secession rights "inconsistent with the Constitution."[52] And Madison penned a letter, read aloud by Hamilton and later published for the benefit of the entire world (which was in real time following the New York cliffhanger with rapt attention, in rather the same way that all eyes are today, as I write these words to you, Sandy, focused on fiscal-cliff negotiations).[53] Here is what Madison wrote: "[T]he Constitution requires an adoption *in toto*, and *for ever*. It has been so adopted by the other States."[54]

OK, now what does all of this mean for our two new books—your new book, *Framed: America's 51 Constitutions and the Crisis of Governance*,[55] and mine, *America's Unwritten Constitution: The Precedents and Principles We Live By*[56]—and for constitutional interpretation more generally?

Here are some similarities and differences between you and me. First, we are both interested in deep and abiding issues of constitutional law, whether or not they are the hot topics of the moment. Some secessionist nonsense did bubble up recently, after Obama's reelection,[57] but you and I have been interested in this issue for decades.

Second, you and I both find federalism fascinating—and in particular both your new book and mine explore ways of thinking about the federal Constitution alongside state constitutions. Your book showcases state

---

Kenneth M. Stampp, and surely qualifies his claim that 'no state convention made the right of secession the subject of extended inquiry.'").

49. AMAR, *supra* note 11, at 38.

50. John P. Kaminski, *New York: The Reluctant Pillar*, *in* THE RELUCTANT PILLAR: NEW YORK AND THE ADOPTION OF THE FEDERAL CONSTITUTION 48, 114 (Stephen L. Schechter ed., 1985).

51. AMAR, *supra* note 11, at 38.

52. *Id.*

53. *Id.*; Kaminski, *supra* note 50, at 112–14; *cf. id.* at 100 ("As the delegates converged on Poughkeepsie, they realized the critical situation of the state and country.").

54. AMAR, *supra* note 11, at 38.

55. LEVINSON, *supra* note 3.

56. AMAR, *supra* note 1.

57. Elizabeth Dias, *Obama's Re-Election Inspires Southern Secessionists*, SWAMPLAND, TIME (Nov. 14, 2012) http://swampland.time.com/2012/11/14/obamas-re-election-inspires-southern-secessionists/.

constitutions alongside the federal charter in its very subtitle—*America's 51 Constitutions*—and my book's concluding chapter features an extended analysis of the similarities and differences between the one federal Constitution and the fifty state constitutions.[58] The secession question tightly focuses on the relationship between state and federal constitutions: Can South Carolina amend its *state constitution* so as to exit from the *U.S. Constitution*? You think this is (or at least was in 1861) an arguable question.[59] I say (among other things) that precisely because the federal "Constitution" is explicitly and self-consciously modeled on pre-existing state constitutions, unilateral secession is impermissible.[60] South Carolina may not unilaterally secede from the Union, just as Spartanburg County may not unilaterally secede from South Carolina.

Third, you and I both care deeply about conscience and legal ethics. In particular, both you and I have written extensively about legal oaths—you in your first book, *Constitutional Faith*[61] (which I reviewed way back when for the *Texas Law Review*[62]) and I in the penultimate chapter of my new book, a chapter entitled "Doing the Right Thing: America's Conscientious Constitution."[63] Lincoln began his time in office by swearing a solemn legal oath to "preserve, protect, and defend the Constitution."[64] Did this oath oblige him to defend "the land" and to "protect" the Constitution and loyal Unionists—and there were such men, Sandy!—in the South? Lincoln thought that his oath did indeed require this, and I think he was plainly correct. Though I do not agree with all of his legal reasoning, I find his actions to be profoundly legal and oath observing on the secession question—perhaps the most momentous question of all of American constitutional law.

Fourth, and related, both you and I are interested in American constitutional *culture*—in what our mutual friend Philip Bobbitt describes as America's constitutional "ethos"[65] and what our mutual co-author Jack Balkin discusses under the rubric of constitutional "narrative."[66] Lincoln of

58. AMAR, *supra* note 1, at 449–77.

59. LEVINSON, *supra* note 3, at 326–28.

60. *See* AMAR, *supra* note 1, at 463–77 (discussing the striking patterns among the states' and the nation's constitutions).

61. SANFORD LEVINSON, CONSTITUTIONAL FAITH 54–55, 57, 91–93, 220 n.6 (1988).

62. Akhil Reed Amar, *Civil Religion and Its Discontents*, 76 TEXAS L. REV. 1153 (1989) (reviewing LEVINSON, *supra* note 61).

63. AMAR, *supra* note 1, ch. 10.

64. U.S. CONST. art. II, § 1, cl. 8; Abraham Lincoln, *First Inaugural Address* (Mar. 4, 1861), *in* ABRAHAM LINCOLN, *supra* note 8, at 215, 224.

65. BOBBIT, *supra* note 5, at 94.

66. JACK M. BALKIN, CONSTITUTIONAL REDEMPTION: POLITICAL FAITH IN AN UNJUST WORLD 2–3 (2011); *see also* Stephen M. Griffin, *How Do We Redeem the Time?*, 91 TEXAS L. REV. 101 (reviewing BALKIN, *supra*, and contrasting Balkin's take on the constitutional narrative with the narrative of discontinuity); Lawrence B. Solum, *Faith and Fidelity: Originalism and the Possibility of Constitutional Redemption*, 91 TEXAS L. REV. 147 (reviewing BALKIN, *supra*,

course is central to our constitutional culture. My 2005 book's opening chapter ended with extended quotations from Lincoln on the secession issue,[67] and my new book features Lincoln prominently throughout its twelve chapters—especially its chapter (Chapter 6, to be specific) on America's cultural icons and symbols. Lincoln is iconic. So are many of his great texts—the First and Second Inaugurals, the Gettysburg Address, the Emancipation Proclamation, and so on. Most of the biggest Supreme Court cases in any given year pivot on the Fourteenth Amendment, an amendment that—along with the adjacent Thirteenth and Fifteenth—memorializes the constitutional vision of Lincoln and his party. My defense of Lincoln aligns me with this defining national narrative—with America's mainstream culture and ethos. But if you are right, Sandy, Lincoln was . . . what? A lawless butcher? A man whose vision in 1861 was no better, legally, than Jeff Davis's?

In turn, this leads to other differences in temperament between us. You are more comfortable playing the role of gadfly. I by contrast aim to offer a more orthodox account—both of what the Constitution rightly read means, and about the proper rules for reading it.

My 2005 book focused more on *what* the "right answers" to various constitutional questions are, and my new book highlights *how* we go about finding those right answers. By temperament, you are more agnostic about just how many right answers there may be, and also more playful about permissible interpretive methods. But I do share your view that the text of the Constitution only gets us so far.[68] That is indeed the unifying methodological theme of my new book—the need to go beyond text in various ways, even as we ultimately remain faithful to the text (rightly read). My antisecession argument is an apt case study. Alongside what the text explicitly says ("Constitution," "more perfect union," "Treason," "supreme Law of the Land, notwithstanding" "Contrary" state rules, etc.), I emphasize many unwritten elements. I highlight what the text pointedly *omits* and portentously does *not* say: "Confederation," "league," state "sovereignty," and so on. I highlight not just what Federalists did say in the ratification conversation, but what they did *not* say—what they NEVER said, namely that states would retain a right of unilateral secession. And I tie all these points together with a *structural* argument requiring us to read the document as a whole instrument, centrally aimed at achieving a geostrategic continental union with defensible borders. The need for structural argumentation—for reading between the lines of various clauses—is the main theme of the opening chapter of my new book.[69]

---

specifically the issue of whether the Balkin's narrative of progressive constitutional faith can be reconciled with constitutional fidelity).

67. AMAR, *supra* note 11, at 51–53.

68. LEVINSON, *supra* note 3, at 17–22.

69. AMAR, *supra* note 1, at 1–5.

Which takes me, finally, to the link between the ideas that I sketch out in the opening chapters of my two most recent books and the ideas about America's current "crisis" that you present in your new opening chapter. I have argued that America has been strong and free for most of its history precisely because the Constitution structured an indivisible geographic union prohibiting unilateral exit. This union kept states from warring against each other (about western land) and enabled them to keep Europe from intermeddling in the American heartland. Liberty thrived because no major standing army in peacetime was necessary for the first 150 years of our constitutional existence. Thanks to the Constitution, the Louisiana Purchase, the Monroe Doctrine, and Manifest Destiny, American liberty was protected first and foremost by our vast oceanic moats (just as England and Scotland, once unified, were safe from foreign invasion thanks to the English Channel). As late as 1945, America benefited from this geostrategic situation. Because of our wide oceans, only Pearl Harbor was bombed—not New York, or San Francisco, or Austin. Meanwhile everyone else in the world fought the war in their own back yards.

Sandy, if the Confederacy had prevailed in the 1860s, heaven help us! Venturing into counterfactual history/science fiction is always perilous, but who knows whether the White Supremacists in a Texas-dominated Confederacy in 1941 would have allied with the USA or with the Nazis? So I continue to think you were mistaken way back when we first started our debate—in Austin, in 1986—and I think you continue to misdiagnose America's constitutional situation today.

You think there is a genuine "crisis" of governance and you tend to blame the Constitution. I see things differently. First, I don't blame the Constitution for most of our problems; I blame the fact that too many of our fellow citizens are kooky—beginning with the Governor of your own state and with his many admirers in Texas and elsewhere. If a huge proportion of Americans have outlandish views, there is only so much that constitutional forms can do. Second, while I admit that foreign parliamentary systems don't have all the same pathologies as presidentialist models—there is less gridlock abroad—I think that foreign parliamentary systems have offsetting pathologies. For example, a plurality party that does not in fact have a genuine mandate for change and that has never won the considered support of the median voter[70] might nevertheless be able to effect major policy change—perhaps for the worse. Parliamentary incumbents can manipulate electoral timing with "snap" elections; and policy can sometimes shift drastically in the wake of a single low-turnout election.

---

70. *See* DUNCAN BLACK, THE THEORY OF COMMITTEES AND ELECTIONS 16–25 (1998) (explaining the median voter theorem); Andrew D. Martin et al., *The Median Justice on the United States Supreme Court*, 83 N.C. L. REV. 1275, 1281 (2005) (noting that the Medium Voter Theorem states that "if voters have single-peaked preferences in a single-dimensional issue space, then the position of the median will prevail under majority rule and various voting procedures" (footnotes omitted)).

So I think that America's situation is not really more critical than, say, Greece's or Italy's or Japan's or France's or Britain's. America is no longer quite the hegemon it was in 1945, but this is because in 1945 America was the only nation standing (thanks to our geostrategic isolation and unification). Today, other nations are now back on their feet—and this is a good thing. Today, many other nations are genuinely democratic—following in America's footsteps!—and this, too, is a good thing for America and the world. True, many of these other nations are parliamentary and not presidential, but both brands of constitutional democracy are viable and attractive. Other countries have workable multiparty systems, but two parties, each vying for the median voter, have worked fairly well for America over the long haul.[71] True, America's Constitution is very hard to amend—but because of this fact, very few bad amendments have prevailed over the course of history, while many good amendments have ultimately cleared the bar. Are state constitutions, which are much easier to amend, generally more functional and more admired by Americans than the federal Constitution?

In short, America is no longer the towering hegemon it was in 1945 because many other countries are beginning to emulate our democratic system; because many other nations have now thankfully demilitarized; and because Americans now spend way more than we did prior to 1945 on national and world defense, and way more than does most of the rest of the world. Plus, our system is truly continental, which poses unique challenges and opportunities. Sandy, even if you think Britain has a better constitutional system than America, does Europe as a whole? C'mon!

In short, Sandy, I am doubtful that there is a genuine crisis in America that would be solved by major constitutional reform. That said, I actually agree with many of the specific reform proposals you favor—direct national election of the President, a less malapportioned Senate, quicker transitions of power after national elections[72]—but I frankly don't think any of these reforms would make a major difference solving America's biggest problems, problems which are, to repeat, no more daunting than the problems facing other nations and regions today. And even more happily, some of the problems that you have identified might be solvable without the need to formally amend the Constitution, as I have explained in the concluding chapter of my new book, and elsewhere. Perhaps in our next go-round we could talk about some of the reforms we would both like to see?

Fondly,

Akhil

---

71. AMAR, *supra* note 1, ch. 10.

72. LEVINSON, *supra* note 3, at 186–90, 154–58, 24–27.

III. Dear Akhil,

So what is the key difference between us concerning secession? By and large, our argument is academic, in both the descriptive and perhaps, for some, pejorative sense. That is, I think it boils down to whether the argument for secession—what you call "unilateral withdrawal" from the Union—is "frivolous," in the sense that no reasonable lawyer could possibly present the argument and, indeed, it would merit sanctions under Rule 11 of the Federal Rules of Civil Procedure[73] if (s)he did so to a federal court. I continue to think that the answer is no, even if I certainly agree with you that almost all contemporary lawyers may well see no merit at all in such a claim and even, more controversially, that most lawyers in 1860 would have been equally dismissive. I once wrote an article on "frivolous cases,"[74] not least because at that time I was teaching "professional responsibility"; one of the most important questions, both practically and jurisprudentially, facing any lawyer is whether there are indeed professional, and not only prudential, limits on what they can argue. Frank Easterbrook, both then and now a distinguished judge on the Seventh Circuit Court of Appeals (and, of course, a member of the University of Chicago Law School faculty as well), proffered the notion that "frivolousness" required near-unanimous rejection by the professional community of lawyers.[75] To adopt the language of our friends Stanley Fish and Jack Balkin, it has to be so "off the wall" that any "interpretive community" of which one is a party would scoff at it and seriously question the competence of the person asserting it.[76] Perhaps it is simply meant as a joke! Jack, Jordan Steiker, and I once wrote a piece in the *Texas Law Review* that questioned whether one could always tell the difference between "serious" arguments and "parodies" of legal argument.[77]

I think you make extremely powerful arguments on why the "better reading" of the Constitution prevents "unilateral withdrawal." But, as I've already suggested, that's not the most basic question. It is, rather, whether the Southern reading, based in part on the "compact theory" of Union enunciated in the Kentucky and Virginia Resolutions[78] written by Jefferson and Madison, respectively, was a "possible" reading. You will tell me, of course, altogether correctly, that Madison rejected, near the end of his life, when John Calhoun and other South Carolina hotheads began bruiting about

---

73. FED. R. CIV. P. 11.

74. Sanford Levinson, *Frivolous Cases: Do Lawyers Really Know Anything at All?*, 24 OSGOODE HALL L.J. 353 (1986).

75. *Id.* at 375.

76. Jack M. Balkin, *How Social Movements Change (or Fail to Change) the Constitution: The Case of the New Departure*, 39 SUFFOLK U. L. REV. 27, 28 (2005); Stanley Fish, *How Come You Do Like You Do? A Response to Dennis Patterson*, 72 TEXAS L. REV. 57 (1993).

77. Jordan Steiker, Sanford Levinson & J.M. Balkin, *Taking Text and Structure* Really *Seriously: Constitutional Interpretation and the Crisis of Presidential Eligibility*, 74 TEXAS L. REV. 237 (1995).

78. LEVINSON, *supra* note 3, at 322–23.

the possibility of secession, the legitimacy of reading the 1798 Resolutions as supporting secession.[79] I have no doubt that this was Madison's sincere belief. But, of course, authors do not control the use that readers make of their arguments.

Also, for what it is worth, I think that the word "unilateral" is doing a lot of work in your argument. I'm curious what you think of the Quebec Secession Reference case decided by the Canadian Supreme Court, which seemingly held both that Quebec had no right, under either Canadian or international law, unilaterally to secede from Canada, *but* that Canada might be under a duty to negotiate with Quebec about possible terms of secession if the province clearly indicated a desire to leave the country.[80]

Finally, although I happen to agree with you that Robert E. Lee and Jefferson Davis might well have been tried (and executed) as traitors to the Union instrumental in causing the death of 750,000 Unionists and Confederates on behalf of the totally abhorrent cause of maintaining chattel slavery, that is not the way our history worked out. To be sure, Lee and Davis were not pardoned during their lives, but it is illuminating—and perhaps discouraging—that Congress in 1975 and 1978 passed Joint Resolutions posthumously restoring full rights of citizenship to both.[81] In signing the 1978 Resolution concerning the would-be President of the Confederacy Davis, President Jimmy Carter, wrote that

> [i]n posthumously restoring the full rights of citizenship to Jefferson Davis, the Congress officially completes the long process of reconciliation that has reunited our people following the tragic conflict between the States. Earlier, he was specifically exempted from resolutions restoring the rights of other officials in the Confederacy. He had served the United States long and honorably as a soldier, Member of the U.S. House and Senate, and as Secretary of War. General Robert E. Lee's citizenship was restored in 1976. It is fitting that Jefferson Davis should no longer be singled out for punishment.
>
> Our Nation needs to clear away the guilts and enmities and recriminations of the past, to finally set at rest the divisions that threatened to destroy our Nation and to discredit the principles on which it was founded. Our people need to turn their attention to the important tasks that still lie before us in establishing those principles for all people.[82]

---

79. DANIEL FARBER, LINCOLN'S CONSTITUTION 62–69 (2003).

80. Reference re Secession of Quebec, [1998] S.C.R. 217, 271–73 (Can.).

81. S.J. Res. 23, 94th Cong., 89 Stat. 380 (1975); S.J. Res. 16, 95th Cong., 92 Stat. 1304 (1978).

82. President Jimmy Carter, Statement on Signing S.J. Res. 16 into Law (Oct. 17, 1978), *available at* http://www.presidency.ucsb.edu/ws/index.php?pid=29993.

It was, of course, William Faulkner, the great Mississippi-born novelist, who wrote "The past is never dead. It's not even past."[83] So it is the great conflagration of 1861–1865 which continues to shape our history and our sense of constitutional possibility a full 150 years after its occurrence. If we really viewed Lee and Davis as "traitors" who committed "treason," then Carter's remarks and the congressional resolution would be as unthinkable as the British placing a monument honoring George Washington in Trafalgar Square.

There is also, let me note, a certain irony in pointing to the British Treaty of Union of 1707 between England and Scotland (that created the United Kingdom), for Scotland will be voting next year on withdrawing from the United Kingdom and thus basically undoing the Treaty.[84] The vote is occurring with the reluctant approval of Her Majesty Queen Elizabeth II's Government,[85] and we certainly do not know what the consequences of an affirmative vote would be. But, again for better and for worse, secessionist movements *are* alive and well in today's world, whether based on abstract notions of "self-determination" and, as The Declaration of Independence put it, the fundamental norm of government by "consent of the governed,"[86] or on more legalistic notions. After all, the oft-derided Soviet Constitution, in its Article 72, explicitly permitted the secession of the constituent Soviet Socialist Republics.[87] To be sure, one can presume that it never occurred to any member of the Soviet elite that Article 72 would be taken seriously. But there it was, to provide a legitimizing rhetoric for the republics wanting to leave the Soviet empire. But enough about secession, which we can both agree is not a live political possibility in the contemporary United States and, therefore, renders legal arguments of no practical interest.

I want to move on to the last part of your letter. It is certainly true that I *do* blame the Constitution for the present pickle we are in regarding the actual inability of the American national government to engage in any serious attempt to resolve the manifest challenges that face us as a polity. One need not look far to read respected analysts and political pundits refer to the present American political system as "dysfunctional" or even, in Thomas L. Friedman's word, "pathological."[88] I quite deliberately begin

---

83. WILLIAM FAULKNER, REQUIEM FOR A NUN 81(1st Vintage Books ed. 1960) (1950).

84. Anthony Faiola, *Britain, Scotland Sign Deal to Allow Independence Vote*, WASH. POST, Oct. 15, 2012, http://articles.washingtonpost.com/2012-10-15/world/35501979_1_scottish-national-party-independence-vote-scottish-referendum.

85. Agreement Between the United Kingdom Government and the Scottish Government on a Referendum on Independence for Scotland, Oct. 15, 2012, *available at* http://www.scotland.gov.uk/Resource/0040/00404789.pdf.

86. THE DECLARATION OF INDEPENDENCE para. 2 (U.S. 1776).

87. KONSTITUTSIIA SSSR (1977) [KONST. SSSR] [USSR CONSTITUTION] art. 72.

88. THOMAS L. FRIEDMAN & MICHAEL MANDELBAUM, THAT USED TO BE US: HOW AMERICA FELL BEHIND IN THE WORLD IT INVENTED AND HOW WE CAN COME BACK 243 (2011) (lamenting the "pathologies of the [American] political system").

Framed with a review of such comments by Friedman and others.[89] If one does not find these descriptions at all plausible, then in a deep sense, we have nothing to discuss. Those who believe that nothing is broken also, obviously, see nothing that may need fixing. But I don't think you are so complacent.

Let me be clear, though, that when I say the Constitution deserves its share of the blame, I do *not* mean to say that it is even remotely the sole cause of our present discontents. All I want to insist is that it is at least a *partial* cause, with the exact weight inevitably unknown. Perhaps it is only 5%, perhaps 25%; even I do not argue for a larger number. But I think it is important to realize that in times of crisis, like being faced with the prospect of driving over a cliff, factors that in calmer and happier times might be nearly irrelevant can suddenly precipitate the ultimate disaster. "I knew I should have checked those brakes, but they always stopped in time before." Or consider the aspirin tablet, a true friend of humankind in many important respects. Like most people my age, I take an aspirin every night before retiring, secure in the knowledge that it will do its part to prevent heart attacks and, apparently, many other diseases. That is the unequivocal good news. *But* it turns out that aspirin can be literally fatal under certain circumstances, either by *interacting* with other drugs in a decidedly negative way or by preventing the body, say, from forming necessary blood clots when bleeding.[90] (This is why one is told not to take aspirin before undergoing surgery, for example.)[91] So it is with the Constitution. I certainly don't have to agree that it is deficient in all respects at all times; that is not my view. Nor do I even have to agree with William Lloyd Garrison that prior to the Reconstruction Amendments, it was a "covenant with Death and an agreement with Hell," though that is in fact my view. That is irrelevant with regard to determining the consequences of living under our Constitution in 2013. All I have to do is to persuade you—and, I hope, other readers—that the framework of government established in 1787 and left remarkably unchanged in too many important respects thereafter has become toxic when interacting with other aspects of our polity and political culture.

Consider, for example, the recent book by Norman Ornstein and Thomas Mann, *It's Even Worse Than It Looks: How the American Constitutional System Collided With the New Politics of Extremism*.[92] It is, as they suggest, "the American constitutional system" that has rendered pathological the capture of one of our two leading political parties by

---

89. LEVINSON, *supra* note 3, at 7–12.

90. *E.g.*, Shauna S. Roberts, *Take Two: Aspirin*, MODERN DRUG DISCOVERY, Oct. 2000, at 23 *available at* http://pubs.acs.org/subscribe/archive/mdd/v03/i08/html/10health.html; *Aspirin Disease Interactions*, DRUGS.COM, http://www.drugs.com/disease-interactions/aspirin.html.

91. *See* Kenneth S. Scher, *Unplanned Reoperation for Bleeding*, 62 AM. SURGEON 52, 52 (1996) (finding that preoperative use of aspirin was associated with heavy bleeding in most patients).

92. THOMAS E. MANN & NORMAN J. ORNSTEIN, IT'S EVEN WORSE THAN IT LOOKS: HOW THE AMERICAN CONSTITUTIONAL SYSTEM COLLIDED WITH THE NEW POLITICS OF EXTREMISM (2012).

ideological extremists who have adopted a scorched-earth approach to politics.[93] James Madison and his friends were desperate to stave off the rise of political parties, which could well be viewed as exemplars of what Madison so memorably called "factions" in *Federalist No. 10*.[94] Quite obviously, they utterly failed, and Madison himself became one of the key leaders of the Jeffersonian Democratic-Republicans who contended with Hamiltonian Federalists. But we have never in our 225-year history truly figured out how to integrate partisan political parties and the "divided governments" they often produce (as has been true during most of the lifetimes of most of our readers) into our system of "separation of powers." As Daryl Levinson (no relation) and Richard Pildes have argued, we must fully confront the implications of "separation of parties" instead of being fixated on eighteenth-century notions of "separation of powers."[95]

The written Constitution is almost wholly devoid of anything helpful in this regard, save the importance of the implicit recognition within the Twelfth Amendment that presidents would be elected as the result of partisan elections and, therefore, that there should be separate tracks for the president and vice president instead of asking electors to vote for the two individuals they thought were best qualified to be president[96] (which gave us the very bad political marriage of John Adams and Thomas Jefferson in 1797–1801).[97] Do you believe that the "unwritten Constitution," properly understood, provides any genuine guidance with regard to preventing the further breakdown of our political system?

Finally, I want to resist your invitation to spend much of our exchange on specific reforms. I certainly have many in mind, some of which I know you agree with. They concern such subjects—and this is only a partial list—as the electoral college, the allocation of voting power in the Senate, life tenure for Supreme Court justices, and perhaps most importantly, changing the Draconian Article V that makes it next to impossible seriously to amend the Constitution with regard to anything of genuine significance (and controversy). But what most dismays me, and I'm afraid sometimes turns me into something of a crank, is that there is no serious conversation at all taking place at the national level about *any* kind of serious constitutional reform. What I strongly desire, as you know, is a new constitutional convention. I think there is much to learn from, and emulate, in the fact that there have been 233 *state* constitutional conventions in our national history and that each of the fifty states has had an average of almost three

---

93. *Id.* at xiii–xiv.

94. THE FEDERALIST NO. 10 (James Madison), *supra* note 22, at 42.

95. Daryl J. Levinson & Richard H. Pildes, *Separation of Parties, Not Powers*, 119 HARV. L. REV. 2311, 2311, 2315 (2006).

96. U.S. CONST. amend. XII.

97. *See* MARK O. HATFIELD, SENATE HISTORICAL OFFICE, VICE PRESIDENTS OF THE UNITED STATES (1789–1993), at 21–23 (1997) (discussing the differing views of Jefferson and Adams).

constitutions.[98] The 1787 Constitution, as you know very well, ruthlessly displaced the six-year-old Articles of Confederation, our first constitution, and what makes the Philadelphia convention and the ratifying process afterwards so inspiring to many of us was precisely the willingness of leading figures of the time to engage in full and frank—often quite brilliant—debate about the adequacy of our political institutions. *That's* what is missing today.

There are two Madisons who haunt our national history. One is the Madison of *Federalist No. 49*, who sharply rejected the advice of his friend Thomas Jefferson, a supporter of frequent conventions and the scrutiny they would bring, in favor of trying to create a national culture of "veneration" of the Constitution.[99] The other is the Madison of *Federalist No. 14*, my favorite of all of the 85 essays that he, Hamilton, and Jay wrote to defend what came out of Philadelphia. I cannot too often reread (or quote) the final paragraph of that essay:

> Is it not the glory of the people of America that, whilst they have paid a decent regard to the opinions of former times and other nations, they have not suffered a blind veneration for antiquity, for custom, or for names, to overrule the suggestions of their own good sense, the knowledge of their own situation, and the lessons of their own experience? . . . Had no important step been taken by the leaders of the Revolution for which a precedent could not be discovered, no government established of which an exact model did not present itself, the people of the United States might, at this moment have been numbered among the melancholy victims of misguided councils, must at best have been laboring under the weight of some of those forms which have crushed the liberties of the rest of mankind. Happily for America, happily, we trust, for the whole human race, they pursued a new and more noble course. . . . They formed the design of a great Confederacy, which it is incumbent on their successors to *improve* and perpetuate.[100]

I have no hesitation, therefore, in embracing *this* Madison by suggesting that the lessons of our own experience, in today's America, suggest that we need to ask hard questions about what aspects of our Constitution, written or unwritten, are worth preserving and which, concomitantly, should be changed.

Finally, I repeat my entreaty from my first letter: How should we change our pedagogy—and perhaps our casebooks as well—in order to instantiate the visions of constitutional discourse that we believe are too often missing within the crabbed confines of the contemporary legal academy? In many ways, you have the easier task, for, as you emphasized, you see your

98. JOHN J. DINAN, THE AMERICAN STATE CONSTITUTIONAL TRADITION 7 (2009).

99. THE FEDERALIST NO. 49 (James Madison), *supra* note 22, at 311, 314.

100. THE FEDERALIST NO. 14, (James Madison), *supra* note 22, at 99–100 (emphasis added).

central task as preparing students to be lawyers, though ones who are sensitized to the importance of making new arguments, based on the "unwritten Constitution," that may be lacking in their present education. My critique, in contrast, is directed far more at training our students to become better *citizens* and leaders, capable of asking how well the Constitution serves us as a structure of governance, rather than better lawyers as such. It is, after all, one of my central arguments that there are two Constitutions, the "Constitution of Conversation" that we obsess about in our classes and articles and the "Constitution of Settlement" that we ignore precisely because there is really nothing to argue about in terms of "interpretation" or "approaches to the Constitution" with respect to the meaning of "two senators," January 20, or the various percentages of the vote needed to override presidential vetoes or propose or ratify constitutional amendments. They raise profound questions about *wisdom*, but legal education shies away from those discussions in favor of endless debates about interpretation—such as the possible legitimacy of secession!

So, on to you, my friend, with, as always, deepest esteem and best wishes,

Sandy

IV. Dear Sandy,

So much to talk about, and so little time! So let me narrow the issues down. I don't know enough about Canadian constitutional law or international law to have strong views about Quebec. Nor do I have much to say about how best to construe a Soviet Constitution born in blood and terror, force and fraud—unlike America's Constitution, which emerged via a continental vote and conversation far more democratic and participatory than anything the world had ever seen. (On this last point, see Chapter 1 of *America's Constitution* and Chapter 2 of *America's Unwritten Constitution.*) For better or worse, I have concentrated on American constitutional law—its text, history, structures, doctrines, traditions, and so on. As for England and Scotland, please note that Britain as a whole has agreed to put various issues of Scottish autonomy/independence on the agenda.[101] So nothing happening there today retroactively supports the right to unilaterally secede claimed by South Carolina in 1860.

On my view, then, were there no legal arguments on South Carolina's side? I needn't say that. I need only insist that whatever legal arguments South Carolina did make and could have made, these arguments were *clearly* worse—legally, as measured by the proper rules and modalities of constitutional interpretation—than the contrasting Unionist arguments. (An analogy: Obama did not need to win every electoral vote. He just needed to

101. *See supra* note 85 and accompanying text.

win more electoral votes than the other guy. He did, and that's why he was *clearly* reelected.)

So why then did South Carolinians and others in 1860 make the legal arguments they did in support of secession? Not, on my view, because these arguments were legally strong when compared to the decisive Unionist refutations. The deal in 1787–1788 was straightforward: The text meant what it said about "the supreme Law of the Land" and virtually no one in that fateful year said that a state could secede unilaterally. But for all its admirable and democratic features, the Founders' Constitution failed to put slavery on a path of extinction, and in fact featured key provisions—especially the Three-Fifths Clause[102]—that ended up giving slavocrats far too much power in antebellum America. With that undue power, Southern slavocrats in the mid-nineteenth century tried to rewrite history, twist law, and suppress truth. Ultimately, this slave-sick Southern society tried to undo a lawful and democratic election by force of arms, and nearly destroyed the last, best hope of earth—government of, by, and for the people.

If my account is right, what follows? Among other things, we can see more clearly where the Founders went right and where they went wrong, and what we must do today—the very issues, Sandy, that you want our fellow citizens to ponder. But our fellow citizens are unlikely to think straight about these issues unless they understand history and law with rigor and precision. Contrary to what you sometimes seem to be saying (or might be misunderstood as saying), the Founders did not go wrong by being unclear about secession. But they did go wrong in failing to credibly address the deadly cancer of slavery in their midst. This failure, and this cancer, almost killed America. And even today, there are remnants of this Founding failure in our existing constitutional regime. The electoral college is a close cousin of the Three-Fifths Clause. It was initially designed in 1787, and redesigned in 1803–1804 (via the Twelfth Amendment), to accommodate southern slaveholding.[103] In a direct-election world, the South would have consistently been outvoted, because of course slaves could not vote.[104] But thanks to the electoral college, the South could count its slaves in the electoral college; the more slaves a slave state had, the more House seats and electoral votes that slave state would get[105]—a truly vicious system.

So a more careful and correct legal story about secession and slavery can actually help our fellow citizens see that perhaps the electoral college should be modified today. The college's roots are in fact intertwined with slavery and the Three-Fifths Clause. And even today, our presidential election system has democracy-dampening features, compared to direct national election. Just as Virginia at the Founding got a fixed number of

---

102. U.S. CONST. art. I, § 2, para. 3, *repealed by* U.S. CONST. amend. XIV.

103. U.S. CONST. art. II, § 1, para. 2–3, *amended by* U.S. CONST. amend. XII.

104. *Wesberry v. Sanders*, 376 U.S. 1, 27 (1964).

105. *Id.*

electoral votes regardless of how many or how few persons it enfranchised—and even though it disfranchised a vast proportion of its population—so Virginia today gets a fixed number of electoral votes whether lots of Virginians or very few show up on election day. But in a reformed, Levinson–Amar world of direct popular election, Virginia's government would have incentives to facilitate voting in Virginia. The more voters from Virginia, the more clout Virginia would have in the overall presidential contest; the fewer voters, the less clout.

In Chapter 12, I explain how a direct-popular system could actually take root without the need for an Article V amendment.[106] (Once this system actually began to operate, it would then become easier to adopt a formal amendment codifying the new status quo.) Chapter 12 also offers up a way to unclog the amendment process more generally via textual amendments that could be voted upon now but that would not go into effect—would not "sunrise"—until far in the future, enabling the current generation to become Rawlsian framers, behind a veil of ignorance, for generations yet unborn.[107] I first floated this idea in my 2005 book,[108] and I was delighted to see that in your first two chapters, you, too, sought to invite "readers [to] detach themselves from the immediate political moment by contemplating the powers they would wish (or at least be willing) to grant the (unknown and unpredictable) president who will be elected in 2016 (or 2020)."[109] To do you one better, Sandy, how about an amendment agreed to now that would only go into effect in 2056 or 2060? With such a time horizon, it might be possible for today's Americans, in your words, to "tame some of the partisan passions almost necessarily present if we focus on known political leaders or groups."[110]

---

106. AMAR, *supra* note 1, at 456–63.

107. *Id.* at 475.

108. AMAR, *supra* note 11, at 428 n.*.

109. LEVINSON, *supra* note 3, at 44.

110. *Id.* Your mention of "partisan[ship]" in this passage does prompt me to respond to what you say about political parties in your most recent letter. As you know, I devote an entire chapter to political parties—Chapter 10, "Joining the Party: America's Partisan Constitution." In that chapter, I try to show how parties are in fact more visible in the text of the written Constitution than is conventionally understood, and how these parties are also an enormous feature of America's unwritten Constitution. You seem to think America's two-party system itself is dysfunctional. LEVINSON, *supra* note 2, at 10. I say that in general a two-party system is an important stabilizing mechanism in actual American governance, and that our present woes are simply due to the fact that at present, a significant portion of the Republican Party has gone kooky. But I predict that the party will straighten itself out because the system creates strong incentives (called elections) for it to do so. And I had you specifically in mind when I wrote the concluding words of Chapter 10:

> Despite all that we have seen, it cannot be said that the Constitution directly addresses political parties in a comprehensive fashion. Is this because, as some scholars have claimed, the document's rules concerning elections and the political process—especially its provisions governing presidential politics and presidential authority—are the petrified fossils of an eighteenth-century world, wholly ill-fitting the political realities of modern America?

And here we come back to our contrasting accounts of secession. You seem to suggest, in your opening chapter, that the Constitution of 1787–1788 was "resolutely silent" on the secession issue because the issue was perhaps "too volatile."[111] Hogwash. The Constitution was clear, but the war came in the 1860s because the Constitution elsewhere pandered to slavery, and slavery corrupted all it touched—including the truth and proper legal argumentation. But on my account, slavery itself did raise difficult political issues: How could free states and slave states in the 1780s find common ground and come together to form the necessary geographically indivisible union? I claim that the Founders failed not because they compromised with slavery: Some compromise was indeed necessary to launch the Union. But the Three-Fifths Clause was the wrong kind of compromise—the kind of compromise that you in your book refer to as a "rotten" compromise, building on the work of Israeli philosopher Avishai Margalit.[112] The reason the Three-Fifths Clause was truly "rotten" is that it surrendered the future. It gave slavocrats extra political clout in every election in perpetuity. It failed to provide a structural mechanism for the long arc of history to bend toward justice. The framers should have instead used more "sunrise" clauses, I argue.[113] Just as the Constitution allowed the transatlantic slave trade to continue for twenty years, but provided that Congress could ban this odious

---

> The evidence suggests otherwise. At the very moment that national parties arose, they began to integrate themselves into the Constitution in both text and deed. America's modern presidency is not the product of eighteenth-century mistakes that later Americans have simply been unable to comprehend or correct. Although the presidency was originally designed for a nonpartisan figure—George Washington—the office was repeatedly redesigned, via many different amendments adopted over the course of many decades, to fit the rise of more partisan chief executives including Thomas Jefferson, Abraham Lincoln, Franklin Roosevelt, and Lyndon Johnson. Most of the rules of presidential power are robust. These rules first worked without an entrenched two-party system and now work within such a system.
>
> To put the point another way, virtually all states have created governorships that look amazingly like the presidency, and most states created these presidential look-alikes after the rise of America's two-party system. Almost no state constitution comprehensively regulates political parties, even though many written state constitutions are quite detailed and relatively easy to amend.
>
> All this evidence suggests that there is a different reason why political parties receive rather spotty treatment in America's fifty-one written constitutions, state and federal. The explanation, quite simply, is that it is far from clear what a more comprehensive constitutional regulatory framework should look like. . . . [V]irtually no state constitution regulates political parties in dramatically different fashion than does the federal Constitution. Unless and until several state constitutions come along and demonstrate a better mousetrap for addressing American-style political parties, most Americans are unlikely to view the federal Constitution as defective in this regard.

AMAR, *supra* note 1, at 415–16.

111. LEVINSON, *supra* note 3, at 19.

112. *See id.* at 43–53 (considering the Constitution in light of Margalit's theory of "rotten compromise").

113. AMAR, *supra* note 1, at 474–76.

traffic in 1808 and thereafter, so the document should have allowed slave states to get extra credit in the House and electoral college (via the Three-Fifths Clause, or some variant) only until 1808, but not thereafter. Likewise, the Framers could and should have provided that slavery as a legal system would need to begin to end by 1808, and that all persons born in America after 1808 would be born free. Or all persons born after 1818, or 1838, or 1888. The specific date is less important than the big, Rawlsian idea: Since slavery was morally wrong, the Founders should have provided for its ultimate extinction. After 1808 or 1838 or whenever, antislavery rules agreed to in 1788 should have been allowed to "sunrise."

And with that idea, Sandy, I come to the final questions that you have posed. When I teach law students, I do indeed stress proper rules of legal interpretation, and I use the secession issue as a case study in proper legal method. Here was perhaps the most momentous constitutional question ever to arise in America, and the standard legal methods do indeed point to a clearly correct legal result! But when I teach undergraduates, Sandy, I am more concerned with larger issues of citizenship preparation. And so I end my undergraduate class, as I end both *America's Constitution* and *America's Unwritten Constitution*, with a sweeping gesture toward the twenty-second century and beyond. Though I do not believe our very constitutional system is currently in crisis, I do find it notably imperfect, and I urge my students—the future leaders of the twenty-first century—to ponder perhaps the most important aspects of America's unwritten Constitution: the amendments still to be written, and the Constitution of 2020, of 2121, and beyond.

I began this letter by saying that we have so much to talk about and so little time. Of course, I meant only that there is little time left in this final round. But I hope there will be lots of time in other venues down the road. Here's to the next quarter century, my friend!

As ever,
Akhil

V. Dear Akhil,

As I write this literally on December 31, 2012, I can't think of a better way to begin the New Year than by continuing, albeit briefly, the discussion—and looking forward, of course, to far more extended conversations in the future.

I suspect there's not much more to say about secession. I think the most important area in which we agree is not only that the original Constitution was significantly blemished by its "rotten compromises" with slavery, but also that the ostensible "reconstruction" of the nation following the slaughter of roughly 750,000 participants in the warfare of 1861–1865[114] was itself

114. J. David Hacker, *A Census-Based Count of the Civil War Dead*, 57 CIV. WAR HIST. 307, 311 (2011). Relying on newly-released census data from the 1850, 1860, 1870, and 1880 censuses,

blemished by the unwillingness or political inability of those who wished to achieve truly fundamental "regime change" to achieve their objectives. As I write in *Framed*, the benefits to the "slavocracy" of the Three-Fifths Clause were followed, perversely, by the even greater benefits of counting the former slaves as full persons, but still, in much of the former Confederacy, not voting participants.[115] Thus what I called the "segregation bonus" entrenched Southern racists in the House (and, for somewhat different reasons, the Senate) until my adult lifetime.[116] I am aware, of course, of your efforts in your chapter on the Warren Court to read Section 2 of the Fourteenth Amendment as an instrument combating this "segregation bonus."[117] I have no trouble accepting your interpretation of Section 2, and one can imagine an alternate history of American constitutional development in which Section 2 would have been vigorously enforced by Congress (or even by the judiciary). Alas, that is not the constitutional development we in fact got, and Section 2 was functionally rendered a dead letter, much to our mutual dismay. True "regime change" was just too radical to contemplate, especially when the American narrative turned to "reconciliation" of the white North and South.[118]

It may be that the most important disagreement concerns the perception of "crisis" confronting the American political system. Again, as I originally wrote these words, the United States seemed poised to go over what has (probably unhelpfully) become labeled as the "fiscal cliff."[119] Even though I now know (on January 2) that a patchwork agreement between Vice

---

demographic and social historian J. David Hacker has challenged the most frequently cited figure of Civil War fatalities, 620,000, estimating that 750,000 men died as a result of the War, with a probable margin of error of plus or minus 100,000. *Id.* at 311–12.

115. LEVINSON, *supra* note 3, at 183–84.

116. *Id.* at 186–187.

117. AMAR, *supra* note 1, at 187–89. The Amendment reads:

> Representatives shall be apportioned among the several States according to their respective numbers, counting the whole number of persons in each State, excluding Indians not taxed. *But when the right to vote at any election for the choice of electors for President and Vice President of the United States, Representatives in Congress, the Executive and Judicial officers of a State, or the members of the Legislature thereof, is denied to any of the male inhabitants of such State, being twenty-one years of age, and citizens of the United States, or in any way abridged, except for participation in rebellion, or other crime, the basis of representation therein shall be reduced in the proportion which the number of such male citizens shall bear to the whole number of male citizens twenty-one years of age in such State.*

U.S. CONST. amend. XIV, § 2 (emphasis added).

118. *See* DAVID W. BLIGHT, RACE AND REUNION: THE CIVIL WAR IN AMERICAN MEMORY 107 (2001) (arguing that "the most vigorous advocates of reconciliation believed they had to banish slavery and race from the discussion").

119. *See* Tom Geoghegan, *Who, What, Why: Who First Called It a 'Fiscal Cliff'?*, BBC NEWS MAG., Nov. 14, 2012, http://www.bbc.co.uk/news/magazine-20318326 (discussing arguments that the term "fiscal cliff"—used by Chairman of the Federal Reserve Ben Bernanke to refer to the expiration of Bush-era tax cuts and automatic reductions in government spending set for Jan. 1, 2013—was an unhelpful metaphor for what would likely be a more gradual process).

President Biden and Senate Minority Leader Mitch McConnell staves it off,[120] no one seriously believes that it will do anything more than kick the can down the road perhaps only for several weeks, let alone months or years.[121] It is all too likely that we continue to face a debacle over raising the national debt ceiling, which in 2011 led to the downgrading of United States financial instruments,[122] *not* because of worries about the American economic system as such, but because of concerns that were expressed about the American political system. I really do think you underestimate the importance of the alarms being sounded by such sober analysts as Mann and Ornstein or Tom Friedman. And if it is true, as I far more than you believe is the case, that the Constitution itself contributes to the dysfunctionality by virtue of the interactive effects of the separation-of-powers system with what Mann and Ornstein call the basically parliamentary party ethos, especially of the Republican Party, then I think we must address these issues sooner rather than later. More than you, I suspect, I am disinclined to believe that most lawyers have much to say that is helpful concerning these issues precisely because, with rare exceptions, they don't call on our well-developed talents for *interpretation*, but rather, ultimately, the redesign of some of our institutions in light of what Madison called the "lessons of experience" and the best teachings of what Hamilton in *Federalist No. 9* was willing to call "[t]he science of politics."[123]

I do believe that the challenge facing any of us teaching about constitutions in America in the 21st century is to figure out ways genuinely to integrate both the United States and American state constitutions, on the one hand, and the Constitution(s) of Settlement with the Constitution(s) of Conversation, on the other. Sufficient integration raises questions not only of pedagogy, but also, of course, of decisions by law school appointment committees as to what skill sets should be sought in potential teachers and,

---

120. Janet Hook et al., *Congress Passes Cliff Deal*, WALL ST. J., Jan. 2, 2013, http://online.wsj.com/article/SB10001424127887323320404578215373352793876.html.

121. *See* David Brooks & Gail Collins, *The Fiscal Riff*, OPINIONATOR, N.Y. TIMES, Jan. 2, 2013, http://opinionator.blogs.nytimes.com/2013/01/02/the-fiscal-riff/?hp ("Not averted. Postponed. We've got another catastrophe coming in a couple of months with the return of sequestration. And by the way how is anybody in government actually supposed to plan a budget when the whole thing may blow up again in 60 days?").

122. In a statement published August 5, 2011, Standard and Poors explained that it had lowered its long-term sovereign credit rating on the United States because it believed that the "prolonged controversy" over raising the debt ceiling and "the related fiscal policy debate" indicated that "further near-term progress containing the growth in public spending, especially on entitlements, or on reaching an agreement on raising revenues is less likely than [it] previously assumed and will remain a contentious and fitful process." *United States of America Long-Term Rating Lowered To "AA+" Due To Political Risks, Rising Debt Burden; Outlook Negative*, STANDARD & POORS (Aug. 5, 2011, 8:13 PM EST), http://www.standardandpoors.com/ratings/articles/en/us/?assetID=1245316529563.

123. THE FEDERALIST NO. 9 (Alexander Hamilton), *supra* note 22, at 72; Letter from James Madison to Thomas Cooper (Mar. 23, 1824), *in* 9 THE WRITINGS OF JAMES MADISON: 1819–1836, 177, 181 (Gaillard Hunt ed., 1910).

ultimately, decisions by curriculum committees as to how much "constitutional law" should be required, as against being left to the market-system of electives. Both of us teach at schools that continue to emphasize the importance of "constitutional law." But Yale and The University of Texas may become increasing outliers in this regard, especially as the legal academy responds to the economic crises that are reshaping legal education in front of our eyes (and about which I know that you have especially strong views).

There is certainly not time now to explore all of the implications raised by various conceptions of what my friend and colleague Gary Jacobsohn calls "constitutional identity."[124] But I certainly hope that we have many years together of active conversation and colleagueship exploring those implications.

Sandy

124. GARY JACOBSOHN, CONSTITUTIONAL IDENTITY (2010).

* * *

Unsettling the Settled: Challenging the Great and Not-So-Great Compromises in the Constitution

FRAMED: AMERICA'S 51 CONSTITUTIONS AND THE CRISIS OF GOVERNANCE. By Sanford Levinson. New York, New York: Oxford University Press, 2012. 448 pages. $29.95.

Reviewed by Robert F. Williams*

Sandy Levinson has, once again, written an extremely interesting and provocative book. It follows rather directly from his 2006 *Our Undemocratic Constitution: Where the Constitution Goes Wrong (And How We the People Can Correct It)*,[1] continuing his "loving criticism"[2] of the American federal Constitution. Levinson's overall thesis is that the United States Constitution was framed in an atmosphere of national crisis, resulting in a number of compromises as to governmental structures that were understandable at the time but which may have become dysfunctional and in need of change after several centuries of operation.[3] He points to the tremendous growth of the American territory and population, together with the unanticipated rise of political parties, as providing a partial explanation for the current "crisis in governance" that he describes in the book.[4] He contends that we are trapped, or "framed," by the view that federal governmental structures that are entrenched in the Constitution cannot (and should not) be changed.[5] He asks "whether fears that made sense in 1787 need control us today."[6]

Levinson reviews the "crisis in governance" at both the national and state levels. He describes the "gridlock" in Washington, D.C., in areas such as major policy initiatives, approval of judicial nominations, ratification of

---

* Distinguished Professor of Law, Rutgers University School of Law, Camden; Associate Director, Center for State Constitutional Studies.

1. SANFORD LEVINSON, OUR UNDEMOCRATIC CONSTITUTION: WHERE THE CONSTITUTION GOES WRONG (AND HOW WE THE PEOPLE CAN CORRECT IT) (2006) [hereinafter LEVINSON, OUR UNDEMOCRATIC CONSTITUTION].

2. SANFORD LEVINSON, FRAMED: AMERICA'S 51 CONSTITUTIONS AND THE CRISIS OF GOVERNANCE 32 (2012) [hereinafter LEVINSON, FRAMED]; *see also* Sanford Levinson, *Introduction: Imperfection and Amendability*, *in* RESPONDING TO IMPERFECTION: THE THEORY AND PRACTICE OF CONSTITUTIONAL AMENDMENT 3 (Sanford Levinson ed., 1995) [hereinafter Levinson, *Introduction*] (contending that the Constitution is imperfect).

3. LEVINSON, FRAMED, *supra* note 2, at 12, 34–40.

4. *Id.* at 7, 9.

5. *Id.* at 8.

6. *Id.* at 215.

treaties, etc.[7] Further, he reminds us of the dysfunction and possible "ungovernability" of states like California.[8] He acknowledges that not all problems arise from the provisions of the formal federal and state constitutions themselves, but contends that the "settled" provisions of these constitutions may, in fact, be the root of a number of these problems.[9]

This book, as Professor Levinson proudly notes, is unusual for several reasons. First, its focus is on the provisions of the federal Constitution that are "settled" and therefore not subject to academic debate or analysis, or to judicial interpretation and litigation.[10] Almost the entire focus of American constitutional law, in both political science and law, is on the great questions of interpretation of the Constitution, with very little attention to its clear provisions, such as the date on which the President will be inaugurated. Levinson refers to these "settled" (and, for the most part, unquestioned and accepted) provisions as the "Constitution of Settlement."[11] By contrast, he refers to the open-textured provisions of the Constitution, subject to scholarly debate and judicial interpretation, as the "Constitution of Conversation."[12] He breaks with almost all American constitutional law scholarship by only considering the former:

> This book is far more concerned with analogues to the Inauguration Day Clause than to the Equal Protection Clause. Though their meaning is indisputable, there is nothing trivial about such clauses. In fact, they may better explain the failures of our political system and fears about governability than the "magnificent generalities" explain its successes. . . .
>
> . . . .
>
> . . . Indeed, this book is predicated on the proposition that almost *all* of the Constitution of Settlement is very much worth talking about by anyone interested in the practicalities of American government[.] However, the nature of the discourse about the Constitution of Settlement is quite different from that generated by the Constitution of Conversation. The latter involves constitutional *meaning*; the former involves the *wisdom* of clear constitutional commands.[13]

Secondly, Professor Levinson includes in his analysis recurring references to the constitutions of the fifty American *states*. Today, most "constitutional law" study and scholarship retains an exclusive focus on the

---

7. *Id.* at 1–5.

8. *Id.* at 4–5.

9. *Id.* at 5–7 ("But the formalities can make a real difference.").

10. *Id.* at 19, 25–26.

11. *Id.* at 19.

12. *Id.*; *see id.* at 6 ("This book is very much about constitutional *structures*, and not, for example, about constitutional *rights*.").

13. *Id.* at 19, 23; *see also id.* at 354 ("One lesson is that constitutions of settlement do not necessarily settle, once and for all, the issue under examination."); *id.* at 146–47, 357–58 (contending that the Constitution cannot and should not "settle" issues for every generation).

federal Constitution.[14] Further, he regularly refers to the constitutions of other countries as shedding light on choices reflected in our federal Constitution, the most obvious being their use of a parliamentary rather than a presidential system.

This book is extremely important and useful for a variety of reasons. First, it provides a clear and understandable analysis of the original reasons (often compromises) for many of the structural and seemingly noncontroversial provisions of the Constitution. Levinson refers to the *Federalist Papers* (which often adverted to *state* constitutions), the arguments of the Anti-Federalists, and the debates at the state ratifying conventions. Then, he places these provisions in modern context. He describes the current serious defects in many of the structural arrangements by reference to actual, fairly recent events, as well as to interesting and troubling hypotheticals about things that might happen in the future. Levinson has therefore provided a fascinating review of the theory behind and the actual operation of our Constitution of Settlement.

Levinson dissects most of the compromises in the original Constitution arising from the familiar small-state/large-state clash as well as the North/South, slave-state/free-state conflict. In a chapter on compromise itself (Chapter 2), Levinson notes that compromise is necessary in most aspects of life and is certainly necessary, as Edmund Burke, James Madison, and others recognized, in constitution making, both federal and state.[15] But some compromises are so "rotten"[16] as "to establish or maintain an inhuman regime . . . of cruelty and humiliation, that is, a regime that does not treat humans as humans."[17] Levinson questions whether the constitutional compromises surrounding slavery were "worth it," particularly from the point of view of the slaves themselves.[18] Most people assume that we would not have had a federal Constitution, at least not in 1787, without (1) the "Great Compromise" where the House of Representatives was based on population (including the "3/5 Compromise" which gave the slave states greater representation in the House; that influence spilled over into the Electoral College, thereby gaining a greater say for the southern states in who would become President and, among other things, appoint Supreme Court Justices)[19] and the Senate was based on equal votes for the states; and (2) the continuation of slavery.[20] The slavery question is a very important question,

14. LEVINSON, FRAMED, *supra* note 2, at 28–29; ROBERT F. WILLIAMS, THE LAW OF AMERICAN STATE CONSTITUTIONS 1 (2009); Sanford Levinson, *America's "Other Constitutions": The Importance of State Constitutions for Our Law and Politics*, 45 TULSA L. REV. 813, 813 (2010).

15. LEVINSON, FRAMED, *supra* note 2, at 33, 40.

16. *Id.* at 43 (citing AVISHAI MARGALIT, ON COMPROMISE AND ROTTEN COMPROMISES (2009)).

17. *Id.* at 44 (quoting MARGALIT, *supra* note 16, at 2).

18. *Id.* at 51.

19. *Id.* at 37–38.

20. *Id.* at 38–40.

albeit academic today. But even asking the question may move readers to let their reverence and veneration for the federal Constitution slip a bit to consider whether the current dysfunction of some of these compromise provisions is "worth it" today.

Levinson shines his analytical light on, among other "settled" provisions, the clauses specifying the date of inauguration,[21] state control of elections,[22] eligibility for public office,[23] bicameralism (with particular criticism of the Senate),[24] the presidential veto,[25] the Electoral College,[26] the presidential as opposed to parliamentary system,[27] the unitary Executive[28]—including powers such as pardoning,[29] making treaties,[30] etc.—length of presidential terms,[31] the role of the Vice President,[32] impeachment (only for misconduct and not for incompetence),[33] divided government,[34] the independent judiciary[35] (including methods of selection and judicial review), federalism,[36] methods of amendment,[37] and emergency powers.[38] It is the *wisdom* of these provisions today in which Professor Levinson is interested. He admits that readers might disagree with him as to the actual negative consequences of these provisions ("*empirical* assumptions"), or whether these consequences are "desirable or undesirable" ("*normative* arguments").[39]

Surprisingly, a number of these "settled" provisions turn out to be problematic under Levinson's critical eye. Just a few examples will indicate the fresh look that he provides for many of the provisions we all take for granted. For example, returning to the Inauguration Day Clause, this results in a several-month "lame duck" period for either a defeated president or one who has served his or her second term—longer than the same period for state governors.[40] During this period presidents have issued many questionable

---

21. *Id.* at 22–24.
22. *Id.* at 100–02.
23. *Id.* at 117–19.
24. *Id.* at 142–44.
25. *Id.* at 164–73.
26. *Id.* at 178–83.
27. *Id.* at 175–78.
28. *Id.* at 239–44.
29. *Id.* at 194–201.
30. *Id.* at 201–02.
31. *Id.* at 209–13.
32. *Id.* at 221–28.
33. *Id.* at 213–19.
34. *Id.* at 229–33.
35. *Id.* at 245–48.
36. *Id.* ch. 14.
37. *Id.* at 331–36.
38. *Id.* at 208, 374–83.
39. *Id.* at 7.
40. *Id.* at 22–25.

pardons of convicted criminals, initiated substantial administrative rule-making processes or repeals, and taken a number of other actions which do not take place in, for example, a number of European countries where transitions of power are quite swift.[41] The American bicameral system, with each house having an absolute veto over lawmaking, is one of the undemocratic features of the United States Constitution that has already been pointed out by Professor Levinson.[42] This situation is exacerbated by the structure of the Senate, which with equal votes for each state, permits "the smallest twenty-six states, which together have approximately 17 percent of the national population, [to] elect a majority of the Senate."[43] The presidential veto, of course, is also undemocratic, and Levinson criticizes it for permitting the veto of legislation enacted by both houses based on the policy preference of the President rather than only on constitutional considerations.[44] He contrasts the federal Constitution's single Executive official, the President, with a number of the state constitutions that provide for a "plural" executive, where a number of officers other than the Governor are elected on a statewide basis.[45] He notes that "forty-eight of the fifty states do *not* give their governors the authority to name the attorney general, perhaps the most important single executive branch official in terms of providing potential oversight of the executive branch with regard to criminal conduct."[46] Levinson also points out that the federal unitary Executive, which gives the President "the power to appoint all executive branch officials," "lends a winner-take-all partisan character to presidential elections."[47] Many of us recognize this when we tell our friends that it is important not simply to vote for a President whom one likes, but to remember that the President who is elected will also likely appoint members of his or her party all the way down to postmaster.

The much-maligned Electoral College, of course, does not escape Levinson's criticism, where he describes the process of choosing the President as "quite [a] spectacularly different process [than that for choosing] any state governor, all of whom are elected in statewide popular elections."[48] Levinson notes further the Electoral College's potential for nonmajority-elected presidents, the possibility of "so-called faithless electors who . . . reject their party's candidate in favor of their own idiosyncratic choices," and the "winner-take-all" problem of state electors and "the one state, one vote

41. *Id.* at 24–25.
42. LEVINSON, OUR UNDEMOCRATIC CONSTITUTION, *supra* note 1, at 30–38.
43. LEVINSON, FRAMED, *supra* note 2, at 150.
44. *Id.* at 164.
45. *Id.* at 240.
46. *Id.*
47. *Id.* at 243.
48. *Id.* at 178.

process by which the House breaks deadlocks."[49] The conclusion that Professor Levinson reaches concerning the Electoral College represents a major theme in this book:

> At the end of the day, the electoral college, perhaps like the specific day the Constitution specifies for the inauguration of a new president, simply exemplifies the importance of *path dependence*, the inertial force possessed by past decisions whether or not we believe they make much sense for us today. One can well doubt that "We the People" would maintain the electoral college if the U.S. Constitution were as easy to amend as most state constitutions. That it persists tells us almost nothing about actual public opinion and much about the difficulty of formal amendment.[50]

The federal Constitution has, of course, endured with very few amendments since its ratification in 1789. That record, Levinson notes, is far beyond the average length of duration for national constitutions.[51] This is a consequence of the reality that the federal Constitution is the most difficult in the world to amend, let alone revise, and is generally a revered and venerated document.[52] The "last truly significant change to the Constitution" was in 1951, limiting presidents to two terms.[53] A constitution under which formal change is extremely difficult leads to more change by interpretative methods, either by the judiciary or through "constitutional moments"[54] accomplished by the Legislative and Executive Branches, with the possible acquiescence of the judiciary. State constitutions, by contrast, are much easier to amend and therefore, as Dr. Alan Tarr has observed, state constitutional change has occurred more often (*too* often, some would say) through formal amendment and revision mechanisms.[55]

Professor Levinson points out that the evolution of the structures of state government, made possible through the availability of formal change, has permitted the states to reevaluate, modify, and improve their governmental structures.[56] As Dr. John Dinan has noted, this availability of

---

49. *Id.* at 188. Some states have taken it upon themselves to try to deal with the "non-majority elected president" problem. *See 888-Word Interstate Compact*, NAT'L POPULAR VOTE, http://www.nationalpopularvote.com/pages/misc/888wordcompact.php. *See generally* Robert W. Bennett, *Possibilities and Problems in the National Popular Vote Movement*, 7 ELECTION L.J. 181 (2008) (reviewing JOHN R. KOZA ET AL., EVERY VOTE EQUAL (1 ed., 8th prtg. 2006), and assessing the "State-Based Plan for Electing the President by National Popular Vote").

50. LEVINSON, FRAMED, *supra* note 2, at 190; *see also id.* at 127 (describing Madison's lack of confidence in the ability of ordinary Americans to "exercise genuine political autonomy").

51. *Id.* at 335–37 (citing ZACHARY ELKINS, TOM GINSBURG & JAMES MELTON, THE ENDURANCE OF NATIONAL CONSTITUTIONS (2009)).

52. *Id.*

53. *Id.* at 210. Levinson notes that this settled provision bars even exceptional presidents from serving more than two terms. *Id.* at 212.

54. *Id.* at 339.

55. G. ALAN TARR, UNDERSTANDING STATE CONSTITUTIONS 23–28 (1998); WILLIAMS, *supra* note 14, at 25, 82–83.

56. LEVINSON, FRAMED, *supra* note 2, at 336.

formal change enabled the people of the states to have an actual constitutional conversation about unsettling governmental structures that had been seemingly settled by earlier generations.[57] This kind of conversation has been virtually impossible, with minor exceptions, at the federal level.

Some caution, however, should be exercised in looking to state constitutional arrangements as models for our federal Constitution.[58] This is because the American state constitutions function within the overall federal constitutional structure, and are to some extent limited by that structure. Furthermore, state constitutions operate with respect to subnational polities, rather than a single national polity. As a consequence, at least in the United States, state constitutions' origins, functions, form, and substance all differ from the federal model.[59] American state constitutions draw their essence from the people themselves, who exercise forms of popular sovereignty in adopting, amending, and revising state constitutions, and further in actually participating in constitutional government through their approval at the polls of matters such as the assumption of debt and the approval of gambling programs.[60] Further, the voices of nonelite people such as women,[61] African Americans, Native Americans, Latinos, plaintiffs, union members, and prison reformers, as well as those of opponents of abortion and same-sex marriage have been heard, and sometimes have prevailed, in the processes of state constitutional change.[62]

State constitutions, in contrast to the federal Constitution's *grants* of power to a limited federal government (albeit one expanded through judicial decision and the practice of "constitutional moments"), function primarily to *limit* the residual power the states retained at the time the United States Constitution was ratified.[63] This different function leads to a differing form and content for the state constitutions. For example, they contain long articles on taxation and finance, education, natural resources, etc.,[64] which are the matters that were retained for state competency. In addition, the state constitutions contain much in the way of policy pronouncements that could be relegated to ordinary statutes within the competence of state legislatures. Consequently, care should be taken when looking to state constitutions as substantive models for the federal Constitution. Further, the matters that will

57. JOHN J. DINAN, THE AMERICAN STATE CONSTITUTIONAL TRADITION 5 (2006) ("[S]tate constitution makers' departures from the federal model are primarily attributable to the flexibility of state amendment processes and the resulting opportunities to benefit from institutional knowledge and experience throughout American history . . . ."); WILLIAMS, *supra* note 14, at 25, 82–83; LEVINSON, FRAMED, *supra* note 2, at 8, 14.

58. LEVINSON, FRAMED, *supra* note 2, at 14, 26.

59. WILLIAMS, *supra* note 14, at 15–36.

60. *Id.* at 31.

61. *Id.* at 34–36.

62. *Id.* at 34–35.

63. *Id.* at 27, 249–55.

64. *See, e.g.*, TEX. CONST. art. VII ("Education"); TEX. CONST. art. VIII ("Taxation and Revenue").

arise in any process of amendment or revision of the federal Constitution will differ materially from those that will arise in the parallel state constitutional processes.[65]

Having issued that note of caution, however, state constitutional arrangements (which are much less "settled" than federal constitutional arrangements) such as an elected judiciary,[66] term limits, a plural executive, and direct democracy, are matters that, despite one's view of them as policy matters, do not seem dependent on the differences between state and federal constitutions. This is particularly true for the mechanisms of change, through amendment or revision, of state constitutions. Those do not necessarily have to differ because they are subnational rather than national.[67] Of course one of the criticisms of state constitutions is that they are too easy to amend or revise.[68]

There is always a tension in constitutions between rigidity and ease of change. Thomas Jefferson supported the idea of easily amended constitutions with review every generation.[69] James Madison, by contrast, supported more permanent constitutions.[70] As I have said, "If state constitutional revision is too difficult, constitutionalism overwhelms democracy; if it is too easy, democracy overwhelms constitutionalism. It is difficult to achieve exactly the right balance, and this point might change over time."[71] If many states (and they vary significantly) are too far toward the democratic end of the continuum, then it seems like the federal constitutional system (at least according to Article V) may be too far toward the "constitutionalism" end.

65. *See generally* 3 STATE CONSTITUTIONS FOR THE TWENTY-FIRST CENTURY: THE AGENDA OF STATE CONSTITUTIONAL REFORM (G. Alan Tarr & Robert F. Williams eds., 2006) (describing how traditionally, amendments to state constitutions concern local issues like state revenue rather than national issues). However, recently a number of *national* issues such as same-sex marriage, labor law, health reform, and others have been reflected in *state* constitutional amendments placed on the ballot in some states. Robert F. Williams, *Why State Constitutions Matter*, 45 NEW ENG. L. REV. 901, 903 (2011).

66. WILLIAMS, *supra* note 14, at 290.

67. *But see* Tom Ginsburg & Eric A. Posner, *Subconstitutionalism*, 62 STAN. L. REV. 1583, 1599–600 (2010) (contending that substate constitutions tend to be easier to amend than federal constitutions).

68. *See, e.g.*, James A. Gardner, *The Failed Discourse of State Constitutionalism*, 90 MICH. L. REV. 761, 818–22 (1992) (enumerating a number of apparently frivolous state constitutional provisions and linking their existence with the relative ease of amending state constitutions). *But see* WILLIAMS, *supra* note 14, at 25 ("Because of their relative ease of amendment, state constitutions could be modified through trial and error over the years concerning matters that, for all practical purposes, remain frozen in the federal Constitution.").

69. LEVINSON, FRAMED, *supra* note 2, at 61; WILLIAMS, *supra* note 14, at 363. Levinson reveals that he has discovered his "inner Jefferson." LEVINSON, FRAMED, *supra* note 2, at 396.

70. WILLIAMS, *supra* note 14, at 363.

71. *Id.*

In the face of the reality that the federal Constitution is virtually unamendable, because of the restrictive requirements of Article V,[72] Professor Levinson makes a radical proposal: Do not follow Article V![73] He invokes the crisis atmosphere of the 1780s, analogous to today's crisis in governance at the state and federal levels, which of course led the Framers of the federal Constitution to engage in a "runaway constitutional convention" in violation of the instructions from Congress and the amendment mechanisms of the Articles of Confederation.[74] He describes the issue facing the Framers in 1787:

> The fate of the country was at stake, and one should hardly feel obliged to conform to a provision of the existing constitution [the requirement of unanimity to amend the Articles] that if followed in its clear, unequivocal, and semantically undebatable meaning would doom the enterprise of what Madison and others viewed as absolutely necessary constitutional revision.[75]

Again, what interests Levinson is not debate over the *meaning* of the Constitution's settled provisions but rather an assessment of their *wisdom* in current times. If the consequences of these settled provisions are bad enough, he suggests a process to change them even if it defies the seemingly settled provisions of Article V. This is serious stuff.

An instructive process took place at the state constitutional level where conflicts arose over whether the rules laid down in the first state constitutions for their amendment and revision actually had to be followed, or rather whether the people in the exercise of their revolutionary popular sovereignty could make extralegal but binding changes in their constitutions.[76] Dr. Christian Fritz explained:

> All Americans agreed that the people created government. They differed over when that collective sovereign might be recognized as having exercised its authority. Some recognized a multitude of ways, none of them exclusive, in which the people could express their will. In their expansive view, the people could use the formal procedures articulated in a constitution to amend or dissolve that document. Such procedures were not indispensable and the people's will could be recognized in other ways. On the other hand, some took a more

---

72. LEVINSON, FRAMED, *supra* note 2, at 11. Actually, the provisions of Article V are a bit less clear than most of us have thought. *See generally* RESPONDING TO IMPERFECTION, *supra* note 2 (discussing the difficulties with the interpretation of Article V in a number of essays). Also, Levinson is not alone in suggesting that the formal requirements of Article V be "side-stepped." LEVINSON, FRAMED, *supra* note 2, at 343–44.

73. LEVINSON, FRAMED, *supra* note 2, at 343–44.

74. *Id.* at 347–58.

75. *Id.* at 354.

76. *See generally* CHRISTIAN G. FRITZ, AMERICAN SOVEREIGNS: THE PEOPLE AND AMERICA'S CONSTITUTIONAL TRADITION BEFORE THE CIVIL WAR (2008) (tracing America's post-Revolution political and constitutional history and the struggle to adopt and implement a collective sovereignty by "the people").

constrained view. For them the sovereign spoke only in conformity with procedures it set forth in advance. That was the exclusive way in which the sovereign's voice would be recognized and heard.

The implications of this divide about when the sovereign had spoken were significant. For instance, one implication was whether the people of a past generation could bind a future one. If the people were, in fact, sovereign, their hands could not be tied and their sovereignty limited by an earlier generation. During this period, many Americans believed that a constitution's expression of fundamental rights and requirements for revisions could not dictate those terms to future generations. The unborn sovereign people of a later period were at liberty—just as the revolutionary generation had been—to express their sovereign will. Thus, each generation of American sovereigns would govern in its own way.[77]

The necessity of following the "rules laid down" ultimately has won out at the state constitutional level, but there were a number of examples of extralegal successes.[78] At the federal level as well, obedience to the rules laid down in Article V has been assumed; that point of view is rejected here by Professor Levinson.

Levinson proposes an unlimited federal constitutional convention, with the delegates chosen at random and compensated adequately, with their proposed revisions being submitted to the people at a national referendum.[79] A similar, although not extralegal (because it was a proposed two-step process, with authorization first provided through an amendment to the state constitution), approach was recently explored in California, but had to be abandoned when fundraising failed to support the necessary steps of amending the state constitution to implement the idea.[80]

Levinson points out that the Constitutional Convention's secrecy made it easier to reach compromise than it would be now, when instant news coverage would bring instant pressure and compromise has become less supported.[81] Also, compromise must often be accomplished in "real time,"[82] with the actors being able to assess the actual partisan impact of their concessions. For this reason, Levinson wisely suggests a Rawlsian "veil of

---

77. *Id.* at 293.

78. *Id.* at 285–88; *see also* LEVINSON, FRAMED, *supra* note 2, at 358 (pointing to Fritz's conclusion that carrying out state constitutional change "even if contrary to established constitutional procedures—is one of the hallmarks of American constitutionalism" (quoting Christian G. Fritz, *Recovering the Lost Worlds of America's Written Constitutions*, 68 ALB. L. REV. 261, 262 (2005))).

79. LEVINSON, FRAMED, *supra* note 2, at 391–92.

80. Evan Halper & Anthony York, *California Constitutional Convention Push Fizzles*, L.A. TIMES, Feb. 13, 2010, http://articles.latimes.com/2010/feb/13/local/la-me-constitutional-convention-2010feb13. *See generally* Symposium, *Rebooting California: Initiatives, Conventions & Government Reform*, 44 LOY. L.A. L. REV. 393 (2011).

81. LEVINSON, FRAMED, *supra* note 2, at 49–50.

82. *Id.* at 26, 34.

ignorance,"[83] or "original position," approach, where any federal constitutional changes are delayed to the point where there is no clear partisan advantage that can be discerned.[84]

Professor Levinson's proposal is radical because it is "extralegal" and beyond the "rules laid down" in Article V of the federal Constitution. It also raises the fears of those who oppose even a constitutional convention within the terms of Article V, not only because they fear what such a convention might propose, but also because they fear a "runaway" convention like that in 1787. These are widespread and deeply held concerns. For that reason, Professor Levinson might have delved even further into the state constitutional experience with the processes of amendment and revision. For example, a number of states have provided in their constitutions for an automatic, periodic vote on whether to call a state constitutional convention.[85] Article V could be amended to provide for this or some variant of it.

In fact, my colleague Alan Tarr and I have pointed out that there is an extremely wide range of state constitutional amendment and revision procedures that have been or could be used in the states to accomplish needed constitutional change.[86] New approaches had to be, and have been, developed at the state level to deal with the problem of state constitutional rigidity. A number of these approaches could be tailored to fit a perceived need for change in the federal Constitution without the fear of a runaway convention. For example, several states have utilized a "two-step process" to achieve needed amendment or revision in their state constitutions.[87] The first, more moderate step is to formally change the "rules laid down" by initially following the established process for a constitutional amendment that authorizes a new, even one-time, process for amendment or revision of the state constitution.[88] Why not consider this approach, now, at the federal

83. *Id.* at 33–34.

84. *Id.* at 26, 33–34.

85. *See generally* John Dinan, *The Political Dynamics of Mandatory State Constitutional Convention Referendums: Lessons from the 2000s Regarding Obstacles and Pathways to Their Passage*, 71 MONT. L. REV. 395 (2010) (discussing the fourteen states that provide for a mandatory convention referendum device in their constitutions and examining the constitutional referendums held in Iowa, Alaska, Missouri, New Hampshire, Rhode Island, Connecticut, Hawaii, and Illinois).

86. *See generally* 1 STATE CONSTITUTIONS FOR THE TWENTY-FIRST CENTURY: THE POLITICS OF STATE CONSTITUTIONAL REFORM, *supra* note 65 (analyzing the political obstacles to state constitutional reform through case studies of reform efforts in Alabama, California, Colorado, Florida, New York, and Virginia); G. Alan Tarr & Robert F. Williams, *Foreword: Getting from Here to There: Twenty-First Century Mechanisms and Opportunities in State Constitutional Reform*, 36 RUTGERS L.J. 1075 (2005) (encouraging state constitutional reformers to take advantage of the numerous available options for reforming their constitutions); Robert F. Williams, *Should the Oregon Constitution Be Revised, and If So, How Should It Be Accomplished?*, 87 OR. L. REV. 867 (2008) [hereinafter Williams, *Oregon*] (examining the Oregon Constitution and efforts to revise it and discussing state constitution revision generally).

87. WILLIAMS, *supra* note 14, at 380.

88. *Id.*

level? For example, in Michigan, a 1960 first-step amendment eased the requirements for calling a constitutional convention, leading to the successful 1961–1962 Constitutional Convention.[89] A similar first-step amendment was adopted in 1950 in Illinois, leading to that state's well-regarded 1970 Constitution.[90] Texas amended its constitution to authorize the Legislature to sit as a constitutional convention for one time only.[91] Although the convention's proposals failed to get the necessary votes to be proposed to the voting public,[92] this was an innovative mechanism. New York used, albeit unsuccessfully, a Temporary Commission that proposed (based on the federal military base closing commission) "a unique action-producing alternative to a state constitutional convention," where the Governor and legislature were urged to act on proposed constitutional amendments and statutes by a date certain.[93]

Again, one of the fears about federal constitutional amendment and revision concerns the legal ability to limit a federal constitutional convention. The experience in the states over the years, however, has indicated that limited state constitutional conventions have been successful in taking "hot button" topics off of the table, and those limits have been seen as legally enforceable.[94] An initial step at the federal level could be to propose an amendment to Article V that clearly provides for a legally enforceable limited constitutional convention, whether on a one-time basis (in response to a perceived crisis or to limit opposition), with its use limited to periodic intervals, or as a permanent amendment to Article V. This would have to be drafted with care, providing a mechanism for determining and enforcing such limitations, processes for choosing delegates,[95] etc. In the states, the objective of a limited convention has been achieved by submitting not only the question whether to have a constitutional convention, but also how such a convention should be limited, to the voters themselves. In this way, the limitations are seen as emanating from the people themselves when they vote to call a constitutional convention, therefore binding their delegates.[96] A similar mechanism could be included in such a limited constitutional convention amendment to Article V, thereby eliminating the possibility of a runaway convention. This two-step approach would solidify the legality of new federal amendment or revision procedures by actually changing the

89. Williams, *Oregon*, *supra* note 86, at 882.

90. *Id.* at 884–85. The Florida Legislature successfully proposed an entirely revised constitution in 1967. *Id.* at 891.

91. *Id.* at 888.

92. *Id.* at 888–89.

93. Tarr & Williams, *supra* note 86, at 1095; Williams, *Oregon*, *supra* note 86, at 894.

94. Tarr & Williams, *supra* note 86, at 1085–92.

95. *See generally* Richard Briffault, *Electing Delegates to a State Constitutional Convention: Some Legal and Policy Issues*, 36 RUTGERS L.J. 1125 (2005) (addressing constitutional requirements for delegate selection and considering alternative delegate selection methods).

96. Tarr & Williams, *supra* note 86, at 1087–88.

"rules laid down" prospectively, before the new procedures are utilized. Then, the second stage would not be extralegal and could be carried out in a much more moderate and less uncertain process than an unlimited federal constitutional convention. An appointed constitutional commission, described below as a matter of state constitutional practice, could make important preparatory recommendations and provide background research and training for the delegates to such a limited federal constitutional convention.

There are a variety of additional techniques that have been developed or suggested in the state constitutional context that might be tailored for use at the federal level. For example, amendments have been proposed with "sunset" provisions limiting their length of effectiveness, shifting the burden to those who want to continue them at their point of expiration from those who want to eliminate them.[97] There is, of course, already federal precedent for this in the clause prohibiting Congress from banning the international slave trade until a date certain.[98] Professor Levinson's suggestion of delaying the effective dates of changes, so that partisan advantage cannot be weighed,[99] is also very important. A variation on the sunset approach would be constitutional amendments that, after a period of time, may be changed by less onerous amendment procedures or even by statute, possibly by supermajority.[100]

During the last century, states have had much success with the use of constitutional *commissions*, which are appointed bodies of experts who prepare proposed changes to the state constitutions and submit them to state legislatures.[101] This commission mechanism, not included in state constitutional amendment and revision procedures, has been developed in the states as an alternative to (or sometimes in preparation for) constitutional conventions, because they cost much less, rely on expertise, and report back to the legislative branch, which can thereby maintain control of the submission of state constitutional amendments or revisions to the electorate.[102] Commissions have been criticized, on the other hand, because they do not rely on the involvement of elected delegates the way constitutional conventions do, and therefore have been described as undemocratic.[103] Despite these drawbacks, the commission mechanism that has been developed successfully in the states could certainly be adapted for

97. *Id.* at 1113–14.

98. U.S. CONST. art. I, § 9, cl. 1.

99. LEVINSON, FRAMED, *supra* note 2, at 26–27.

100. Tarr & Williams, *supra* note 86, at 1114–15.

101. *Id.* at 1094–100.

102. *See id.* at 1094–95 (characterizing commissions as providing expert opinions while preserving ultimate authority with the legislature).

103. *Id.* at 1099.

use at the federal level. This would be a moderate approach and would not require an amendment to Article V or operate in defiance of it.

Almost all state constitutional commissions have operated without any formal change to the "rules laid down" for state constitutional amendment or revision. This is an unnecessary step because the commissions' proposed amendments or revisions are submitted to the legislative branch for its consideration pursuant to the formal processes of state constitutional change that are already in place.[104] So, the use of an appointed commission, broadly representative but utilizing expertise, might be able to examine some of the "settled" provisions of the federal Constitution that Professor Levinson describes as dysfunctional and contributing to the current gridlock in our federal government. Compromise would be necessary here, as it is in all constitution making. Thinking of constitutional commissions somewhat differently, one could be utilized to advise Congress on how to propose an amendment or amendments to Article V that would authorize limited constitutional conventions, what the limits should be, and how to make such limits legally enforceable. A commission was used recently in New Jersey for this purpose.[105]

One of the keys to the success of state constitutional revision has been moderation.[106] State constitutional conventions and commissions that have attempted to do too much, or to accomplish radical change, have often ended in failure.[107] Therefore, any proposed method of amendment or revision of the federal Constitution should aim for moderation. Some improvement is better than none. Levinson recognizes that "the best works as an enemy of the good."[108] It may be that reasonable and moderate adjustments to some of the "settled" provisions of the federal Constitution, such as the Inauguration Clause (probably not two Senators, per state), would not be nearly as controversial as proposed changes to other parts of the Constitution of Conversation. A limited federal constitutional convention, or constitutional commission, might be structured to focus only on the Constitution of Settlement and not be permitted to consider the more controversial Constitution of Conversation.

---

104. The one exception to this is Florida, where the state constitution creates two appointed commissions that meet periodically and can submit their proposed revisions directly to the voters. Rebecca Mae Salokar, *Constitutional Revision in Florida: Planning, Politics, Policy, and Publicity*, *in* 1 STATE CONSTITUTIONS FOR THE TWENTY-FIRST CENTURY: THE POLITICS OF STATE CONSTITUTIONAL REFORM, *supra* note 65, at 19, 19; Robert F. Williams, *The Florida Constitution Revision Commission in Historic and National Context*, 50 FLA. L. REV. 215, 220 (1998); Robert F. Williams, *Foreword: Is Constitutional Revision Success Worth Its Popular Sovereignty Price?*, 52 FLA. L. REV. 249, 252 (2000); Williams, *Oregon*, *supra* note 86, at 891–93.

105. Tarr & Williams, *supra* note 86, at 1104–05; *see also* Williams, *Oregon*, *supra* note 86, at 884 (describing similar use of a commission in Illinois).

106. WILLIAMS, *supra* note 14, at 378.

107. Salokar, *supra* note 104, at 39–40; Williams, *Oregon*, *supra* note 86, at 892.

108. LEVINSON, FRAMED, *supra* note 2, at 391.

A careful evaluation of the defects in the Constitution of Settlement, described by Professor Levinson, must consider whether the individual defects could be remedied by a series of unrelated amendments, or rather the defects are so interrelated as to render the federal Constitution incoherent and in need of more extensive revision rather than mere amendment. Alan Tarr noted that:

> Of course, it is possible to introduce significant constitutional reform without calling a convention or adopting a new constitution—amendments proposed by constitutional commissions, by initiative, or by state legislatures may also produce constitutional reform. But in thinking about constitutional reform, it is important to distinguish it from the ordinary constitutional change that is so prevalent in the states. Any alteration of a state constitution, no matter how technical or minor, qualifies as constitutional change. In contrast, constitutional reform involves a more fundamental reconsideration of constitutional foundations. It introduces changes of considerable breadth and impact, changes that substantially affect the operation of state government or the public policy of the state. The replacement of one constitution by another obviously qualifies as constitutional reform. So too may major constitutional amendments or interconnected sets of amendments. However, most constitutional change in the states does not qualify.[109]

Many people, as Levinson acknowledges, "are basically terrified" of a federal constitutional convention.[110] This fear also now manifests itself at the state constitutional level, where political scientists Gerald Benjamin and Thomas Gais have observed what they call "conventionphobia."[111] Calls for state constitutional conventions are now routinely defeated by the voters. I have said:

---

109. G. Alan Tarr, *Introduction* to 1 STATE CONSTITUTIONS FOR THE TWENTY-FIRST CENTURY: THE POLITICS OF STATE CONSTITUTIONAL REFORM, *supra* note 65, at 1, 2 (footnote omitted). As Bruce Cain has noted:

> In theory, constitutional revision should be more comprehensive and qualitatively more significant than a constitutional amendment. But what if revision occurs increasingly through amendment: What is gained and what is lost? The most important advantage should lie in the ability of a Revision Commission to consider how all the pieces fit together. Where the amendment process is piecemeal and sequential, the revision process affords the opportunity to logically relate proposals to goals, and to make the entire package of proposal[s] coherent.

Bruce E. Cain, *Constitutional Revision in California: The Triumph of Amendment over Revision*, *in* 1 STATE CONSTITUTIONS FOR THE TWENTY-FIRST CENTURY: THE POLITICS OF STATE CONSTITUTIONAL REFORM, *supra* note 65, at 59, 64.

110. LEVINSON, FRAMED, *supra* note 2, at 392.

111. Gerald Benjamin & Thomas Gais, *Constitutional Conventionphobia*, 1 HOFSTRA L. & POL'Y SYMP. 53, 69–70 (1996); Thomas Gais & Gerald Benjamin, *Public Discontent and the Decline of Deliberation: A Dilemma in State Constitutional Reform*, 68 TEMPLE L. REV. 1291, 1303–05 (1995).

> The public seems to view a constitutional convention as political business as usual by the "government industry." Constitutional conventions seem to have lost their *legitimacy* in the public mind. At the time many states' original constitutions were drafted, the politicians and special interests were afraid of the *people* acting through constitutional conventions. Now, by contrast, the people are afraid of *politicians and special interests* acting through constitutional conventions.[112]

This is certainly an attitude that will provide additional resistance to Professor Levinson's proposal, but which might not reject more moderate approaches out of hand.

Sandy Levinson has made important, and often convincing, criticisms of provisions of our Constitution that are not often debated. His proposed remedy, however, is radical, and in many people's view, dangerous to our federal constitutional system. For readers who agree with some of his criticisms, but worry about an extralegal, unlimited federal constitutional convention (or even a legal convention under Article V), the lessons learned from state constitutional amendment processes may be much more practical, moderate, and comforting.

Those seriously seeking to resolve at least some of the difficulties we currently experience because of the "settled" provisions of the federal Constitution would be wise to pick and choose among the lessons from the states to develop realistic possibilities for moderate change at the federal constitutional level. After all, despite the fact that most people think our Constitution has served us very well, it seems clear now that it could certainly be improved upon. Possibly now is the time that Article V should be made ("framed") to serve us rather than us having to serve (be "framed" by) Article V.

---

112. WILLIAMS, *supra* note 14, at 388 (footnote omitted).

The Limits of Antitrust Scholarship

THE GLOBAL LIMITS OF COMPETITION LAW. Edited by Ioannis Lianos & D. Daniel Sokol. Stanford University Press, 2012. 288 pages. $50.00.

Reviewed by Barak Orbach*

About thirty years ago, Professor Frank Easterbrook published his seminal article, *The Limits of Antitrust*, in the *Texas Law Review.*[1] Easterbrook declared that "[t]he goal of antitrust is to perfect the operation of competitive markets,"[2] and concluded that "[a]ntitrust is an imperfect tool for the regulation of competition."[3] It is an imperfect tool, he explained, "because we rarely know the right amount of competition there should be, because neither judges nor juries are particularly good at handling complex economic arguments, and because many plaintiffs are interested in restraining rather than promoting competition."[4]

Since Easterbrook published his article, the intellectual resources invested in antitrust in the United States have been in decline (see Figure 1). Easterbrook wrote about institutional and conceptual limits of antitrust—the *internal limits* of antitrust. Others have addressed the *extrinsic limits* of antitrust—the relationships of antitrust with other areas of law, such as intellectual property and regulation.[5] The decrease in depth of antitrust writing introduced a new form of limits in antitrust: diminishing critique and intellectual development in the field. This is the *depth limit* of antitrust. Of course, one may argue that there is no need for antitrust enforcement or antitrust scholarship, or at least no need for much.[6] Such arguments, however, tend to reflect general objections to regulation that have their own

---

* Professor of Law, The University of Arizona College of Law.

1. Frank H. Easterbrook, *The Limits of Antitrust*, 63 TEXAS L. REV. 1 (1984).

2. *Id.* at 1.

3. *Id.* at 39.

4. *Id.*

5. *See, e.g.*, Dennis W. Carlton & Randal C. Picker, *Antitrust and Regulation*, *in* ECONOMIC REGULATION AND ITS REFORM: WHAT HAVE WE LEARNED? (Nancy L. Rose ed., forthcoming) (on file with the author); Louis Kaplow, *The Patent-Antitrust Intersection: A Reappraisal*, 97 HARV. L. REV. 1813, 1816–18 (1984) (proposing a new framework for balancing the inherent conflict between trade restrictions provided for by patent law and prohibited by antitrust law).

6. *See, e.g.*, Robert W. Crandall & Clifford Winston, *Does Antitrust Policy Improve Consumer Welfare? Assessing the Evidence*, 17 J. ECON. PERSP. 3, 4 (2003) ("We find little empirical evidence that past [antitrust] interventions have provided much direct benefit to consumers or significantly deterred anticompetitive behavior."); Thomas W. Hazlett, *Is Antitrust Anticompetitive?*, 9 HARV. J.L. & PUB. POL'Y 277, 336 (1986) ("Given the long history of antitrust law and its contempt for true market rivalry, perhaps the most effective proconsumer program would be to consider federal enforcement of the antitrust laws to be a per se restraint of trade.").

social costs.[7] They often rely on "fire of truth" theories that no knowledge or analysis can possibly challenge.[8] The oversimplicity of such "fire of truth" arguments has been burdening antitrust for too long.

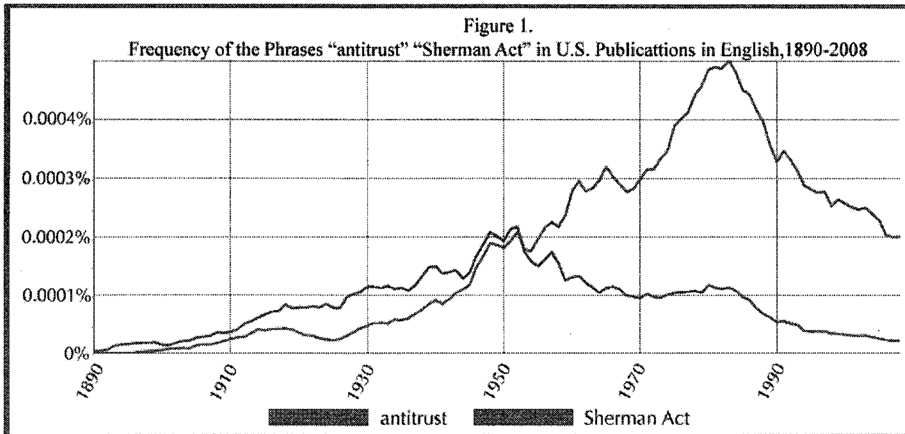


Figure 1.
Frequency of the Phrases "antitrust" "Sherman Act" in U.S. Publicattions in English,1890-2008

Source: Google Ngram.[9]

Ioannis Lianos and Daniel Sokol's *The Global Limits of Competition Law* (*GLCL*) is the first book in a series intending to develop antitrust scholarship.[10] *GLCL*'s "starting point [is] the intrinsic limits of competition law that Judge Frank Easterbrook highlighted."[11] The purpose of the book is to explore a broad set of limits to competition laws, "some intrinsic to antitrust, others extrinsic."[12] By definition, antitrust scholarship, including scholarship about the limits of antitrust, expands the depth limits of antitrust.

---

7. For a critique of those general objections to regulation, see generally BARAK ORBACH, REGULATION: WHY AND HOW THE STATE REGULATES (2012). *See also* Barak Orbach, *What Is Regulation?*, 30 YALE J. ON REG. ONLINE 1, 9–10 (2013) [hereinafter Orbach, *What Is Regulation?*] (arguing that the preponderance of irrational, ideological positions on regulation obstruct a more productive approach that acknowledges regulation's strengths and weaknesses); Barak Orbach, *How Antitrust Lost Its Goal*, 81 FORDHAM L. REV. 2253 (2013) [hereinafter Orbach, *How Antitrust Lost Its Goal*] (describing the reflection of the trend against regulation in antitrust); Thomas Philippon & Ariell Reshef, *Wages and Human Capital in the U.S. Finance Industry: 1909–2006*, 127 Q.J. ECON. 1551 (2012) (finding correlation between deregulation and substantial wage premiums in the financial industry).

8. *See* Edmund W. Kitch, *The Fire of Truth: A Remembrance of Law and Economics at Chicago 1932–1970*, 26 J.L. & ECON. 163, 183 (1983) (reproducing a discussion among leaders of the Chicago School, who compare their experiences with the movement to "religious conversion"); Herbert Hovenkamp, *Antitrust Policy After Chicago*, 84 MICH. L. REV. 213, 216 (1985) ("[N]othing—not even an intellectual structure as imposing as the Chicago School—lasts forever.").

9. For the methodology and its limitations, see Jean-Baptiste Michel et al., *Quantitative Analysis of Culture Using Millions of Digitized Books*, 331 SCIENCE 176 (2011).

10. THE GLOBAL LIMITS OF COMPETITION LAW (Ioannis Lianos & D. Daniel Sokol eds., 2012) [hereinafter GLCL].

11. Ioannis Lianos & D. Daniel Sokol, *Introduction*, *in* GLCL, *supra* note 10, at 1, 1.

12. *Id.*

GLCL is a book about competition laws, known in the United States as antitrust.[13] The book examines "competition laws" as a concept. It consists of fifteen essays written by prominent antitrust scholars. As a collection of essays, *GLCL* presents complementary perspectives of today's limits of antitrust.

The understanding of *GLCL* requires some appreciation of the simplicity and hospitality traditions in antitrust. The simplicity tradition refers to the tendency of individuals to view markets, businesses, and business practices either as competitive or as anticompetitive. That is, the individual simplifies facts and realities in a manner that allows her to consistently reach the same conclusion about competitiveness. In *The Limits of Antitrust*, Frank Easterbrook described the "inhospitality tradition" as "[t]he tradition [in which] judges view each business practice with suspicion, always wondering how firms are using it to harm consumers. If the defendant cannot convince the judge that its practices are an essential feature of competition, the judge forbids their use."[14]

GLCL's essays shed light about the complexities of reality effectively rejecting the simplicity and hospitality traditions. As Herbert Hovenkamp sums up in his essay: "extremes [in] antitrust policy should [be] avoid[ed]."[15]

Easterbrook's *The Limits of Antitrust* is a seminal article that, unlike ordinary academic works, has survived developments in time and is still relevant.[16] Being a Chicago School disciple, Easterbrook presented a taxonomy of antitrust errors, arguing that if we "let some socially undesirable practices escape, the cost is bearable,"[17] while the "costs of deterring beneficial conduct (a byproduct of any search for the undesirable examples) are high."[18] This taxonomy was insightful as an instrument against the "inhospitality tradition of antitrust."

Easterbrook's taxonomy, however, is simple and may backfire under the antithetical tradition where judges perceive the marketplace as the cure for all problems and government intervention as the source of all problems. The taxonomy was helpful for certain things, but it reflects the simplicity tradition in antitrust. To illustrate, consider judges who endorse the Chicago

13. For the significance of the label "antitrust," see Barak Orbach & Grace Campbell Rebling, *The Antitrust Curse of Bigness*, 85 S. CAL. L. REV. 605 (2012).

14. Easterbrook, *supra* note 1, at 14. Easterbrook attributed the phrase to Donald Turner. *Id.* at 4. While heading the Antitrust Division at the Department of Justice, Turner described the "inhospitality in the tradition of antitrust law." Donald F. Turner, *Some Reflections on Antitrust*, 1966 N.Y. ST. B.A. ANTITRUST L. SYMP. 1, 1–2. The phrases "a tradition of inhospitality" and "inhospitality in a tradition" are not equivalent, of course.

15. Herbert Hovenkamp, *Antitrust and the Close Look: Transaction Cost Economics in Competition Policy*, *in* GLCL, *supra* note 10, at 66, 80.

16. *See, e.g.*, Fred S. McChesney, *Easterbrook On Errors*, 6 J. COMPETITION L. & ECON. 11, 12–13 (2010) (highlighting the large number of citations to Easterbrook's article in cases and secondary sources as evidence of its success).

17. Easterbrook, *supra* note 1, at 14.

18. *Id.* at 15.

tradition at its extreme.[19] For example, in 1979, relying on a dubious study written in the Chicago tradition, the Supreme Court stated that the goal of antitrust is "consumer welfare."[20] The U.S. Supreme Court has never examined the standard, yet it keeps using it.[21]

While antitrust has many limits, studies show that human overconfidence tends to have fewer limits. People tend to be dismissive of and reject information that conflicts with their own beliefs.[22] Specifically, people who hold strong opinions are likely to evaluate facts and empirical evidence in a biased manner.[23] This well-documented tendency has profound effects on communication and political polarization.[24] It also explains the simplicity tradition in antitrust. An influential article, which argues that action is costly but inaction is bearable, may become immortal for those who want to believe that, notwithstanding evidence to the contrary, government action is costly. That is, a study of policy limits may establish limits, if the study forms a belief that problems solve themselves, while policies cause problems.

---

19. For a discussion of this culture in courts, see generally Thomas W. Merrill, *Capture Theory and the Courts: 1967–1983*, 72 CHI.-KENT L. REV. 1039 (1997).

20. Reiter v. Sonotone Corp., 442 U.S. 330, 343 (1979). *See generally* Barak Orbach, *The Antitrust Consumer Welfare Paradox*, 7 J. COMPETITION L. & ECON. 133 (2011) (chronicling the academic and judicial confusion with respect to the meaning of consumer welfare).

21. *See generally* Orbach, *How Antitrust Lost Its Goal*, *supra* note 7.

22. *See generally* Albert H. Hastorf & Hadley Cantril, *They Saw a Game: A Case Study*, 49 J. ABNORMAL & SOC. PSYCHOL. 129 (1954) (explaining divergent perceptions of violence in a football game corresponding to team identification with the proposition that from the total array of possible perceptions, viewers select those which they understand as significant); Hugo Mercier & Dan Sperber, *Why Do Humans Reason? Arguments for an Argumentative Theory*, 34 BEHAV. & BRAIN SCI. 57 (2011) (reevaluating the function of human reasoning to account for the fact that people typically ignore arguments that do not support their own views); Raymond S. Nickerson, *Confirmation Bias: A Ubiquitous Phenomenon in Many Guises*, 2 REV. GEN. PSYCHOL. 175 (1998) (marshaling evidence demonstrating the strength of confirmation bias).

23. *See generally* Charles G. Lord et al., *Biased Assimilation and Attitude Polarization: The Effects of Prior Theories on Subsequently Considered Evidence*, 37 J. PERSONALITY & SOC. PSYCHOL. 2098 (1979) (finding that people who hold strong opinions on complex social issues fail to give equal weight to confirming and disconfirming evidence).

24. *See generally* James Andreoni & Tymofiy Mylovanov, *Diverging Opinions*, 4 AM. ECON. J.: MICROECONOMICS 209 (2012) (noting the limited effectiveness of communication due to persistence of disagreement even in the face of sufficient information to reach agreement); Roland Bénabou & Jean Tirole, *Self-Confidence and Personal Motivation*, 117 Q.J. ECON. 871 (2002) (examining how self-serving beliefs factor into internal, intrapersonal communication); Avinash K. Dixit & Jörgen W. Weibull, *Political Polarization*, 104 PROC. NAT'L ACAD. SCI. 7351 (2007) (offering an explanation of how voter processing of information can result in polarization of the electorate); Barak Orbach & Frances R. Sjoberg, *Excessive Speech, Civility Norms, and the Clucking Theorem*, 44 CONN. L. REV. 1 (2011) (arguing that extraneous societal communication imposes costs that impede beneficial changes); Barak Orbach, *On Hubris, Civility, and Incivility*, 54 ARIZ. L. REV. 443 (2012) (questioning the effectiveness of certain social norms purportedly designed to encourage openness to differing viewpoints); Rajiv Sethi & Muhamet Yildiz, *Public Disagreement*, 4 AM. ECON. J.: MICROECONOMICS 57 (2012) (suggesting that communication in certain societies can serve to magnify existing biases and to create new biases where none previously existed).

Scholarship, antitrust scholarship included, has its own limits. This is why we should be concerned about the depth limits of antitrust. *GLCL* is an important book for its survey of the constraints of antitrust. The works in the book explain how antitrust can work in the real world when constraints exist. Under the simplistic tradition, the government is the primary source of constraints, and we should "let some socially undesirable practices escape [because] the cost is bearable."[25] In the real world, however, there are several sources of constraints: transaction costs, inadequate information, preferences for differentiation, bounded rationality, and fallibility. Antitrust as a form of government regulation is needed to address the way market participants utilize these constraints or to evaluate the effects of these constraints on business conduct. The imperfections of antitrust, its intrinsic limits, are a byproduct of these constraints.[26]

GLCL explores some of the complexities of managing antitrust in the real world. The book includes three essays about the institutional design of antitrust. Javier Tapia and Santiago Montt describe the relationships between courts and competition agencies.[27] Frédéric Jenny analyzes the significance of independency and advocacy for the work of competition agencies.[28] Ioannis Lianos describes the misunderstanding of antitrust remedies.[29]

GLCL includes four essays on the intrinsic limits of antitrust. George Priest explains the intellectual foundation of Easterbrook's *The Limits of Antitrust* in the Chicago School tradition.[30] Herbert Hovenkamp describes the relationships among several schools of thoughts in antitrust, focusing on transaction cost economics, which is related to the Chicago School of antitrust.[31] Hovenkamp's review stresses the need for depth in antitrust, describing several points of stagnation.[32] Jeffrey Harrison presents the challenges of competition policy in addressing the powerful buyers.[33] Anne-Lise Sibony reviews the challenges of the legal institution to properly utilize

25. Easterbrook, *supra* note 1, at 14.

26. Orbach, *How Antitrust Lost Its Goal*, *supra* note 7. *See generally* Orbach, *What Is Regulation?*, *supra* note 7 (observing that imperfect regulation is a necessary consequence of human limitation).

27. Javier Tapia & Santiago Montt, *Judicial Scrutiny and Competition Authorities: The Institutional Limits of Antitrust*, *in* GLCL, *supra* note 10, at 141.

28. Frédéric Jenny, *Competition Authorities: Independence and Advocacy*, *in* GLCL, *supra* note 10, at 158.

29. Ioannis Lianos, *Competition Law Remedies: In Search of a Theory*, *in* GLCL, *supra* note 10, at 177.

30. George L. Priest, *The Limits of Antitrust and the Chicago School Tradition*, *in* GLCL, *supra* note 10, at 15.

31. Hovenkamp, *supra* note 15, at 66.

32. *Id.*

33. Jeffrey L. Harrison, *Complications in the Antitrust Response to Monopsony*, *in* GLCL, *supra* note 10, at 54.

economics.[34] In some ways, her essay explains the spirit of the simplistic tradition.

Like most distinctions, the distinction between the intrinsic and extrinsic limits of antitrust is somewhat artificial. These limits are interrelated and define each other. The internal strengths and weaknesses of competition policies influence their reach, while other legal regimes and policies external to antitrust influence the effectiveness of competition policies.

GLCL's essays on the extrinsic limits of antitrust illustrate this point. Daniel Sokol reviews the effects of government policies on the strength of antitrust policy, focusing on regulatory regimes that weaken antitrust.[35] Damien Gerard's essay describes one form of such regulatory policies: the state action doctrine in Europe.[36] Daniel Crane describes choices of litigants between antitrust and intellectual property, considering the limits of each area of law.[37] And Paolisa Nebbia describes the relationships between competition law and consumer protection.[38] One of the central debates in antitrust is whether the law serves (or ought to serve) as a means for consumer protection, or whether consumer protection is outside the limits of the field.[39]

GLCL does not explore or even present all the limits of antitrust. The book offers fifteen perspectives of certain limits. It should be understood as a book that seeks to challenge the present depth limits of antitrust by offering important antitrust contributions. As such, *GLCL* is indeed an important antitrust book.

Contrary to Easterbrook's statement, the goal of antitrust is *not* "to perfect the operation of competitive markets."[40] Perfection has never been the goal of antitrust, and it should not be the goal of any policy.[41] Perfection is also not a trait of scholarship. But antitrust scholarship is too often a messanger of the tradition of hospitality in antitrust. *GLCL* is a successful collective effort to explain the depth needed in antitrust.

---

34. Anne-Lise Sibony, *Limits of Imports from Economics into Competition Law*, *in* GLCL, *supra* note 10, at 39.

35. D. Daniel Sokol, *Anticompetitive Government Regulation*, *in* GLCL, *supra* note 10, at 83.

36. Damien M.B. Gerard, *A Global Perspective on State Action*, *in* GLCL, *supra* note 10, at 99.

37. Daniel A. Crane, *IP's Advantages over Antitrust*, *in* GLCL, *supra* note 10, at 117.

38. Paolisa Nebbia, *Competition Law and Consumer Protection Against Unfair Commercial Practices: A More-than-Complementary Relationship?*, *in* GLCL, *supra* note 10, at 127.

39. For the controversy on consumer protection in antitrust, see John B. Kirkwood, *Protecting Consumers and Small Suppliers from Anticompetitive Conduct: The Goal With the Widest Support*, 81 FORDHAM L. REV. 2425 (2013); Robert H. Lande, *Consumer Choice as the Ultimate Goal of Antitrust*, 62 U. PITT. L. REV. 503 (2001); Robert H. Lande, *A Traditional and Textualist Analysis of the Antitrust Statutes: Efficiency, Wealth Transfers, and Consumer Choice*, 81 FORDHAM L. REV. 2349 (2013); and Joshua Wright & Douglas Ginsburg, *The Goals of Antitrust: Why Welfare Trumps Choice*, 81 FORDHAM L. REV. 2405 (2013).

40. Easterbrook, *supra* note 1, at 1.

41. *See generally* STEPHEN BREYER, BREAKING THE VICIOUS CIRCLE: TOWARD EFFECTIVE RISK REGULATION (1993).

Stirring the Melting Pot: A Recipe for Immigrant Acceptance

THE IMMIGRATION CRUCIBLE: TRANSFORMING RACE, NATION, AND THE LIMITS OF THE LAW. By Philip Kretsedemas. New York, New York: Columbia University Press, 2012. 232 pages. $28.00.

Reviewed by Michael Scaperlanda*

> The interstate highway "made distant what had been close, and close what had been distant."[1]

In *The Immigration Crucible*, Philip Kretsedemas hopes to break the "habit of developing arguments that are simply reactions to the 'other side'"[2] and desires to "map a political, cultural, and economic terrain that . . . provides some new insights into why so many noncitizens are in a difficult situation"[3] while drawing "attention to the limitations of the mainstream proimmigration position."[4] Toward this end, he seeks "an engagement across lines of difference that has the potential to transform the perspectives of all parties . . . involved in the encounter."[5] In this spirit, I offer my critique of this challenging book. I share Kretsedemas's sentiment—if my Review "is successful in getting people to think about U.S. immigration policy in a new way, . . . I will be more than pleased. Either way, I have put forward my best effort."[6]

Kretsedemas ultimately fails in his task because as much as he tries to escape—to transcend—liberal anthropology with its peculiar notions of the state and the state's relationship to immigrants and other denizens, he remains within liberalism's orbit, pulled in by its unseen gravitational forces. Instead of providing "a paradigm shift" that leads to "an entirely new understanding,"[7] he offers a particular view of the terrain from a worn and aging neoliberal spacecraft.

This Review will proceed in five stages. First, I will provide a brief summary of the book. Second, I will offer three critiques: (a) Kretsedemas's

---

* Gene and Elaine Edwards Family Chair in Law and Associate Dean for Academics, University of Oklahoma College of Law.

1. WENDELL BERRY, JAYBER CROW 281 (2000).
2. PHILIP KRETSEDEMAS, THE IMMIGRATION CRUCIBLE XII (2012).
3. *Id.*
4. *Id.*
5. *Id.* at 148. "Instead of retreating from the public debate—or simply tolerating the status quo—more effort should be made to open up and pluralize this field of debate." *Id.* at 147.
6. *Id.* at XV.
7. *Id.* at 151.

creation of a stereotyped "Other," which he marginalizes and stigmatizes, undermines his call to transformational dialogue; (b) while decrying both Executive discretion and state control over immigration, he fails to recognize and therefore leaves unresolved the question of how immigration policy ought to be adopted and implemented; and (c) although he desires a "stronger ethical foundation" for the pro-immigration discourse, he offers none.[8] Finally, I will offer a brief response to the central theme of his book, which is a desire "to address the problem of immigrant marginality."[9]

I. The Book: A Summary

> Even though Jim Crow is now a closed chapter in U.S. legal history, there is still a romantic attachment within the popular culture to images of national community that stem from this era.[10]

Kretsedemas believes that images of national community formed in the Jim Crow era drive immigration policy, fostering structures and institutions that create immigrant alienation. He hopes his book project will serve as a vehicle for "transforming the political culture to make it more inclusive of new immigrant populations."[11] To succeed, his project "requires a critical race analysis . . . that is not just oriented toward fixing racial inequalities" but also displays "a willingness to examine and reconstruct popular ideas about whiteness and the cultural difference of immigrants."[12]

Three key factors enter into Kretsedemas's equation: the marginal immigrant, the state with its broad discretionary powers, and the acquiescence of a broad spectrum of intellectuals—"liberal, conservative, and Marxist"—in the status quo.[13] The introductory chapter provides a broad overview of his case stating that both pro- and anti-immigrant forces have worked to expand the "extralegal (or marginally legal) discretionary powers" of the state, which sometimes favor "liberalization of migrant flows" and at other times serve "to control racial minority populations."[14]

---

8. *Id.* at 4.

9. *Id.*

10. *Id.* at 151. He uses Jim Crow as a rhetorical device recognizing that there are significant distinctions between Jim Crow and the current immigration landscape. *E.g.*, *id.* at 87 ("[I]t does not appear that Latino migrants are being treated like a separate racial caste, as was the case for black populations during Jim Crow."); *id.* at 83 ("The exclusion of the black person under Jim Crow was justified by their so-called racial difference. The exclusion of the undocumented migrant, on the other hand, is justified by the fact of their unauthorized entry.").

11. *Id.* at 150. He also hopes that this project is "connected to a broader project of regenerating a political culture that does a better job of including and safeguarding the rights of the entire U.S. population." *Id.*

12. *Id.* "Unfortunately, the postracial rhetoric of the Obama era has made this already difficult task even more daunting" because the "subtle message sent by Obama's campaign speeches is that systemic racial inequalities can be addressed, in a way that avoids divisive racial politics." *Id.* at 150–51.

13. *Id.* at 130.

14. *Id.* at 8.

Rejecting—or at least deemphasizing—formal "legal categories . . . defined by the state,"[15] Kretsedemas uses Chapter Two to reimagine many noncitizens, many nonimmigrants, the undocumented,[16] and even some immigrants,[17] as "de facto stateless."[18] With respect to nonimmigrants, he emphasizes the changing nature of arrivals to the United States: in the early twentieth century the immigration flow dwarfed the nonimmigrant flow but today the nonimmigrant flow is thirty times greater than the immigration flow.[19] The increase in the number of nonimmigrants, Kretsedemas suggests, makes it possible to expand the pool of noncitizen workers without the political cost of increasing immigration[20] with the corollary benefit that this population can be controlled through the government's "security-enforcement apparatus."[21] De facto statelessness befalls a segment of nonimmigrant visa holders because many of these "persons enter with dependents and with an intent to settle," with the nonimmigrant visa serving as "a probationary legal status."[22]

---

15. *Id.* at 13.

16. *Id.* at 44 ("'[I]llegality'—a more dire kind of statelessness."). "[I]llegality has become the organizing framework for recruiting and regulating the workforce." *Id.* at 31.

17. *Id.* at 43 (discussing a case of denaturalization and then deportation).

18. *E.g.*, *id.* at 19 ("The recent literature on statelessness has made a deliberate effort at complicating the relationship between statelessness as a formal, legal-juridical status and statelessness as a sociopolitical condition."). Toward what end? In this view, "victims of hurricane Katrina" and those citizens "subjected to warrantless searches" are de facto stateless because they "are subject to the law but not protected by the law." *Id.* at 20–21. This deliberate jettisoning of legal-juridical categories obfuscates the plight of and duties owed to two very different kinds of marginalized persons. In this era of the nation–state, the truly stateless person finds herself without the benefit of a nation–state that in some sense owes allegiance to her just as she, in a reciprocal manner, owes allegiance to her country. The claim of the Katrina victim or the citizen subjected to warrantless search is very different. It is that the nation–state that owes her its allegiance and protection has failed in its duty to administer justice.

19. *Id.* at 17.

20. *Id.* at 21.

21. *Id.* at 34–35. Kretsedemas focuses on the wrong ratios in the narrative because visitors—overwhelmingly tourists—make up the vast majority of nonimmigrant arrivals. His analysis would be tighter if he focused on the ratios between temporary workers and immigrants. In 2009, the last year he deals with, 1.13 million people were granted permanent residence. U.S. DEP'T OF HOMELAND SEC., 2010 YEARBOOK OF IMMIGRATION STATISTICS 5 (2011), *available at* http://www.dhs.gov/xlibrary/assets/statistics/yearbook/2010/ois_yb_2010.pdf. In that year, 1.7 million temporary workers (and their families) entered the United States out of 36.2 million nonimmigrant I-94 admissions. *Id.* at 65. After refocusing, the ratio is reduced from 30:1 to 1.5:1. In drawing his conclusions, Kretsedemas also fails to address the changes in travel between 1909 and 2009 and how that alone may account for a shift in the ratio.

22. KRETSEDEMAS, *supra* note 2, at 17–18. Kretsedemas fails to address two obvious sets of questions. First, is a nonimmigrant who can return to his country of citizenship really "stateless?" If he is, as Kretsedemas seems to assume, why? Second, most nonimmigrant visa holders must by law have nonimmigrant intent and have a foreign residence that they have "no intention of abandoning." *E.g.*, 8 U.S.C. § 1101(a)(15)(H)(ii)(a) (2006). What responsibility does the nonimmigrant visa holder bear for her situation? Should fraud in the visa application have consequences, including deportation?

Chapter Three reviews "the expansion of executive authority within the U.S. presidency" to "explain how this expanded authority has been used to craft immigration policy."[23] Kretsedemas notes that Executive discretion grew out of necessity to manage crises but that beginning with President Theodore Roosevelt the Executive began to exercise discretion "in an active, creative way" and not merely in reaction to external conditions or events.[24] Discretion unleashed can be used to expand or restrict rights.[25] Kretsedemas's thesis plays out, sometimes in unexpected ways, in the arena of immigration enforcement. For example, "Despite the fact that Democratic administrations are often viewed as being more proimmigrant than Republican administrations, [they] have actually been tougher on border control and immigration enforcement."[26] Far from being an accumulation of power in the Executive, expanded discretion—at least in American history—has a devolutionary component where private entities and local governments share in this discretion.[27]

In Chapter Four, Kretsedemas links the expansion of Executive discretionary authority with recent growth in local enforcement of immigration law, which "has produced a situation in which police officers, landlords, election booth workers, and health care workers have been given more freedom to participate in enforcement practices that used to be regarded as the exclusive preserve of the federal immigration system."[28] Drawing a

---

23. KRETSEDEMAS, *supra* note 2, at 49. The Obama administration's decision to provide "certain young people sometimes called 'Dreamers'" legal protection despite their undocumented status provides a perfect example of an Executive's exercise of discretion. *See* President Barack Obama, Remarks by the President on Immigration (June 15, 2012), *available at* http://www.whitehouse.gov/the-press-office/2012/06/15/remarks-president-immigration (highlighting that the announced policies on immigration enforcement came "[i]n the absence of any immigration action from Congress"); *see also* Memorandum from Janet Napolitano, Sec'y, U.S. Dep't of Homeland Sec., to David Aguilar, Acting Comm'r, U.S. Customs & Border Prot. et al. (June 15, 2012), *available at* http://www.dhs.gov/xlibrary/assets/s1-exercising-prosecutorial-discretion-individuals-who-came-to-us-as-children.pdf (setting forth guidelines for enforcing the Administration's policy of providing protection for those who came to the United States as children, thus "exercis[ing] discretion within the framework of the existing law").

24. KRETSEDEMAS, *supra* note 2, at 54 ("In these situations, the executive officer is no longer reacting to unforeseen, calamitous events. It begins to craft the definition of 'emergency conditions' in ways that complement its ideology, strategic interests, and specific policy objectives.").

25. *Id.* at 63. For example the immigrant-welcoming Bracero program and the immigration-restrictionist Operation Wetback "were both creatures of discretionary executive authority." *Id.* at 68 (arguing that both programs are more about the labor market and less about transnational migration).

26. *Id.* at 66.

27. *Id.* at 61.

28. *Id.* at 73 ("[L]ocal immigration laws have allowed the authority of the federal government to be parceled out to a variety of state and nonstate actors."). Kretsedemas's argument would have been more powerful if he had acknowledged the nuanced and complex historical relationship between federal, state, and nonstate actors in enforcing United States immigration laws. *See, e.g.*, Gerald L. Neuman, *The Lost Century of American Immigration Law (1776–1875)*, 93 COLUM. L.

parallel between local enforcement of immigration laws and Jim Crow laws, he argues that both provided a strategy for integrating the marginalized group into the economy while maintaining their status as inferior beings.[29] And he sees masked racism behind much of the local immigration law enforcement movement.[30]

For Kretsedemas, immigration laws, like Jim Crow laws, provide an example of the ways that local governance is "bound up with the symbolic politics of the majority group identity."[31] Chapter Five, therefore, leaves the practical world of laws and governance to explore the intellectual world where scholars from multiple ideological schools have converged to lay a foundation that fails to support adequately legal and cultural structures that welcome the immigrant.[32] Through a convergence of "liberal, conservative, and Marxist intellectuals," governing strategies that lead to immigrant alienation have "been quietly reinforced by both sides of the debate over immigration policy."[33] Strands of liberalism, for example, unwittingly collude with "racist ideologies" by promoting a pragmatic "cost-benefit rationality that becomes invested in perpetuating" inequality.[34]

The book's conclusion reveals the marginalized immigrant as a symbolic representative of all who live on the margins of American life.[35] As an important step toward eliminating immigrant (and others') alienation, Kretsedemas desires an "informed public dialogue" on immigration, "the meaning of democracy, national identity, and the continuing legacy of race in

---

REV. 1833, 1873–74 (1993) (noting that the states, enlisting the aid of private individuals, including ship captains, controlled immigration for the United States' first century).

29. KRETSEDEMAS, *supra* note 2, at 84–85. "Like Jim Crow, these laws may be viewed as an exercise in coercive integration that secures a place for most noncitizens in the U.S. socioeconomic order, but in a way that underlines the inferior legal and social status of unauthorized and low wage migrant populations." *Id.* at 85.

30. *Id.* at 88 ("[O]utrage against the illegal alien appears to offer a legitimate—that is, nonracist—way of redefining the scope and limits of an increasingly complex society.").

31. *Id.* at 103.

32. *Id.* at 130–32.

33. *Id.* at 130.

34. *Id.* at 131. "Liberal individualism" is weak on creating structural equality because it views interventions of this type "as illiberal impositions on the individual rights and freedoms of others." *Id.* at 107. Liberal cultural pluralism mutes the discussion of racial inequality because it seeks "to deemphasize the continuing significance of race." *Id.* at 109–10. Rather than focus on the structural causes of inequality, "liberal pluralist theory has tended to place most of the responsibility for overcoming social barriers to integration in the hands of immigrants themselves." *Id.* at 112. With a hands-off approach, these liberal tendencies reinforce "a field of racial-ethnic dividing lines" that cultural preservationists insist "must not be transgressed." *Id.* at 113. Marxists like Slavoj Zizek offer "a new kind of egalitarian unity that liquidates all differences" between persons, forcing the marginalized immigrant (and all others) "to conform to an already established set of ideals." *Id.* at 127.

35. *See id.* at 150 ("[I]mmigrant marginality is just one manifestation of a type of legal-political marginality that is shared by a growing number of legal residents and native-born persons. Simply put, the same policy developments that have weakened immigrant rights have also weakened the social, civil, and legal rights of all U.S. residents.").

the United States."[36] To harness the potential "to transform the perspectives of all parties . . . involved in the encounter," the dialogue must engage people "across lines of difference."[37] To advance this debate, he wants the public to "be exposed to arguments that demonstrate that conservative populists are not the only ones who are frustrated with the current immigration situation."[38] In the end, he advocates a paradigm shift that will encourage the public "to become actively involved in a discussion about 'who we are' as a national people."[39]

II. Critique One: Extinguishing the Possibility of Dialogue

> Just as the Jim Crow laws were designed to exclude those of African descent from American society, the laws excluding Asian immigrants upheld in *Chae Chan Ping* and *Fong Yue Ting* betray a belief in racial separation.[40]

Race has played a large and infamous role in United States immigration law and policy, gaining the imprimatur of the Supreme Court a generation after the end of the Civil War. In *The Chinese Exclusion Case*,[41] the Court upheld the exclusion of Chae Chan Ping, a twelve-year resident of the United States, stating:

> If . . . the government of the United States, through its legislative department, considers the presence of foreigners of a different race in this country, who will not assimilate with us, to be dangerous to its peace and security, their exclusion is not to be stayed because at the time there are no actual hostilities with the nation of which the foreigners are subjects.[42]

Although immigration laws are now facially race neutral, there are undoubtedly lingering present effects of these past practices.[43] And it would

36. *Id.* at 147–48.

37. *Id.* at 148.

38. *Id.* at 147. Some of these conservative populists engage in "incendiary rhetoric" emanating from an "outrage" that "rests on a foundation of white privilege," which justifies "incitement to violence . . . as an expression of patriotism." *Id.* at 145–46.

39. *Id.* at 147.

40. Gabriel J. Chin, *Segregation's Last Stronghold: Race Discrimination and the Constitutional Law of Immigration*, 46 UCLA L. REV. 1, 2 (1998).

41. Chae Chan Ping v. United States, 130 U.S. 581 (1889).

42. *Id.* at 606. California had asked Congress to take action barring Chinese immigration:

> In December, 1878, the convention which framed the present constitution of California, being in session, took this subject up, and memorialized Congress upon it, setting forth, in substance, that the presence of Chinese laborers had a baneful effect upon the material interests of the State, and upon public morals; that their immigration was in numbers approaching the character of an Oriental invasion, and was a menace to our civilization; that the discontent from this cause was not confined to any political party, or to any class or nationality, but was well-nigh universal . . . .

Id. at 595.

43. *See* Chin, *supra* note 40, at 38–50 (noting that, as a result of past immigration laws, Asian Americans are underrepresented in the political process, have fewer connections through which to

deny human experience to suggest that covert or subconscious racial classifications have been eliminated from the minds and hearts of all denizens of the United States.[44] But does this mean that all immigration restrictionists and all persons who favor stopping the flow of illegal migration are motivated by racial prejudice? Might they have some legitimate nonracial concerns?

Kretsedemas dreams of an America where structural and institutional barriers to immigrant inclusion are eliminated. To get there from here, he desires a robust and transformative dialogue involving individuals across the ideological and racial spectra.[45] Given our history, he sees "critical race" analysis as necessary to this dialogue.[46] Despite recognizing the difficulty of this task, Kretsedemas's method of using race to frame his argument ultimately disserves his stated objective, undermining rather than advancing the dialogue on immigrant marginality.

In any transformative dialogue, rigorous truth telling and truth recognition is vital.[47] But dialogue is difficult if not impossible when one of the dialogue partners—especially the one calling for the dialogue—assumes the worst of another partner with little or no evidence. Yet this is the path chosen by Kretsedemas. He, for example, argues that "[t]here is evidence that the contemporary anti-immigrant movement is still steeped in the racial ideologies of the Jim Crow era and that this has carried over into the movement to expand local enforcement laws."[48] His evidence? The KKK "tried to use local complaints about illegal immigration as a recruitment tool."[49] This would be like saying that the current labor movement is steeped in Marxist ideology because the Communist Party is using the "assault" on public unions as a recruiting tool. He further suggests that "outrage against the illegal alien appears to offer a legitimate—that is, nonracist—way of redefining the scope and limits of an increasingly complex society."[50] This conservative "populist outrage rests on a foundation of white privilege, which makes it possible for the incitement to violence to be viewed as an

---

assimilate into new occupations and geographic regions, and still face the "stigma of discrimination").

44. *See generally* Charles R. Lawrence III, *The Id, the Ego, and Equal Protection: Reckoning with Unconscious Racism*, 39 STAN. L. REV. 317 (1987) (describing the effects of subconscious racism on the legislative process, particularly as it relates to the distinction between discriminatory purpose and disproportionate impact standards of discrimination).

45. KRETSEDEMAS, *supra* note 2, at 148.

46. *Id.* at 150.

47. *See* DANIEL PHILPOTT, JUST AND UNJUST PEACE: AN ETHIC OF POLITICAL RECONCILIATION 183 (2012) ("Acknowledgment of past injustice . . . aims to achieve intrinsically valuable primary restorations, redressing wounds that are wider and deeper than is often recognized.").

48. KRETSEDEMAS, *supra* note 2, at 87.

49. *Id.*

50. *Id.* at 88.

expression of patriotism—and not as a threat to the public safety."[51] Ironically, he creates a stereotyped "Other" who is to be stigmatized and marginalized.[52] As a scholar with an admittedly pro-immigrant bent,[53] I would find it impossible to have a transformative dialogue with those I have stereotyped and associated with the KKK, Jim Crow,[54] white privilege, and violence unless they were superhumanly able to turn a blind eye and forgive me my prejudices.

III. Critique Two: Some Institution Must Govern

According to Kretsedemas, "structural-institutional conditions ... produce immigrant marginality."[55] Eliminating these conditions will require dialogue among an inclusive populace "about 'who we are' as a national people" so that our "popular concept of the nation can be interrogated and transformed."[56] To be effective though, the fruits of this dialogue will need to be implemented by some governing authority. And, it is here, at the stage

51. *Id.* at 146. In stark contrast to Kretsedemas's approach, Kevin Johnson provides a challenging but nonaccusatory call to dialogue on the racial implications of immigration law and policy. *See, e.g.*, Kevin R. Johnson, *Race Matters: Immigration Law and Policy Scholarship, Law in the Ivory Tower, and the Legal Indifference of the Race Critique*, 2000 U. ILL. L. REV. 525, 534 ("[O]ne cannot categorically state that the U.S. immigration laws are 'racist.' Nonetheless, a greater percentage of immigrants would be people of color without the many screening devices that disparately impact potential immigrants from developing nations.").

52. *See* KRETSEDEMAS, *supra* note 2, at 145–46 ("[S]ome conservative populists [assert] the moral authority to take the life of the Other.").

53. *See, e.g.*, Michael A. Scaperlanda, *More on the Ethics of Immigration*, FIRST THINGS, June/July 2008, at 16, 18, *available at* http://www.firstthings.com/article/2008/07/002-more-on-the-ethics-of-immigration-43 (arguing, as part of a dialogue with William W. Chip, that religious principles, particularly the idea of welcoming foreigners who are seeking a more prosperous life, should be considered as part of the modern immigration debate); Michael A. Scaperlanda, *The Ethics of Immigration: An Exchange*, FIRST THINGS, May 2008, at 40, 46, *available at* http://www.firstthings.com/article/2008/04/005-the-ethics-of-immigration-an-exchange-5 (same); Michael A. Scaperlanda, *Reflections on Immigration Reform, the Workplace and the Family*, 4 U. ST. THOMAS L.J. 508, 510 (2007) (proposing a resolution wherein the United States adopts a guest worker program, legalizes most of the undocumented immigrants currently in the country, and provides aid to foreign countries to reduce the economic disparities drawing people towards the United States); Michael Scaperlanda, *Who Is My Neighbor?: An Essay on Immigrants, Welfare Reform, and the Constitution*, 29 CONN. L. REV. 1587, 1589 (1997) (critiquing the immigrant stripping provisions of the Welfare Reform Act as antithetical to American ideals); Michael Scaperlanda, *Partial Membership: Aliens and the Constitutional Community*, 81 IOWA L. REV. 707, 707 (1996) ("[A] strong anti-immigrant undertow threatens to pull us from our constitutional commitment to equality and from our national mythology of open arms and golden doors."); Michael Scaperlanda, *Polishing the Tarnished Golden Door*, 1993 WIS. L. REV. 965, 966 (arguing that the constitutional protections nominally granted to aliens have largely proven illusory).

54. Throughout the book, Kretsedemas draws parallels between the sentiment that motivated Jim Crow and the sentiment that motivates today's immigration restrictionist arguments. Jim Crow is mentioned on 24 of the book's 151 pages (more than 15%). *See* KRETSEDEMAS, *supra* note 2, at 209 (listing pages where the term Jim Crow appears).

55. *Id.* at 134; *see also id.* at 101 ("[S]tructural inequalities and forms of institutional discrimination . . . may [a]ffect the life chances of [immigrant] children.").

56. *Id.* at 147.

of governing, that an unstated ambivalence appears in the argument, leaving an unresolved tension as to who ought to govern with respect to immigration law and policy.

Kretsedemas devotes two of his six chapters to executive and local governance of immigration law. In Chapter Three ("The Secret Life of the State"), he decries the expansion of Executive discretion in the American political landscape. In Chapter Four ("Concerned Citizens, Local Exclusions: Local Immigration Laws and the Legacy of Jim Crow"), he criticizes the devolution of authority over immigration to state and local governments. Does he want more congressional or judicial involvement?[57] He doesn't say! Do Executive discretion and local immigration lawmaking contribute to the structural and institutional defects that "produce" immigrant marginality? If so, what is the solution? He doesn't offer one!

By criticizing the "who" without offering a viable alternative, Kretsedemas misses out on the "what." Our history makes clear that both good and bad policy can be made at all levels of government and in all branches of government. And it makes equally clear that both fair and arbitrary implementation of that policy can be made at all levels of government. No one level or branch of government has a monopoly on the virtues or the vices. Therefore, exposing the vices of two groups of policy makers/implementers without an argument as to why those groups are particularly ill suited to address the immigration issues confronting them is singularly unhelpful in advancing the dialogue.

In the end, Kretsedemas seems to favor a strong centralized governing authority that would bind itself and others to "predefined rules and regulations" that diminish the ability to exercise discretion in the face of "unfolding contingencies."[58] He decries "deregulation and federal devolution" because they "create spaces of decision-making authority—which free the individual from binding legalities—that can be granted to a variety of private and public actors."[59] But he does not tell us who this centralized authority is, why he has confidence that it will—at least in his opinion—get the rules and regulations right, or how various actors are to respond to unforeseen contingencies in the absence of discretion.[60]

---

57. If he were to add chapters on congressional and judicial decision making, he would need to include their racially laced actions, including passage of and judicial acquiescence in the Chinese Exclusion laws. *See supra* notes 41–42 and accompanying text.

58. KRETSEDEMAS, *supra* note 2, at 61. *But see* Arizona v. United States, 132 S. Ct. 2492, 2499 (2012) ("Discretion in the enforcement of immigration law embraces immediate human concerns.").

59. *Id.* With devolution and deregulation, "[p]rivate corporations are given more freedom to relocate their manufacturing centers and recruit (and terminate) workers as needed." *Id.* Philosopher Bertrand de Jouvenel articulately sums up what I take to be Kretsedemas's ideal governing authority: "[I]t aims at being the organizer-in-chief of society, and at making its monopoly of this role ever more complete." BERTRAND DE JOUVENEL, ON POWER 236 (1949).

60. Kretsedemas views broad discretion as "a forbidden continuity that connects the power practices of modern governments to those of the feudal monarchy." KRETSEDEMAS, *supra* note 2,

IV. Critique Three: In Search of an Ethical Foundation

To address the problem of "immigrant marginality," Kretsedemas wants to move beyond an "emphasis on the utility of the immigrant worker" and put the pro-immigration case on "a stronger ethical foundation."[61] But, he fails to offer one! A stronger ethical foundation requires an answer to this critical question: Why should we as individuals or as a nation care about the immigrant, marginal or not?[62] He does not ask, much less answer, the question.

Although liberalism's framework has its own set of problems,[63] its framework is not available to Kretsedemas on his stated terms because he wants "to look beyond the neoliberal common sense that has dominated federal policy for the past several decades."[64] In addition to liberalism, he discusses the approach of two other rival intellectual traditions—conservatism and Marxism—to the problem of immigrant marginality, finding all three traditions deficient.[65] But he never proposes an alternative. Although his critical race analysis shines a particular light on immigrant marginality, it does not provide a framework for answering the foundational question of why we should care about the immigrant.

Kretsedemas suggests that American "ideals could be revitalized by an agnostic engagement with [a] wider world of ideas"[66] as it searches for a

---

at 53. De Jouvenel offers a different perspective: "The assents of people or assembly, so far from fettering for the rulers a freedom to act which they never had, made possible an extension of governmental authority." DE JOUVENEL, *supra* note 59, at 208–09 ("The power to legislate is not an attribute which was taken from Power by the establishment of an assembly or by popular consultation. It is an addition to Power, of so novel a kind that without an assembly or without popular consultation it would have been impossible.").

61. KRETSEDEMAS, *supra* note 2, at 4.

62. Kretsedemas does ask important foundational questions: "[E]ven if we all agree that we are in favor of a more just and democratic society, who gets to define what those terms mean? . . . And at what point does this willingness to be perpetually open to new voices . . . run the risk of collapsing into an incoherent relativism?" *Id.* at 148. Great questions—especially given his desire for a stronger ethical foundation—but he punts, saying that "thorough exploration of these questions is beyond the scope of this book." *Id.* Without this exploration, Kretsedemas leaves us without criteria for assessing the answers.

63. *See* Michael A. Scaperlanda, *Immigration Justice: Beyond Liberal Egalitarian and Communitarian Perspectives*, 57 REV. SOC. ECON. 523, 527 (1999) (noting that liberal theory is ill equipped to justify immigration restrictions).

64. KRETSEDEMAS, *supra* note 2, at 4.

65. *Id.* at 130. Kretsedemas does not explore whether religious traditions—Jewish, Christian, or Muslim, for instance—have resources that might aid in solving the problem of immigrant marginality. *Cf.* Soskin v. Reinertson, 353 F.3d 1242, 1265 n.1 (10th Cir. 2004) (Henry, J., dissenting) ("If an alien will reside with you in your land, you shall not persecute him. *The alien who resides with you shall be to you like a citizen of yours*, and you shall love him as yourself, because you were aliens in the land of Egypt. I am the YWWH, your God." (quoting *Leviticus* 19:33–34)).

66. *Id.* at 129–30. American identity "is better described as an open-ended project (propelled by an agnostic dialogue between social equals)." *Id.* at 135. His "agnostic engagement" like Bruce Ackerman's "neutral dialogue" plants him firmly within liberalism's orbit. *See generally* BRUCE A. ACKERMAN, SOCIAL JUSTICE IN THE LIBERAL STATE (1981).

"forward-looking vision" for immigration policy.[67] Unfortunately, lacking a foundation, an "agnostic engagement" cannot produce a "forward-looking vision." When I get into my car, I would be negligent to approach the gears with agnostic engagement. Foundations—including ethical foundations—require a clear sense of where one stands, what is behind, and what is in front. In short, ethical foundations require a criterion for judgment.

Standing on a firm ethical foundation with a clear vision forward does not necessarily entail arrogance or close mindedness. With proper humility, anyone standing on what she thinks is the strongest ethical foundations will be open to taking direction and even correction from those with whom she is engaged. Instead of an agnostic engagement, I suggest a thickly pluralistic engagement, where each participant in the dialogue brings herself, including her intellectual tradition with its ethical core, to the conversation. In this difficult dialogue, there will be, as Kretsedemas understands, multiple and contested visions of forward.

In our pluralistic society, the difficulty lies in the fact that we have rival intellectual traditions with different ethical foundations and "there is no neutral way of characterizing . . . the standards by which their claims are to be evaluated."[68] Alasdair MacIntyre suggests a difficult two-stage process for engagement under these conditions. First, each participant "characterizes the contentions of its rival in its own terms."[69] Second, after recognizing the inability of one's own tradition to solve intractable problems, the participant looks to another tradition to see if it has the resources to solve these problems in a more satisfactory fashion.[70] In short, I advocate an openness without agnosticism.

V. Response: Immigration and the Human Experience

Kretsedemas argues that "neoliberal priorities guiding U.S. immigration policy have been actively creating the structural-institutional conditions that *produce* immigrant marginality."[71] Structural-institutional conditions can exacerbate or mitigate immigrant marginality, but they do not *produce* it.

---

67. *Id.* at 136. Despite his calls for agnosticism, it is clear that Kretsedemas is not an immigration-policy agnostic, having rejected the ideas of "cultural conservatives" like Huntington. KRETSEDEMAS, *supra* note 2, at 129.

68. ALASDAIR MACINTYRE, WHOSE JUSTICE? WHICH RATIONALITY? 166 (1988).

69. *Id.* In doing this, each tradition must "mak[e] explicit the grounds for rejecting what is incompatible with its own central theses" while recognizing what it might learn from its rival "on marginal and subordinate questions." *Id.*

70. *Id.* at 166–67. As MacIntyre notes:

> In controversy between rival traditions the difficulty in passing from the first stage to the second is that it requires a rare gift of empathy as well as of intellectual insight for the protagonists of such a tradition to be able to understand the theses, arguments, and concepts of their rival in such a way that they are able to view themselves from such an alien standpoint . . . .

Id. at 167.

71. KRETSEDEMAS, *supra* note 2, at 134 (emphasis added).

Immigrant marginality is a reality, inherent to the human condition. Believing that institutional or structural changes can eliminate it is simply fanciful utopian thinking.

Even if the term "alien" is in some sense pejorative in labeling an immigrant,[72] in a very real sense "alien" is an appropriate term for describing the relationship between the immigrant and his new country.[73] Language, culture, history, and tradition often create a wide gulf between the migrant and the native. They do not yet belong to each other. Each may view the other with suspicion. The migrant may wonder whether he can trust the police and other local authorities, whether he will be discriminated against in the workplace, and ultimately whether he will be accepted. The native may wonder whether the migrant can be trusted to obey the laws, whether he will deplete precious community resources—including jobs—and whether he will disrupt and perhaps destroy the embedded language, culture, and traditions. To borrow a popular phrase from Kretsedemas, each looks at the other as "Other."

The difference between the migrant and the native is that the migrant is alone, or at least more alone, having left her community—her language, culture, history, and tradition—to begin life anew in another community. Almost by definition, the alien will reside on the margins of that new community. No change in institutional structure or condition can change this fact. In his autobiographical account, *Next Year in Cuba: A Cubano's Coming-of-Age in America*, Gustavo Pérez Firmat describes refugees as:

> amputees. . . . Just as people who lose limbs sometimes continue to ache or tingle in the missing calf or hand, the exile suffers the absence of the self he left behind. I feel the loss of that Cuban boy inside me. He's my phantom limb, at times dogging me like a guilty thought, at other times accompanying me like a guardian angel.[74]

Although Firmat draws a distinction between a person in "exile" and an "immigrant,"[75] his metaphor applies to economic immigrants as well, especially those who leave their country of origin because of an inability to support themselves or their families.

---

72. *See* STEPHEN LEGOMSKY & CRISTINA RODRÍGUEZ, IMMIGRATION AND REFUGEE LAW AND POLICY 1–2 (5th ed. 2009) ("[T]he word 'alien,' even when not adorned with the modifier 'illegal,' has always struck a disturbing chord. Many feel that the term connotes dehumanizing qualities of strangeness or inferiority (space aliens come readily to mind) and that its use builds walls, strips human beings of their essential dignity, and needlessly reinforces an 'outsider' status.").

73. The *Oxford English Dictionary* defines "alien" as "[b]elonging to another person, place, or family; not of one's own; from elsewhere, foreign." *See Alien*, OXFORD ENG. DICTIONARY, http://www.oed.com/view/Entry/4988?isAdvanced=false&result=1&rskey=7xc6tn& (last updated Sept. 2012).

74. GUSTAVO PÉREZ FIRMAT, NEXT YEAR IN CUBA: A CUBANO'S COMING-OF-AGE IN AMERICA 22 (1995).

75. *Id.* at 121–22.

The migrant suffers great loss and often bears the burden alone. But it would be a mistake to ignore the loss—real, perceived, or potential—suffered by the native population. In advocating open borders, Joseph Carens acknowledges that "immigration . . . might destroy old ways of life, highly valued by some, but it would make possible new ways of life, highly valued by others."[76] As Wendell Berry's character, Jayber Crow, recognizes, change can make "distant what had been close, and close what had been distant."[77] Speaking of himself, Crow said:

> If you have lived in Port William a little more than two years, you are still, by Port William standards, a stranger, liable, to have your name mispronounced. . . . [T]hough I was only twenty-two when I came to the town, many . . . would call me "Mr. Cray" to acknowledge that they did not know me well. . . . Once my customers took me to themselves, they called me Jaybird, and then Jayber. Thus I became, and have remained, a possession of Port William.[78]

Integration of migrants takes time. Although governmental institutions and structures do not cause migrant marginality and cannot eliminate it, they might serve to mitigate it. Congress could enact legislation that more effectively closes the backdoor of illegal migration by giving the Executive the discretionary authority to match the number of nonimmigrant laborers to the rise and fall in the demand for labor coupled with effective sanctions for employing unauthorized workers. Giving nonimmigrants job portability might reduce the incidences of employer exploitation. Any enforcement officer—whether federal or state—ought to be held accountable if they fail to treat noncitizens with dignity and respect. And states ought to remove barriers that raise the cost for individuals and communities to care for the immigrant as she adjusts to her new life in a new country.[79]

The government will assign the nonimmigrant an identifying number but will not learn the nonimmigrant's name, much less how to pronounce the name. The government will not take a personal interest in the nonimmigrant's family, culture, or history. Immigrant marginality recedes and immigrant integration begins at the backyard barbecue, the pub, and the church as families celebrate births, graduations, marriages, deaths, and holidays together. The migrant will not be at home in her adopted country until she is known and loved in her new community. And that takes time.

---

76. Joseph H. Carens, *Aliens and Citizens: The Case for Open Borders*, 49 REV. POL. 251, 271 (1987). Although Carens is addressing open immigration, his thesis holds with more limited migration.

77. *See* BERRY, *supra* note 1, at 281.

78. *Id.* at 11. Crow, *the stranger*, was born in the town of Goforth a couple of miles from Port William. *Id.* at 11–12.

79. *See, e.g.*, Michael Scaperlanda, *Religious Freedom in the Face of Harsh State and Local Immigration Laws*, 15 TULSA J. COMP. & INT'L L. 165, 166 (2008) (discussing an Oklahoma law, "which the Dallas Morning News referred to as 'the nation's toughest law on illegal immigrants, making it a felony to harbor, transport, shelter or conceal undocumented immigrants'").

* * *

Making Sense of the Marriage Debate

FROM THE CLOSET TO THE ALTAR: COURTS, BACKLASH, AND THE STRUGGLE FOR SAME-SEX MARRIAGE. By Michael J. Klarman. New York, New York: Oxford University Press, 2013. 276 pages. $27.95.

Reviewed by Jane S. Schacter*

When are courts justified in trumping a majority's will? Can countermajoritarian decisions produce meaningful social change? Which minority groups command special judicial protection from the depredations of the majority? These are classic questions of constitutional law and theory and have shaped the scholarly literature for two generations. The ongoing movement for marriage equality features all of these questions and has, since its inception in the early 1990s, spawned a national debate about the role of courts.

Michael Klarman's *From the Closet to the Altar: Courts, Backlash, and the Struggle for Same-Sex Marriage*[1] comprehensively traces the marriage debate with a special eye on the role of courts in propelling it. Among its many gifts is that of exquisite timing. The book was published only a few months before the Supreme Court announced in late 2012 that it would hear constitutional challenges to the federal Defense of Marriage Act and to California's Proposition 8.[2] If the marriage debate were a symphony whose first movement began with an unexpected Hawaii decision in 1993,[3] one might say that the Supreme Court's twin grants of certiorari in these cases foreshadowed a crescendo of sorts. Or maybe not. In fact, as the book reflects, the Supreme Court will enter this debate after some twenty years of groundbreaking litigation around the country,[4] noisy debates in state and federal legislative chambers,[5] and scores of hotly contested ballot measures.[6] What the Supreme Court decides to do will be significant and highly watched. But one of the points the book communicates so effectively is that

---

* William Nelson Cromwell Professor of Law, Stanford Law School.

1. MICHAEL J. KLARMAN, FROM THE CLOSET TO THE ALTAR: COURTS, BACKLASH, AND THE STRUGGLE FOR SAME-SEX MARRIAGE (2013).

2. United States v. Windsor, 699 F.3d 169 (2d Cir. 2012), *cert. granted*, 133 S. Ct. 786 (2012); Perry v. Brown, 671 F.3d 1052 (9th Cir. 2012), *cert. granted sub nom.* Hollingsworth v. Perry, 133 S. Ct. 786 (2012); Adam Liptak, *Justices to Hear Two Challenges on Gay Marriage*, N.Y. TIMES, Dec. 7, 2012, http://www.nytimes.com/2012/12/08/us/supreme-court-agrees-to-hear-two-cases-on-gay-marriage.html?ref=adamliptak&_r=0.

3. Baehr v. Lewin, 852 P.2d 44 (Haw. 1993); KLARMAN, *supra* note 1, at 48–60.

4. KLARMAN, *supra* note 1, at 48–89, 90–91.

5. *Id.* at 108–09, 143–55.

6. *Id.* at 65–66, 84, 112, 120–26, 144–46, 193.

the trajectory of public opinion strongly favors marriage equality, with young people vastly more supportive than older citizens.[7] The proverbial writing seems to be on the wall. Thus, the Court's first foray into this national debate may well tell us more about how the justices want their role in it to be remembered than it does about how the issue will be substantively settled in American society.

Klarman's book will, in any event, equip its readers to reflect thoughtfully about whatever the Court decides to do. The book sets the stage for the Court's action by offering a readable history, in chapter and verse, of the developments that have shaped the marriage equality movement. Klarman closely follows the legal trajectory from the 1993 Hawaii decision that made same-sex marriage appear imminent,[8] through the 2003 Massachusetts decision that actually legalized same-sex marriage for the first time in the United States,[9] through many other state court decisions, as well as the more recent federal cases. But this history goes far beyond any narrow charting of judicial decisions or doctrinal developments. Klarman also closely explores the fierce backlash around the country in the form of dozens of anti-same-sex-marriage measures on the state and federal level,[10] as well as the political context that shaped this backlash.[11] Throughout, he deftly explores the key dynamics in the social, political, and cultural environment that have both fueled and thwarted the claim in favor of same-sex marriage. For those who have pressed for marriage equality, this history has been full of soaring victories and bruising defeats, along with plenty of political mobilization and countermobilization. But through it all, there has been a steady growth of public support[12] for what was once seen as the marginal and socially implausible idea of state-recognized same-sex marriage.

The book sets out not only to tell, but to understand, this deeply mixed history and to consider what lessons we might draw from it. In this review, I first assess Klarman's rendering of the story and the conclusions he reaches. I then consider what the story he tells might suggest about some enduring questions in American constitutional law and scholarship.

* * *

About two-thirds of the book tells the story of the movement for marriage equality. The remaining third reflects on the causes and implications of the backlash against the equalizing efforts of courts. The

---

7. *Id.* at 199–200, 218.

8. *Id.* at 57–60.

9. *Id.* at 90–93.

10. *Id.* at 26–29, 68, 80–83, 95–98, 144–46, 175–76.

11. *Id.* at 31–36, 60–63.

12. *Id.* at 105–06, 135–37, 161, 166–69, 178–98; *see also* Frank Newport, *Religion Big Factor for Americans Against Same-Sex Marriage*, GALLUP (Dec. 5, 2012), http://www.gallup.com/poll/159089/religion-major-factor-americans-opposed-sex-marriage.aspx (discussing the results of a Gallup poll conducted in November 2012 showing that a majority of young Americans support same-sex marriage).

careful historical chapters are fascinating in their own right, and also set the stage for the later reflections on the role played by litigation.

The principal historical narrative stretches from the 1950s and 1960s to mid-2011, when the New York legislature enacted marriage-equality legislation.[13] The radically disparate periods that bookend the historical portion of the book speak volumes about one of Klarman's principal themes: the enormous and ongoing social change in the LGBT-rights[14] arena. In the time period addressed in Chapter 1, every state criminalized consensual sexual activity between partners of the same gender, the medical profession saw homosexuality as a disease, and even the ACLU saw no problems criminalizing behavior that it called "socially heretical or deviant."[15] Early attempts to protest or organize against a pervasively repressive status quo were fraught with danger.[16] The contrast with 2011 could hardly be starker. When New York enacted its marriage legislation with the enthusiastic support of Governor Andrew Cuomo, not only did it join five other states and the District of Columbia in offering full marriage equality, but an additional twelve states offered civil union or domestic partnership protection, twenty-one states had added sexual orientation to their antidiscrimination statutes, the Supreme Court had ruled bans on consensual sodomy unconstitutional, and it had become common for LGBT persons to come out and to be widely featured in popular culture, to name just a few developments of note.[17]

Klarman's first eight chapters touch on many of the key developments—large and small—that put such a great distance between the 1950s and 2011. None are bigger for his story of the marriage-equality movement than the 1993 Hawaii Supreme Court decision in *Baehr v. Lewin*.[18] *Baehr* was a case that LGBT-rights litigators had declined to bring, fearing that it was premature.[19] It was brought by a private attorney.[20] Much to the surprise of virtually all observers, the decision held that the Hawaii Constitution mandated the application of strict scrutiny to the state's traditional marriage laws—making it highly likely that the state law restricting marriage to a man and a woman would be found unconstitutional.[21] The specter of same-sex couples getting married in

13. *Id.* at 163–64.

14. I will use the inclusive term "LGBT" (lesbian, gay, bisexual, and transgender), though the issues relating to bisexuality and transgender do not figure much in the book.

15. *Id.* at 3–6.

16. *Id.* at 7.

17. *Id.* at 163–64; *see also Same-Sex Marriage, Civil Unions, and Domestic Partnerships*, N.Y. TIMES, Dec. 30, 2012, http://topics.nytimes.com/top/reference/timestopics/subjects/s/same_sex_marriage/index.html (providing an overview of recent legal, political, and public opinion changes with regard to same-sex marriage).

18. 852 P.2d 44 (Haw. 1993).

19. KLARMAN, *supra* note 1, at 55.

20. *Id.*

21. *Id.* at 56.

Hawaii and seeking recognition in other states was the big bang, as it were, of a debate that has been roiling ever since. As Klarman and others have noted, soon after *Baehr*, LGBT-rights litigators felt they had little choice but to hop on a train that they themselves had not thought ready to leave the station.[22]

While history will record 1993 as the key start date, Klarman's narrative reflects that it was, in fact, only a few years after the Stonewall uprising kicked off the modern gay-rights movement in 1969 that the first marriage-equality lawsuits were launched.[23] They were not taken terribly seriously, though it was, interestingly, one of these early suits—pressed by two gay students at the University of Minnesota who had unsuccessfully sought a marriage license—that led to the Supreme Court's summary affirmance in *Baker v. Nelson*.[24] That ruling has regularly shown up in briefs opposing marriage equality.[25] It seems unlikely the justices deciding the pending *Windsor*[26] and *Perry*[27] cases will be too concerned with a forty-year-old summary affirmance issued before any of the contemporary gay-rights cases in constitutional law, but it will surely be enlisted for support by those defending DOMA and Prop 8.

Much of the story Klarman tells will be familiar to students of the LGBT-rights movement. Indeed, because a lot of it is very recent and has been the subject of extensive media coverage, some will be familiar even to those who have not immersed themselves in the history of LGBT rights. Still, the history is quite well told and is synthesized in ways likely to engage a general audience. One might wish that Klarman had devoted more

---

22. *Id.* at 55; *see also* Scott L. Cummings & Douglas NeJaime, *Lawyering for Marriage Equality*, 57 UCLA L. REV. 1235, 1245 (2010) (explaining that LGBT lawyers "did not affirmatively pursue litigation to achieve the right to marry in Hawaii," but instead joined the *Baehr* effort after the fact to help shape legal strategy); Jane S. Schacter, *Courts and the Politics of Backlash: Marriage Equality Litigation, Then and Now*, 82 S. CAL. L. REV. 1153, 1165–66 (2009) [hereinafter Schacter, *Politics of Backlash*] (discussing how the major gay-rights litigators refused to take on the *Baehr* action, believing it to be premature).

23. *See* Schacter, *Politics of Backlash*, *supra* note 22, at 1165 (explaining that neither *Goodridge* nor *Baehr* was the first lawsuit to challenge different-sex-only marriage and pointing to early test cases in Kentucky, Minnesota, and Washington that the government won). *See generally* Jones v. Hallahan, 501 S.W.2d 588 (Ky. 1973) (affirming the lower court's refusal to issue a marriage license to a same-sex couple); Baker v. Nelson, 191 N.W.2d 185 (Minn. 1971) (holding that Minnesota law limited marriage to different-sex couples, which did not violate the United States Constitution), *appeal dismissed*, 409 U.S. 810 (1972); Singer v. Hara, 522 P.2d 1187 (Wash. Ct. App. 1974) (holding that the statutory prohibition on same-sex marriage did not violate state or federal constitutional rights).

24. 191 N.W.2d 185 (Minn. 1971).

25. *E.g.*, Defendant-Intervenors-Appellants' Opening Brief 32 n.7, Perry v. Brown, 671 F.3d 1052 (9th Cir. 2012) (No. 11-16577), 2011 WL 6117216, at *32 n.7, *cert. granted sub nom.* Hollingsworth v. Perry, 133 S. Ct. 786 (2012). For a recent decision accepting defense arguments that *Baker* should be accorded precedential effect, see Jackson v. Abercrombie, No. 11-00734, 2012 U.S. Dist. LEXIS 111376, at *45–51 (D. Haw. Aug. 8, 2012).

26. Windsor v. United States, 699 F.3d 169 (2d Cir. 2012), *cert. granted*, 133 S. Ct. 786 (2012).

27. *Hollingsworth*, 133 S. Ct. 786.

attention to certain stories that profoundly capture the kinds of personal struggles that happen when the way people live is not matched by an available legal infrastructure. For example, Klarman only briefly touches on the stories of two women—both, coincidentally, named Sharon—who became iconic within the LGBT community as their legal battles unfolded in the 1980s and '90s.[28] Sharon Kowalski was in a disabling auto accident in 1983, and her longtime partner was shut out of her life for years by Sharon's family of origin.[29] Ultimately, a court allowed Sharon to choose her own guardian and she chose her partner.[30] In 1993, in Virginia, Sharon Bottoms lost custody of her son Tyler to her own mother after Sharon came out as a lesbian and her mother alleged she was an unfit parent.[31] In this instance, the courts did not rule in her favor.[32] Much more could have been said about both, though, in fairness to Klarman, he does not purport to offer a detailed exposition of all important LGBT legal battles.

Klarman also weaves into his narrative some information that is less widely known. Three examples of such stories are illustrative, and each ties to a larger theme that characterizes the movement for marriage equality. One example is when Klarman tells of an early attempt by an unnamed male couple to secure a same-sex marriage license in Colorado in 1975.[33] After receiving advice from a local district attorney that the state marriage law did not clearly outlaw same-sex marriage, a county clerk granted licenses to this couple and a few others.[34] About a month later, the state Attorney General shut down the clerk by issuing an opinion that same-sex marriage was prohibited.[35] While this episode of on-the-ground activism garnered some publicity and seems to have exposed the couples to some hostile reactions, its relatively modest public profile contrasts starkly with the climate over the last two decades. In that more recent climate—one in which the internet has turbocharged the flow of information—all things same-sex marriage have been a magnet for media attention and have quickly become part of a polarized national political debate.[36]

A second example sheds some light on precisely the absence of a polarized national political debate in the 1970s. Klarman explores the role of

---

28. KLARMAN, *supra* note 1, at 50–51.

29. *Id.* at 50.

30. *Id.*

31. *Id.* at 51.

32. *Id.* at 50–51.

33. *Id.* at 20–21.

34. *Id.* at 21.

35. *Id.*

36. Indeed, Richard Adams, one of the men seeking to get a marriage license in Colorado recently died, and the New York Times published an obituary. Margalit Fox, *Richard Adams, Same-Sex Spouse Who Sued U.S., Dies at 65*, N.Y. TIMES, Dec. 24, 2012, https://www.nytimes.com/2012/12/25/us/richard-adams-who-sued-us-after-1975-gay-marriage-dies-at-65.html?_r=0. One wonders if his death would have drawn the same public notice in the absence of the high-profile contemporary marriage debate.

LGBT issues in helping to propel the rise of the religious right in the late 1970s and early 1980s. He notes that the first-ever advertisement on gay issues to run in a presidential campaign was offered up by a group called "Christians for Reagan" in the 1980 election.[37] Whereas President Carter had carried evangelical voters in 1976, Klarman notes, Reagan won them by a two-to-one margin in 1980.[38] Reagan's election has proven to be something of a prototype for what has now become utterly routine and familiar: the close intersection of LGBT and other social issues with electoral politics, and a steady and predictable partisan alignment. That linkage has played a big role in the marriage debate. By 1993, when *Baehr* was decided, the forces who would oppose same-sex marriage had long since mobilized against LGBT rights and made themselves a vocal part of national politics. That political organization helped to shape—and quickly nationalize—the backlash by positioning cultural conservatives to respond quickly to developments like the surprise ruling in Hawaii.[39]

A third example relates to another main theme in the book: the veritable chasm of an age divide in the general public on the same-sex marriage issue.[40] Klarman emphasizes the pronounced difference in support for same-sex marriage, as between older and younger segments of the electorate.[41] At one point, though, he probes an interesting variant of this phenomenon with poll data reflecting a pronounced age effect even *within* the LGBT community. In 2003, 18-year-old gay respondents were 31% more likely to support same-sex marriage than 65-year-old gay respondents.[42] Generational differences in this context suggest a change in both expectations and priorities in the LGBT community.

All in all, Klarman's telling of the story is well done in the way it weaves together the interacting legal, political, social, and cultural forces, and connects small details to larger developments. He makes clear, moreover, that while the same-sex marriage movement began with a Hawaii lawsuit, its dynamics have ranged far beyond the judicial domain and have proven quite complex.

Having said that, though, there are places where assertions are made that seem puzzling or unpersuasive. For example, in the course of introducing the debate over same-sex marriage, Klarman observes that the "[a]rguments for and against gay marriage have not changed much over the past two decades."[43] While he may be correct that some core concepts have

---

37. KLARMAN, *supra* note 1, at 33.

38. *Id.*

39. *See id.* at 55–57 (chronicling the circumstances leading up to and the backlash against the Hawaii ruling).

40. *Id.* at 199–200.

41. *Id.*

42. *Id.* at 51.

43. *Id.* at 52.

persisted through these years—say, the much debated link between procreation and marriage—in fact, opponents of marriage equality have markedly refined and moderated their arguments over the years. For example, whereas the congressional debates about DOMA in 1995 featured plenty of references to "perversion" and "lust," the campaign for Prop 8 in 2008 stayed studiously away from such incendiary rhetoric and focused, instead, on claims about what children would have to be taught in schools.[44] The perceived strategic advantage in toning down arguments is a point worth pausing to note because it captures the dynamic nature of the debate, and corresponds to the rise in pro-gay public opinion that Klarman makes a central point of his analysis.

A more significant issue in the book, though, is with some problematic claims of cause and effect. At one point, for example, Klarman suggests that some wondered if the rash of pro-marriage-equality developments in what he calls the "gay marriage spring" of 2009 would affect the California Supreme Court justices deliberating on a state constitutional challenge to Prop 8.[45] Having raised that possibility, he then concludes that the developments did not, in fact, influence them.[46] He reaches that conclusion, presumably, because the state supreme court went on to uphold Prop 8.[47] But it is only a very narrow concept of "influence" that would reason to that conclusion from the outcome of the case. It could well be that the justices *were* influenced, but in the other direction. That is, it is plausible that they were influenced to turn down the challenge because they could see the trajectory of public opinion and were less inclined to believe that judicial intervention would be necessary to overturn Prop 8. In any event, the question of how outside developments actually "influence" judges is a difficult one to study. Even assuming that judges themselves could correctly identify what influences their decisions, they are not likely to recite or reveal it.

Consider another example: Klarman's treatment of the marriages performed in 2004 in San Francisco, as directed by then-Mayor Gavin Newsom.[48] Newsom acted without legal authority, and the marriages he permitted were later declared invalid.[49] There is no question that, as ably described by Klarman, the Newsom weddings were quite controversial, and

---

44. On the changes since the DOMA debate, see Ariane De Vogue, *Congress Evolves on DOMA, Same-Sex Marriage*, ABC NEWS (Dec. 6, 2012), http://abcnews.go.com/Politics/congress-evolves-doma-sex-marriage/story?id=17888075#.UNYwnHdU3. On the character of the arguments stressed in the Prop 8 campaign, see Melissa Murray, *Marriage Rights and Parental Rights: Parents, the State, and Proposition 8*, 5 STAN. J. C.R. & C.L. 357, 366–67 (2009) (characterizing Prop 8 proponents' campaign as presenting their position less as "homophobia and discrimination" and more as "reasonable dissent").

45. KLARMAN, *supra* note 1, at 119, 134.

46. *Id.* at 134.

47. Strauss v. Horton, 207 P.3d 48, 122 (Cal. 2009).

48. KLARMAN, *supra* note 1, at 189–90.

49. *See id.* at 99 ("Newsom's action was largely symbolic, as experts were certain that the state would not recognize such licenses.").

many thought they were counterproductive. But it is puzzling for Klarman to assert, without any obvious way to prove it, that the rogue weddings in San Francisco (as well as similar weddings in Oregon) "more than the *Goodridge* decision that inspired them, ignited the powerful political backlash of 2004."[50] That statement begs any number of questions—how could one know which action caused more backlash, or make fine calibrations, given that they took place within months of each other? How could or would one test this proposition? The facts, cited by Klarman, that legislators like Barney Frank and Dianne Feinstein lamented what Newsom did, that opponents of marriage equality thought it helped them, and that Karl Rove seemed to feel that President Bush derived political benefit from it,[51] do not supply that proof.

At one point, Klarman speculates that these weddings backfired because observers had a "visceral[ly]" negative reaction to seeing the celebrating couples.[52] That is plausible and, for some who watched the coverage, likely to be true. But the suggestion is in some tension with the point stressed elsewhere that a key dynamic in boosting public support for gay rights has been the increased visibility of gay people. Klarman explicitly discusses the proliferation of gay television characters in the 1990s, as well as the effect of more gay people coming out to friends, family, and others.[53] It is, then, unclear how particular images of weddings would be in a totally different category. To the extent the backlash he associates with the west-coast weddings involves couples kissing, in particular, perhaps there is a distinguishing characteristic there.[54] But, of course, not all couples on line to get married kissed one another, not all who saw those images would have reacted the same way, and—in general—the proof remains elusive.

None of these individual points is overwhelmingly important, and the point is not to nitpick. The point is, instead, to notice that it is difficult to make confident assessments of causation when there are so many complex dynamics in play, and so many different individuals and subcommunities taking it all in. The scholarly impulse to reach causal conclusions is understandable, but the facts are often too messy to warrant sure conclusions.

Indeed, one of the most salutary aspects of the book is that, on the large issue of assessing backlash, Klarman demonstrates an admirable ability to capture this messiness. In fact, this is a significant way in which the book compares favorably with Klarman's own earlier work on same-sex marriage

50. *Id.* at 192; *cf. id.* at 189 (arguing that these weddings "early in 2004 generated at least as much backlash against gay marriage as had *Goodridge* itself").

51. *Id.* at 192.

52. *Id.* at 175.

53. *Id.* at 73.

54. *Id.* at 175–76.

and backlash. In a 2005 article, he compared *Brown*,[55] *Lawrence*,[56] and *Goodridge*[57] as he explored what causes antijudicial backlash.[58] There, he offered up discrete criteria for predicting backlash and made stronger, more categorical pronouncements about the negative effects of launching litigation before public opinion is sufficiently supportive.[59] His proffered criteria looked to whether a court ruling made an issue more salient, generated anger over "outsider interference" or "judicial activism," and pursued social change in a different order than what majoritarian institutions would do.[60] He argued in 2005 that *Goodridge* fit these criteria, as did the Supreme Court's decision in *Lawrence*, which became more controversial than might otherwise have been the case because its invalidation of sodomy laws was assessed as part of the ongoing controversy about marriage.[61] His conclusion in the article was summed up as: "By outpacing public opinion on issues of social reform, such rulings mobilize opponents, undercut moderates, and retard the cause they purport to advance."[62] In the course of this argument, he attributed to *Goodridge* several consequences that undermined LGBT interests—not only the enactment of many anti-same-sex-marriage measures, but possibly delivering the 2004 election to George Bush; providing the margin of difference to several Republican Senate candidates in close races and thus making it more difficult for LGBT-supportive Democrats to block the appointment of conservative federal judges; and giving cultural conservatives an enduring political issue to use to great effect.[63]

Although Klarman covers much of the same ground in the book and alludes to the same factors in explaining the backlash, there is a noticeable change of tone and conclusion from the earlier article. In the book, Klarman is much less committed to a negative assessment of litigating for same-sex marriage at a time when public opinion was not supportive. Indeed, having explored both the costs and benefits of litigation, he concludes in the book that, "[o]n balance, litigation has probably advanced the cause of gay marriage more than it has retarded it."[64] And, to a much greater degree than he did in his earlier work, Klarman recognizes that "[l]itigation put gay marriage on the table," and that, had early litigation not made marriage

55. Brown v. Bd. of Educ., 347 U.S. 483 (1954).

56. Lawrence v. Texas, 539 U.S. 558 (2003).

57. Goodridge v. Dep't of Pub. Health, 798 N.E.2d 941 (Mass. 2003).

58. Michael J. Klarman, Brown *and* Lawrence *(and* Goodridge*)*, 104 MICH. L. REV. 431 (2005).

59. *Id.* at 473.

60. *Id.*

61. *See id.* at 459–73 (discussing backlash against *Goodridge*); *see also id.* at 459 (connecting adverse reaction to *Lawrence* to the marriage debate).

62. *Id.* at 482.

63. *Id.* at 459–73.

64. KLARMAN, *supra* note 1, at 218. Here, he lists the costs as impeding progress on other gay-rights priorities, causing Senate candidates to lose reelection and state judges to lose their positions, and perhaps affecting the outcome of the 2004 election, which in turn led to a more conservative court. *Id.* at 218–19.

salient, it is "unlikely that more than 50 percent of Americans would support gay marriage in 2012."[65] To his credit, Klarman notes expressly in the book that some of his views have changed.[66] Klarman is not alone in having perspectives on the marriage controversy that have "evolved,"[67] and I think his candor about it is admirable. Indeed, the fact that the marriage debate has moved so quickly, and public support for marriage equality risen so rapidly, has created a challenge for scholars analyzing the debate in real time. At a minimum, the fast pace of change means that it is wise for anyone studying the issues to revisit and reassess, rather than clinging to earlier expressed opinions.

The more ambivalent assessment he offers in the book strikes me as on much more solid ground than the earlier writing. I have argued elsewhere that approaches to backlash that make categorical assumptions about the involvement of courts in contested policy issues can be both too generalized (in positing that court decisions will reliably generate backlash under a relatively general set of circumstances)[68] and too particularized (in treating backlash against courts as different in kind from other kinds of political backlash).[69] Moreover, the idea of backlash itself must be disaggregated. As we see in the context of the marriage debate, the widespread *policy* backlash reflected in DOMA and scores of anti-marriage-equality measures in the states was not accompanied by a similar *public opinion* backlash. To the contrary, favorable opinion has grown sharply over time.[70] As we think about the role of courts, then, it is crucial to remember that the Hawaii courts started the debate at a time when the issue of same-sex marriage was nowhere near the political or cultural radar.[71] Courts entered the marriage debate years before any majoritarian institution would have. It would be erroneous to say that courts therefore "caused" the skyrocketing public support for marriage equality over the last several years, but it is fair to say that courts crucially *ignited* a movement that otherwise looked to be years away. Decisions like *Goodridge* and those in the next few states that adopted same-sex marriage as a result of a court decision are, moreover, responsible for another effect: the *reality*, as opposed to the frightening *possibility*, of married same-sex couples. What has the effect of that reality been? It seems safe to say that it has not had the same effect on all observers, but it is reasonable to hypothesize that it has increased public support because those marriages have simply not had the kind of palpable and catastrophic social effects that some opponents had predicted.

---

65. *Id.* at 208.

66. *Id.* at 223.

67. *Cf. id.* at 196 (noting that Barack Obama had said several times that his views on same-sex marriage were "evolving").

68. Schacter, *Politics of Backlash*, *supra* note 22, at 1217.

69. *Id.* at 1218.

70. *Id.* at 1219–23.

71. *Id.* at 1220.

The central point here is that it is very difficult to draw clean causal arrows from point A to point B when exploring something as complex as the same-sex-marriage debate, which has involved multiple institutions (courts, legislatures, direct democracy, and electoral politics), multiple venues (local, state, and federal), and multiple domains (cultural, political, and social). The challenges of mapping actions to consequences in such circumstances lie at the heart of a book committed to better understanding the dynamics of antijudicial backlash, yet those challenges are formidably difficult. For the most part, Klarman skillfully acknowledges the complexity and the ambiguous picture of simultaneous progress and retrenchment for supporters of marriage equality. At points, as he enumerates the adverse developments for the LGBT community following *Baehr* and *Goodridge*, one is left with a vague sense that he would like to return to the more critical stance he took in 2005 toward early litigation. But his conclusions at the end of the book are more balanced and nuanced, and ultimately more persuasive, than was his earlier analysis.

Are there larger lessons here for the way scholars think about constitutional law and theory? My answer is: on some points, yes; on others, maybe. The way the marriage debate has unfolded can be read to suggest that we take a fresh look at some staples of constitutional law. But on some points, there are reasons to wonder if the marriage debate is too idiosyncratic to warrant much generalization.

First, as I have suggested above, the marriage debate illustrates the perils of reductionism in explaining cause and effect in the context of court decisions. Too often, debates about the consequences of controversial constitutional cases devolve into misleading questions about whether courts "can" or "cannot" produce meaningful social change. Take Gerald Rosenberg's well known book, *The Hollow Hope*,[72] in which he pitted the romantic myth of a "Dynamic Court" (one able and willing to pursue needed change even when elected officials won't) against his revisionist reality of the "Constrained Court" (one unable to do so).[73] Though controversial in some of its particulars, the book is a leading work on litigation as a means of social change. In 2008, Rosenberg published a second edition of *The Hollow Hope* that incorporated the same-sex marriage debate into his analysis.[74] The original edition of Rosenberg's book in 1991 emphasized *Brown v. Board of Education*[75] and *Roe v. Wade*,[76] arguing that observers misattribute to those decisions (and others) more impact than they actually had, and that changes

---

72. GERALD N. ROSENBERG, THE HOLLOW HOPE: CAN COURTS BRING ABOUT SOCIAL CHANGE? (1991).

73. *Id.* at 10–27.

74. GERALD N. ROSENBERG, THE HOLLOW HOPE: CAN COURTS BRING ABOUT SOCIAL CHANGE? (2d ed. 2008) [hereinafter ROSENBERG 2008].

75. Brown v. Bd. of Educ., 347 U.S. 483 (1954).

76. Roe v. Wade, 410 U.S. 113 (1973).

brought about by political institutions are both necessary to achieve lasting change, and less likely to backfire.[77] In his second edition, Rosenberg ardently defended this same view about same-sex marriage litigation:

> Ultimately, the use of litigation to win the right to same-sex marriage lends further support to the argument that courts are severely limited in their capacity to further the interests of the relatively disadvantaged . . . . By litigating when they did, proponents of same sex marriage moved too far and too fast ahead of the curve, leaping beyond what the American public could bear. The lesson here is a simple one: those who rely on the courts absent significant public and political support will fail to achieve meaningful social change, and may set their cause back.[78]

In support of this conclusion, Rosenberg relied on the fact that, as of the time he wrote, full marriage equality had not migrated beyond Massachusetts.[79] He also considered the case for several possible "indirect" benefits of litigation, but rejected most of them.[80] He concluded, for example, that there was more media coverage of same-sex marriage as a result of litigation, but that much of it was negative; that there had been no rise in contributions to gay rights groups that could be attributed to the marriage litigation; and that, on his reading, public opinion about same-sex marriage had not changed substantially between 1992–2006.[81]

Rosenberg's work has been influential and is impressive in many ways, but he seems far too committed to the purity of his institutional claim to acknowledge the complexity and ultimate ambiguity of the dynamics in play. I have argued elsewhere that Rosenberg's approach to courts fails to appreciate the murkiness of what might constitute social change.[82] I have argued, as well, that when applied in the area of LGBT rights, his approach fails to account for significant instances in which judicial action supporting equality has escaped backlash, and the actions of politically accountable institutions have provoked it. Examples, among others, are the successful litigation to secure adoptive rights for same-sex partners in nearly half the states in the country (producing no backlash),[83] and newly elected President Bill Clinton's attempt to open the military to gays in 1993 (producing strong backlash).[84]

The marriage debate strongly suggests the need for a less dogmatic, more pragmatic approach—one that recognizes the ways in which judicial

---

77. ROSENBERG, *supra* note 72, at 107–56, 228–46.

78. ROSENBERG 2008, *supra* note 74, at 419.

79. *Id.* at 353–54.

80. *Id.* at 355–419.

81. *Id.* at 360–61, 382–407.

82. Jane S. Schacter, *Sexual Orientation, Social Change, and the Courts*, 54 DRAKE L. REV. 861, 868 (2006) [hereinafter Schacter, *Social Change*].

83. *Id.* at 875–78.

84. Schacter, *Politics of Backlash*, *supra* note 22, at 1218–19.

action can generate both progress and backlash at the same time. Indeed, one takeaway from Klarman's book is that *how* one judges the wisdom of beginning a battle in court can depend critically on *when* the judging takes place. The aftermath of litigation can look very different based on when it is assessed. The time that elapsed between the *Baehr* decision in 1993 and *Goodridge* in 2003 would support a fairly bleak assessment. Most of the activity had been in the form of anti-equality, backlash measures on the federal and state level.[85] In the pursuit of marriage equality, only the Vermont Supreme Court's 1999 decision leading to civil unions marked any significant progress during this time and, of course, by today's standards, it looks fairly retrograde.[86] While *Goodridge* marked a stunning victory, it was quickly followed by another round of state ballot measures designed to head off *Goodridge*-clone rulings in other states.[87] With the marriage issue achieving new salience in the 2004 election, anxieties about backlash were perhaps at their peak. Indeed, it was in the wake of this election that Klarman, in his 2005 article, seemed to come down more on the Rosenberg side of the ledger.

Looked at from 2012, though, the picture is dramatically different. Indeed, Rosenberg himself said in his second edition that his analysis might be "overtaken by events."[88] And so it seems to have been. It is instructive to consider what happened between 2007 (the last year for which Rosenberg reported new developments)[89] and February 2012 (the end of the period addressed at all by Klarman).[90] These events alone might explain why Klarman is, justifiably, more restrained in his critique of litigation than is Rosenberg. Eight states plus the District of Columbia adopted full marriage equality, some by judicial action, others by legislative action. These were Connecticut (2008), California (2008), Iowa (2009), Vermont (2009), New Hampshire (2010), District of Columbia (2010), New York (2011), Maryland (2012), and Washington (2012).[91] Even though the California state supreme court's ruling was wiped out by Prop 8 later in 2008,[92] and the legislative actions by both Maryland and Washington were put to voter referenda later

85. *See* KLARMAN, *supra* note 1, at 57–68 (recounting federal and state legislation after *Baehr* limiting recognition of gay marriages and defining marriage as between a man and a woman).

86. Baker v. Vermont, 744 A.2d 864 (Vt. 1999).

87. KLARMAN, *supra* note 1, at 105.

88. ROSENBERG 2008, *supra* note 74, at 341.

89. *Id.* at 351.

90. Klarman's last historical chapter ends in 2011, but his conclusion addresses some developments in early 2012. KLARMAN, *supra* note 1, at 223.

91. Connecticut, California, and Iowa legalized same-sex marriage by judicial action. Vermont, New Hampshire, the District of Columbia, New York, Maryland, and Washington did so by legislative action. For details, see *States*, FREEDOM TO MARRY, http://www.freedomtomarry.org/states/ (last updated Nov. 8, 2012) and *Relationship Recognition for Same-Sex Couples in the U.S.*, NAT'L GAY & LESBIAN TASK FORCE, http://www.thetaskforce.org/downloads/reports/issue_maps/rel_recog_11_7_12_color.pdf (last updated Nov. 7, 2012).

92. *California*, FREEDOM TO MARRY, http://www.freedomtomarry.org/states/entry/c/california.

in 2012,[93] the eruption of marriage equality in several different parts of the country in this time period is quite striking. In addition to states moving to marriage equality, four states began to recognize same-sex marriages performed in other states;[94] four states adopted comprehensive civil unions;[95] and three states adopted limited relationship protections for same-sex couples.[96]

There was, to be sure, some further retrenchment between 2007 and February 2012. In addition to the enactment of Prop 8, both Arizona and Florida passed anti-same-sex-marriage amendments in 2008.[97] But even taking account of these, the national map on relationship recognition for same-sex couples came to look vastly different between the Rosenberg second edition and the Klarman book. Also conspicuously "overtaken" was Rosenberg's claim that public opinion on marriage had not changed much since 1992. That claim was questionable even as of 2007,[98] but by the time of Klarman's book, it simply fell outside any range of plausibility.

The trend continues, moreover, for the picture has changed substantially *even since* the end of the period covered by Klarman. Consider a few data points. Not unreasonably, Klarman rated it unlikely that President Obama would announce support for same-sex marriage before the election,[99] yet the President did exactly that in May 2012.[100] In addition, for the first time, supporters of marriage equality prevailed at the ballot box on Election Day 2012, as measures in four states that opposed marriage equality were all rejected by voters.[101] True, an anti-marriage amendment had carried in North Carolina by a large margin in June 2012,[102] but the Election Day four-state sweep reflected major change and might one day be seen as a tipping point.

93. Frank Bruni, *A Big Test for Gay Marriage*, N.Y. TIMES (Oct. 8, 2012, 8:36 PM), http://bruni.blogs.nytimes.com/2012/10/08/a-big-test-for-gay-marriage/.

94. These were Rhode Island, Maryland, New Mexico, and Illinois. *Relationship Recognition for Same-Sex Couples in the U.S.*, *supra* note 91.

95. These were Washington, Nevada, Illinois, and Delaware. *Id.*

96. These were Colorado, Maryland, and Wisconsin. *Id.*

97. *Arizona*, FREEDOM TO MARRY, http:/www.freedomtomarry.org/states/entry/c/arizona; *Florida*, FREEDOM TO MARRY, http://www.freedomtomarry.org/states/entry/c/florida.

98. Schacter, *Politics of Backlash*, *supra* note 22, at 1193–94.

99. KLARMAN, *supra* note 1, at 223.

100. Noam Cohen, *The Breakfast Meeting: Obama Stops 'Evolving' on Same-Sex Marriage*, MEDIA DECODER, N.Y. TIMES (May 10, 2012, 8:56 AM), http://mediadecoder.blogs.nytimes.com/2012/05/10/the-breakfast-meeting-obama-stops-evolving-on-same-sex-marriage/?ref=samesexmarriage.

101. Stuart Elliott, *After Success on Same-Sex Marriage, Gay Rights Group Uses Ad to Keep Pressure On*, MEDIA DECODER, N.Y. TIMES (Nov. 25, 2012, 5:59 PM), http://mediadecoder.blogs.nytimes.com/2012/11/25/after-success-on-same-sex-marriage-gay-rights-group-uses-ad-to-keep-pressure-on/.

102. Bruni, *supra* note 93.

And the upward trend in public support for same-sex marriage is, if anything, seemingly accelerating.[103]

What has followed litigation in *Baehr* and *Goodridge*, then, is both dramatic retrenchment and dramatic progress. At the very least, this ambiguous picture challenges any simple faceoff like the one Rosenberg posits between a romantic and a revisionist notion of courts. It suggests that it was neither a brilliant tactic nor a grave mistake that the campaign for marriage equality began with litigation. In showing that judicial decisions can both further and undermine social change, and can do these two things simultaneously, one point comes across clearly: courts rarely act in a vacuum. What courts do is necessarily mediated and communicated through politics, social movements, media, popular culture, and any number of other forces. How those forces interact, and the trajectory that interaction creates for the social change sought, is likely to be complex and deeply contextual, and to defy easy mapping. It also cannot necessarily be predicted in advance by those who see litigation as all virtue or all vice. Finally, the trajectory will not be the same for all. Consider the regional and cultural differences that have long characterized the marriage debate and help to explain why marriage equality has come to some states far sooner than others and, conversely, why some states have been more prone to backlash than others.[104]

A second point driven home by the marriage debate is that academic inquiries about the capacity of courts to generate social change have often been excessively focused on the United States Supreme Court. *Brown* and *Roe* are canonical examples, but they are not the only ones.[105] As the marriage debate now moves to the Supreme Court, perhaps the names *Perry* and *Windsor* may be added to that pantheon. But the virtue of Klarman's book (and other studies of same-sex marriage) being published before the Court issues any pronouncements on the issue is that it chronicles the two decades of judicial developments, overwhelmingly in state courts, that preceded the Court's entry. This was by the express design of LGBT-rights litigators, who elected to stay out of federal court for nearly twenty years. True, *Lawrence* was decided only a few months before *Goodridge*, and several of the justices' opinions gestured in some way toward same-sex

---

103. For one 2012 poll with dramatic results, see Neil King Jr., *WSJ/NBC Poll: Majority Now Backs Gay Marriage*, WALL ST. J. (Dec. 13, 2012, 12:38 PM), http://blogs.wsj.com/washwire/2012/12/13/wsjnbc-poll-majority-now-backs-gay-marriage/ (showing the NBC/Wall Street Journal poll reflecting a twenty-one point rise in support since 2004).

104. For example, Virginia, Texas, and Utah each passed three separate anti-marriage-equality measures at various points in time. *Texas*, FREEDOM TO MARRY, http://www.freedomtomarry.org/states/entry/c/texas; *Utah*, FREEDOM TO MARRY, http://www.freedomtomarry.org/states/entry/c/utah; *Virginia*, FREEDOM TO MARRY, http://www.freedomtomarry.org/states/entry/c/virginia.

105. *See* ROSENBERG 2008, *supra* note 74, at 292–93, 303–04 (discussing the Supreme Court's role in the "[r]eapportionment [r]evolution" and in extending procedural rights to criminal defendants).

marriage.[106] But the drivers of the debate were *Baehr* and *Goodridge* and, to a lesser extent, some of the state supreme court cases that followed those crucial firsts. What is missed by an excessive focus on the Supreme Court? One important point is that this focus offers only one very familiar picture of the well-known countermajoritarian difficulty.[107] In that picture, the justices are appointed with life tenure, and, most of the time, they are reviewing the actions of elected officials, whether legislative or executive. While Hawaii and Massachusetts likewise have appointed judges, California and Iowa, for example, have systems in which judges have to face the voters in some way.[108] In many states, moreover, the voters have recourse to direct democracy to enact policy and to counter judicial actions[109] and, of course, in the Prop 8 case, the marriage ban under review was passed by voters, not by a legislature. These institutional differences do not eliminate the countermajoritarian difficulty, but they do recast it in certain ways. The fact that some judges face voters might, on some views, mitigate the anxieties of countermajoritarianism.[110] The fact that voters have enacted laws on same-sex marriage directly might either exacerbate or mitigate that difficulty, depending on the normative posture one takes about direct democracy. In any event, these institutional factors merit notice and study.

The role of state courts in the marriage debate does, however, reflect one respect in which the marriage cases might be somewhat idiosyncratic. Unlike many other matters litigated in state courts, this one was nationalized very quickly.[111] The Supreme Court did not decide a marriage case between 1993–2012, but as Klarman effectively conveys, the issue nevertheless commanded the national stage and triggered a debate about judicial activism comparable to the one triggered by major Supreme Court cases. That owes, at least in part, to how nationalized the larger debate about LGBT rights

---

106. *See* Lawrence v. Texas, 539 U.S. 558, 585 (O'Connor, J., concurring in judgment) (arguing that Texas's sodomy laws are not supported by a legitimate state interest such as "preserving traditional marriage"); *id.* at 604 (Scalia, J., dissenting) (warning that the Court's decision will lead to the "judicial imposition of homosexual marriage").

107. *See* ALEXANDER M. BICKEL, THE LEAST DANGEROUS BRANCH: THE SUPREME COURT AT THE BAR OF POLITICS 16–17 (1962) (acknowledging that when the Supreme Court declares a legislative or executive act unconstitutional, it "thwarts the will of representatives" elected by the majority). For a historical perspective, see BARRY FRIEDMAN, THE WILL OF THE PEOPLE: HOW PUBLIC OPINION HAS INFLUENDCED THE SUPREME COURT AND SHAPED THE MEANING OF THE CONSTITUTION 258–62 (2009).

108. *Methods of Judicial Selection*, AM. JUDICATURE SOC'Y, http://www.judicialselection.com/judicial_selection/methods/selection_of_judges.cfm?state=.

109. *See State-by-State List of Initiative and Referendum Provisions*, INITIATIVE & REFERENDUM INST., http://www.iandrinstitute.org/statewide_i%26r.htm (listing twenty-seven states that allow voters to place measures on the ballot through initiatives and/or referenda).

110. *See* Jed Handelsman Shugerman, *Economic Crisis and the Rise of Judicial Elections and Judicial Review*, 123 HARV. L. REV. 1061, 1129 (2010) (offering a historical perspective on countermajoritarianism and elected judges).

111. *See* Schacter, *Politics of Backlash*, *supra* note 22, at 1183–93 (discussing national scope of debate following the *Baehr* ruling in 1993).

already was when the marriage controversy first appeared. It is also related to the idea that a couple married in one state would seek recognition in others.[112] With interstate travel being routine, a similar dynamic might have some role to play in other contested areas—say, abortion or public benefits—but the idea of a chain reaction in the realm of marriage has a particularized purchase of its own.[113]

Finally, the marriage debate also poses some very fundamental questions about standard doctrinal approaches to constitutional law. A staple of federal equal protection and due process review has been the issue of choosing the appropriate standard of review. This has played out, as well, in the marriage arena, with virtually all the state courts adopting some version of the relevant federal constitutional doctrine and attending to the standard of review.[114] But the state cases, taken as a whole, suggest that this inquiry does not count for all that much. They are all over the map on standard of review. For example, cases overturning state marriage statutes have been decided at every level of equal protection review—rational basis, intermediate, and strict scrutiny.[115] Similarly, the two circuits that have struck down DOMA on equal protection grounds employed different levels of review.[116] This variability suggests that all the attention paid to level of review, and all the thousands of pages written about it in briefs about marriage equality, may prove to have been mostly a sideshow.

Moreover, significant conceptual problems with one particular aspect of the scrutiny issue are thrown into high relief in the marriage litigation. One prong of the analysis traditionally used to decide whether to heighten scrutiny is an inquiry into whether the group is politically powerless, such that aggressive judicial review is necessary to protect the group's interest.[117] This issue was, in fact, the subject of extended expert testimony in the federal court trial on Prop 8's constitutionality.[118] As I have suggested elsewhere, issues about the political power or powerlessness of the LGBT community reveal enigmatic aspects of this part of the doctrine. Among the vexing questions made salient by the marriage debate are how to measure political power, how to account for the fact that groups may develop some measure of political power only because they are subjected to special discrimination in

112. *Id.* at 1185.

113. *Id.*

114. *See, e.g.*, Goodridge v. Dep't of Pub. Health, 798 N.E.2d 941, 961 (holding that because a same-sex marriage prohibition fails a rational basis test the court need not determine whether a strict scrutiny test is warranted).

115. Jane S. Schacter, *Ely at the Altar: Political Process Theory Through the Lens of the Marriage Debate*, 109 MICH. L. REV. 1363, 1381 (2011).

116. The First Circuit employed a version of rational basis review. Mass. U.S. Dep't of Health, 682 F.3d 1, 10–11 (1st Cir. 2012). The Second Circuit employed intermediate scrutiny. Windsor v. United States, 699 F.3d 169, 176 (2d Cir. 2012), *cert. granted* 133 S. Ct. 786 (Dec. 7, 2012).

117. Schacter, *supra* note 115, at 1372.

118. *Id.* at 1383.

the first instance, how to assess a history that includes both political victories and political defeats, and how to understand the extension of heightened scrutiny to race- and sex-based classifications notwithstanding the fact that racial minorities and women have won significant political and legislative battles.[119]

Indeed, the election of 2012 is likely to pose further questions about what it means for LGBT persons to have or to lack political power. Recall that supporters of the marriage equality side won on four of four ballot measures on Election Day (having lost in North Carolina in June of 2012). Recall, as well, that the President endorsed marriage equality before the election and paid no obvious political price for doing so. Indeed, the issue was not raised by his Republican opponent—a stunning contrast to the election of 2004, in which President Bush used his opposition to marriage equality as a prominent issue, and Republicans perceived great strategic advantage in getting the issue on the ballot in thirteen states that year.[120] Finally, in 2012, Representative Tammy Baldwin, an openly lesbian candidate in Wisconsin, was elected to the Senate.[121]

The events of the 2012 election are likely to be aggressively argued as evidence of the growing political power of the LGBT community. This will not and should not resolve the doctrinal question of political powerlessness at a time when thirty states still have laws banning same-sex couples from marrying in their constitutions and several other states have statutory bans,[122] most states do not include sexual orientation in their antidiscrimination laws,[123] and antigay hate crimes statistics are on the rise.[124] But the election results are likely to complicate the conversation. And that is consistent with what I take to be a central lesson from the marriage debate and from Klarman's book: There are no easy institutional answers or lessons here. Embrace the complexity.

---

119. *See id.* at 1390–96 (describing the problems with political process theory in the context of the marriage debate).

120. Jonathan Capehart, *Silence Is Golden on Gay Issues*, WASH. POST (Oct. 22, 2012), http://www.washingtonpost.com/blogs/post-partisan/post/silence-is-golden-on-gay-issues/2012/10/21/60427dea-1baa-11e2-ba31-3083ca97c314_blog.html.

121. Michael M. Grynbaum, *Fickle Wisconsin Sends a Trusty Progressive to the Senate*, N.Y. TIMES, Nov. 9, 2012, http://www.nytimes.com/2012/11/10/us/politics/fickle-wisconsin-sends-tammy-baldwin-to-senate.html?_r=0.

122. *Where State Laws Stand*, FREEDOM TO MARRY, http://www.freedomtomarry.org/pages/where-state-laws-stand.

123. *State Nondiscrimination Laws in the U.S.*, NAT'L GAY & LESBIAN TASK FORCE, http://www.thetaskforce.org/downloads/reports/issue_maps/non_discrimination_1_12_color.pdf.

124. Danielle Ryan, *Hate Crimes Down in 2011, But Anti-Gay Violence Is Up, FBI Says*, L.A. TIMES, Dec. 10, 2012, http://articles.latimes.com/2012/dec/10/nation/la-na-fbi-hate-crimes-20121211.

Notes

Jury Unanimity and the Problem with Specificity: Trying to Understand What Jurors Must Agree About by Examining the Problem of Prosecuting Child Molesters*

Introduction 1203
I. A Brief History and Overview of CSA Statutes 1208
II. Current Supreme Court Doctrine 1210
III. An Analysis of State Law 1212
A. An Overview of California 1212
1. *California and Jury Unanimity* 1212
2. *The History and Constitutionality of California's CSA Statute* 1213
B. An Overview of Texas 1214
1. *Texas and Jury Unanimity* 1214
2. *The History and Constitutionality of Texas's CSA Statute* 1215
IV. The Problem with a Lack of Specificity in Jury Agreement 1217
A. General Problems 1218
1. *Patchwork Verdicts* 1218
2. *A Strategic Advantage for Prosecutors* 1220
B. Constitutional Problems 1221
V. Where Do We Go from Here? 1222
A. How to Handle Current CSA Statutes 1222
B. How CSA Statutes Can Be Changed to Better Address the Problems with Prosecuting Child Molesters 1223
1. *A Return (in Part) to the Status Quo* 1223
2. *Replacing the Anti-Unanimity Provision* 1224
Conclusion 1226

Introduction

Debates about jury unanimity in criminal trials currently focus on whether the requirement should be applied to states through the Due Process Clause of the Fourteenth Amendment.[1] Although the answer to this has been

* Thanks to everyone who discussed this topic with me and provided feedback on earlier drafts. Thanks as well to the members of the *Texas Law Review* for their help preparing this Note for publication. I am especially grateful for the constant support from my family and friends.

1. Joshua Dressler and Alan Michaels provide a description of the current status of the requirement of jury unanimity in state criminal trials. *See* JOSHUA DRESSLER & ALAN C.

"no" ever since the Supreme Court decision in *Apodaca v. Oregon*,[2] academics continue to focus on the topic.[3] Despite this attention on incorporation, a looming background question remains mostly untouched by academics[4]—one that matters regardless of whether the requirement for juries is unanimity or a supermajority. This is the question of what exactly jurors have to agree about. To be more precise: What is the level of specificity that must be agreed on by jurors?

MICHAELS, UNDERSTANDING CRIMINAL PROCEDURE, VOLUME 1: INVESTIGATION 47 & n.32 (5th ed. 2010) (discussing *Apodaca v. Oregon*, 406 U.S. 404 (1972), in which "the Court split 4–4 on whether the Sixth Amendment requires that jury verdicts be unanimous" and noting that Justice Powell's concurring opinion, which "concluded that the Sixth Amendment requires unanimous verdicts, but that the Fourteenth Amendment does not impose this requirement on the states" produced "the anomalous result that, although eight of the justices would have applied the same constitutional rule regarding jury unanimity to the states and to the federal government, the unanimity requirement in fact applies to the latter but not to the former"). For a brief discussion of the history of the incorporation of the Bill of Rights, see *id.* at 41–48.

2. 406 U.S. 404, 410–12 (1972).

3. Jury unanimity advocates were given a potential constitutional opening after *McDonald v. City of Chicago*, 130 S. Ct. 3020 (2010), a Second Amendment incorporation case where a majority of the Court expressed support for full incorporation of the Bill of Rights. *See id.* at 3047 (plurality opinion) ("[T]his Court decades ago abandoned 'the notion that the Fourteenth Amendment applies to the States only a watered-down, subjective version of the individual guarantees of the Bill of Rights.'" (citation omitted)); *id.* at 3058–59 (Thomas, J., concurring in part and concurring in the judgment) (expressing agreement with the plurality regarding the scope of incorporation but preferring that incorporation of the Bill of Rights go through the Fourteenth Amendment's Privileges or Immunities Clause, rather than the Fourteenth Amendment's Due Process Clause). But advocates of jury unanimity have so far been rebuffed in their recent attempts to get a case before the Supreme Court. The most recent example of the Supreme Court denying a petition for writ of certiorari on this issue is *Herrera v. Oregon*, 131 S. Ct. 904 (2011). *See* Petition for a Writ of Certiorari at i, *Herrera*, 131 S. Ct. 904 (No. 10-344) (framing the question before the Court as "[w]hether the Sixth Amendment, as incorporated against the States by the Fourteenth, likewise requires a unanimous jury verdict to convict a person of a crime"). Seventeen professors were accounted for in four amicus briefs, and Eugene Volokh wrote the petition for certiorari for Herrera. *See* Brief of Oregon Criminal-Law and Criminal-Procedure Professors as *Amici Curiae* in Support of Petitioner at 1–2, *Herrera*, 131 S. Ct. 904 (No. 10-344) (listing the eight law professors who prepared the brief); Brief of *Amicus Curiae* Professor Jeffrey B. Abramson in Support of Petitioner at 1, *Herrera*, 131 S. Ct. 904 (No. 10-344) (introducing the amicus as a professor of law and government); Brief of *Amicus Curiae* Professor Kate Stith at 1, *Herrera*, 131 S. Ct. 904 (No. 10-344) (identifying the amicus as a professor at Yale Law School); Brief of Shari Seidman Diamond, Valerie Hans, Kenneth S. Klein, Stephan Landsman, Michael Saks, Rita Simon, and Neil Vidmar as Amici Curiae in Support of Petitioner at 1, *Herrera*, 131 S. Ct. 904 (No. 10-344) (identifying the seven amici as "law school professors"); Petition for a Writ of Certiorari, *Herrera*, 131 S. Ct. 904 (No. 10-344) (listing the author as Eugene Volokh on the cover page of the petition).

4. For a sampling of the few scholars who have addressed this question more directly, see generally Brian M. Morris, *Something Upon Which We Can All Agree: Requiring a Unanimous Jury Verdict in Criminal Cases*, 62 MONT. L. REV. 1 (2001); Jessica A. Roth, *Alternative Elements*, 59 UCLA L. REV. 170 (2011); Peter Westen & Eric Ow, *Reaching Agreement on When Jurors Must Agree*, 10 NEW CRIM. L. REV. 153 (2007); Eric S. Miller, Note, *Compound-Complex Criminal Statutes and the Constitution: Demanding Unanimity as to Predicate Acts*, 104 YALE L.J. 2277 (1995).

One way to approach questions about the level of specificity required for jury agreement[5] is by examining current state laws and their various jury agreement doctrines.[6] Continuous course of conduct offenses in particular can be helpful in examining these types of questions. Continuous course of conduct offenses can allow defendants to be found guilty without jury agreement on particular events, acts, or dates, as long as other requirements are met.[7] This Note will focus on one particular type of continuous course of conduct offenses—"continuous sexual abuse of a child" (CSA) statutes—that has become popular in the last two decades.

CSA statutes are meant to battle a difficulty in convicting child molesters: many of these cases revolve around alleged repeated sexual abuse with only generic evidence available since the child in question has difficulty providing event-specific evidence.[8] I use the term "generic evidence" to refer to evidence regarding abuse in general that is not specific to any one particular event in time. There is a tension present between trying to enforce offenses like this and traditional notions imbedded in our criminal law system.[9] CSA statutes have been written by states to increase prosecutors'

---

5. I use the term "jury agreement" to mean the jury decision-making rule in general—whether it is a supermajority requirement or a unanimity requirement. Given that forty-eight out of fifty states still have a unanimity requirement for criminal trials, jury agreement doctrines will often be referred to as unanimity doctrines. Petition for a Writ of Certiorari, *supra* note 3, at 12 (recognizing that forty-eight states "require unanimity for a criminal jury verdict").

6. Another approach in analyzing the level of specificity required is to focus on the nature and limits of *actus reus* requirements. This approach might actually be more theoretically consistent with the underlying question of specificity. This Note, though, will focus on state and Supreme Court doctrines that examine the specificity question under jury unanimity doctrines.

7. California courts have defined continuous course of conduct offenses as "aris[ing] in two contexts." People v. Thompson, 206 Cal. Rptr. 516, 518 (Cal. Ct. App. 1984). The first context "is when the acts are so closely connected that they form part of one and the same transaction, and thus one offense"—for example, repeated acts of rape occurring within one hour. *Id.* The second context "is when . . . the statute contemplates a continuous course of conduct of a series of acts over a period of time"—for example, "pimping, . . . pandering, . . . failure to provide for a minor child, contributing to the delinquency of a minor . . . and child abuse." *Id.* at 518–19.

8. *See, e.g.*, People v. Jones, 792 P.2d 643, 648 (Cal. 1990) (observing that criminal cases involving alleged child molestation "frequently involve difficult, even paradoxical, proof problems" since "[a] young victim . . . assertedly molested over a substantial period by a parent or other adult residing in his home, may have no practical way of recollecting, reconstructing, distinguishing or identifying by 'specific incidents or dates' all or even any such incidents").

9. *See, e.g.*, Dixon v. State, 201 S.W.3d 731, 737 (Tex. Crim. App. 2006) (Cochran, J., concurring). In her concurrence, Judge Cochran explained:

> [Texas] criminal procedures are intended to protect a defendant from being tried for being a "bad" person who acts in conformity with a criminal character propensity. Our rules are intended to give the defendant advance notice of precisely what criminal act he is alleged to have committed and when it occurred. Our state constitution requires the jurors to make a unanimous decision on the occurrence of one specific criminal act. The law focuses the advocates, judge, and jurors on whether the person charged is guilty of this one, very specific criminal act that he is charged with having committed.

Id.

ability to convict child molesters,[10] yet to do this CSA statutes actually focus on proxies for this sort of continuous abuse, rather than having to prove the individual bad acts themselves.[11] Thus, CSA statutes allow a defendant to be convicted without jury agreement on particular specific acts committed, as long as the requisite number of acts can be agreed on, along with other requirements depending on the state.[12]

The two states that will receive the most focus in this Note are California and Texas. In California, the CSA statute states:

> (a) Any person who either resides in the same home with the minor child or has recurring access to the child, who over a period of time, not less than three months in duration, engages in three or more acts of substantial sexual conduct with a child under the age of 14 years at the time of the commission of the offense . . . is guilty of the offense of continuous sexual abuse of a child and shall be punished by imprisonment in the state prison for a term of 6, 12, or 16 years.
>
> (b) To convict under this section the trier of fact, if a jury, need unanimously agree only that the requisite number of acts occurred not on which acts constitute the requisite number.[13]

Texas's CSA statute states:

> A person commits an offense if: . . . during a period that is 30 or more days in duration, the person commits two or more acts of sexual abuse, regardless of whether the acts of sexual abuse are committed against one or more victims . . . .
>
> . . . .
>
> . . . [M]embers of the jury are not required to agree unanimously on which specific acts of sexual abuse were committed by the defendant or the exact date when those acts were committed. The jury must agree unanimously that the defendant, during a period that is 30 or more days in duration, committed two or more acts of sexual abuse.[14]

Current Supreme Court doctrine allows wide latitude for how states wish to construct their criminal offenses and define what level of specificity must be agreed upon by jurors.[15] States have added their own standards for what the level of specificity needs to be, but continuous course of conduct

10. Jeffrey A. Sandquist, *Continuous Child Sexual Abuse*, 26 ARIZ. ST. L.J. 317, 317–18 (1994).

11. *See infra* Part I.

12. *See infra* Part I.

13. CAL. PENAL CODE § 288.5(a)–(b) (West 2008).

14. TEX. PENAL CODE ANN. § 21.02(b)(1), (d) (West Supp. 2012).

15. *See* JOSHUA DRESSLER & ALAN C. MICHAELS, UNDERSTANDING CRIMINAL PROCEDURE, VOLUME 2: ADJUDICATION 291–93 (4th ed. 2006) (explaining that jury unanimity is constitutionally required as to "elements" but not as to "means").

offenses are typically an exception to these state requirements.[16] Whether an offense is to be regarded as a continuous course of conduct offense is usually a job for the state legislature as well, based on how the offense is worded.[17]

Allowing little to no restrictions on what level of specificity is required by jurors has a significant impact on the jury decision-making rule in general. For a unanimity or supermajority requirement to have any force, there must also be set limits on what the level of specificity should be. Allowing state legislatures almost plenary power to define this level of specificity can turn a decision-making requirement, whether unanimity or supermajority, into a shell of a requirement.

The difficulty in prosecuting child molestation cases is acknowledged.[18] But this attempt by states to solve this problem is too broad. This criticism will be applied to all CSA statutes, but the harshest criticism will be for states that have a CSA statute without the requirement that the victim be the same or the requirement that the defendant have the same residence or continuous access.[19] A certain level of specificity in jury agreement has been built into states' jury unanimity doctrines, as will be shown by looking at Texas and California in particular, and to allow this level of specificity to be ignored by simply terming an offense as a continuous course of conduct offense threatens to lead to a new wave in criminal statutes that work around specificity requirements. I will argue that the continuous course of conduct exception can be a needed one, but to continue to allow this exception to be stretched to whatever limit that is wished by legislatures will risk allowing this exception to swallow the rule.

The discussion in this Note proceeds in the following way: Part I provides a general overview of CSA statutes and their history. Part II briefly discusses the current constitutional doctrinal framework. In Part III, I focus on California and Texas to show what their respective case law says about jury unanimity, specificity requirements for unanimity, and the continuous course of conduct exception. Part IV provides an analysis of CSA statutes and will attempt to answer the questions of (1) whether these statutes are staying true to the purpose behind specificity in jury agreement and (2) whether these statutes may bump up against constitutional problems in the future. Part V discusses some of the possible ways to address current

16. *See, e.g.*, People v. Gear, 23 Cal. Rptr. 2d 261, 265 (Cal. Ct. App. 1993) (noting that a continuous course of conduct exception exists in California regarding "the requirement of jury unanimity on which specific acts the defendant committed").

17. *See, e.g.*, *infra* note 67 and accompanying text.

18. *See, e.g.*, Pennsylvania v. Ritchie, 480 U.S. 39, 60 (1987) (acknowledging that child abuse cases are extremely difficult to prove because the child tends to be the only witness); Lloyd Leva Plaine, Comment, *Evidentiary Problems in Criminal Child Abuse Prosecutions*, 63 GEO. L.J. 257, 258 (1974) (explaining that "[m]ost child abuse cases do not result in criminal prosecution" because the offenses usually occur in the home with no other witnesses besides the victim and the perpetrator).

19. Among the CSA statutes mentioned in Part I, Texas retains the sole honor of excluding both these requirements. TEX. PENAL CODE ANN. § 21.02 (West Supp. 2012); *see infra* Part I.

CSA statutes and then provides possible solutions for how to better address the problem of prosecuting child molesters. Part VI provides concluding remarks.

I. A Brief History and Overview of CSA Statutes

California was the first state to adopt a CSA statute,[20] but other states have since followed suit.[21] All the state statutes examined in this Note include a clause—which I will refer to as the anti-unanimity clause—similar to California's CSA statute, which states: "To convict under this section the trier of fact, if a jury, need unanimously agree only that the requisite number of acts occurred not on which acts constitute the requisite number."[22] These statutes require proof that (1) a minimum number of sexual acts against a child occurred (usually two[23] or three[24]) and (2) these acts occurred over a certain time period.[25] Most states also require that the alleged victim be the

---

20. CAL. PENAL CODE § 288.5 (West 2008).

21. *E.g.*, ARIZ. REV. STAT. ANN. § 13-1417 (2010); DEL. CODE ANN. tit. 11, § 776 (Supp. 2010); N.Y. PENAL LAW § 130.75 (McKinney 2009); N.D. CENT. CODE § 12.1-20-03.1 (2012); TEX. PENAL CODE ANN. § 21.02 (West Supp. 2012); WIS. STAT. ANN. § 948.025 (West 2005 & Supp. 2011). It is important to note that many states have judicially constructed analogues to CSA statutes. *See, e.g.*, People v. Reynolds, 689 N.E.2d 335, 343 (Ill. App. Ct. 1997) (relying on the theory of a continuous course of conduct to conclude that only one transaction and not multiple transactions were at issue even though multiple counts were charged, and concluding that "a general verdict form is sufficient when the various counts state the same transaction"). A pattern can be seen with some of these states: a state's high court laments the clash between sexual abuse statutes written as specific, single offenses and how this clashes with jury unanimity, and then the court appeals to the state legislature to fix this problem. *See, e.g.*, State v. Arceo, 928 P.2d 843, 880 (Haw. 1996) (Nakayama, J., dissenting) (urging the state legislature to enact a CSA statute in order "to cure the problems inherent in the criminal prosecution of sexual abuse cases involving a minor of tender years who is unable to specifically recall dates, instances or circumstances surrounding the abuse"); Baker v. State, 948 N.E.2d 1169, 1174–75 (Ind. 2011) ("[T]he Indiana legislature has not adopted a statute criminalizing an ongoing pattern of sexual abuse when the victim is unable to reconstruct the specific circumstances of any one incident. We encourage the General Assembly to consider this issue."); Dixon v. State, 201 S.W.3d 731, 737 (Tex. Crim. App. 2006) (Cochran, J., concurring) ("Perhaps the Texas Legislature can address this conundrum and consider enacting a new penal statute that focuses upon a continuing course of conduct crime—a sexually abusive relationship that is marked by a pattern or course of conduct of various sexual acts.").

22. CAL. PENAL CODE § 288.5(b) (West 2008). Though not the focus of this Note, there are states that have passed statutes that are very similar to the CSA statutes previously mentioned, but without an explicit anti-unanimity clause. *See, e.g.*, State v. Fortier, 780 A.2d 1243, 1251 (N.H. 2001) (interpreting New Hampshire's felonious sexual assault statute, which allows for conviction when a person engages in a pattern of sexual assault against a child, to not require jurors to "agree on the particular acts, provided that they find the requisite number of acts occurred during the statutory time period").

23. *E.g.*, N.Y. PENAL LAW § 130.75(1)(a) (McKinney 2009) (requiring "two or more acts"); TEX. PENAL CODE ANN. § 21.02(b)(1) (West Supp. 2012) (same).

24. *E.g.*, ARIZ. REV. STAT. ANN. § 13-1417(a) (2010) (requiring "three or more" acts); CAL. PENAL CODE § 288.5(a) (West 2008) (same); DEL CODE ANN. tit. 11, § 776(a) (Supp. 2010) (same); N.D. CENT. CODE § 12.1-20-03.1(1) (2012) (same); WIS. STAT. ANN. § 948.025(1) (West 2005 & Supp. 2011) (same).

25. *See, e.g.*, ARIZ. REV. STAT. ANN. § 13-1417(a) (2010) (describing the period of time as needing to be three months or more); CAL. PENAL CODE § 288.5(a) (West 2008) (same); DEL.

same for all claimed acts.[26] Additionally, in some states a shared residency or routine access is also required.[27] Since this is a continuous course of conduct offense, these requirements are what juries are required to agree about, and not the specific acts of abuse.[28]

In all the states that have passed a CSA statute, either the state supreme court or the lower courts have upheld these statutes as constitutional—under the state constitution and the United States Constitution[29]—with one notable exception. The Supreme Court of Hawaii struck down Hawaii's CSA statute as unconstitutional based on the state's constitutional requirements of due process and separation of powers.[30] The end result in Hawaii was still the same: a state constitutional amendment was passed after the court's decision that granted the legislature explicit authority in defining "[w]hat behavior constitutes a continuing course of conduct" and "[w]hat constitutes the jury unanimity that is required for a conviction."[31] Thus, in the one state that found a CSA statute unconstitutional on state constitutional grounds, the state's constitution was subsequently amended.

CODE ANN. tit. 11, § 776(a) (Supp. 2010) (same); N.Y. PENAL LAW § 130.75(1) (McKinney 2009) (same); N.D. CENT. CODE § 12.1-20-03.1 (2012) (same); TEX. PENAL CODE ANN. § 21.02(b)(1) (West Supp. 2012) (describing the required time as "during a period that is 30 or more days in duration").

26. Some states are explicit that the victim must be the same for all acts. *See, e.g.*, WIS. STAT. ANN. § 948.025(1) (West 2005 & Supp. 2011) (providing that the acts must "involv[e] the same child"). Some states are more ambivalent. *See, e.g.*, ARIZ. REV. STAT. ANN. § 13-1417(a) (2010) (providing that the perpetrator has to engage in the acts with "a child"). Texas is the only state that explicitly mentions that multiple victims may qualify. *See* TEX. PENAL CODE ANN. § 21.02 (b)(1) (West Supp. 2012) ("A person commits an offense if . . . the person commits two or more acts of sexual abuse, *regardless of whether the acts of sexual abuse are committed against one or more victims* . . . ." (emphasis added)).

27. *See, e.g.*, CAL. PENAL CODE § 288.5 (West 2008) (requiring that the perpetrator "reside[] in the same home with the minor child or ha[ve] recurring access to the child"); DEL. CODE ANN. tit. 11, § 776(a) (Supp. 2010) (same).

28. Also, like other continuous course of conduct offenses, CSA statutes include protection for the defendant: the defendant cannot be convicted of both the CSA offense and any of the offenses that constituted the underlying acts. *See, e.g.*, People v. Johnson, 47 P.3d 1064, 1069 (Cal. 2002) (describing various strategic options that prosecutors may pursue but noting that if prosecutors decide to "charge discrete sexual offenses and continuous sexual abuse in the alternative . . . they may not obtain multiple convictions").

29. *See, e.g.*, People v. Calloway, 672 N.Y.S.2d 638, 642–43 (N.Y. Cnty. Ct. 1998) (holding that "allowing jurors to convict upon unanimous agreement that two sexual assaults were committed during a certain time period, without requiring agreement as to which two" was not "an infringement on defendant's right to a unanimous verdict" given the rationale behind the New York CSA statute); *see also* Charles M. Jones II, Comment, *Guilty of What? Unanimous Verdicts in Texas: Developing a Test to Distinguish Between Acts Constituting One Offense and Acts Constituting Separate Offenses*, 40 TEX. TECH L. REV. 391, 412 (2008) (noting that Arizona, California, New York, and Wisconsin, among others, have all passed statutes that criminalize the ongoing sexual abuse of children, and all have withstood judicial scrutiny).

30. State v. Rabago, 81 P.3d 1151, 1169 (Haw. 2003).

31. HAW. CONST. art. I, § 25.

II. Current Supreme Court Doctrine

The most recent Supreme Court cases involving jury unanimity are *Schad v. Arizona*[32] and *Richardson v. United States*.[33] In *Schad*, a plurality found that charging first-degree murder disjunctively as "premeditated murder or felony murder" without a jury unanimity instruction was not unconstitutional.[34] First, the Court stated that a jury need not be unanimous about the mere means of committing an offense.[35] Then, the Court made clear which branch of government was typically responsible for deciding what are the means of committing an offense:

> Decisions about what facts are material and what are immaterial, or, in terms of [*In re Winship*], what "fact[s] [are] necessary to constitute the crime," and therefore must be proved individually, and what facts are mere means, represent value choices more appropriately made in the first instance by a legislature than by a court.[36]

How far can a state legislature go in defining criminal conduct in their statutes? That became the main issue before the Court—whether Arizona's legislature violated the Due Process Clause of the Fourteenth Amendment.[37]

The plurality in *Schad* relied on two tests for their conclusion that "the jury's options in this case did not fall beyond the constitutional bounds of fundamental fairness and rationality"[38]: (1) historical or contemporary practice[39] and (2) moral equivalence.[40] Justice Scalia joined with the plurality in their conclusion but not in much of their reasoning.[41] To Justice Scalia, the deciding factor was history alone.[42]

---

32. 501 U.S. 624 (1991).

33. 526 U.S. 813 (1999). For an excellent in-depth analysis of these decisions that I cannot hope to replicate, see Westen & Ow, *supra* note 4, at 160–83.

34. *Schad*, 501 U.S. at 627.

35. *See id.* at 631 ("We have never suggested that in returning general verdicts in such cases the jurors should be required to agree upon a single means of commission, any more than the indictments were required to specify one alone.").

36. *Id.* at 638 (citation omitted).

37. *See id.* at 637 (characterizing the central issue in the case as whether Arizona's choice to allow a general verdict as to first-degree murder is constitutional and explaining that, in determining whether Arizona's choice meets the constitutional requirements for definitional and verdict specificity, the best measure is "a distillate of the concept of due process with its demands of fundamental fairness").

38. *Id.* at 645.

39. *See id.* at 643 ("[H]istory and current practice are significant indicators of what we as a people regard as fundamentally fair and rational ways of defining criminal offenses, which are nevertheless always open to critical examination.").

40. *See id.* at 643–44 (explaining that "[t]he proper critical question is not whether premeditated murder is necessarily the moral equivalent of felony murder in all possible instances of the latter" but that "[t]he question, rather, is whether felony murder may ever be treated as the equivalent of murder by deliberation, and in particular whether robbery murder as charged in this case may be treated as thus equivalent").

41. *Id.* at 649 (Scalia, J., concurring in part and concurring in the judgment).

42. *See id.* at 650. Justice Scalia stated,

In *Richardson*, a 5–4 majority settled a circuit split in deciding that the federal continuing criminal enterprise (CCE) statute's[43] language of a "series of violations" should be interpreted to require unanimity for the underlying offenses.[44] (While not explicitly defined, "series" in this statute had been previously considered by some circuits to be three or more violations.)[45] This decision was based on the Court's interpretation of the term "violations,"[46] as well as the breadth of the statute.[47] The Court interpreted the term "violations" in the statute to be defined as typically meaning more than simply an act or conduct—rather, the term carried with it a legal connotation that had to do with "an act or conduct that [was] contrary to [the] law."[48] And, for the Court, it was consistent with tradition to require agreement among jurors when the issue was whether conduct violated the law.[49]

The Court also viewed the breadth as an important consideration since the CCE statute applied to so many different kinds of behavior involving varying degrees of seriousness, which could "cover up wide disagreement among the jurors about just what the defendant did, or did not, do."[50] The Court viewed this as raising similar due process concerns as addressed in *Schad*, and chose the interpretation of the statute that avoided that risk.[51] The dissent was critical of this employment of constitutional avoidance. The dissent, written by Justice Kennedy and joined by Justice Ginsburg and Justice O'Connor, argued that interpreting the CCE as a continuous conduct

---

> It is precisely the historical practices that *define* what is "due." "Fundamental fairness" analysis may appropriately be applied to *departures* from traditional American conceptions of due process; but when judges test their individual notions of "fairness" against an American tradition that is deep and broad and continuing, it is not the tradition that is on trial, but the judges.

Id.

43. 21 U.S.C. § 848 (2006).

44. Richardson v. United States, 526 U.S. 813, 824 (1999).

45. *See* United States v. Young, 745 F.2d 733, 747 (2d Cir. 1984) ("[T]here is a consensus of authority that to establish a 'series' the government must prove at least three felony violations."). In *Richardson*, the Court "assume[d], but [did] not decide, that the necessary number is three." *Richardson*, 526 U.S. at 818.

46. *Richardson*, 526 U.S. at 818–19 ("To hold that each 'violation' here amounts to a separate element is consistent with a tradition of requiring juror unanimity where the issue is whether a defendant has engaged in conduct that violates the law.").

47. *Id.* at 819.

48. *Id.* at 818 (citing BLACK'S LAW DICTIONARY 1570 (6th ed. 1990)).

49. *Richardson*, 526 U.S. at 818–19.

50. *Id.* at 819. This breadth, combined with treating violations as means, could also increase the amount of violations a defendant was charged with, "significantly aggravat[ing] the risk (present at least to a small degree whenever multiple means are at issue) that jurors, unless required to focus upon specific factual detail, will fail to do so, simply concluding from testimony, say, of bad reputation, that where there is smoke there must be fire." *Id.*

51. *Id.* at 820.

offense was what Congress intended and such an interpretation did not raise due process concerns.[52]

Thus, the Supreme Court has not provided clear guidance for how specificity should be addressed. There appears to be significant differences of opinions among the Justices when it comes to these sorts of questions, which do not trace any of the more typical political alignments. Instead, these scattered opinions might be the result of the difficulty of the questions being asked and the lack of an apparent solution.

III. An Analysis of State Law

In terms of expansiveness, California and Texas might be the two best examples of the end of each pole, with California's CSA statute being the most limited and Texas's CSA statute being the most broad. California's CSA statute has a "shared residency or routine access" requirement, a same victim requirement, and the acts of abuse must happen over a period of time that is three months or longer;[53] Texas's CSA statute has no "shared residency or routine access" requirement, does not require the same victim, and the acts of abuse can happen over a shorter period of time (thirty days or more).[54] A look at the case law of these two states' respective CSA statutes is helpful to show how each state has constructed its own specificity requirements (through jury unanimity doctrines), how each state's CSA statute came about, and how each state's courts have protected its CSA statute against challenges of state and federal unconstitutionality.

A. An Overview of California

1. California and Jury Unanimity.—The right to a unanimous jury in criminal cases in California originates in the California Constitution.[55] As stated by the court in *People v. Sutherland*,[56] "where the evidence proved several distinct episodes, any of which could have supported the defendant's conviction of a single count of bribery, a unanimity instruction was required."[57] The court provided the following example to illuminate this point: "[A] unanimity instruction was required where the defendant was charged in one count with passing 35 bad checks, because each check represented a potentially separate and independent offense, creating the

52. *See id.* at 829 (Kennedy, J., dissenting) ("The continuing series element reflects Congress' intent to punish those who organize or direct ongoing narcotics-related activity. . . . The continuing series element . . . is directed at identifying drug enterprises of the requisite size and dangerousness, not at punishing drug offenders for discrete drug violations.").

53. CAL. PENAL CODE § 288.5(a) (West 2008).

54. TEX. PENAL CODE ANN. § 21.02(b)(1) (West Supp. 2012).

55. People v. Jones, 792 P.2d 643, 658 (Cal. 1990) (citing CAL. CONST. art. I, § 16).

56. 21 Cal. Rptr. 2d 752 (Cal. Ct. App. 1993).

57. *Id.* at 757.

possibility the jury might not have unanimously agreed that the defendant committed any single offense."[58]

In contrast, difficulties may arise like those raised by CSA statutes when a single offense is defined as being able to be committed in alternative ways.[59] California courts have established a two-step analysis for these types of situations: (1) "whether under established principles announced in California cases a unanimity instruction was required in this case" and then (2) "whether due process nonetheless required one."[60]

2. The History and Constitutionality of California's CSA Statute.—In the 1980s, California had a line of cases that held "purely generic testimony outlining repeated and continuous molestations without distinguishing between time, place or circumstance" was insufficient to sustain a conviction because, in part, "the jury cannot unanimously agree beyond a reasonable doubt that any such act occurred."[61] The leading decision in this line of cases was *People v. Van Hoek*,[62] which was decided in 1988. In 1989, as a response to these decisions, California was the first state to pass a CSA statute, California Penal Code § 288.5.[63]

The constitutionality of § 288.5 was addressed in *People v. Higgins*[64] and *People v. Gear*.[65] *Higgins* held that § 288.5 did not violate California's specificity requirements because it was a course of conduct offense: "So-called continuous-course-of-conduct crimes, generally committed against the same victim who sustains cumulative injury, require neither allegations nor unanimous findings of specific acts."[66]

So how does an offense become classified as a continuous course of conduct offense? According to the court in *Gear*, "The continuous-course-of-conduct exception 'arises . . . when, as here, the statute contemplates a continuous course of conduct of a series of acts over a period of time.'"[67]

Thus, the answer to California's first question of jury unanimity analysis is that a unanimity instruction is not required when the legislature makes clear that the offense is to be considered a continuous course of

---

58. *Id.*
59. *Id.*
60. *Id.* at 761.
61. People v. Jones, 792 P.2d 643, 650 (Cal. 1990).
62. 246 Cal. Rptr. 352 (Cal. Ct. App. 1988).
63. *Jones*, 792 P.2d at 652.
64. 11 Cal. Rptr. 2d 694 (Cal. Ct. App. 1992).
65. 23 Cal. Rptr. 2d 261 (Cal. Ct. App. 1993).
66. *Higgins*, 11 Cal. Rptr. 2d at 698; *see also id.* ("[W]hen the issue presented to the jury is whether a defendant committed a course of conduct and not whether he committed a specific act on a specific day, the prosecutor does not have to elect a specific act and the jury need not unanimously agree on a specific act.").
67. *Gear*, 23 Cal. Rptr. 2d at 265 (citation omitted).

conduct offense.[68] The second question of whether the Due Process Clause of the Fourteenth Amendment is violated by the legislature defining the criminal conduct in this way has largely been ignored by California courts. The court in *People v. Cissna*[69] provided a cursory due process analysis of § 288.5:

> [C]iting and distinguishing the *Gear* decision, the *Richardson* court recognized that this constitutional concern did not necessarily apply to state statutes that involved difficult problems of proof. *Richardson* noted that state statutes that permit conviction for sexual abuse of a minor based on a continuous course of conduct "may well respond to special difficulties of proving underlying criminal acts, which difficulties are absent here." Thus, *Richardson* supports the constitutionally [sic] of the continuous-course-of-conduct exception applied by the Legislature in section 288.5, subdivision (b).[70]

California courts had mostly decided their view of the issues in the California CSA statute before *Richardson* was decided by the Supreme Court. While dicta in *Richardson* may provide some basis for concluding that CSA statutes do not violate due process, a full analysis of California's CSA statute has not been done by California courts. This dicta from *Richardson* is one paragraph offering distinguishing features between the federal statute being analyzed (the CCE) and California's CSA statute[71]—the Supreme Court did not and was not intending to provide an in-depth analysis of the constitutionality of CSA statutes under the Fourteenth Amendment's Due Process Clause.[72] Further analysis is needed on this particular issue, but unfortunately California courts have avoided doing so.

B. *An Overview of Texas*

1. Texas and Jury Unanimity.—In Texas's seminal case on jury unanimity, *Ngo v. State*,[73] the court described jury unanimity as a requirement in felony cases under the state constitution and as a requirement in all criminal cases under state statutes.[74] Texas courts have described one of the main purposes of jury unanimity as complementing and helping to effectuate the "beyond a reasonable doubt" standard of proof.[75] As said by the court in

---

68. *See Higgins*, 11 Cal. Rptr. 2d at 700 ("The Legislature has the prerogative to proscribe a course of conduct, rather than specific acts, a prerogative exercised by adoption of Penal Code section 288.5.").

69. 106 Cal. Rptr. 3d 54 (Cal. Ct. App. 2010).

70. *Id.* at 70 (citations omitted).

71. Richardson v. United States, 526 U.S. 813, 821 (1999).

72. *See infra* subpart IV(B).

73. 175 S.W.3d 738 (Tex. Crim. App. 2005).

74. *Id.* at 745 & n.23 (citing Francis v. State, 36 S.W.3d 121, 126 (Tex. Crim. App. 2000) (Womack, J., concurring), which in turn cites TEX. CONST. art. V, § 13; TEX. CODE CRIM. PROC. ANN. arts. 36.29(a), 37.02, 37.03, 45.034–.036 (West 1997)).

75. *Id.* at 745 n.23 (citing United States v. Gipson, 553 F.2d 453, 457 n.7 (5th Cir. 1977)).

Ngo, "Unanimity in this context means that each and every juror agrees that the defendant committed the same, single, specific criminal act."[76] As for the level of specificity that jurors are required to agree about, the court in *Ngo* stated: "Stealing a credit card on Monday is not the same specific criminal offense as receiving a stolen credit card on Tuesday or presenting a stolen credit card to a bartender on Wednesday."[77]

2. *The History and Constitutionality of Texas's CSA Statute.*—As mentioned earlier, this formulation of jury unanimity presents problems for child molestation cases.[78] Judge Cochran of the Texas Court of Criminal Appeals (CCA) acknowledged as much in her concurrence in *Dixon*, which included a plea to the state legislature to solve this problem.[79] Judge Cochran's opinion was that "[w]e are headed for a train wreck in Texas law because our bedrock procedural protections cannot adapt to the common factual scenario of an ongoing crime involving an abusive sexual relationship of a child under current penal provisions."[80]

Judge Cochran felt that the Texas Legislature could address this problem by enacting a penal statute with this issue (a pattern of a sexually abusive relationship with a child) that was a continuous course of conduct offense.[81] She felt that a statute like this would "have advantages and disadvantages for both the prosecution and defense, but it might well assist in preserving our bedrock criminal-procedure principles of double jeopardy, jury unanimity, due-process notice, grand-jury indictments, and election law."[82]

Less than a year later, Texas Penal Code § 21.02 was enacted.[83] This CSA statute is notably more far-reaching than prior states' CSA statutes. A reason for this might be political in nature—the statute was originally called Jessica's Law, which was based on a Florida statute originally called by the same name in honor of a child who was raped and murdered by a previously convicted child molester.[84] After this incident in 2005, various states passed their own version of Jessica's Law, and these laws were notable for harsh

---

76. *Id.* at 745.

77. *Id.*

78. *See supra* note 8 and accompanying text.

79. Dixon v. State, 201 S.W.3d 731, 737 (Tex. Crim. App. 2006) (Cochran, J., concurring).

80. *Id.*

81. *See id.* (writing that "[p]erhaps the Texas Legislature can address this conundrum and consider enacting a new penal statute that focuses upon a continuing course of conduct crime—a sexually abusive relationship that is marked by a pattern or course of conduct of various sexual acts").

82. *Id.*

83. TEX. PENAL CODE ANN. § 21.02 (West Supp. 2012).

84. FLA. STAT. ANN. § 800.04 (West 2007); *see also* Sarah Shekhter, Note, *Every Step You Take, They'll Be Watching You: The Legal and Practical Implications of Lifetime GPS Monitoring of Sex Offenders*, 38 HASTINGS CONST. L.Q. 1085, 1086 (2011) (noting that Jessica's Law refers to Jessica Lunsford, a 9-year-old girl who was abducted, raped, and murdered by a 47-year-old man).

penalties, along with varied punishments (e.g., GPS tracking devices for released sex offenders and harsher zoning restrictions).[85] Thus, Texas passed a bill that contained within it a CSA statute as well as changes to the Texas Code of Criminal Procedure as influenced by other states' versions of Jessica's Law.

The statute has yet to come before the CCA, but lower courts have ruled on its constitutionality under both the Texas Constitution and the U.S. Constitution. Texas lower courts have been handling this question in two parts: (1) What did the legislature intend the jury be unanimous about?[86] (2) Does defining § 21.02 in this way violate the Due Process Clause of the Fourteenth Amendment?[87]

The first question has been answered in a similar fashion as California courts have answered it. In *Jacobsen v. State*,[88] the court found that the "statute is clear"—"it is the commission of two or more acts of sexual abuse over the specified time period—that is, the pattern of behavior or the series of acts—that is the *actus reus* element of the offense as to which the jurors must be unanimous in order to convict."[89]

As for the second question regarding due process, the court in *Jacobsen* held that § 21.02 did not violate due process by not requiring jury unanimity as to the individual acts that made up the course of conduct in § 21.02.[90] The court did not provide much rationale for this conclusion, besides citing that other courts in other states have ruled in favor of upholding CSA statutes.[91]

Thus, Texas courts have reasoned that the Texas legislature has the power to define what is a course of conduct offense, and no due process violation occurs if an offense is a course of conduct offense. This reasoning seems quite circular. It effectively claims that there are no due process concerns under the United States Constitution because the Texas legislature has decided that this offense should be a continuous course of conduct offense. Like California, Texas spends little time analyzing due process issues—as discussed in *Schad* and *Richardson*—that could possibly arise given how CSA statutes are defined.

---

85. *See generally* Shekhter, *supra* note 84 (discussing the various state laws patterned after Florida's Jessica's Law). Texas's version of Jessica's Law originally called for real-time GPS tracking of sex offenders and capital punishment for those whose victims were younger than fourteen years old. H.B. 8, 80th Leg., Reg. Sess. (Tex. 2007). The version that was enacted, though, did not include these harsher provisions. TEX. PENAL CODE ANN. § 21.02 (West Supp. 2012).

86. Jacobsen v. State, 325 S.W.3d 733, 736 (Tex. App.—Austin 2010, no pet.).

87. *Id.* at 737.

88. 325 S.W.3d 733 (Tex. App.—Austin 2010, no pet.).

89. *Id.* at 737.

90. *Id.* at 739.

91. *Id.* (citing People v. Cissna, 106 Cal. Rptr. 3d 54, 68–70 (Cal. Ct. App. 2010); State v. Sleeper, 846 A.2d 545, 550–51 (N.H. 2004); State v. Johnson, 627 N.W.2d 455, 460–64 (Wis. 2001)).

IV. The Problem with a Lack of Specificity in Jury Agreement

As discussed in Part III, Texas and California justify the potential jury unanimity issues with their CSA statutes on the basis that the offense is written to be a continuing course of conduct offense. This allows an exception for jury unanimity regarding the underlying acts/offenses. The Texas and California legislatures were clear that these offenses were written to be continuous course of conduct offenses, which has helped these state courts to easily answer the first question of their analyses.[92]

Describing an offense as not a single specific act but rather as a series of acts that constitute a continuous conduct offense is not a new invention. And treating these generic evidence cases as such an offense fits very cleanly with what is occurring in one type of these cases—cases that involve ongoing harm over a period of time. Thus, what a jury should agree on in these cases is that an ongoing abusive relationship has occurred.

These statutes were constructed imperfectly though. Instead of simply addressing the situation of an ongoing abusive relationship that only has generic evidence, these CSA statutes can be applied much more expansively. There are three variations of CSA statutes I will examine that have the following explicit requirements, to varying degrees: (1) a set number of acts, (2) a time component, (3) the victim be the same child in all acts, and (4) same residence or recurring access. The most expansive CSA statutes only have requirements 1 and 2.[93] Average expansiveness CSA statutes have requirements 1, 2, and 3.[94] The least expansive CSA statutes contain requirements 1, 2, 3, and 4.[95]

The problem, even in the least expansive variety of CSA statutes, is that these requirements are simply not appropriate enough proxies for the original problem sought to be solved, and the end result makes it possible that applications of CSA statutes will stretch far beyond what solving the original

---

92. *See supra* text accompanying notes 66–68 & 88–89.

93. *E.g.*, TEX. PENAL CODE ANN. § 21.02 (West Supp. 2012) (defining the offense of continuous sexual abuse as two or more acts of sexual abuse during a period of thirty days or more, "regardless of whether the acts of sexual abuse are committed against one or more victims").

94. *E.g.*, ARIZ. REV. STAT. ANN. § 13-1417 (2010) (defining continuous sexual abuse of a child as requiring three or more acts over the period of three months or more against a child); N.Y. PENAL LAW § 130.75 (McKinney 2009) (defining the "[c]ourse of sexual conduct against a child" offense as two or more acts of sexual conduct with a child over a period of three months or more); N.D. CENT. CODE § 12.1-20-03.1 (2012) (requiring three or more sexual acts "with a minor under the age of fifteen . . . during a period of three or more months"); WIS. STAT. ANN. § 948.025 (West 2005 & Supp. 2011) (requiring three or more violations within a specified period of time "involving the same child").

95. *E.g.*, CAL. PENAL CODE § 288.5 (West 2008) (requiring three or more acts with a child over a period of three or more months by a person who "either resides in the same home with the minor child or has recurring access to the child"); DEL. CODE ANN. tit. 11, § 776 (Supp. 2010) (stating that a person commits an offense by engaging in three or more acts of sexual conduct with a minor over a period of three or more months while "either residing in the same home with the minor child or having recurring access to the child").

problem called for. I will discuss these various issues by using the following broad categories: general problems and constitutional problems.

A. *General Problems*

1. Patchwork Verdicts.—What level of specificity is needed when it comes to jury unanimity? Different courts and commentators have different answers to this question.[96] As previously noted, continuous course of conduct offenses change the specificity level from requiring unanimity on the underlying acts to only requiring unanimity on a much more broadly defined criminal offense.[97] This introduces the problem of patchwork verdicts.[98] Patchwork verdicts exist when there is agreement "on the general verdict but disagree[ment] on the particulars behind the verdict."[99] Patchwork verdicts are not necessarily a problem. Justice Scalia provides such an example in his concurrence in *Schad*:

> When a woman's charred body has been found in a burned house, and there is ample evidence that the defendant set out to kill her, it would be absurd to set him free because six jurors believe he strangled her to death (and caused the fire accidentally in his hasty escape), while six others believe he left her unconscious and set the fire to kill her.[100]

But patchwork verdicts can also be problematic. In Justice Scalia's example, there is no doubt among the jury that murder was committed by the defendant, just how the murder actually occurred. Not too many people would feel uneasy about this result. But can the same be said about jurors disagreeing about the underlying acts that comprise a broader (more amorphous) offense? I will examine patchwork verdicts in the context of the types of CSA statutes, and readers can decide for themselves.[101]

---

96. *See, e.g.*, Morris, *supra* note 4, at 51–57 (reviewing state and federal court decisions requiring a specific unanimity jury instruction but noting that defining the precise circumstances in which such instruction is required "remains an elusive exercise"); Westen & Ow, *supra* note 4, at 153 (explaining that Supreme Court Justices have, in two cases addressing the issue, agreed that the Constitution sometimes requires unanimity on how offenses are committed but disagreed "about nearly everything else"); Miller, *supra* note 4, at 2279 (proposing a framework that would require unanimity as to the elements of a crime but permit divergence as to "mere alternate means of fulfilling the elements").

97. *See supra* note 7 and accompanying text.

98. *See generally* Hayden J. Trubitt, *Patchwork Verdicts, Different-Jurors Verdicts, and American Jury Theory: Whether Verdicts Are Invalidated by Juror Disagreement on Issues*, 36 OKLA. L. REV. 473 (1983).

99. Elizabeth A. Larsen, Comment, *Specificity and Juror Agreement in Civil Cases*, 69 U. CHI. L. REV. 379, 381 n.10 (2002).

100. Schad v. Arizona, 501 U.S. 624, 650 (1991) (Scalia, J., concurring in part and concurring in the judgment).

101. For the sake of brevity (and to avoid repetitiveness), I will only look at the most expansive and the least expansive CSA statutes. This follows along with my focus on Texas and California.

a. Most Expansive.—Take the following example: Adam is alleged to have sexually assaulted child A in December 2010, sexually assaulted child B in June 2011, and sexually assaulted child C in December 2011. In Texas, Adam can now be charged with violating Texas's CSA statute. Assuming that each alleged sexual assault occurred once as one specific event (and the evidence being presented only pertains to one event occurring and not to all three events occurring), it is hard to see how this situation is like the original problem intended to be solved by CSA statutes. There is no problem with generic evidence since it involves three distinct events with specific evidence only relating to each event.

This results in the burden of proof being effectively lowered for each specific assault since patchwork verdicts are now possible. It is possible Adam can be found guilty if only six jurors believe Adam is guilty of sexually assaulting Child A and Child B (but not Child C) and the other six jurors believe Adam is guilty of sexually assaulting Child B and Child C (but not Child A). Note that the jury was only unanimous about Adam sexually assaulting Child B and there was not even a simple majority on whether Adam sexually assaulted Child A or Child C. Thus, if the prosecutor had been forced to try and convict Adam under three charges of sexual assault, only one successful conviction would have resulted. Instead, under the CSA statute, Adam can be convicted of a crime that carries with it a much greater punishment than sexual assault.[102]

This hypothetical can be made even more extreme. Imagine if Adam was charged with sexually assaulting four different children in four different, specific events. Now, it is possible that there is not even a majority of jurors who believe Adam is guilty of committing one of the four specific acts of abuse.[103] Or at its most extreme: if Adam is charged with sexually assaulting twenty-four different children, it is possible that a patchwork verdict could be reached when Adam is guilty under the CSA statute but only a single juror believes he is guilty under each individual alleged sexual assault.

b. Least Expansive.—While less egregious, patchwork verdicts are still possible under the least expansive CSA statutes. The requirement of "shared residence or recurring access" can do a significant amount of work in capturing only the fact patterns that are like the problem originally intended to be addressed by CSA statutes—generic evidence cases of an ongoing abusive relationship. In application though, this has not occurred in

102. In Texas, sexual assault can be a felony of the second degree while continuous sexual abuse of a young child or children is a felony of the first degree with a minimum sentence of twenty-five years. *Compare* TEX. PENAL CODE ANN. § 22.011 (West 2011) (providing the punishment for sexual assault), *with* TEX. PENAL CODE ANN. § 21.02 (West Supp. 2012) (providing the punishment for continuous sexual abuse of a child).

103. This could occur if Jurors 1–6 believe Adam is guilty of sexually assaulting only Child A and Child B and Jurors 7–12 believe Adam is guilty of sexually assaulting only Child C and Child D.

California. Instead, "recurring access" has been given a very broad meaning by California's courts with the very intent of making sure that California's statute is as expansive as possible: instead of California's CSA statute acting as a limited offense aimed at addressing the problem of generic evidence, California courts see this statute as an expansion of power for prosecutors going after alleged child molesters.[104]

Thus, with a broader meaning given to "recurring access," it is not hard to come up with hypotheticals similar to the ones seen in the previous subsection. For example, imagine a coach accused of sexually assaulting one of the boys on his team four times over a span of six months. Each event could be a discrete occurrence with different facts involved (and thus specific evidence related to each event), and a jury could mix and match verdicts on each event to satisfy the requirement of three acts within the time period.

While the patchwork verdicts that are possible under California's CSA statute are significantly less dramatic than ones that are possible in Texas, the following needs to be reiterated: this hypothetical does not represent the original problem of generic evidence that CSA statutes were meant to address, and the situation does not seem like a typical continuous course of conduct offense.

2. A Strategic Advantage for Prosecutors.—Along with patchwork verdicts, another problem arises with CSA statutes, no matter which version is considered: an inequity is produced by CSA statutes granting prosecutors the ability to strategically choose between charging someone with a continuing course of conduct offense or the specific, underlying offenses. Allowing the prosecutor the opportunity to choose what to charge the defendant with allows the prosecutor to avoid the disadvantages of working with a continuous course of conduct statute (which only allows one conviction for all acts) and to maximize the advantages (by making a conviction easier by allowing patchwork verdicts).

To provide an example: if a prosecutor feels like he has very strong evidence of a person sexually assaulting a child on five separate occasions, the prosecutor could simply pursue five separate charges of sexual assault. Alternatively, if the prosecutor feels like he has strong evidence for one of the offenses, but for the other occasions the evidence is more conflicted, the prosecutor could try to maximize his chances of a conviction by taking the lower burden of the CSA statute and allowing the jurors to mix and match verdicts (i.e., produce patchwork verdicts).

---

104. *See* People v. Rodriguez, 49 P.3d 1085, 1089 (Cal. 2002) ("[A]s made clear by the legislative declarations accompanying it, section 288.5 was enacted to broaden, not narrow, the reach of this state's child molestation laws.").

B. *Constitutional Problems*

It is unlikely that the Supreme Court will address CSA statutes any time soon. While the Court indicated there could be due process problems with the CCE federal statute if interpreted in a way that would match how CSA statutes are used, the Supreme Court decided to dodge the issue by choosing an alternative interpretation that faced no due process dangers.[105] Dicta in *Richardson* also indicated that the Supreme Court may view the CCE statute and California's CSA statute as different, since "[t]he state practice may well respond to special difficulties of proving individual underlying criminal acts, which difficulties are absent here."[106]

But these difficulties only exist with generic evidence in situations of ongoing sexual abuse of a young child. In the situations outlined in the patchwork verdicts section, no such special difficulties existed. The patchwork verdicts possible under an expansive CSA statute also look very similar to a situation described by Justice Scalia in his concurrence (and cited approvingly by the *Richardson* majority),[107] in which Justice Scalia believes due process would be violated: "We would not permit, for example, an indictment charging that the defendant assaulted either X on Tuesday or Y on Wednesday, despite the 'moral equivalence' of those two acts."[108] Yet, expansive CSA statutes allow just that, with the extra window dressing that the acts must have occurred "during a period that is 30 or more days in duration."[109]

After mentioning California's CSA statute, the Court also mentions a difference between federal law, which requires jury unanimity under the Sixth Amendment, as compared to states, which are allowed supermajority verdicts under the Fourteenth Amendment.[110] I submit that this is a misstep in analysis on the part of the Court in *Richardson*—the issue with the CCE statute and CSA statutes is the level of specificity required by jury agreement

---

105. *See* Richardson v. United States, 526 U.S. 813, 820 (1999) ("We have no reason to believe that Congress intended to come close to, or to test, those constitutional limits [imposed by the Due Process Clause] when it wrote [the CCE] statute.").

106. *Id.* at 821 (citation omitted). Another distinction not mentioned by the Court that might be important is the difference in breadth. CSA statutes normally contain less than a dozen qualifying predicate acts. *See, e.g.*, TEX. PENAL CODE ANN. § 21.02 (West Supp. 2012) (listing eight different underlying offenses that qualify for purposes of satisfying the two act requirement). The CCE statute, on the other hand, is drawn from "[t]he two chapters of the Federal Criminal Code setting forth drug crimes contain[ing] approximately 90 numbered sections, many of which proscribe various acts that may be alleged as 'violations' for purposes of the series requirement in the statute." *Richardson*, 526 U.S. at 819.

107. *Richardson*, 526 U.S. at 820.

108. Schad v. Arizona, 501 U.S. 624, 651 (1991) (Scalia, J., concurring in part and concurring in the judgment).

109. *E.g.*, TEX. PENAL CODE ANN. § 21.02(b)(1) (West Supp. 2012).

110. *See Richardson*, 526 U.S. at 821 (recognizing that states are not bound by the federal jury unanimity standard); *see also supra* note 1.

(regardless of whether that agreement must be unanimous or a supermajority).

V. Where Do We Go from Here?

A. How to Handle Current CSA Statutes

CSA statutes have raised unnecessary problems in trying to address the problems of prosecuting child molesters. The problems caused by current CSA statutes could be addressed in one of three ways: (1) by a challenge under the U.S. Constitution; (2) by state constitutional challenges; or (3) by the state legislatures amending these statutes to closer resemble the original problem trying to be addressed—the problem of generic evidence when prosecuting child molesters.

A successful challenge under the U.S. Constitution is unlikely in the near future.[111] Six Supreme Court Justices were concerned enough by similar issues with the federal CCE statute to employ statutory construction as a means of constitutional avoidance in *Richardson*.[112] But no such option would be present with these state statutes; the Supreme Court would be asked to rule a state statute unconstitutional. The Supreme Court generally does not like to do this, especially with criminal statutes.[113] Further complications would be that *Schad* and *Richardson* have shown this area of the law to be a doctrinal mess, with a lack of agreement among Justices on what the actual rules should be.[114]

Striking down these statutes through state courts is also unlikely to be effective. As previously mentioned, only one state court has tried this approach: Hawaii.[115] Other states' highest courts either have already found the respective state's CSA statute to be constitutional (under both the state

111. Such a challenge would be an as-applied constitutional challenge, since CSA statutes do not always trigger constitutional concerns in their application. For a discussion of how such a challenge would proceed, see generally Michael C. Dorf, *Facial Challenges to State and Federal Statutes*, 46 STAN. L. REV. 235 (1994). "Conventional wisdom holds that a court may declare a statute unconstitutional in one of two manners: (1) the court may declare it invalid on its face, or (2) the court may find the statute unconstitutional as applied to a particular set of circumstances." *Id.* at 236.

112. *Richardson*, 526 U.S. at 820.

113. *See* Patterson v. New York, 432 U.S. 197, 201 (1977) ("It goes without saying that preventing and dealing with crime is much more the business of the States than it is of the Federal Government, and that we should not lightly construe the Constitution so as to intrude upon the administration of justice by the individual States." (citation omitted)). For a list of the 935 state statutes ruled unconstitutional between 1809 and 2002, see CONG. RESEARCH SERV., LIBRARY OF CONG., THE CONSTITUTION OF THE UNITED STATES OF AMERICA: ANALYSIS AND INTERPRETATION 2163–323 (Johnny H. Killian et al. eds., 2004), *available at* http://www.gpo.gov/fdsys/pkg/GPO-CONAN-2002/pdf/GPO-CONAN-2002.pdf.

114. *See* Westen & Ow, *supra* note 4, at 167 ("The twelve Justices in *Schad* and *Richardson* together advance five constitutional tests, four of which come from *Schad*, and one of which inheres in Justice Kennedy's dissent in *Richardson*.").

115. *See supra* notes 30–31 and accompanying text.

constitution and U.S. Constitution) or have decided not to review lower courts' decisions upholding these statutes.[116] And even in Hawaii, where the Supreme Court of Hawaii ruled that their state's CSA statute violated the state constitution, the end result was a state constitutional amendment to allow the return of Hawaii's CSA statute.[117] Such a result is not surprising, due to the highly charged topic of these statutes: it is much easier to garner public support for protecting children from child molesters than it is to garner support for protecting alleged child molesters from overly aggressive state statutes.

The last approach—state legislatures amending their respective CSA statutes—has the best chance of succeeding, though garnering support in state legislatures would be a difficult task, especially for such an unpopular cause. Such an amendment could replace the anti-unanimity provision in these CSA statutes with a more narrowly constructed provision,[118] or legislatures could choose to remove the anti-unanimity provision entirely.[119]

B. *How CSA Statutes Can Be Changed to Better Address the Problems with Prosecuting Child Molesters*

1. A Return (in Part) to the Status Quo.—One possible solution is for state legislators to either just remove the anti-unanimity provision in their CSA statute or get rid of the CSA statute completely. This could allow state judges to rule as a matter of law, based on the particular facts of the case, whether the CSA (or a more traditional sexual abuse statute, if there is no CSA statute) should be applied as a continuous course of conduct offense.

There is some irony to that fact that California has judicially created a continuous course of conduct exception in its traditional sexual abuse statute (California Penal Code § 288), which has been the case since *People v. Jones*,[120] California's seminal decision on jury unanimity. *Jones* was decided shortly after California's CSA statute, California Penal Code § 288.5, was enacted.[121] The court in *Jones* could have treated § 288 just like *Van Hoek*, which did not allow a conviction based on generic evidence,[122] since a CSA statute now existed in California to handle the problems of child molesters and generic evidence. Instead, the court in *Jones* decided to overrule the *Van Hoek* line of cases.[123] Thus, while the legislature relied on a CSA statute, § 288.5, to provide a statute to fill in the gap created by the *Van Hoek* line of

116. *See supra* notes 29, 64–66, 90–91 and accompanying text.
117. *See supra* note 31 and accompanying text.
118. *See infra* section V(B)(2).
119. *See infra* section V(B)(1).
120. 792 P.2d 643 (Cal. 1990).
121. *Id.* at 652.
122. *See supra* notes 61–62 and accompanying text.
123. *Id.* at 659.

cases, the California Supreme Court felt that they could just overturn *Van Hoek*, making § 288.5 unnecessary.

In retrospect, the California Supreme Court's solution may have been the better one. As said by the court in *Jones*,

> [W]e reject the contention that jury unanimity is necessarily unattainable where testimony regarding repeated identical offenses is presented in child molestation cases. In such cases, although the jury may not be able to readily distinguish between the various acts, it is certainly capable of unanimously agreeing that they took place in the number and manner described.[124]

The court provided the following example:

> [I]f an information charged *two* counts of lewd conduct during a particular time period, the child victim testified that such conduct took place *three times* during that same period, and the jury believed that testimony in toto, its difficulty in differentiating between the various acts should not preclude a conviction of the two counts charged, so long as there is no possibility of jury disagreement regarding the defendant's commission of any of these acts.[125]

Thus, in *Jones* the California Supreme Court set forth a way of treating generic evidence as an all-or-nothing argument: either the jury believed beyond a reasonable doubt that the abuse occurred because of the generic evidence or the jury did not believe beyond a reasonable doubt that the abuse occurred.[126] This represents an exception to California's jury unanimity doctrine, but it is a narrow one. Since it is dependent on generic-evidence-type situations, the reach for this exception is less broad than California's CSA statute.

California's approach to § 288 could be applied to § 288.5 as well if the anti-unanimity provision was removed. A problem with this approach is that many states turned to CSA statutes precisely to avoid this approach: allowing criminal offenses to become continuous course of conduct offenses, based on the facts, raised concerns about principles of separation of powers and traditional concepts of criminal law.[127] While this approach might be better than the supposed cure (CSA statutes), it does not represent the most comforting solution.

2. Replacing the Anti-Unanimity Provision.—The best possible solution might be for state legislatures to fix CSA statutes themselves. There is a way to craft a CSA-type offense in a narrower way that more accurately represents the original problem with prosecuting child molesters. It could be

124. *Id.* at 658.

125. *Id.*

126. *Id.* at 658–59.

127. *See supra* notes 9, 30 and accompanying text.

done in most cases by replacing the anti-unanimity clause in CSA statutes with the following clause in each of these statutes:

> If there is generic evidence for multiple acts occurring, a jury should receive a jury instruction that they may rely on this generic evidence to conclude that the requisite number of acts were committed.
>
> For purposes of this section, generic evidence is defined as evidence that is not particular to one event that is relied upon to attempt to prove the occurrence of more than three [or "three or more"] events occurring.

Such a provision would help resolve issues that some courts have in allowing generic evidence to prove the state's more traditional sexual abuse statute, while being narrowly tailored to resolve the specific problem at hand (and not reaching far beyond the original problem, like current CSA statutes do).

A drawback to CSA statutes constructed in this way is that there is no clear solution on how to address situations where there is both generic evidence of abuse and specific evidence relating to a particular event where abuse occurred. An argument can be made that letting in generic evidence, when there is specific evidence for each event of abuse trying to be proved, allows prosecutors to pile on the generic evidence, even if the generic evidence is too weak for a jury to conclude that the defendant is guilty on the basis of the generic evidence alone. Allowing generic evidence into a trial triggers concerns that some courts have had about traditional principles of criminal law based around proving a discrete act or conduct.[128] Opening up the gates to generic evidence, even in this more limited way, can allow prosecutors to present generic evidence, even when they realize that the generic evidence, by itself, will not be strong enough for a conviction.[129] A rebuttal to this argument is that the very nature of these crimes is the reason for greater latitude in allowing in evidence, and thus there really is no problem with allowing prosecutors to do such a thing. Alternatively, this provision could be even narrowly constructed to state that judges should make a determination whether the generic evidence is adequate to convince a jury beyond a reasonable doubt that the defendant committed the requisite number of acts in the statute: if it is, then it should be allowed in as evidence; if it is not, then it should be excluded.

Ultimately, I believe the sample provision I have provided in this section represents the best solution in constructing a CSA statute that is not dangerously broad but is still able to remedy the original problem meant to be addressed in prosecuting child molesters.

---

128. *See, e.g.*, Dixon v. State, 201 S.W.3d 731, 737 (Tex. Crim. App. 2006) (Cochran, J., concurring) (stating that jurors must agree on whether the defendant committed one very specific criminal act and that a purpose of criminal procedure is to provide a defendant with "advance[d] notice of precisely what criminal act he is alleged to have committed and when it occurred").

129. *See supra* section IV(A)(2).

Conclusion

In this Note I argue that legislatures in certain states have gone too far in adopting CSA statutes. These statutes attempt to solve very real problems in our criminal justice system, but the end result is an approach that is much too broad and far-reaching. If states wish to better honor their respective case law doctrine on jury unanimity and the level of specificity required, they need to rethink their devotion to these types of statutes. The problem attempted to be solved by CSA statutes can be addressed in more narrow ways that pay more respect to jury unanimity, due process, and separation of powers. I believe the best solution would be for state legislators to amend their current CSA statutes and replace the anti-unanimity provision in these statutes with a more narrowly constructed provision that directly addresses the original problem intended to be solved by these statutes.

—Brian Bah

U-Turns on the One-Way Street: Public Rights and Representation Theory in Patent Validity Litigation*

Introduction 1227
I. Joinder and Consolidation 1229
A. Mandatory Joinder 1229
B. Permissive Joinder and Consolidation 1231
II. Class Actions 1234
A. Numerosity 1235
B. Commonality 1236
C. Typicality 1237
D. Adequacy of Representation 1238
E. Rule 23(b)(2) 1239
III. Virtual Representation 1242
IV. Is Preclusion on Validity Good Public Policy? 1245
A. Challenge "Bad Patents" 1245
B. Prevent Patent Devaluation 1247
C. Notice 1247
D. Conserve Judicial Resources 1248
E. Fairness to the Patentee 1249
Conclusion 1251

Introduction

Blonder-Tongue[1] is a rough "one-way street" for a patent owner.[2] If the patent owner loses a judgment on validity just one time, his patent becomes invalid against the world.[3] Conversely, if he wins, absentee infringers can

---

* I am deeply honored and humbled to find my work in these pages. Many contributed to the development of this Note. I offer a few special thanks here at the risk of omitting others without whose help I could never have completed this project. First, thank you to Professor Robert G. Bone and Adjunct Professor Robert Turner of the University of Texas School of Law. This paper incorporates much that I learned in their classes, and both took far more time than I ever should have requested to discuss the project and improve it. Second, thank you to the Volume 91 Editorial Board and staff of the *Texas Law Review*, especially its outstanding Notes editors Michael Selkirk, Monica Hughes, Ross MacDonald, and Lauren Ross, who honored me with this platform and strengthened this piece immeasurably with their sharp editing and diligent research. Finally, thank you to my wonderful family and friends, especially my beautiful Emilija, for all the advice, patience, support, and companionship.

1. Blonder-Tongue Labs., Inc. v. Univ. of Ill. Found., 402 U.S. 313 (1971).

2. *See In re* Yarn Processing Patent Litig., 56 F.R.D. 648, 654 (S.D. Fla. 1972) ("The one-way street of *Blonder-Tongue* . . . means that if the patentee loses, the other putative class members win by collateral estoppel default. If the patentee prevails in one suit, that does not control any subsequent adjudication. A new defendant is allowed to retry the validity of the patent anew.").

3. *Blonder-Tongue*, 402 U.S. at 350.

challenge the same patent's validity again and again in subsequent suits.[4] At the core of this rule are the due process right and the "deep-rooted historic tradition that everyone should have his own day in court."[5] Although a once-victorious patent owner already had his day in court, the absentee infringer did not. Thus, even though issues relevant to validity do not differ between infringers, the patentee must defend his patent anew against every subsequent infringer that arises.

But does due process require that every infringer have a day in court to challenge patent validity?[6] After all, an infringer that challenges patent validity stands not for an individual right, but for "*the public's right* to retain knowledge already in the public domain."[7] Because patent validity challenges vindicate public rights, the same procedural entitlements that attach to individual rights may not attach for successive challenges of validity.[8] Moreover, the law's allowance of successive challenges to patent validity frustrates the notice function and valuation of patents, provides disincentives to actually litigate the validity of bad patents, and impairs judicial economy.[9] Therefore, this Note contends that a vigorous challenge against a patent's validity fully vindicates the public right to access ideas in the public domain and should preclude successive challenges by future infringers.

Part I looks at the foundational tools of group litigation—joinder and consolidation—as applied in patent law, including recently adopted restrictions in the Leahy-Smith America Invents Act of 2011 (AIA).[10] Part II asks whether the Rule 23 class action allows certification of present and future infringers for validity determinations. Part III examines whether the "public law" nature of the validity inquiry reinvigorates the notion of "virtual representation" rejected in *Taylor v. Sturgell*.[11] Part IV analyzes whether procedural rules that facilitate the establishment of patent validity further the substantive goals of patent law and the imperatives of wise judicial administration.

---

4. *See id.* (rejecting the doctrine of mutuality for collateral estoppel).

5. Taylor v. Sturgell, 553 U.S. 880, 892–93 (2008). In *Taylor*, the Supreme Court identified only six categories of nonparty preclusion permissible under the Due Process Clause. *Id.* at 893–95.

6. *See* Thillens, Inc. v. Cmty. Currency Exch. Ass'n of Ill., Inc., 97 F.R.D. 668, 674 (N.D. Ill. 1983) ("It is the hallmark of our system of justice that personal rights cannot be compromised without due process.").

7. Pfaff v. Wells Elecs., Inc., 525 U.S. 55, 65 (1998) (emphasis added); *see also* Precision Instrument Mfg. Co. v. Auto. Maint. Mach. Co., 324 U.S. 806, 816 (1945) ("A patent by its very nature is affected with a public interest."); Merrill v. Yeomans, 94 U.S. 568, 573 (1876) ("The public [would] be deprived of rights supposed to belong to it . . . .")

8. *See* Robert G. Bone, *The Puzzling Idea of Adjudicative Representation: Lessons for Aggregate Litigation and Class Actions*, 79 GEO. WASH. L. REV. 577, 610 (2011) (arguing that process-based rights do not attach when only group, rather than individual, rights are at stake).

9. *See infra* Part IV.

10. Leahy-Smith America Invents Act, Pub. L. No. 112-29, 125 Stat. 284 (2011) (codified in scattered sections of 35 U.S.C.).

11. 553 U.S. 880, 886 (2008).

I. Joinder and Consolidation

The ordinary mechanism for binding a party to a judgment is joinder.[12] However, for a patentee seeking to fortify his patent from validity challenges, simply suing all available infringers has several practical limitations. First, future infringers—at least, those not imminently near the act of infringement—cannot be precluded because they cannot be sued.[13] Second, a patentee may not have the resources to identify and sue all current infringers.[14] Third, a host of jurisdictional problems can arise.[15] Finally, recently adopted provisions in the AIA restrict joinder and consolidation for trial of defendants simply on the basis of their alleged infringement of the same patent.[16]

A. Mandatory Joinder

Absentee infringers are not necessary parties under any category of Rule 19. Rule 19 provides that parties must be joined if their presence in the litigation is necessary to provide complete relief between the parties, the litigation threatens to "impair or impede" their legal rights, or there is a risk that conflicting judgments will create inconsistent obligations for the opposing party in litigation.[17] None of these apply to absentee infringers on the issue of validity. First, because the court can resolve patent disputes without the involvement of unrelated infringers, these additional parties are

---

12. Martin v. Wilks, 490 U.S. 755, 765 (1989).

13. *See* Vieux Carre Prop. Owners, Residents & Assocs., Inc. v. Brown, 875 F.2d 453, 457 (5th Cir. 1989) ("[I]t is implicit in Rule 19(a) itself that before a party . . . will be joined as a defendant the plaintiff must have a cause of action against it."); *see also* Charles Silver, *Comparing Class Actions and Consolidations*, 10 REV. LITIG. 495, 502 (1991) (explaining that consolidation is restricted to pending cases and cannot encompass future parties as can a class action).

14. *See* Claude Crampes & Corinne Langinier, *Litigation and Settlement in Patent Infringement Cases*, 33 RAND J. ECON., 258, 258–60 (2002) (discussing the high costs of monitoring for patent infringement and its effects on enforcement).

15. *See generally* Joan Steinman, *The Effects of Case Consolidation on the Procedural Rights of Litigants: What They Are, What They Might Be, Part I: Justiciability and Jurisdiction (Original and Appellate)*, 42 UCLA L. REV. 717 (1995) (examining the effects of consolidation on the procedural rights of litigants and focusing on jurisdictional issues). In patent suits, venue is appropriate in any jurisdiction in which the defendant is subject to personal jurisdiction. *See* VE Holding Corp. v. Johnson Gas Appliance Co., 917 F.2d 1574, 1584 (Fed. Cir. 1990) ("[T]he first test for venue under § 1400(b) with respect to a defendant that is a corporation . . . is whether the defendant was subject to personal jurisdiction in the district of suit at the time the action was commenced.").

16. *See* Leahy-Smith America Invents Act, Pub. L. No. 112-29, § 299, 125 Stat. 284, 332–33 (2011) (codified at 35 U.S.C. § 299 (2006 & Supp. V 2012) ("[A]ccused infringers may not be joined in one action as defendants or counterclaim defendants, or have their actions consolidated for trial, based solely on allegations that they each have infringed the patent or patents in suit."); *see also infra* notes 36–46 and accompanying text.

17. FED. R. CIV. P. 19(a); *see also* Donald E. Burton, *The Metes and Bounds of the Defendant Class Action in Patent Cases*, 5 J. MARSHALL REV. INTELL. PROP. L. 292, 295–96 (2006) (considering the use of class actions to avoid the issue preclusion asymmetry caused by *Blonder-Tongue* by joining all potential defendants in the first action if possible).

not necessary to provide complete relief.[18] Second, because a prior judgment does not affect absentees' rights beyond the effect of mere negative legal precedent, a court's denial of an infringer's invalidity arguments does not "impair or impede" the rights of absentees.[19] Finally, conflicting judgments do not create inconsistent obligations for the patent owner because the patent is presumed valid against all and becomes invalid against all if ever found invalid.[20] Thus, as it relates to patent validity under current law, unrelated infringers are not necessary parties.[21]

Still, Rule 19 and its conception of necessary parties serve as a useful frame for the day-in-court right. Consider the counterfactual: if a patent owner could conclusively establish patent validity by winning in the first litigation, absentee infringers might be considered necessary parties because a binding judgment of validity would "impair" the absentee's personal legal rights by denying him the opportunity to challenge the patent's validity in successive actions.[22] If so, the absentee's due process rights would entitle the absentee to a day in court to contest the patent's validity, unless he is joined in the first action.[23] But this argument assumes that the infringer has a personal legal right to challenge patent validity, a due process right.

For public rights like the right to access ideas in the public domain, attributing due process rights to each of the public right's individual beneficiaries enables essentially infinite challenges and frustrates the establishment of binding judgments.[24] As a result, Rule 19 requirements are often relaxed in cases involving public rights, allowing for final judgments without necessitating joinder of (arguably) impaired absentees.[25] According to this view, an adverse judgment justifiably impairs an absentee's ability to

18. Edward Hsieh, Note, *Mandatory Joinder: An Indirect Method for Improving Patent Quality*, 77 S. CAL. L. REV. 683, 689 (2004).

19. *E.g.*, Janney Montgomery Scott, Inc. v. Shepard Niles, Inc., 11 F.3d 399, 406–07 (3d Cir. 1993) (finding that a possibility of a negative persuasive precedent does not require joinder under FED. R. CIV. P. 19(a)).

20. *Cf.* EEOC v. Peabody W. Coal Co., 610 F.3d 1070, 1080–81 (9th Cir. 2010) ("If the Secretary is not made a party to the suit, he may ignore the court's judgment and place conflicting demands upon the Nation who will be required by res judicata to honor the judgment.").

21. *See* Burton, *supra* note 17, at 314 ("[T]he 23(b) requirements will be a major obstacle to convincing a court to certify a class of infringers.").

22. *See, e.g.*, *Janney Montgomery Scott*, 11 F.3d at 406 & n.8 (holding that to be a necessary party, there must be a reasonably likely outcome that can preclude the absent party with respect to an issue material to the absent party's rights or duties).

23. *See* Note, *Preclusion of Absent Disputants to Compel Intervention*, 79 COLUM. L. REV. 1551, 1560–61 (1979) (criticizing a case interpreted as binding absentee infringers to the outcome of patent litigation when the infringers had not been considered necessary parties, thus potentially violating their due process right to be fully heard).

24. *See* Tyus v. Schoemehl, 93 F.3d 449, 456 (8th Cir. 1996), *abrogated by* Taylor v. Sturgell, 553 U.S. 880 (2008) ("If parties were allowed to continually raise issues already decided, public law claims would assume immortality." (internal quotation marks omitted)).

25. *See* Carl Tobias, *Rule 19 and the Public Rights Exception to Party Joinder*, 65 N.C. L. REV. 745, 745–46 (1987) (analyzing the judicially created public-rights exception to joinder rules, which sometimes permits public-rights suits to continue despite prejudice to absentees).

invoke shared public rights when an adequate representative fully vindicates the public right on behalf of all interested absentees.[26] But what makes a representative "adequate"? What does it mean to "fully vindicate" the public right? How do we identify "interested absentees"? These questions clash practicality against theory, and I attempt to address them below.

First, however, I explore two tools that Congress recently removed from the patent owner's procedural tool bag: permissive joinder and consolidation.

B. *Permissive Joinder and Consolidation*

Unaltered, Rule 20(a)(2) permits joinder of defendants where the plaintiff asserts rights to relief against all defendants arising out of the same transaction or occurrence and concerning a common question of law or fact.[27] "The purpose of the rule is to promote trial convenience and expedite final determination of disputes, thereby preventing multiple lawsuits."[28] Joinder is "strongly encouraged" by the "impulse [] toward entertaining the broadest possible scope of action consistent with fairness to the parties."[29] In light of this impulse, the "same transaction" requirement is given a "flexible meaning [that] may comprehend a series of many occurrences, depending not so much upon the immediateness of their connection as upon their logical relationship."[30]

Similarly, Rule 42 allows a federal district court to "join for hearing or trial any or all matters at issue . . . [or to] consolidate" actions that involve "a common question of law or fact."[31] In addition, Rule 42 allows the court to bifurcate the litigation so that particular issues can be heard in separate trials.[32] Unlike joinder, consolidation "does not merge the suits into a single cause, or change the rights of the parties, or make those who are parties in one suit parties in another."[33] But as a practical matter, very little differentiates joinder and consolidation at the validity phase of patent litigation. Under either mechanism, the validity inquiry focuses only on the patent, its prosecution, and the prior art.[34] Wrestling with these rules, courts

---

26. *Id.* at 746, 776.

27. FED. R. CIV. P. 20(a)(2).

28. Mosley v. Gen. Motors Corp., 497 F.2d 1330, 1332 (8th Cir. 1974).

29. United Mine Workers of Am. v. Gibbs, 383 U.S. 715, 724 (1966).

30. Moore v. N.Y. Cotton Exch., 270 U.S. 593, 610 (1926); *see also Mosley*, 497 F.2d at 1333 ("[A]ll 'logically related' events entitling a person to institute a legal action against another generally are regarded as comprising a transaction or occurrence.").

31. FED. R. CIV. P. 42(a).

32. FED. R. CIV. P. 42(b).

33. Johnson v. Manhattan Ry. Co., 289 U.S. 479, 496–97 (1933).

34. *See* Markman v. Westview Instruments, Inc., 517 U.S. 370, 373–74 (1996) (explaining the requirements of a valid patent under U.S. law).

have experimented with a variety of trial procedures in order to achieve fair and efficient resolution of patent disputes.[35]

On September 16, 2011, President Obama signed into law the Leahy-Smith America Invents Act, a sweeping revision of the patent laws.[36] Among the changes wrought by the AIA, § 299 of the Patent Act now restricts the ability of patent owners to join or consolidate for trial actions against unrelated infringers "solely on allegations that they each have infringed the patent or patents in suit."[37] In theory, these restrictions impede so-called "patent trolls" from filing nuisance suits against numerous defendants in the hopes of extracting unwarranted settlements.[38] In particular, the restrictions aimed to curtail a unique practice in the patent-friendly Eastern District of Texas, which prior to passage of the Act applied a liberal "logical relationship test" to allow joinder of all defendants alleged to have infringed the same patent even if those defendants had no other connections to each other.[39]

Although appropriate for infringement phases, the AIA restrictions are too broad because they fail to provide a categorical exception for aggregated litigation of validity. It makes sense to restrict joinder and consolidation on the issue of infringement because the infringement inquiry focuses on

35. *See, e.g.*, Ceats, Inc. v. Cont'l Airlines, Inc., No. 6:10cv120, 2011 WL 2971243, at *1–2 (E.D. Tex. July 21, 2011) (applying a variety of different factors used by courts to determine proper trial procedure in a patent case); *see also* Phillips v. AWH Corp., 415 F.3d 1303, 1313–24 (Fed. Cir. 2005) (describing various ways courts resolve patent disputes).

36. Press Release, The White House, President Obama Signs America Invents Act, Overhauling the Patent System to Stimulate Economic Growth, and Announces New Steps to Help Entrepreneurs Create Jobs (Sept. 16, 2011), *available at* http://www.whitehouse.gov/the-press-office/2011/09/16/president-obama-signs-america-invents-act-overhauling-patent-system-stim.

37. Leahy-Smith America Invents Act, Publ. L. No. 112–29, § 19(d)(1), 125 Stat. 332 (2011) (codified at 35 U.S.C § 299 (2006 & Supp. V 2012)).

38. In general, the term "troll" refers to nonpracticing entities that own, assert, and enforce patents. Although derided by some scholars as "patent extortionists," others see these nonpracticing entities in a more favorable light. *Compare* James F. McDonough III, Comment, *The Myth of the Patent Troll: An Alternative View of the Function of Patent Dealers in an Idea Economy*, 56 EMORY L.J. 189, 189–90, 193 (2006) (arguing that patent trolls benefit society by acting as intermediaries that provide liquidity, market clearing, and increased efficiency in patent markets), *with* Victoria E. Luxardo, *Towards a Solution to the Problem of Illegitimate Patent Enforcement Practices in the United States: An Equitable Affirmative Defense of "Fair Use" in Patent*, 20 EMORY INT'L L. REV. 791, 793, 800 (arguing that patent trolling is illegitimate and "socially undesirable for the same reasons as extortion").

39. *See* 157 CONG. REC. S5402, S5429–31 (daily ed. Sept. 8, 2011) (statement of Sen. Kyl) (noting that the AIA restrictions simply codify current law as applied everywhere outside of the Eastern District of Texas); *see also* Rudd v. Lux Prods. Corp., No. 09-cv-6957, 2011 WL 148052, at *3 (N.D. Ill. Jan. 12, 2011) (identifying the Eastern District of Texas as the outlier with respect to joinder standards). Largely as a result of these restrictions, post-AIA patent filings in the Eastern District of Texas have decreased dramatically. JAMES C. PISTORINO & SUSAN CRANE, 2011 TRENDS IN PATENT CASE FILINGS: EASTERN DISTRICT OF TEXAS CONTINUES TO LEAD UNTIL AMERICA INVENTS ACT IS SIGNED 11 (2012), *available at* http://www.perkinscoie.com/files/upload/PL_12_03PistorinoArticle.pdf.

individual conduct by accused infringers.[40] Validity, however, concerns only the patent, its prosecution, and the prior art.[41] Because the jury must consider the exact same issues for all infringers, consolidation for the validity phase does not prejudice defendant infringers and is far more efficient than repetitive, separate adjudication. However, the AIA allows for consolidation only if the defendant agrees to waive the § 299 restrictions.[42] Apparently, Congress never considered restricting consolidation to the validity issue alone and requiring separate trials for infringement.[43]

Unfortunately, not just trolls, but all patent owners, are prejudiced by the AIA's joinder and consolidation restrictions. The restrictions especially disadvantage small inventors holding low-value patents because the added costs of relitigating validity deter enforcement.[44] On the other hand, the restrictions may do little to deter trolls that traffic in high-dollar patents because these entities often have sufficient resources and damages incentives to bring multiple suits.[45] Thus, by restricting consolidation for all issues in the patent case, the AIA restrictions cause costly, duplicative litigation of validity in high-value troll suits, while erecting potentially overwhelming enforcement barriers to small inventors. This result is neither efficient nor faithful to the substantive goals of patent law.[46]

Limited by these statutory and practical constraints, joinder and consolidation cannot provide binding judgments of validity good against the world. I turn then to a more powerful procedural device not addressed in the AIA: the Rule 23 class action.

---

40. *See* 35 U.S.C. § 271(a) (2006) (providing that "whoever without authority makes, uses, offers to sell, or sells any patented invention . . . infringes the patent").

41. Markman v. Westview Instruments, Inc., 517 U.S. 370, 373–74 (1996).

42. 35 U.S.C. § 299(c) ("A party that is an accused infringer may waive the limitations set forth in this section with respect to that party.").

43. Some supporting senators expressed skepticism that the restrictions on consolidation were necessary. *E.g.*, 157 CONG. REC. S5402, S5429 (daily ed. Sept. 8, 2011) (statement of Sen. Kyl) ("[T]he bill extends the limit on joinder to also bar consolidation . . . . When this change was first proposed, I was skeptical that it was necessary."). Ultimately, however, it was decided that, "if a court that was barred from joining defendants in one action could instead simply consolidate their cases for trial under Rule 42, section 299's purpose of allowing unrelated patent defendants to insist on being tried separately would be undermined." *Id.*

44. *See* Jean O. Lanjouw & Mark Schankerman, *Protecting Intellectual Property Rights: Are Small Firms Handicapped?*, 47 J.L. & ECON. 45, 48 (2004) (arguing that the ability to mount a credible threat of litigation is necessary to enforce intellectual property rights).

45. *See* Colleen V. Chien, *From Arms Race to Marketplace: The Complex Patent Ecosystem and Its Implications for the Patent System*, 62 HASTINGS L.J. 297, 318 (2010) (noting a higher proportion of "crown jewels" amongst patents owned by nonpracticing entities); *see also* Timo Fischer & Joachim Henkel, Patent Trolls on Markets for Technology—An Empirical Analysis of Trolls' Patent Acquisitions 6, 18 (April 2011) (unpublished manuscript), *available at* http://www2.druid.dk/conferences/viewpaper.php?id=501834&cf=43 (finding that "troll" patents were, on average, of higher quality than patents owned by other entities).

46. S. COMM. ON THE JUDICIARY, 90TH CONG., TO PROMOTE THE PROGRESS OF USEFUL ARTS: REPORT OF THE PRESIDENT'S COMMISSION ON THE PATENT SYSTEM, S. DOC. NO. 5, at 10–11 (1st Sess. 1966) (identifying four major justifications for patent law, including providing incentives to creators to invent by guaranteeing the right to exclude others from using the invention).

II. Class Actions

Although some authorities find patent cases wel suited to class action litigation, few courts have certified an "infringer class."[47] This Part makes the argument for class actions in patent litigation. For the purposes of this analysis, an infringer class is defined as all current and future entities potentially liable for direct, indirect, or contributory infringement of the patent. In addition, because Rule 23(c)(4) allows the court to isolate particular issues for class treatment,[48] this analysis assumes certification only for the isolated issue of validity.

The origins of the class action can be traced to the representative suit of the medieval court of equity.[49] During this period, a group representative could bring suit "on behalf of" the larger group, often entire town or village communities, to enforce a collective right belonging to the group as a whole.[50] A paradigmatic example of a collective right enforceable through representative suits was the right of townspersons to access water sources for drinking and fishing.[51] A plaintiff advocating on behalf of such group rights sought to vindicate not only his own individual rights, but those of the group as a whole.[52] Indeed, the representative's individual rights depend in such cases on membership in the group—if the individual moved out of the town with drinking-water rights, for example, he would lose his individually enjoyed right to the drinking water.[53]

The famous case of *Hansberry v. Lee*[54] marked the transition of the class action from a device for defending shared group rights to a device for aggregating common individual rights.[55] Under the modern revised Rule 23,

---

47. *See* Burton, *supra* note 17, at 292–96, 299–314 (explaining that class actions could be useful for patent litigation but describing how meeting the requirements of Rule 23 can be an obstacle to class certification in these cases). For a strong argument in support of class certification, see Matthew K.K. Sumida, Comment, *Defendant Class Actions and Patent Infringement Litigation*, 58 UCLA L. REV. 843, 844–45 (2011) (stating that defendant class actions are appropriate for patent cases because they reduce litigation costs and ensure consistency of adjudication; however, they are uncommon because Rule 23 and a Supreme Court case are obstacles to certification).

48. FED. R. CIV. P. 23(c)(4).

49. *See* STEPHEN C. YEAZELL, FROM MEDIEVAL GROUP LITIGATION TO THE MODERN CLASS ACTION 4 (1987) (discussing the emergence of representative group litigation in medieval England).

50. *Id.* at 38–39.

51. *See id.* at 180–81 (discussing *Mayor of York v. Pilkington*, 26 Eng. Rep. 180 (Ch. 1737), an early English case revolving around group fishing rights).

52. *See* Robert G. Bone, *Personal and Impersonal Litigative Forms: Reconceiving the History of Adjudicative Representation*, 70 B.U. L. REV. 213, 230 (1990) (reviewing YEAZELL, *supra* note 49) ("The group-in-court model pictures a class as a legal actor which presents itself in court when some of its members seek legal recognition of the group as an entity and of themselves as its representatives . . . ." (internal quotation marks omitted)).

53. *Cf. id.* at 230–31 (stating that the legal rights of the individual representative had to be the same as those belonging to the absent class members).

54. 311 U.S. 32 (1940).

55. *See id.* at 43 (stating that when members of a class are present and entitled to stand in judgment for those who are not, it can be assumed that the procedure protects the parties who are

class certification entails four prerequisites: numerosity, commonality, typicality, and adequacy of representation.[56] In addition, the Rule requires that the litigation concern one of three types of action: suits where joinder is mandatory but impracticable,[57] suits for declaratory or injunctive relief,[58] or suits predominated by questions of law or fact common to all class members.[59] According to some jurists and scholars, the Rule 23 class action and its procedural safeguards are the exclusive means for providing adequate representation of absentees.[60]

A. *Numerosity*

Rule 23 first requires that "the class [be] so numerous that joinder of all members is impracticable."[61] The rationale for this requirement is that, if feasible, litigants should have the opportunity to argue their own case rather than be precluded from their day-in-court right by a representative's suit. However, the Rule provides no hard lower limit on the number of class members required to certify a class.[62]

Despite being an easy hurdle in most class actions, insufficient numerosity is often fatal to infringer-class certifications because frequently only a few potential infringers exist in the market at the time of suit.[63] In one of the few cases supporting infringer certification, the court in *Dale Electronics, Inc. v R.C.L. Electronics, Inc.* sua sponte certified a class of just thirteen infringers.[64] The *Dale* court noted that "a determination of practi-

---

absent but represented, satisfying the requirements of the Due Process and Full Faith and Credit Clauses); *see also* Bone, *supra* note 52, at 225 ("In [*Hansberry v. Lee*], the Supreme Court constitutionalized the interest representation theory by holding that adequate representation of interests could meet constitutional due process requirements for binding nonparties.").

56. FED. R. CIV. P. 23(a).

57. FED. R. CIV. P. 23(b)(1). Because unrelated infringers are not necessary parties under Rule 19, this class action category is not applicable to patent suits. *See supra* Part I.

58. FED. R. CIV. P. 23(b)(2).

59. FED. R. CIV. P. 23(b)(3). This provision would not work in patent suits because defendants would likely opt out of the class in order to avoid preclusion.

60. *See, e.g.*, Tice v. Am. Airlines, Inc., 162 F.3d 966, 972–73 (7th Cir. 1998) (rejecting a "common-law kind of class action" and "de facto class actions"); *see also* Howard M. Erichson, *Informal Aggregation: Procedural and Ethical Implications of Coordination Among Counsel in Related Lawsuits*, 50 DUKE L.J. 381, 458–62 (2000) (arguing that "nonparty preclusion based on informal aggregation should be rejected to avoid circumvention of protections built into formal aggregation mechanisms, especially the class action rule").

61. FED. R. CIV. P. 23(a)(1).

62. Generally, the minimum for numerosity is considered to be about forty class members. Oplchenski v. Parfums Givenchy, Inc., 254 F.R.D. 489, 495 (N.D. Ill. 2008). *But see* Dale Elecs., Inc. v. R.C.L. Elecs., Inc., 53 F.R.D. 531, 534–36 (D.N.H. 1971) (certifying a class of thirteen alleged patent infringers).

63. *E.g.*, Sperberg v. Firestone Tire & Rubber Co., 61 F.R.D. 70, 76 (N.D. Ohio 1973) (finding insufficient numerosity where purported class consisted of twenty-one infringers).

64. 53 F.R.D. 531, at 534–36.

cability should depend upon all the circumstances."[65] The *Dale* court's conclusion reflects an intuition that infringers bringing invalidity suits stand for collective, rather than individual, rights. Even if there are not presently numerous infringers, any member of the public can infringe by using the purportedly patented invention. Parties challenging validity argue that the purportedly patented ideas, in fact, remain in the public domain.[66] Because the right to access the ideas is public, numerosity is manifest.

However, under the joinder restrictions in the AIA, infringers cannot be joined unless they are involved in related acts of infringement, not merely infringing the same patent.[67] With respect to certification of classes of unrelated infringers, these restrictions could be interpreted to mean that numerosity is easily met because joinder is indeed "impracticable," that is, because it is statutorily prohibited. On the other hand, it could mean that numerosity is never met because joinder would not be permissible even if it were feasible. If the latter reading is correct, then certification of infringer classes may violate the America Invents Act, absent waiver by the defendants.

B. Commonality

Rule 23 requires commonality in the class.[68] In the past, this requirement posed a low threshold to certification because any common question of law or fact sufficed to show commonality. The Supreme Court's recent decision in *Wal-Mart Stores, Inc. v. Dukes*,[69] however, recast the commonality inquiry: "What matters to class certification . . . [is] the capacity of a classwide proceeding to generate common *answers* apt to drive the resolution of the litigation. Dissimilarities within the proposed class are what have the potential to impede the generation of common answers."[70]

In short, the Court's emphasis on common answers supports certification in invalidity suits because a declaration indivisibly resolves the contentions of the class. Furthermore, because validity only concerns issues intrinsic to the patent and its prosecution, no "dissimilarities within the proposed class"[71] impede the generation of common answers. In this respect, the inquiry differs greatly from that in *Wal-Mart*. In *Wal-Mart*, individual questions predominated because the alleged acts of discrimination against

---

65. *Id.* at 534 (quoting Demarco v. Edens, 390 F.2d 836, 845 (2d Cir. 1968) (internal quotation marks omitted).

66. *See id.* at 535–36 (recognizing the practical and public benefits of certifying even small classes in patent litigation).

67. Leahy-Smith America Invents Act, Pub. L. No. 112-29, § 299(a)–(b), 125 Stat. 284, 332–33 (2011) (codified at 35 U.S.C. § 299(a)–(b) (2006 & Supp V 2012)).

68. FED. R. CIV. P. 23(a)(2).

69. 131 S. Ct. 2541 (2011).

70. *Id.* at 2551 (quoting Richard A. Nagareda, *Class Certification in the Age of Aggregate Proof*, 84 N.Y.U. L. REV. 97, 132 (2009) (internal quotation omitted).

71. *Id.* (quoting Nagareda, *supra* note 70, at 132).

each class plaintiff occurred in isolated incidents by independent regional managers.[72] Each of these incidents required different, particularized evidentiary proof. By contrast, in validity inquiries, no such individual questions exist with respect to each infringer, and only evidence related to the patent is relevant to the inquiry.[73]

C. *Typicality*

Rule 23 requires typicality and adequacy of the representative with respect to the class.[74] The gist of these requirements is that the class representative should be similarly situated and capable of vigorously pursuing the mutual interests of the class members.[75] Typicality ensures that "the class representatives are sufficiently *similar* to the rest of the class—in terms of their legal claims, factual circumstances, and stake in the litigation—so that certifying those individuals to represent the class will be fair to the rest of the proposed class."[76] In a patent validity class action, all infringers are identically situated.[77] The class representative's legal claims and factual circumstances perfectly align with the class because all oppose the same patent.

However, a court considering certification of an infringer class should carefully examine whether the class representative has sufficient interest in obtaining a declaration of patent invalidity to ensure a vigorous challenge. Absent this inquiry, patentees could game the system by bringing suit against a weak, disinterested infringer and certifying that infringer as class representative. One procedural hurdle that ensures the interestedness of the representative is the cases-and-controversies limitation in Article III. A party does not have constitutional standing to argue the issue of validity if it is not imminently subject to suit for infringement.[78] Beyond this, typicality probably requires the court to consider the magnitude of the purported infringer's

72. *Id.* at 2552.

73. *See* Microsoft Corp. v. i4i Ltd. P'ship, 131 S.Ct. 2238, 2242 (2011) (holding that invalidity defenses to infringement suits must be proven by clear and convincing evidence of patent invalidity).

74. FED. R. CIV. P. 23(a)(3)–(4). These requirements, along with commonality, "tend to merge." Gen. Tel. Co. v. Falcon, 457 U.S. 147, 157 n.13 (1982).

75. *See, e.g.*, *Falcon*, 457 U.S. at 157 n.13 (noting the importance of a finding that the plaintiff's claim and the class claim are "so interrelated" that the interests of class members will be adequately represented); Schlesinger v. Reservists Comm. to Stop the War, 418 U.S. 208, 216 (1974) (holding that it is "essential" that a plaintiff "possess the same interest and suffer the same injury shared by all members of the class he represents").

76. *In re* Schering Plough Corp. ERISA Litig., 589 F.3d 585, 597 (3d Cir. 2009).

77. Burton, *supra* note 17, at 301–02.

78. *See* Sandisk Corp. v. STMicroelectronics, Inc., 480 F.3d 1372, 1381 (Fed. Cir. 2007) (holding that an Article III controversy arises "where a patentee asserts rights under a patent based on certain identified ongoing or planned activity of another party, and where that party contends that it has the right to engage in the accused activity without [a] license").

business interest in invalidating the patent.[79] If the patented technology is only marginally beneficial to the defendant, then that defendant is probably not a good representative for the public interest.

On the other hand, the problem of ensuring a sufficiently interested representative may resolve itself because the patent owner has an interest in suing major infringers and in fortifying any class action judgment against potential collateral attacks. The patentee is likely to sue the infringer with the largest available damages reward, which is likely to correlate strongly with the infringer's business interest. Furthermore, in order to protect the judgment from collateral attack on the grounds that the representative was inadequate,[80] the patentee has a strong incentive to ensure that the representative infringer can and will mount a vigorous challenge. Thus, the patentee may be inclined to bring suit against a strong defendant that unquestionably has high stakes in the patent's validity.

D. *Adequacy of Representation*

For adequacy, courts ask whether any conflicts of interest divide the representative from other class members and whether class counsel has the resources, ability, and experience to vigorously litigate.[81]

For infringer classes, the first impediment to adequacy involves potential conflicts in the class over claims construction. Because litigants argue for constructions that bolster their individual cases, different infringers may have divergent views on proper claims interpretation.[82] One might say these variances amount only to disagreement about trial strategy, since all infringers ultimately seek the same equivalent declaration of patent invalidity. On this level, such strategic disagreements are endemic to the class action and underscore its day-in-court tradeoffs. But more importantly, if the patent survives the validity challenge, the same constructions will apply against the infringers in the individualized proceedings to follow, arguably creating class conflicts.[83]

These conflicts can be overcome because claims construction is a matter of law.[84] In *Markman v. Westview Instruments, Inc.*,[85] the Supreme Court

---

79. *See* Burton, *supra* note 17, at 302 ("[I]f there is enough at stake, [the defendant class] will vigorously defend *themselves*, and will thus defend the class . . . .").

80. *See, e.g.*, Van Gemert v. Boeing Co., 590 F.2d 433, 440 n.15 (2d Cir.1978), *aff'd*, 444 U.S. 472 (1980) ("Judgment in a class action is not secure from collateral attack unless the absentees were adequately and vigorously represented.").

81. Wetzel v. Liberty Mut. Ins. Co., 508 F.2d 239, 247 (3d Cir. 1975).

82. *See* Burton, *supra* note 17, at 302 ("Adequacy of the class representative ought not to be difficult for the patentee to show . . . as long as there are truly no intra-class conflicts of interest . . . ." (internal citations omitted)).

83. *Cf.* Markman v. Westview Instruments, Inc., 517 U.S. 370, 390–91 (1996) (arguing that the importance of uniformity in patent cases makes it necessary for judges to interpret the terminology included in patents as a matter of law).

84. *Id.*

diminished the importance of evidence in claims construction, suggesting that the inquiry is susceptible of an objectively correct outcome.[86] Thus, as a matter of law, a court can provide a proper interpretation of what a skilled artisan with knowledge of the prior art would understand the patent to mean. Furthermore, "treating interpretive issues as purely legal . . . promote[s] . . . intrajurisdictional certainty" because the legal character of the inquiry mimics the "principles of issue preclusion [that] ordinarily foster uniformity."[87] In addition, even if certification is granted, current infringers remain free to intervene and supply the court with additional views on interpretation.[88] Because *Markman* contemplates objectively proper claims constructions as a matter of law, possible disagreements about proper interpretation should not impede certification.

In addition to identifying conflicts, the court must ensure that the representative's "attorney is capable of prosecuting the instant claim with some degree of expertise."[89] In some cases, the court may need to appoint able class counsel with experience in patent litigation. To compensate the representative and class counsel, the court might also establish a reward if the patent is invalidated. A reward in this scenario operates on a restitution theory: because absentee infringers benefit from the invalidity judgment, they should have to compensate the representative and class counsel. However, the difficulty is in calculating a proper amount for the reward that incentivizes vigorous challenges, maximizes social utility, and is economically rational for absentee "beneficiaries" such that the restitution theory remains viable.[90]

E. *Rule 23(b)(2)*

Rule 23(b)(2) applies to class actions for declaratory or injunctive relief.[91] "[C]ertification under [this rule] is appropriate only if members of

85. 517 U.S. 370 (1996).

86. *See id.* at 388 (quoting 2 WILLIAM C. ROBINSON, THE LAW OF PATENTS FOR USEFUL INVENTIONS § 732, 481–83 (1890)) ("A patent is a legal instrument, to be construed, like other legal instruments, according to its tenor . . . .").

87. *Id.* at 391.

88. The prospect of future individualized litigation suggests the court should order notice for identifiable current infringers that may wish to intervene on claims construction. Although Rule 23(b)(2) does not require notice, the court may order it under Rule 23(c)(2)(A). FED. R. CIV. P. 23(c)(2)(a). Perhaps notice in the Federal Register or an online database would suffice, assuming interested parties could monitor these resources with minimal difficulty. Intervenors could supply the court with more information about prior art and argue for preferred claims constructions.

89. Richard Marcus, *Revising Judicial Gatekeeping of Aggregation: Scrutinizing the Merits on Class Certification*, 79 GEO. WASH. L. REV. 324, 343 n.113 (2011).

90. Perhaps the reward could be pegged to the patent's value as measured by certain intrinsic characteristics, such as the number of prior art citations. *See* John R. Allison et al., *Valuable Patents*, 92 GEO. L.J. 435, 436–37 (2004) (profiling patent value based on a number of empirical factors).

91. Rule 23(B)(2) class actions are allowed when "the party opposing the class has acted or refused to act on grounds that apply generally to the class, so that final injunctive relief or

the proposed class would benefit from the injunctive [or declaratory] relief they request."[92] Furthermore, certification should be granted only where the relief sought would be "both reasonably necessary and appropriate" were the class to succeed on the merits.[93] These hurdles appear to be easily cleared in patent validity litigation because each infringer would benefit and the declaration of invalidity would be both necessary and appropriate if the infringer succeeds on the merits.

When problems arise under Rule 23(b)(2), they normally concern the attachment of individualized claims for relief.[94] In the case of invalidity suits, however, only a declaration that would provide relief evenly to all in the class is sought. Because the declaration is binary (either the patent is invalidated or not, with no middle ground), no discrepancies exist between the relief sought by the representative and the relief sought by others in the class. In addition, no money damages attach to claims for patent invalidity. Although some infringers may bring counterclaims for damages or other individualized relief during the course of the litigation, these individual issues are not related to patent validity and can be adjudicated during the later infringement phases.[95]

Some courts categorically refuse to certify defendant classes in (b)(2) class actions because the rule's express reference to "the party opposing the class [acting] or [refusing] to act" arguably suggests that only plaintiff classes can be certified.[96] In addition, generally speaking, defendant class actions exacerbate the due process concerns implicated by class actions.[97] Infringer classes have typically been treated as defendant classes because the infringers are defendants on the issue of infringement. However, infringers bringing claims or counterclaims for declaration of patent invalidity are plaintiffs with respect to those claims. Therefore, the language of Rule

---

corresponding declaratory relief is appropriate respecting the class as a whole." FED. R. CIV. P. 23(b)(2).

92. *In re* Monumental Life Ins. Co., 365 F.3d 408, 416 (5th Cir. 2004), *cert. denied*, 543 U.S. 870 (2004).

93. Robinson v. Metro-North Commuter R.R. Co., 267 F.3d 147, 164 (2d Cir. 2001).

94. For example, the fact that the class action also sought damages for backpay, rather than merely an injunction, was a significant factor in the Court's *Wal-Mart* certification denial. *See* Wal-Mart Stores, Inc. v. Dukes, 131 S.Ct. 2541, 2557 (2011) (holding that claims for backpay may not be certified under Rule 23(b)(2), "at least where . . . the monetary relief is not incidental to the injunctive or declaratory relief").

95. *See* Burton, *supra* note 17, at 301 (noting that "[t]he invalidity determination (almost always counterclaimed by the alleged infringer), focused as it is on the patent itself, will depend on common issues and facts" and therefore "should be strong in the patent context").

96. *E.g.*, Tilley v. TJX Cos., 345 F.3d 34, 39–40 (1st Cir. 2003) ("In cases involving garden-variety defendant classes, there will be no single act or refusal to act on the part of the plaintiff . . . that makes injunctive or declaratory relief appropriate. Rather, it will be the defendants—the members of the putative class—who allegedly have acted in the same tortious or unlawful way (here, by selling [copyright-]infringing articles).").

97. *See* Thillens, Inc. v. Cmty. Currency Exch. Ass'n. of Ill., 97 F.R.D. 668, 674 (N.D. Ill. 1983) ("The crux of the distinction is: the unnamed plaintiff stands to gain while the unnamed defendant stands to lose.").

23(b)(2) permits certification of infringer classes on the isolated issue of validity.[98] Moreover, because the infringer stands for public rights, enhanced due process issues are not a major concern.

Finally, any discussion of class actions must acknowledge the reality that a majority of suits settle before judgment. What should be the effect of such a settlement on non-party infringers? In *Lear Inc. v. Adkins*,[99] the Supreme Court found that barring a licensee's validity challenge "undermine[s] the strong federal policy favoring the full and free use of ideas in the public domain."[100] Because licensees are often the only parties with incentives to challenge patent validity, if they were barred, "the public [would] continually be required to pay tribute to would-be monopolists without need or justification."[101] Thus, to enable the defense of the public right in a court of law, the Court held that a licensee can challenge validity even if contractually bound to observe the patent.[102] Since even the settling licensee is free to challenge patent validity, nonparties to the license should remain free to do so as well.

Similarly, in *Foster v. Hallco Manufacturing Co.*,[103] the parties entered into a consent decree that stipulated the validity of the patents at issue.[104] Later, the licensee sought to invalidate the patents, but the Federal Circuit held that normal principles of res judicata barred subsequent challenges to the decree by that licensee.[105] The court explained the distinction with *Lear*:

> When a consent decree is to be given *res judicata* effect, litigants are encouraged to litigate the issue of validity rather than foreclosing themselves by a consent decree. If they were given a second chance to litigate the issue of validity, alleged infringers might well accept a license under a consent decree and forego an attack on validity . . . . By giving *res judicata* effect to consent decrees this court protects the public interest in that an alleged infringer is deprived of a judicial device which could be used to postpone and delay a final adjudication.[106]

---

98. On the other hand, issues related to enforcement probably should remain as individualized affirmative defenses that continue to be available to absentee infringers in subsequent litigation. Many of these defenses, such as patent misuse and estoppel, contain elements specific to particular infringers, rather than focusing only on the patent and the patentee's conduct. In addition, defenses such as misrepresentation should always be open to relitigation if a sufficient showing of the patentee's prior fraud is made. *But see* Sumida, *supra* note 47, at 845 (arguing that enforcement issues should be included in the class action litigation).

99. 395 U.S. 653 (1969).

100. *Id.* at 674.

101. *Id.* at 670.

102. *Id.* at 674.

103. 947 F.2d 469 (Fed. Cir. 1991).

104. *Id.* at 472.

105. *Id.* at 475.

106. *Id.* at 476–77 (quoting Schlegel Mfg. Co. v. USM Corp., 525 F.2d 775, 781 (6th Cir. 1975)).

The object, then, of precluding challenges after validity stipulations is the same as that for allowing licensees to challenge validity: to encourage vigorous litigation of validity on behalf of the public. Only when validity is vigorously litigated is the public right vindicated. Thus, if the representative enters into a consent decree on validity, res judicata should not bind future infringers from challenging validity.

Like the townsperson suing on behalf of the town's right to access waters, an infringer challenging patent validity stands for a public right. If class actions were common practice, a winning verdict for the patentee would conclusively establish the patent's validity and make it unassailable to challenge. However, the procedural requirements of the modern-day rule raise impediments to certification. Because close adherence to Rule 23 may prevent class certification, I turn now to a more flexible theory of aggregate litigation: virtual representation.

III. Virtual Representation

The theory of virtual representation rests on roughly the same theoretical grounds as the medieval representative suit.[107] The fundamental insight of both is that, in certain circumstances, a litigant standing for group rights can fully vindicate the right and preclude successive challenges by other group members. This Part analyzes whether virtual representation theory remains viable in patent validity litigation.

In *Taylor v. Sturgell*, the Supreme Court seemingly rejected virtual representation theory and recognized the formal requirements of Rule 23 as essential safeguards of due process in aggregate litigation.[108] In particular, the Court disapproved of the manner in which virtual representation "authorize[d] preclusion based on identity of interests and some kind of relationship between parties and nonparties, shorn of the procedural protections prescribed in . . . Rule 23."[109] Because the prior suit did not observe these procedural protections, the plaintiff Taylor could bring an identical request for information under the Freedom of Information Act.[110] In reversing the lower court's finding of preclusion by virtual representation, the Court derided virtual representation as, "in effect, a common-law kind of class action . . . that allowed courts to create *de facto* class actions at will."[111]

But courts have not always been so hostile to the theory of virtual representation in public rights cases. Applying virtual representation theory

---

107. *See* Bone, *supra* note 52, at 275–77 (explaining how the doctrine of virtual representation applied to representative suits).

108. *Taylor v. Sturgell*, 553 U.S. 880, 898–901 (2008).

109. *Id.* at 901.

110. *Id.* at 885, 890.

111. *Id.* at 901 (quoting Tice v. Am. Airlines, Inc., 162 F.3d 966, 972–73 (7th Cir. 1998)) (internal quotation marks omitted).

in a case prior to *Taylor*,[112] the Eighth Circuit held that a few African-American residents were barred from bringing a Voting Rights Act challenge to the aldermanic district boundaries of St. Louis by a prior summary judgment for defendant against earlier challengers on the same issue.[113] Because the plaintiffs alleged "that the strength of the black vote in general ha[d] been diluted," rather than that their individual voting rights had been violated, they possessed "a special commonality of interests" with the prior plaintiffs.[114] Viewed differently, both the prior and subsequent plaintiffs sought to vindicate the same group, rather than individual, rights.

The notion of group rights aligns closely with public rights cases. The *Tyus* court noted that public law cases raise certain policy concerns favoring preclusion:

> [I]n public law cases, the number of plaintiffs with standing is potentially limitless. If parties were allowed to continually raise issues already decided, public law claims "would assume immortality." . . . [I]n the public law context, if the plaintiff wins, by definition everyone benefits. Holding preclusion inapplicable in this context would encourage fence-sitting . . . .[115]

Based on this analysis, the judgment in the first suit precluded subsequent challengers from relitigating the district-lines issue again because the prior plaintiffs fully vindicated the exact same public right.[116]

In *Richards v. Jefferson County*,[117] the Supreme Court seemed to endorse a distinction between private and public rights. In that case, the Court identified two types of taxpayer challenges: first, those in which the taxpayer complained about misuse of public funds or other public action that had only an indirect impact on his interests; and second, those in which the taxpayer challenged the state's attempt to levy on his personal funds.[118] The former scenario implicated only public rights and therefore enabled "wide latitude to establish procedures not only to limit the number of judicial proceedings . . . but also to determine whether to accord a taxpayer any standing at all."[119] On the other hand, the latter scenario demanded the full panoply of due process, day-in-court rights because individual rights were directly at stake.[120] But in *Taylor*, the Supreme Court repudiated this distinction as applied in that case.[121]

---

112. Tyus v. Schoemehl, 93 F.3d 449 (8th Cir. 1996), *abrogated by* Taylor v. Sturgell, 553 U.S. 880 (2008).
113. *Id.* at 458.
114. *Id.* at 457.
115. *Id.* at 456 (citation omitted).
116. *Id.* at 458.
117. 517 U.S. 793 (1996).
118. *Id.* at 803.
119. *Id.*
120. *Id.* at 805.
121. *Taylor*, 553 U.S. at 902–04.

So is virtual representation theory dead after *Taylor*? In fact, the Court's logic may have left the door open for virtual representation in patent suits. First, the FOIA request at issue in *Taylor*, although vindicating a public interest, provided a remedy only to the individual plaintiff rather than benefiting the public at large (to wit, only the plaintiff received the requested documents).[122] Second, the Court found no evidence of "abusive FOIA suits" in the circuits that had rejected virtual representation theory and thus no ill consequence from the perverse "fence-sitting" incentives mentioned in *Tyus*.[123] If a party could show such abusive or socially deleterious relitigation, a stronger argument could be made for virtual representation. Finally, as the Court explained, there exist "instances in which the first judgment foreclose[s] successive litigation by other plaintiffs because . . . [the suit can] be brought only on behalf of the public at large," but such instances are confined to "special statutory schemes that expressly limit subsequent suits."[124] This latter scenario is known as the "statutory scheme" exception to the nonparty preclusion rule.

The "statutory scheme" exception presents the biggest opening for virtual representation in public rights litigation, such as that concerning patent invalidity. If individual due process rights are indeed the barrier to virtual representation preclusion imposed by the courts, then that same barrier should restrict the legislature from imposing preclusion on absentees as well. In other words, neither judicial nor legislative fiat can abridge individual due process rights, if such rights in fact exist. Therefore, what justifies preclusion in a statutory scheme must be something more than the mere fact that Congress ordains it. Might this something more be the idea that certain public law situations do not truly implicate individual due process rights?[125]

As discussed above, patent validity challenges invoke the public's right to access ideas in the public domain. Unlike a FOIA request, a declaration of patent invalidity both vindicates a public interest *and* provides relief to the public at large. Furthermore, contrary to *Taylor*'s undemonstrated showing of abusive FOIA suits, repetitive proceedings on patent validity impose significant costs on the patent system, the courts, and the public.[126] Finally, because patent validity challenges vindicate the public's right to access ideas in the public domain, "there is no, or at best a very weak, process-based right to participate [because] the lawsuit and the judgment do not single out any

122. *Id.* at 902.

123. *Id.* at 904.

124. *Id.* at 903 (internal quotation marks omitted). According to the *Taylor* Court, examples of such schemes include bankruptcy and probate proceedings, as well as quo warranto actions. *Id.* at 903 n.12.

125. *See* Bone, *supra* note 8, at 610–11 ("[I]ndividualized participation is not required when legislation acts on a general class and does not focus on any individual personally.").

126. *See infra* Part IV.

person for individual treatment."[127] Therefore, even if Congress refuses to mandate a "special statutory scheme," courts arguably have the capacity to preclude successive validity challenges under a theory of virtual representation that survives *Taylor*.

However, even a virtual representation theory requires some notion of adequate representation to ensure that the representative challenger is capable and interested enough to vigorously challenge validity on behalf of the public.[128] As in the class action typicality analysis, the cases-and-controversies limitation in Article III helps ensure that only interested parties have standing to bring suit.[129] Still, the court should inquire whether the representative is sufficiently invested in obtaining validity that it can be expected to pursue the judgment vigorously. Similarly, the court must ensure that the representative retains capable counsel and has sufficient resources to litigate. These adequacy requirements are the touchstone of any theory of representative litigation. If the representative is incapable of vigorous challenge, the litigation will not properly vindicate the public right and therefore cannot justifiably preclude subsequent challenges.

However, when an adequate representative vigorously litigates validity, the public right is fully vindicated and absentees are justifiably barred from challenging the patent again. I turn now to the policy question of whether precluding successive challenges to patent validity furthers the goals of patent law and wise judicial administration.

IV. Is Preclusion on Validity Good Public Policy?

This Part argues that preclusion of successive challenges is good policy because it encourages challenges to bad patents, increases patent values, improves the notice function of patents, conserves judicial resources, and provides fairness for the patentee.

A. Challenge "Bad Patents"

The problem of uncertain patent validity begins at the overburdened Patent and Trademark Office (PTO). In an average year, the PTO reviews 350,000 patent applications and issues about 180,000 patents, both figures that have steadily risen in recent years.[130] Because of the overwhelming deluge of patent applications, the understaffed PTO is unable to adequately

---

127. Bone, *supra* note 8, at 610.

128. *See* Taylor v. Blakey, 490 F.3d 965, 973–75 (D.C. Cir. 2007), *vacated and remanded by* Taylor v. Sturgell, 553 U.S. 880 (2008) (evaluating adequacy of representation while applying the virtual representation doctrine).

129. Simon v. E. Ky. Welfare Rights Org., 426 U.S. 26, 39 (1976) ("The necessity that the plaintiff who seeks to invoke judicial power stand to profit in some personal interest remains an Art. III requirement.").

130. Kimberly A. Moore, *Worthless Patents*, 20 BERKELEY TECH. L.J. 1521, 1521 & nn.1–2 (2005).

review each application that comes before it.[131] Thus, numerous "bad patents" that have been issued may actually be invalid.[132] Because of the PTO's deficiencies, "the courts are an integral part of the patent system and serve as an institutional mechanism not only for protecting and enforcing valid patent rights, but also for maintaining the integrity of the process used to grant those rights."[133] Underscoring just how significant the problem is and what an important role courts can play, one study found that over 46% of patents litigated to a judgment were ultimately invalidated.[134]

The problem is that, under current law, neither patent owners nor accused infringers have incentives to litigate patent validity all the way to judgment.[135] Even if the patentee prevails, he gains little because the issue must be relitigated in future suits against different infringers. On the other hand, for defendant infringers, pursuit of patent invalidation suffers from a free-rider problem.[136] Because a patent declared invalid in any one action becomes invalid against the world, infringers can wait and hope that another infringer will invalidate the patent for everyone's benefit. Furthermore, many infringers find it economically preferable to simply pay for a license even if they believe the patent is probably invalid.[137] As a result, many bad patents remain unchallenged.

Preclusion of future challenges to patent validity seriously mitigates these problems. For patent owners, the promise of absentee preclusion provides an incentive to litigate validity. Furthermore, because the judgment may be vulnerable to collateral attack if the representative is inadequate, the

---

131. *See* Arti Rai, *Addressing the Patent Gold Rush: The Role of Deference to PTO Patent Denials*, 2 WASH. U. J.L. & POL'Y 199, 218 (2000) (proposing a simple reform that increases the number and quality of patent examiners at the PTO).

132. *See* Jennifer K. Gregory, *The Troll Next Door*, 6 J. MARSHALL REV. INTELL. PROP. L. 292, 295 (2007) ("[T]here are too many patents of questionable quality that have been granted in the recent past, giving the trolls an opportunity to acquire and enforce vague patents against countless companies.").

133. Jay P. Kesan & Gwendolyn G. Ball, *How Are Patent Cases Resolved? An Empirical Examination of the Adjudication and Settlement of Patent Disputes*, 84 WASH. U. L. REV. 237, 242 (2006).

134. John R. Allison & Mark A. Lemley, *Empirical Evidence on the Validity of Litigated Patents*, 26 AIPLA Q.J. 185, 205 (1998).

135. *See* Kesan & Ball, *supra* note 132, at 243 ("[W]hen it becomes obvious that a patent is very likely to be invalidated, it is in the best interests of the patent holder to offer a cheap license to keep the patent rights intact, and it is in the best interests of the defendant to accept such an offer rather than incur further significant legal costs.").

136. *Id.* at 244.

137. *See* Blonder-Tongue Labs., Inc. v. Univ. of Ill. Found., 402 U.S. 313, 342 (1971) ("Even in cases where [the infringer] feels strongly that the patent would ultimately be held invalid, when he considers the hundreds of thousands of dollars in complex cases that could be involved in defending a suit, he may conclude that the best course of action is to settle for less to get rid of the problem." (quoting *Patent Law Revision: Hearing on S. 2, S. 1042, S. 1377, and S. 1691 Before the Subcomm. on Patents, Trademarks, and Copyrights of the S. Comm. on the Judiciary*, 90th Cong. 103 (1968) (statement of James W. Birkenstock, Vice President, IBM Corp.; Member, President's Comm'n on the Patent Sys.))).

infringer is motivated to bring suit against an infringer capable of mounting a vigorous challenge. On the infringer side, the free-rider problem disappears because all infringers suffer from an adverse judgment on validity. Thus, absentee infringers are encouraged to intervene in the first action. Furthermore, the involvement of additional infringers provides economies of scale that help overcome the negative-value problem in separate litigation. Thus, in a preclusion regime, bad patents are more likely to be challenged, litigated to judgment, and invalidated. In addition, because patent owners know that assertion of a bad patent is more likely to lead to invalidation, they are discouraged from asserting these patents in the first place.

B. *Prevent Patent Devaluation*

The uncertain validity of issued patents frustrates the valuation of patents, which in turn weakens the monopoly incentives at the core of patent law. Efficient market theory holds that, where investors have full and accurate information about a commodity, the market price of that commodity accurately reflects its value.[138] Although patents are not traded on securities markets, they are transferable commodities subject to similar forces.[139] Uncertainty about patent validity, not to mention claims construction, is antithetical to full and accurate information. Therefore, such uncertainty undermines the establishment of stable and efficient patent markets. Furthermore, uncertain validity heightens the risk associated with ownership of any patent. Because increased risk associated with an investment diminishes its value, questions about a patent's validity impair the patent's value.[140]

Establishment of patent validity through preclusion erases this uncertainty. In addition, litigating to a judgment ensures the court an opportunity to construe the claims for the benefit of all interested parties. In turn, markets are better able to evaluate patent value, which allows for more robust and informed trading. Furthermore, because uncertainties are removed, the overall value of patents should rise. This effect bolsters the innovation incentives at the core of patent law.

C. *Notice*

In addition to establishing incentives for inventors, patent law promotes innovating by encouraging inventors to share their inventions with the public. This notice serves twin purposes: first, it allows innovators to build on

---

138. *See* Eugene F. Fama, *Efficient Capital Markets: A Review of Theory and Empirical Work*, 25 J. FIN. 383, 387 (1970) (positing that with free exchange of information and perfect agreement on information's market effects, market prices "'fully reflect[]' all available information").

139. *See* Amy L. Landers, *Liquid Patents*, 84 DENV. U. L. REV. 199, 207–11 (2006) (describing trading practices of various actors in the patent system).

140. Emily Michiko Morris, Res *or Rules? Patents and the (Uncertain) Rules of the Game*, 18 MICH. TELECOMM. & TECH. L. REV. 481, 491 (2012).

inventions already patented; and second, it facilitates the "clearance" process wherein innovators identify prior art and compensate the owners of patented inventions that are necessary to new innovation. Ideally, the system should reliably inform innovators about the scope and validity of patented inventions, and it should be easy to search. Unfortunately, uncertain patent validity inhibits this provision of notice to the public.

Even when an innovator identifies potentially implicated patents, the validity and meaning of the patents' claims remain uncertain. As a result, clearance suffers. Innovation may stall for extended periods while information about prior patents is painstakingly sought.[141] Alternatively, innovators may ignore patents altogether and proceed blithely to infringe.[142] Furthermore, unscrupulous patent owners may conceal their patents until the most lucrative opportunity to sue, which will often be once the infringer has largely committed itself to development of the infringing device.[143]

Establishment of patent validity enhances notice to innovators. First, it informs innovators about which of the millions of existing patents is certainly valid. This expedites the innovators' negotiation of licensing arrangements with patent holders and provides clear guidance on the patents that most need to be designed around. Second, it informs the public about the "metes and bounds" of the patent because the court must interpret the claims during litigation. Third, because legal databases are familiar and relatively simple to use, finding the validated patents becomes easier. Thus, final binding judgments of validity improve the public-notice function of patent law.[144]

D. *Conserve Judicial Resources*

Invalidity is an expensive and time-consuming phase of patent litigation. In one study, the average number of days to a ruling on validity was 874 and the average number of filed documents was 221.[145] Even cases that settled took an average of 392 days.[146] During that period, a court might

---

141. *See* JAMES BESSEN & MICHAEL J. MEURER, PATENT FAILURE: HOW JUDGES, BUREAUCRATS, AND LAWYERS PUT INNOVATORS AT RISK 50 (2009) (hypothesizing that the complexity and uncertainty of patents may lead businesses to conclude that the costs of clearance may outweigh the benefits).

142. *See* Mark A. Lemley, *Ignoring Patents*, 2008 MICH. ST. L. REV. 19, 21 (2008) (arguing that researchers and companies in component industries often simply ignore patents rather than go through the hassle of clearance).

143. *See* Chien, *supra* note 45, at 319 ("[S]ecrecy serves a 'troll' business model, in which patentees wait until companies are already practicing an invention to 'surprise' them with a suit."). The doctrine of laches limits this practice to some extent. Laches has two elements: (1) "the patentee's delay in bringing suit was unreasonable and inexcusable," and (2) "the alleged infringer suffered material prejudice attributable to the delay." A.C. Aukerman Co. v. R.L. Chaides Constr. Co., 960 F.2d 1020, 1028 (Fed. Cir. 1992).

144. *See* BESSEN & MEURER, *supra* note 140, at 7 (explaining the importance of public notice in protecting the right to exclude in property law generally, and noting problems with this function in patent law).

145. Kesan & Ball, *supra* note 133, at 310.

146. *Id.* at 288.

issue dozens of orders and host several pretrial hearings.[147] This requires a significant expenditure of judicial resources, not to mention litigation costs to the parties. Requiring a particular patent to pass through this crucible multiple times is a waste of judicial resources. Therefore, preclusion of successive challenges saves resources of both the courts and the litigants.

Perhaps this argument tries to have it both ways: can preclusion of future challenges both encourage parties to litigate validity to a judgment, and conserve judicial resources? Indeed, some patents will be litigated under a preclusion regime that would not otherwise have been litigated. Thus, judicial resources will be consumed that otherwise would not have been. On the other hand, a prerequisite to litigation is uncertainty. When the parties have similar expectations about the likely outcome of litigation, they are likely to settle.[148] It follows that certainty about patent validity and claims construction promotes settlement of future cases. Thus, although some extra validity litigation is likely to occur in a preclusion regime, it is offset by reduced litigation in future cases concerning prior-litigated patents.

E. Fairness to the Patentee

After the patentee has defended the patent against a vigorous challenge by a capable adversary, he should not have to do so again. As the Supreme Court found in *Blonder-Tongue*, "[p]ermitting repeated litigation of the same issue as long as the supply of unrelated defendants holds out reflects either the aura of the gaming table or a lack of discipline and of disinterestedness on the part of the [courts]."[149] A second jury is not bound to follow the findings of a prior jury on validity. Thus, statistically speaking, a patentee is likely to lose if litigation is repeated several times. And if the patentee loses just one time, the patent is invalidated forever.[150] This is unfair to the patentee and especially damaging to small inventors.

Even if infringers are unlikely to spend resources challenging a patent that has already survived numerous challenges, the threat of repeated challenges nevertheless damages the patentee's bargaining leverage. Perhaps limiting patentee bargaining leverage is needed to contain trolls, which some

---

147. *See, e.g.*, Docket List, 01 Communique Lab., Inc. v. Citrix Sys., Inc., No. 6CV00253 (N.D. Ohio Feb. 1, 2006) (containing 268 docket proceedings as of Oct. 20, 2012).

148. *See* Bruce L. Hay & Kathryn E. Spier, *Litigation and Settlement* 5 (Harvard Law Sch. John M. Olin Ctr. for Law, Econ. & Bus., Discussion Paper No. 218, 1997) (stating that divergent trial expectations reduce the likelihood of settlement).

149. Blonder-Tongue Labs., Inc. v. Univ. of Ill. Found., 402 U.S. 313, 329 (1971) (citations omitted) (internal quotation marks omitted).

150. If the patentee wins, *inter partes* review remains as a fail-safe mechanism to protect the public against "would-be monopolists" holding a rightfully invalid patent. Thus, if a challenger can raise a "new and substantial" issue of invalidity and demonstrate a reasonable likelihood of success, the patent can still be invalidated. *See* 35 U.S.C. § 311 (2006) (establishing the procedures for requesting *inter partes* re-examination of a Patent Office decision).

commentators see as pariahs of the patent system.[151] Indeed, this was the thinking behind the joinder and consolidation restrictions in the AIA.[152] But separated validity proceedings probably little deter trolls that traffic in mostly valid, high-value patents. In addition, if more small inventors sell their patents to nonpracticing entities in order to avoid higher enforcement costs, the AIA's restrictions may elevate the role of trolls in patent markets. Finally, other mechanisms for deterring trolls are better suited to the task: require separate litigations for infringement in keeping with traditional understandings of prejudice under Rule 42; sanction frivolous suits under Rule 11; and allow vigorous defenses of unenforceability where the patentee engages in unseemly conduct.

Is repetitive litigation of validity needed to protect the public interest? As the Supreme Court has noted, "The far-reaching social and economic consequences of a patent . . . give the public a paramount interest in seeing that patent monopolies . . . are kept within their legitimate scope."[153] Perhaps this "paramount interest" justifies repetitive litigation of validity in order to be absolutely certain that the patent is valid. But as the *Blonder* Court asked years ago: "[W]hat reason [is there] to expect that a second district judge or court of appeals would be able to decide the issue more accurately" than the first?[154] Indeed, the second (or third, or fourth, or tenth) court that hears the case may decide the case wrongly as well. And although infringers retain recourse against a falsely upheld patent in the form of *inter partes* review, the defeated patentee is plain out of luck.[155]

Nevertheless, any changes to the finality of a validity judgment may need to come from Congress, not the courts. In this light, what distinguishes a statutory scheme from virtual representation theory is that Congress, as elected representatives of the people, is better situated to delineate the scope of individuals' enjoyment of the public right. In patent law, Congress has determined that separate infringers should not be joined or consolidated for any purpose absent their consent.[156] Whether these restrictions are wise or not, it may exceed the proper judicial role for courts to define the public right in different terms than Congress. If this is correct, then Congress should act to amend the America Invents Act. Driven by demagoguery against trolls, its current restrictions disrupt judicial economy, harm patentees unfairly, and undermine substantive goals of patent law.

---

151. *See supra* note 38 and accompanying text.

152. *See supra* notes 36–46 and accompanying text.

153. *See* Precision Instrument Mfg. Co. v. Auto. Maint. Mach. Co., 324 U.S. 806, 816 (1945).

154. *Blonder-Tongue*, 402 U.S. at 331–32.

155. *See supra* note 149.

156. *See supra* notes 36–46 and accompanying text.

Conclusion

The procedural rules for challenging patent validity should reflect that the right to access ideas in the public domain is a public right fully vindicated by a vigorous challenge. If an adequate challenger vigorously and capably argues for a patent's invalidity, the public's right is fully vindicated. Because the right is public, absentee infringers have little to no due process right to a day in court on the issue. Group litigation—either via the formal Rule 23 class action or informal virtual representation theory—provides a procedural mechanism for precluding future challenges by absentees. Moreover, such preclusion furthers the substantive goals of patent law and promotes wise judicial administration. Therefore, if the patent owner prevails against a robust challenge to his patent's validity, the matter should be a "thing decided," and res judicata should bind the public from challenging the patent again.

—*Brett Rosenthal*

* * *

Pursuing Academic Freedom After *Garcetti v. Ceballos*[*]

In 1915, the American Association of University Professors' (AAUP) *Declaration of Principles* sought to usher in an as yet undeveloped notion of academic freedom in the United States.[1] Drawing on the traditional notion of *Lehrfreiheit*, the AAUP proposed "not the absolute freedom of utterance of the individual scholar, but the absolute freedom of thought, of inquiry, of discussion and of teaching."[2] The court system gradually seized onto the concept of academic freedom.[3] In 1957, the Supreme Court showed its commitment to academic freedom in *Sweezy v. New Hampshire*,[4] declaring "[t]he essentiality of freedom in the community of American universities . . . almost self-evident;" the Court cautioned that "[n]o one should underestimate the vital role in a democracy that is played by those who guide and train our youth."[5] Ten years later, the Court reaffirmed its commitment to academic freedom in *Keyishian v. Board of Regents of the University of the State of New York*.[6] *Keyishian* contained what has become one of the most often quoted defenses of academic freedom:

> Our Nation is deeply committed to safeguarding academic freedom, which is of transcendent value to all of us and not merely to the teachers concerned. That freedom is therefore a special concern of the First Amendment, which does not tolerate laws that cast a pall of orthodoxy over the classroom. The vigilant protection of constitutional freedoms is nowhere more vital than in the community of American schools.[7]

Given this ringing endorsement, one could easily think that academic freedom is securely entrenched in American jurisprudence. A 2006 Supreme

---

* I am incredibly grateful to Professor David Rabban for his guidance and encouragement while writing this Note. I would also like to thank Monica Hughes, Ross MacDonald, and Michael Selkirk for their comments and suggestions on earlier drafts of this Note. Finally, thanks to my family for all of their support.

1. *See* AM. ASS'N OF UNIV. PROFESSORS, 1915 DECLARATION OF PRINCIPLES ON ACADEMIC FREEDOM AND ACADEMIC TENURE (1915), *reprinted in* POLICY DOCUMENTS AND REPORTS app. I, at 291, 291–92 (10th ed. 2006) [hereinafter 1915 DECLARATION] ("At the meeting of the American Association of University Professors in January 1915, it was decided to take up the problem of academic freedom . . . .").

2. *Id.* at 292, 300.

3. *See generally* William W. Van Alstyne, *Academic Freedom and the First Amendment in the Supreme Court of the United States: An Unhurried Historical Review*, 53 LAW & CONTEMP. PROBS. 79 (1990) (providing a detailed examination of how the principles of academic freedom were assimilated by American courts).

4. 354 U.S. 234 (1957).

5. *Id.* at 250.

6. 385 U.S. 589 (1967).

7. *Id.* at 603 (internal quotation marks omitted).

Court decision, *Garcetti v. Ceballos*,[8] however, has drawn into stark relief the problems inherent in our notion of academic freedom. In *Garcetti*, the Court restricted the right to free speech for public employees who speak pursuant to their official duties but left open the question of how *Garcetti* would impact academic freedom.[9] This Note attempts to distill the problems with academic freedom, particularly as demonstrated in cases grappling with the application of *Garcetti*, and ultimately argues that the Court should recognize a constitutional, individual right to academic freedom moving forward.

This Note begins by updating and reposing William Van Alstyne's concept of a specific theory of academic freedom, distinguishing the right to academic freedom from a more general notion of First Amendment rights. It then turns to *Garcetti v. Ceballos*, providing historical context in order to understand why *Garcetti* marks such a change in our understanding of free speech. The next portion of this Note considers how courts have applied *Garcetti* to cases raising academic freedom issues. Using the problems revealed in the post-*Garcetti* decisions, this Note then suggests the Court should officially recognize a right to academic freedom and offers thoughts on what that right should encompass.

I. The Specific Theory of Academic Freedom After *Garcetti v. Ceballos*

The specific theory of academic freedom seeks to draw a distinction between the right to academic freedom and a more universal, general concept of the right to free expression.[10] The initial concept of academic freedom in the United States was closely tied to the distinctive rights and functions of professors; professors needed the ability to express themselves, without fear of sanction, in order to fulfill their important social role of fostering discussion and advancing knowledge.[11] Quickly, however, given the lack of general free speech rights, teachers seized on the idea of academic freedom and tried to force the concept to serve as a more general right to free speech.[12]

Van Alstyne argues this tendency to blur academic freedom into the general civil liberty of free expression has made it more difficult to recognize legitimate claims of academic freedom, delayed the assimilation of academic freedom into full constitutional protection, and left professors with less

8. 547 U.S. 410 (2006).

9. *Id.* at 421, 425.

10. William Van Alstyne, *The Specific Theory of Academic Freedom and the General Issue of Civil Liberty*, *in* THE CONCEPT OF ACADEMIC FREEDOM 59, 59–60 (Edmund L. Pincoffs ed., 1975).

11. *See* 1915 DECLARATION, *supra* note 1, at 295–96 (describing the three purposes of universities as "promot[ing] inquiry and advanc[ing] the sum of human knowledge; . . . provid[ing] general instruction to the students; and . . . develop[ing] experts for various branches of the public service").

12. Van Alstyne, *supra* note 10, at 62.

protection than other public employees.[13] Though in the wake of *Garcetti* Van Alstyne's final point no longer holds true, his other concerns have become even more pressing. As we will see in the review of cases after *Garcetti*, courts have difficulty recognizing legitimate claims of academic freedom.[14] The Supreme Court still has not articulated a fully developed theory of academic freedom that establishes what is, and what is not, protected under that specific right. And, because the free speech rights of all public employees are more limited now,[15] there is an even greater risk that academics will try to use academic freedom to justify what are more properly characterized as free speech claims.

Van Alstyne defines academic freedom as a "personal liberty to pursue the investigation, research, teaching, and publication of any subject as a matter of professional interest."[16] It is not, however, the absolute freedom to pursue academic inquiry without any standards; instead, it is a freedom that is subject only to standards of professional integrity.[17] This has both advantages and disadvantages for professors.[18] In some respects, professors have greater rights than other citizens; they claim a right to protection even if they express an unpopular view or a perspective that might have negative repercussions for their university employer.[19] Academic freedom, however, also gives professors less freedom. In claiming that they should only be evaluated based on professional standards, professors in effect claim a greater burden of ensuring that their speech complies with those professional standards.[20] Without this protection, though, professors would not be able to fulfill their essential functions. Professors are expected to be dedicated to "[examining] received learning and values critically," and it would make no sense to levy that expectation on professors without giving them the freedom to meet it.[21]

Under the specific theory of academic freedom, the academic context of the speech "may well be determinative."[22] Judging professors by

---

13. *Id.* at 63–69.

14. *See infra* Part III.

15. *See infra* Part II.

16. Van Alstyne, *supra* note 10, at 71.

17. *See id.* (explaining that, in exercising this "personal liberty to pursue the investigation, research, teaching, and publication of any subject," professors should be free from sanction unless they commit "an inexcusable breach of professional ethics in the exercise of that freedom").

18. For a more detailed explanation of how academic freedom gives professors both more and fewer rights, see *id.* at 75–77.

19. *See id.* at 74 ("[T]he charge of [faculty committees] is strictly limited: it is to ignore the particular impact of any teacher's exercise of his academic freedom upon the institution and to concern itself solely with the question of whether the teacher . . . has been guilty of . . . an inexcusable breach of professional ethics . . . .").

20. *See id.* at 75–76 ("[I]n respect to his academic freedom, the teacher or scholar is simultaneously under *more* constraint as well as under less constraint than would ordinarily obtain.")

21. *Id.* at 77.

22. *Id.* at 78.

professional standards means they are not subject to the same restrictions as other public employees. If it cannot be shown that a professor has violated professional ethics with his or her speech, "the law or institutional rule that operates to abridge the exercise of that academic freedom should be held invalid as applied to the particular case."[23] This characteristic of academic freedom shows that academic freedom is best understood as a subset of free speech rights; academic freedom, while related to speech, makes some speech of professors subject to a different standard and requires that speech be evaluated based only on professional standards.[24]

Separating a specific theory of academic freedom from a more general right to free speech also demands that professors develop a clearer understanding of the two rights.[25] Claims for redress based on extraprofessional speech should be grounded in a general free speech right.[26] Van Alstyne concedes that some "personal conduct" of the faculty member may be so integrally related to the functioning of the university that it too should receive additional protection.[27] In light of the Court's decision not to safeguard the speech of public employees made pursuant to their official duties, I would phrase this somewhat differently. The question we now face is determining what, if any, responsibilities of a professor are so connected to the university's mission of critical inquiry that they deserve additional constitutional protection.

II. *Garcetti v. Ceballos*: Changing First Amendment Rights and the Unanswered Questions for Academic Freedom

In 2006, the Supreme Court issued its opinion in *Garcetti v. Ceballos*, dramatically restricting the free speech rights of public employees and effecting an overall change in First Amendment jurisprudence.[28] Pre-*Garcetti* case law required courts to balance the interests of public employees with the interests of the government, but with the decision in *Garcetti*, the Court erected a roadblock before courts can even reach the issue of how to balance the divergent interests.[29] If plaintiffs cannot meet this threshold

23. *Id.*

24. *Id.*

25. *See id.* at 84 ("What needs to be done . . . is . . . to make clearer that a faculty member may not properly be held to answer to an institution for the integrity of his general utterances by the same professional standard by which he may have to account for his academic freedom . . . .").

26. *See id.* (explaining that "[f]or an alleged abuse of one's ordinary freedom of speech, general provisions of law are available to provide for measures of redress").

27. *Id.* at 84–85.

28. *See* Robert J. Tepper & Craig G. White, *Speak No Evil: Academic Freedom and the Application of* Garcetti v. Ceballos *to Public University Faculty*, 59 CATH. U. L. REV. 125, 157 (2009) (characterizing *Garcetti* as a "watershed case").

29. *Compare* Garcetti v. Ceballos, 547 U.S. 410, 421 (2006) (holding that if public employees speak "pursuant to their official duties," their speech is not protected), *with* Pickering v. Bd. of Educ., 391 U.S. 563, 568 (1968) (explaining the need to "balance . . . the interests of the teacher, as

inquiry, courts will categorically deny First Amendment protection to the speech in question.[30] To truly understand *Garcetti*, however, one needs to understand the test in place before this landmark decision. Familiarity with the prior case law will not only help to contextualize the impact of *Garcetti*, but it will also prove vital to understanding how courts have addressed issues of academic freedom in its aftermath. I will first discuss two important pre-*Garcetti* cases for free speech protection of public employees, *Pickering v. Board of Education*[31] and *Connick v. Myers*,[32] before turning to *Garcetti* itself. In addition to covering the *Garcetti* decision, my discussion will pay particular attention to the issue of academic freedom, which the Court deliberately left unresolved.[33]

A. Pickering v. Board of Education

Pickering concerned the First Amendment rights of a public school teacher who claimed he had been unfairly terminated for exercising his right to free speech.[34] Pickering, a school teacher, wrote to the editor of the local newspaper sharply and somewhat unfairly (several of his statements in the letter were proven false) criticizing the Board of Education's handling of a school bond and tax issue.[35] The Board of Education "dismissed Pickering for writing and publishing the letter," a decision it later affirmed in a hearing mandated under state law.[36] Pickering brought suit in state court claiming that he could not be fired for writing such a letter, but the Illinois Supreme Court rejected his claim.[37]

On appeal, the Supreme Court reversed.[38] The Court immediately acknowledged that teachers could not "be compelled to relinquish the First Amendment rights they would otherwise enjoy as citizens to comment on matters of public interest" and framed the issue as one of balancing those interests of the teacher to speak on matters of public concern with the interests of the state as an employer in operating efficiently.[39] In a decision clearly influenced by its recent decision in *New York Times Co. v. Sullivan*,[40]

a citizen, in commenting upon matters of public concern and the interest of the State, as an employer, in promoting the efficiency of the public services it performs").

30. *Garcetti*, 547 U.S. at 421.

31. 391 U.S. 563 (1968).

32. 461 U.S. 138 (1983).

33. *See Garcetti*, 547 U.S. at 425 ("We need not, and for that reason do not, decide whether the analysis we conduct today would apply in the same manner to a case involving speech related to scholarship or teaching.").

34. *Pickering*, 391 U.S. at 564–65.

35. *Id.* at 566–67, 570.

36. *Id.* at 566; Pickering v. Bd. of Educ., 225 N.E.2d 1, 2 (Ill. 1967).

37. *Pickering*, 225 N.E.2d at 6–7.

38. *Pickering*, 391 U.S. at 575.

39. *Id.* at 568.

40. 376 U.S. 254 (1964); *see id.* at 270 ("[W]e consider this case against the background of a profound national commitment to the principle that debate on public issues should be uninhibited,

the Court held that Pickering's speech was protected under the First Amendment.[41] His false statements had not interfered with his ability to do his job or the Board's ability to perform its duties, and without proof that Pickering had made his false statements knowingly or recklessly, the Board could not punish him for writing the letter.[42] As summarized in later cases, the *Pickering* balancing test presents courts with two questions: First, did the employee speak on a matter of public concern?[43] And second, did the State, as employer, have "adequate justification for treating the employee differently from any other" citizen speaking on an issue of public concern?[44]

B. Connick v. Myers

In *Connick*, the Court refined the *Pickering* analysis, clarifying the question of what constituted a matter of public concern. Myers, an assistant district attorney, received an order transferring her to a different division, a decision she protested.[45] After being informed the transfer order was official despite her complaints, Myers drafted and distributed a questionnaire to her colleagues, seeking employee opinions on a variety of topics.[46] After her initial protests, Myers indicated she would consider the transfer, but Myers's supervisor, Connick, terminated her after she distributed the survey, telling her that she was being dismissed for her refusal to accept the transfer.[47] When Myers brought suit alleging retaliation for exercising her First Amendment right to free speech, the district court concluded Myers had been terminated because of the survey and agreed that her rights had been violated,[48] a decision that the court of appeals affirmed.[49]

The Supreme Court, however, reversed that decision.[50] In analyzing whether Myers's survey dealt with a matter of public concern, the Court

---

robust, and wide-open, and that it may well include vehement, caustic, and sometimes unpleasantly sharp attacks on government and public officials."); *see also* Stanley S. Arkin & Luther A. Granquist, *The Presumption of General Damages in the Law of Constitutional Libel*, 68 COLUM. L. REV. 1482, 1486–87 (1968) (explaining that, just as in the *Sullivan* case, the *Pickering* Court was unwilling to find liability where no actual harm could be shown).

41. *Pickering*, 391 U.S. at 574–75.

42. *Id.* at 572–75.

43. *See* Garcetti v. Ceballos, 547 U.S. 410, 418 (2006) (noting that an analysis under *Pickering* first asks if "the employee spoke as a citizen on a matter of public concern"); Connick v. Myers, 461 U.S. 138, 145 (1983) ("Our cases following *Pickering* also involved safeguarding speech on matters of public concern.").

44. *Garcetti*, 547 U.S. at 418.

45. *Connick*, 461 U.S. at 140.

46. *Id.* at 140–41. The survey asked for employee opinions on the "transfer policy, office morale, the need for a grievance committee, the level of confidence in supervisors, and whether employees felt pressured to work in political campaigns." *Id.* at 141.

47. *Id.* at 140–41.

48. Myers v. Connick, 507 F. Supp. 752, 760 (E.D. La. 1981).

49. Myers v. Connick, 654 F.2d 719, 719 (5th Cir. 1981).

50. *Connick*, 461 U.S. at 154.

considered "the content, form, and context of [the] statement."[51] Only one question, regarding pressure to participate in political campaigns, touched on a matter of public concern, and despite this finding, the Court held that the government's interest in efficiency and effectiveness outweighed Myers's interest in speaking.[52] Despite the Court's interest in protecting the rights of public employees to speak on matters of public concern, it did not "attempt to constitutionalize the employee grievance."[53]

Connick added two things to the understanding of the free speech rights of public employees. First, it confirmed that *Pickering* required a true balancing of interests and that the government's interest deserved consideration equal to that shown to the interests of the employee.[54] Second, it clarified that *Pickering* did not make every aspect of working for a public employer a matter of public concern; though a public employee did not lose his or her constitutional right to comment on a matter of public concern, the employee also did not have a First Amendment right to criticize his or her employer on purely internal matters.[55]

C. Garcetti v. Ceballos

Garcetti changed the evaluation process for the First Amendment claims of public employees by emphasizing a different aspect of *Pickering*: the Court directed its attention to the first prong of the *Pickering* balancing test, "the interests of the [public employee], as a *citizen*, in commenting upon matters of public concern."[56] Ceballos, a deputy district attorney, conducted an investigation that revealed "serious misrepresentations" in an affidavit in one of his cases, and he wrote two memos reporting the mistakes in the affidavit to his supervisors.[57] After a contentious meeting held to discuss Ceballos's memos, the district attorney decided to move ahead with the case despite Ceballos's concerns.[58] After he was "reassign[ed] from his calendar deputy position to a trial deputy position, transfer[red] to another courthouse, and deni[ed] . . . a promotion," Ceballos brought suit alleging his employer had retaliated against him and violated his First Amendment rights.[59]

---

51. *Id.* at 147–48.
52. *Id.* at 149–50, 154.
53. *Id.* at 154.
54. *Id.* at 149–50.
55. *See id.* at 146 (stating that a government employer does not violate First Amendment rights when the employee's expression could not "be fairly considered as relating to any matter of political, social, or other concern to the community").
56. Garcetti v. Ceballos, 547 U.S. 410, 417–18 (2006) (emphasis added) (quoting Pickering v. Bd. of Educ., 391 U.S. 563, 568 (1968)).
57. *Id.* at 413–14.
58. *Id.* at 414.
59. *Id.* at 415.

The district court held that Ceballos's speech was not protected,[60] a decision that the court of appeals later reversed.[61] Applying a *Pickering–Connick* balancing test, the U.S. Court of Appeals for the Ninth Circuit concluded that Ceballos spoke on a matter of public concern.[62] The District Attorney's Office did not allege that Ceballos had disturbed its ability to operate effectively and efficiently, and as such, the balance of the competing interests favored Ceballos.[63]

The Supreme Court granted certiorari and reversed the Ninth Circuit's decision.[64] The Court's decision turned on the fact that Ceballos spoke in his official capacity.[65] The Court further elaborated that:

> [T]he fact that Ceballos spoke as a prosecutor fulfilling a responsibility to advise his supervisor about how best to proceed with a pending case—distinguishes Ceballos' case from those in which the First Amendment provides protection against discipline. We hold that when public employees make statements pursuant to their official duties, the employees are not speaking as citizens for First Amendment purposes, and the Constitution does not insulate their communications from employer discipline.[66]

This changed the *Pickering–Connick* analysis by mandating that courts first consider whether the plaintiff spoke as a citizen or as part of his official duties. Though official duties might seem ambiguous, the Court explained that "[t]he proper inquiry is a practical one"; in evaluating free speech claims in light of *Garcetti*, courts need not confine themselves solely to the employee's official job description but instead should look at "the duties an employee actually is expected to perform."[67]

Garcetti was a 5–4 decision, and one of the major points of disagreement between the majority and the dissent concerned the decision's potential impact on academic freedom.[68] If the First Amendment no longer protects speech made pursuant to one's official duties, as Justice Souter observed in his dissent, "even the teaching of a public university professor" could lose the protection of the First Amendment.[69] Justice Souter specifically expressed his "hope that today's majority does not mean to

---

60. Ceballos v. Garcetti, No. CV0011106AHMAJWX, 2002 WL 34098285, at *6 (C.D. Cal. Jan. 30, 2002).

61. Ceballos v. Garcetti, 361 F.3d 1168, 1172–73 (9th Cir. 2004).

62. *Id.* at 1178–80.

63. *Id.* at 1180.

64. *Garcetti*, 547 U.S. at 426.

65. *See id.* at 421 ("The controlling factor in Ceballos' case is that his expressions were made pursuant to his duties as a calendar deputy.").

66. *Id.*

67. *Id.* at 424–25.

68. *Id.* at 412; *see id.* at 425 ("Justice Souter [who wrote the dissent] suggests today's decision may have important ramifications for academic freedom . . . .").

69. *Id.* at 438 (Souter, J., dissenting).

imperil First Amendment protection of academic freedom in public colleges and universities."[70] Citing the language regarding academic freedom in *Grutter v. Bollinger*,[71] *Keyishian v. Board of Regents of the University of the State of New York*, and *Sweezy v. New Hampshire*, Justice Souter noted the Court's tendency to treat academic freedom and the work of universities as "special."[72]

The majority paused briefly to acknowledge Justice Souter's concerns. The majority conceded that "[t]here is some argument that expression related to academic scholarship or classroom instruction implicates additional constitutional interests . . . not fully accounted for by this Court's customary employee-speech jurisprudence."[73] The majority, however, declined to resolve the question because Ceballos's situation did not directly raise the issue.[74]

III. The Status of Academic Freedom Post-*Garcetti*

The Supreme Court still has not addressed the question it left open in *Garcetti*, namely how the rule should apply to academic freedom. Even framing the question left undecided, though, poses some difficulty—is it a question of how *Garcetti* impacts the right to academic freedom? A question of whether *Garcetti* applies to cases alleging "speech related to *scholarship* or *teaching*?"[75] Or does the unresolved issue have broader implications, suggesting an exception to the rule outlined in *Garcetti* that would apply to all speech of teachers? And, if it does raise those broader concerns, would it protect the rights of all teachers or just those of professors?

Perhaps unsurprisingly given the variety of ways to frame these questions, courts have taken radically different approaches in applying *Garcetti* to cases raising free speech and academic freedom concerns of professors and teachers. I will discuss a representative sample of cases below, but before beginning, an overview may be helpful in order to observe themes and trends as they appear in the following cases.

I divide these post-*Garcetti* cases into three categories. First, I only found one case where the court applied *Garcetti* without even acknowledging the potential that the analysis should be different because the court was

---

70. *Id.*

71. 539 U.S. 306 (2003).

72. *See Garcetti*, 547 U.S. at 438–39 (Souter, J., dissenting) ("[U]niversities occupy a special niche in our constitutional tradition." (quoting *Grutter*, 539 U.S. at 329)); *id.* ("[Academic] freedom is therefore a special concern of the First Amendment . . . ." (quoting Keyishian v. Bd. of Regents of the Univ. of the State of N.Y., 385 U.S. 589, 603 (1967))); *id.* (characterizing academic freedom as an "area[] in which government should be extremely reticent to tread" (quoting Sweezy v. New Hampshire, 354 U.S. 234, 250 (1957))).

73. *Id.* at 425 (majority opinion).

74. *Id.*

75. *Id.* (emphasis added).

considering the First Amendment retaliation claim of a professor;[76] I will term this the no-exception category. The majority of courts to consider claims of academic freedom post-*Garcetti* have acknowledged the dissent's concerns, and their approaches then fall into two categories: first, apply an alternate test, or second, apply *Garcetti* but decline to resolve the open issues fully. In the first group, the court declines to apply *Garcetti*'s pursuant-to-official-duties test and instead uses a *Pickering–Connick* balancing test; I will refer to these cases as the *Pickering–Connick* exception. The second group of cases shows courts acknowledging that some exception to *Garcetti* may exist but then declining to place the case within that exception. I will characterize these as the limited-exception group and further refine them based on whether they are limited based on the class of person to whom they apply or limited by the court's understanding of official duties.

A. No Exception

1. Case Law in the No-Exception Category.—Renken v. Gregory,[77] decided by the U.S. Court of Appeals for the Seventh Circuit, appears to be the only case where the court applied *Garcetti* to a professor's First Amendment claim without discussing whether a different standard should or does apply to cases related to academic freedom. It is also arguably a case that belongs in the category of cases that the Supreme Court referred to when it noted that "expression related to academic scholarship . . . implicates additional constitutional interests."[78]

Renken, a tenured engineering professor, applied for and received a grant from the National Science Foundation (NSF), which required that the university match the NSF funding.[79] Gregory, the dean of the engineering school, then sent Renken a proposal about the university's contribution, including specifics about "equipment, salaries . . . and laboratory space"; he promised to release the university's contribution once Renken agreed to these conditions.[80] Renken complained that the conditions were inadequate (insufficient lab space, delays in paying students working on the project, etc.) and "that Gregory's fund proposal contravened NSF regulations regarding

76. *See* Renken v. Gregory, 541 F.3d 769, 773–74 (7th Cir. 2008) (applying *Garcetti* to a professor's claims of retaliation for speech related to his research without mentioning the possible exception discussed in *Garcetti*). I use the term "First Amendment claim" to describe the professor's allegations deliberately, because that is how the court refers to them; the term "academic freedom" does not appear in the opinion. *Hong v. Grant*, 516 F. Supp. 2d 1158, 1161, 1165 (C.D. Cal. 2007) also applied *Garcetti* to a professor's speech claims without acknowledging the academic freedom exception, but I do not include *Hong* here because, on appeal, the Ninth Circuit found the defendants were protected by immunity. *Hong v. Grant*, 403 F. App'x 236, 237–38 (9th Cir. 2010).

77. 541 F.3d 769 (7th Cir. 2008).

78. *Garcetti*, 547 U.S. at 425.

79. *Renken*, 541 F.3d at 770–71.

80. *Id.* at 771.

matching funds."[81] Gregory provided a revised proposal that changed the lab space requirements and cautioned that, if Renken did not accept the proposal, he would cancel the grant.[82] When Renken refused to sign the revised proposal and instead replied with his own requests, Gregory informed him that the university had begun the process of returning NSF funds.[83] After Renken filed several internal complaints and refused to sign a compromise proposal, the university made the official decision to return the NSF funds.[84]

Renken filed suit alleging that the university retaliated against him for complaining about the university's handling of the grant funding, speech he claimed the First Amendment protected.[85] The district court disagreed and granted the university's motion for summary judgment, concluding that Renken "spoke pursuant to his official duties" and so his speech was not protected under the First Amendment.[86]

On appeal, the Seventh Circuit applied *Garcetti* and held that Renken spoke as an employee, not a citizen.[87] Renken argued his work with the grant was "in the course of his job" but not "a *requirement* of his job."[88] The court, however, observed that Renken's job required teaching, research, and service, and characterized his work with the NSF grant as related to both service (part of the stated purpose of the grant was improving student's undergraduate experience) and teaching (the research grant reduced his teaching responsibilities).[89] Renken complained about the grant "pursuant to his official duties as a University professor," and so had no First Amendment protection for his speech.[90]

2. *The Problems with the No-Exception Category.*—As noted above, the phrase "academic freedom" does not appear anywhere in the *Renken* opinion. The court does not mention the unsolved question of *Garcetti* when it states the rule from *Garcetti* or when it analyzes Renken's case. In addition, Renken never used academic freedom as a potential protection for his speech. He made a more generalized First Amendment claim, and his primary defense rested on distinguishing himself from the *Garcetti* rule and

81. *Id.*

82. *Id.* at 772.

83. *Id.*

84. *Id.* at 772–73.

85. *Id.* at 770. Renken specifically alleged that Gregory had "delay[ed]/refus[ed]: Personnel Action Forms[,] . . . External Requisitions to purchase equipment, supplies and materials, and other expenditure items" related to this grant and others. *Id.* at 772.

86. Renken v. Gregory, No. 04-C-1176, 2007 WL 2220210, at *1–3 (E.D. Wis. July 31, 2007).

87. *Renken*, 541 F.3d at 774.

88. *Id.* at 773.

89. *Id.* at 771, 773–74.

90. *Id.* at 775.

attempting to argue that he was not acting in an official capacity.[91] This is striking because his situation seems a strong one for arguing that additional considerations are at play; unlike other cases where courts distinguished *Garcetti* based on the specific language used,[92] Renken was working at a public university, as a professor, who spoke on a matter related to both scholarship and teaching.[93] Although it is difficult if not impossible to know why academic freedom was not part of the discussion in *Renken*, this case is notable for that exact reason.

Although the Seventh Circuit rejected them, Renken's arguments in his defense are also interesting because they seem to suggest a possible limitation to *Garcetti*: restrict the scope of the decision by interpreting pursuant to official duties to mean only duties required as part of one's job. The Court's decision in *Garcetti* seems to leave this possibility open.[94] Neither Garcetti nor his employer disputed that Garcetti had written the memo pursuant to his official duties, and so the Court had "no occasion to articulate a comprehensive framework for defining the scope of an employee's duties."[95] Given this second unresolved question of how to define official duties, it may seem reasonable to believe that post-*Garcetti* decisions could use Renken's arguments to protect speech.

An example makes the appeal of this argument even clearer. A typical job description for a professor is likely to require teaching, research, and service. So take, for example, the case of an environmental sciences professor who gives a speech on campus; the professor reports that, in the course of his research, he has concluded that global warming does not exist. School officials, frustrated with the professor's unpopular opinion, refuse to renew his contract. Technically, the professor's job description did not require him to make that speech; he was only required to teach, conduct research, and perform some service to the university. If a court were to accept this argument, the professor could escape analysis under *Garcetti* because his speech was not pursuant to his official (required) duties.

Ultimately, however, courts seem unlikely to limit *Garcetti*'s "pursuant to . . . official duties"[96] to "required by official duties." First, though only in dicta, the *Garcetti* majority suggests that is not an appropriate interpretation

---

91. *See id.* at 773–74 ("Renken argues that the tasks that he conducted in relation to the grant were implemented at his discretion '*while* in the course of his job and not as a *requirement* of his job.'").

92. *See infra* section III(C)(1) (observing that some courts refuse to apply an academic freedom exception for high school teachers because Justice Souter's dissent specifically mentions "the teaching of a public university professor," *Garcetti v. Ceballos*, 547 U.S. 410, 438 (2006) (Souter, J., dissenting), and also noting that other courts refuse an academic freedom exception if the speech of the university professor is not, in the words of the *Garcetti* majority, "related to scholarship or teaching," *id.* at 425 (majority opinion)).

93. *Renken*, 541 F.3d at 770–73.

94. *Garcetti*, 547 U.S. at 424–25.

95. *Id.* at 424.

96. *Id.* at 421.

of pursuant to official duties.[97] Responding to the dissent's concerns that employers, in an attempt to restrict the free speech of public employees, will craft overly broad job descriptions in the wake of *Garcetti*, the majority cautions that job descriptions are not the best tool for determining what actions are pursuant to official duties.[98] Formal job descriptions often fail to match the reality of what an employee is actually expected to perform, and so, "[t]he proper inquiry is a practical one."[99] Second, decisions after *Garcetti* suggest that courts have not interpreted pursuant to official duties to mean only those duties required by the job.[100] Though Renken's argument to apply *Garcetti* only to required duties might be appealing, the argument seems unlikely to succeed.

B. Pickering–Connick *Exception*

The U.S. Court of Appeals for the Fourth Circuit has been the most consistent court in applying what I have called the *Pickering–Connick* exception; given the uncertainty regarding how to evaluate the First Amendment claims of professors and teachers in the wake of *Garcetti*, the Fourth Circuit has repeatedly chosen to apply the *Pickering–Connick* balancing test to cases raising this issue.[101] District courts in the Sixth Circuit have used the *Pickering–Connick* test to evaluate the First Amendment claims of professors post-*Garcetti*, but a more recent decision from the Sixth Circuit, *Evans-Marshall v. Board of Education*,[102] casts some doubt as to whether that will continue in the future.[103]

1. Case Law in the Pickering–Connick *Category.*—The Fourth Circuit evaluated a high school teacher's claim that the school had violated his First

97. *Id.* at 424–25.

98. *Id.*

99. *Id.* at 424.

100. *See, e.g.*, Gorum v. Sessoms, 561 F.3d 179, 185 (3d Cir. 2009) (rejecting Gorum's argument that his speech was not pursuant to his official duties simply because it "went beyond his specified responsibilities in the Collective Bargaining Agreement"); Savage v. Gee, 716 F. Supp. 2d 709, 710, 717 (S.D. Ohio 2010) (holding that Savage was acting pursuant to his official duties when he made recommendations as part of a faculty committee and that it made "no difference that he was not strictly required to serve on the committee").

101. *See, e.g.*, Adams v. Trs. of the Univ. of N.C.–Wilmington, 640 F.3d 550, 564 (4th Cir. 2011) ("utiliz[ing] the *Pickering–Connick* analysis for determining whether [Adams's speech] was that of a public employee, speaking as a citizen upon a matter of public concern"); Lee v. York Cnty. Sch. Div., 484 F.3d 687, 694 n.11 (4th Cir. 2007) (explaining that, because the Supreme Court did not decide whether *Garcetti* "would apply in the same manner to a case involving speech related to teaching," the Fourth Circuit would apply the *Pickering–Connick* standard).

102. 624 F.3d 332 (6th Cir. 2010).

103. *Compare* Kerr v. Hurd, 694 F. Supp. 2d 817, 843 (S.D. Ohio 2010) ("[A]bsent Sixth Circuit or further Supreme Court guidance . . . this Court will continue to apply the . . . *Pickering–Connick* approach to cases involving in-class speech by primary and secondary school teachers."), *with Evans-Marshall*, 624 F.3d at 340 (applying *Garcetti*'s pursuant-to-official-duties test to the First Amendment claims of a high school teacher).

Amendment rights in *Lee v. York County School Division*.[104] Lee, a high school Spanish teacher, posted several religious pieces that discussed prayer in politics and missionary work on his class bulletin board.[105] After the school received a complaint, the principal removed the items from Lee's bulletin board because "he 'could not find any reason why [these items] would be posted in a classroom.'"[106] Lee protested the removal, and when the school board refused him permission to repost the articles, he claimed his free speech rights had been violated.[107]

The district court used *Pickering–Connick* and Fourth Circuit precedent to evaluate Lee's claims, ultimately concluding that his speech was not protected under the First Amendment.[108] The district court concluded that Lee's speech was curricular and, as such, "not a matter of public concern" and unprotected speech;[109] on appeal, the Fourth Circuit affirmed.[110] Noting that the Supreme Court had not determined how *Garcetti* would apply to the speech of teachers, the Fourth Circuit, like the district court, used *Pickering–Connick* and *Boring*.[111] The court's analysis relied heavily on *Boring*'s holding that curricular speech is not protected under the First Amendment.[112] It also placed particular emphasis on a more institutional understanding of First Amendment rights in schools; a school board needs some control over the speech of teachers within the classroom in order to effect its mission of properly educating students.[113] Within the school environment, teachers do not have "First Amendment free speech rights . . . 'automatically coextensive with the rights of adults in other settings.'"[114]

Similarly, the Fourth Circuit took an institutional view of academic freedom when evaluating a professor's claims that his First Amendment rights had been violated in *Adams v. Trustees of the University of North Carolina–Wilmington*.[115] Adams, a tenured associate professor of criminology at the University of North Carolina–Wilmington, applied for a promotion to a full professor position.[116] As part of his application, Adams (who had recently converted to Christianity) listed his work advising

---

104. 484 F.3d 687, 689 (4th Cir. 2007).

105. *Id.* at 689–90.

106. *Id.*

107. *Id.* at 691–92.

108. Lee v. York Cnty. Sch. Div., 418 F. Supp. 2d 816, 821–27 (E.D. Va. 2006). For the Fourth Circuit precedent used to evaluate Lee's allegations, see Boring v. Buncombe Cnty. Bd. of Educ., 136 F.3d 364, 368 (4th Cir. 1998) (holding that a teacher's choice of a school play was curricular and, therefore, not protected under the First Amendment).

109. *Lee*, 418 F. Supp. 2d at 825, 827–28.

110. *Lee*, 484 F.3d at 689.

111. *Id.* at 694 n.11; Boring v. Buncombe Cnty. Bd. of Educ., 136 F.3d 364 (4th Cir. 1998).

112. *Lee*, 484 F.3d at 696–97.

113. *Id.* at 695–96.

114. *Id.* at 695 (quoting Hazelwood Sch. Dist. v. Kuhlmeier, 484 U.S. 260, 266 (1988)).

115. 640 F.3d 550 (4th Cir. 2011).

116. *Id.* at 553.

Christian groups on campus and his external writings and speeches on Christianity.[117] The selection committee chose not to promote Adams, citing its concern about his "scholarly research productivity."[118] Adams then filed a complaint against the university for violating his speech rights, and the district court granted the defendant's motion for summary judgment because the speech at issue was pursuant to Adams's duties as a faculty member.[119]

On appeal, the Fourth Circuit reversed after applying *Pickering–Connick*.[120] The court first gave context to its decision, noting that colleges and universities are best equipped to evaluate their own employment decisions and that courts should only make a limited review of those decisions.[121] Turning then to the district court's decision, the court concluded the lower court had "misread *Garcetti*" and erred in using the standard from "*Garcetti* without acknowledging, let alone addressing, the clear language in that opinion that casts doubt on whether the *Garcetti* analysis applies in the academic context of a public university."[122] The court went on to state that it did not believe *Garcetti* intended to make speech related to a professor's academic work and service unprotected.[123] While the court acknowledged that a professor's speech "declaring or administering university policy" might be subject to a *Garcetti* analysis, it concluded "that *Garcetti* would not apply in the *academic context* of a public university as" seen in this case.[124] Adams spoke on matters of public concern,[125] and though the court did not elaborate significantly on his role while writing, the court ultimately held that Adams's speech was protected as "that of a citizen . . . on a matter of public concern."[126]

Though the Sixth Circuit's decision in *Evans-Marshall* may have set a new standard for cases decided in the circuit moving forward, it is worth noting *Kerr v. Hurd*,[127] where the U.S. District Court for the Southern District of Ohio applied *Pickering–Connick* to assess a professor's claims that his First Amendment rights had been violated.[128] Kerr, a medical school professor, alleged the school had harassed him, disciplined him, and falsely

---

117. *Id.* at 553–54.

118. *Id.* at 555–56.

119. *Id.* at 556, 561.

120. *Id.* at 560–66.

121. *See id.* at 557 ("It is with this well-established understanding of the limited review courts may undertake in cases involving employment decisions of academic institutions that we consider Adams' claims.").

122. *Id.* at 561.

123. *Id.* at 564.

124. *Id.* at 562–63 (emphasis added).

125. *See id.* at 565 (observing that Adams's external writings concerned "academic freedom, civil rights, campus culture, sex, feminism, abortion, homosexuality, religion, and morality," which were all "plainly" matters of public concern).

126. *Id.*

127. 694 F. Supp. 2d 817 (S.D. Ohio 2010).

128. *Id.* at 828, 840, 843.

accused him of professional misconduct because he insisted on teaching certain methods of delivery.[129] Hurd, relying on *Garcetti*, countered that Kerr's speech was unprotected because Kerr spoke in his capacity as a professor in the medical school.[130]

In deciding the case, the court chose to use *Pickering–Connick* to evaluate Kerr's claims.[131] It noted that the issue of how to apply *Garcetti* in an academic context was still undecided and drew on *Lee* to support its decision to apply *Pickering–Connick*; the court also relied on another Ohio district court opinion that had used *Pickering–Connick* to determine whether a teacher's classroom speech was protected.[132] The *Kerr* court took an even stronger stance, however, writing,

> Even without the binding precedent, this Court would find an academic exception to *Garcetti*. Recognizing an academic freedom exception to the *Garcetti* analysis is important to protecting First Amendment values. Universities should be the active trading floors in the marketplace of ideas. Public universities should be no different from private universities in that respect.[133]

Kerr spoke in the classroom and his lessons on delivery methods fell "well within the range of accepted medical opinion," and so the court denied summary judgment to Hurd.[134]

2. Problems with the Pickering–Connick *Category.*—Despite the different outcomes and views on academic freedom seen above, these three cases show one possible way to deal with the question left open under *Garcetti*: use *Pickering–Connick* to determine whether a teacher's or professor's speech is protected. Even these cases, though, show that this standard alone will not necessarily create a consistent rule or give teachers or professors confidence in what speech is protected and what speech is unprotected. For example, the teacher in *Lee* was not protected because, according to circuit precedent, his curricular speech was not a matter of public concern, yet in *Adams* the professor's classroom speech was protected. It is clearly possible to distinguish these cases based on who was speaking (a high school teacher as compared to a medical school professor),

129. *Id.* at 834.

130. *Id.* at 843.

131. *Id.*

132. *Id.* (discussing Evans-Marshall v. Bd. of Educ., No. 3:03cv091, 2008 WL 2987174 (S.D. Ohio July 30, 2008)). Though the other Ohio district court case, *Evans-Marshall v. Board of Education*, was affirmed on appeal, the Sixth Circuit's analysis differed from the district court's. The Sixth Circuit applied *Garcetti*'s pursuant-to-official-duties test to hold that the high school teacher's classroom speech was not protected under the First Amendment. Evans-Marshall v. Bd. of Educ., 624 F.3d 332, 340 (6th Cir. 2010). For more explanation, see *infra* section III(C)(1).

133. *Kerr*, 694 F. Supp. 2d at 843–44.

134. *Id.* at 834, 844, 848.

but I believe this points to a bigger problem with simply substituting *Pickering–Connick* in academic speech cases.

Garcetti shows the need for a specific, constitutional right to academic freedom, and simply relying on *Pickering–Connick* as an exception to the rule in *Garcetti* would not meet that need. Using *Pickering–Connick* instead of *Garcetti* still does not mean courts have made a principled decision about how to apply *Garcetti* in the academic workplace; Lee lost protection for his speech not because he was a teacher but because the court concluded his speech was not a matter of public concern. As noted by Judith Areen, neither *Pickering* nor *Connick* "takes adequate account of the distinctive nature of the academic workplace."[135] Courts simply apply the same balancing test used to evaluate the free speech claims of all other public employees, regardless of the fact that the academic workplace should balance concerns other than public concern or government efficiency.

The specific theory of academic freedom is rooted in the idea that there is something distinctive about the speech of professors that warrants additional constitutional protection,[136] but the *Pickering–Connick* exception would not take the unique nature of the speech of professors into account or provide professors with additional constitutional safeguards. *Pickering* and *Connick* were decided on general First Amendment principles. Neither case mentions the concept of academic freedom, nor does the Court rely on academic freedom in either case when determining what speech is or is not protected. Because *Pickering* and *Connick* are part of general First Amendment jurisprudence, using *Pickering–Connick* as an exception to *Garcetti* would do nothing to incorporate or recognize the concept of academic freedom.

Cases raising questions of academic freedom that were decided before *Garcetti* further reveal the problems with relying on a *Pickering–Connick* analysis to protect the free speech of professors. In *Urofsky v. Gilmore*,[137] university professors challenged a state statute that limited their ability to access sexually explicit material on state-owned computers.[138] These professors claimed that, even though the statute allowed professors to gain permission to access sexually explicit material for academic purposes, it violated their right to academic freedom.[139] In reversing the district court's decision and holding that the statute did not violate the professors' First Amendment rights, the court reviewed case law to determine what, exactly, academic freedom meant.[140] The Fourth Circuit concluded that "the best that

135. Judith Areen, *Government as Educator: A New Understanding of First Amendment Protection of Academic Freedom and Governance*, 97 GEO. L.J. 945, 989 (2009).

136. *See supra* Part I.

137. 216 F.3d 401 (4th Cir. 2000).

138. *Id.* at 404.

139. *Id.* at 405–06.

140. *Id.* at 410–16.

can be said for [the professors'] claim that the Constitution protects . . . academic freedom . . . is that teachers were the first public employees to be afforded the now-universal protection against dismissal for the exercise of First Amendment rights."[141] In essence, the First Amendment rights of professors were coincident with the First Amendment rights of other public employees.[142] The *Pickering–Connick* exception does nothing to prevent courts from reaching this same conclusion and continues to leave the free speech and academic freedom rights of professors unprotected.

Finally, the *Pickering–Connick* analysis also fails to give a definitive answer about whether speech related to school administration or other professional duties outside of the classroom would be protected within the *Garcetti* exception. I believe this is one of the most troubling questions left unanswered by *Garcetti*; any attempt to resolve *Garcetti*'s application to academic cases must give clear guidance on how courts are to address speech that occurs outside the classroom.

C. *Limited Exception*

The majority of cases after *Garcetti* fall somewhere in between refusing to acknowledge the potential for a different standard (such as *Renken*) and applying *Pickering–Connick* in lieu of guidance as to how to apply *Garcetti*. These cases all recognize that the question of how to apply *Garcetti* in the academic setting is unsettled, but they only recognize a limited exception to the *Garcetti* holding. I divide these cases into two separate categories: (1) decisions limiting *Garcetti* based on who the right to academic freedom applies to, and (2) decisions limiting *Garcetti* based on the scope of the academic's official duties.

1. *Case Law in the Limited-Exception Category.*

a. Exception Limited by Class of Persons.—Several courts have considered whether the right to academic freedom referenced in *Garcetti* extends to those other than university professors; these courts have all determined that any exception to *Garcetti* would extend only to professors. In *Mayer v. Monroe County Community School Corp.*,[143] the court considered an elementary school teacher's claims that her school district had violated her rights under the First Amendment.[144] In the course of a current events lesson on political demonstrations, Mayer told her students that she

---

141. *Id.* at 415.

142. It is perhaps worth noting that the decision in *Urofsky* was not unanimous. Judge Wilkinson's concurrence recognizes the right of academic freedom and the importance of protecting it from both a legal and philosophical perspective. *Id.* at 426–35 (Wilkinson, J., concurring).

143. 474 F.3d 477 (7th Cir. 2007).

144. *Id.* at 478.

had honked when she saw a sign that said "Honk for Peace."[145] Parents complained to the school, and when Mayer's contract was not renewed, she brought suit against the school board.[146] The district court ruled in favor of the school board; it found that, although Mayer spoke on a matter of public concern, she was acting as an employee, and so her speech was not protected under the First Amendment.[147] In affirming the district court, the appellate court held that, because the teacher spoke pursuant to her official duties, her speech was not protected under the First Amendment.[148] Though Mayer claimed academic freedom protected her speech, the court disagreed because "[t]he Constitution does not entitle teachers to present personal views to captive audiences against the instructions of elected officials."[149]

Likewise, in *Evans-Marshall*, the Sixth Circuit concluded that the First Amendment did not protect the in-class, curricular speech of a high school teacher.[150] Evans-Marshall came under fire after she gave a lesson on censorship that required her students to read and report on a commonly banned book and assigned Hermann Hesse's *Siddhartha* to her class.[151] After a tumultuous year in which the principal repeatedly criticized her curricular choices and teaching methods, Evans-Marshall's contract was not renewed.[152] She claimed that the firing was in retaliation for exercising her First Amendment rights, but the district court granted the school board's motion for summary judgment.[153] In affirming that decision on appeal, the court relied on *Mayer* and held that "*Garcetti's* caveat offer[ed] no refuge" because she was not a professor.[154]

b. Exception Limited by Scope of Official Duties.—Other courts have limited the potential for an academic freedom exception under *Garcetti* with a narrow understanding of what types of speech the exception might cover. These decisions have shown a strict reliance on the words of the *Garcetti* majority, and if the speech in question is not related to scholarship or teaching, the courts have rejected the claims.

A district court in Ohio considering the academic freedom claims of a university librarian found that his speech was unprotected under *Garcetti* in

145. *Id.*

146. *Id.*

147. Mayer v. Monroe Cnty. Cmty. Sch. Corp., No. 1:04-CV-1695-SEB-VSS, 2006 WL 693555, at *12 (S.D. Ind. Mar. 10, 2006).

148. *Mayer*, 474 F.3d at 480.

149. *Id.* at 479–80.

150. Evans-Marshall v. Bd. of Educ., 624 F.3d 332, 343 (6th Cir. 2010).

151. *Id.* at 334–35.

152. *Id.* at 335–36.

153. Evans-Marshall v. Bd. of Educ., No. 3:03cv091, 2008 WL 2987174, at *17 (S.D. Ohio July 30, 2008).

154. *Evans-Marshall*, 624 F.3d at 342–43.

Savage v. Gee.[155] Savage served on a university committee tasked with selecting a book for incoming freshman to read; Savage (allegedly to make a "sarcastic point" in response to one of his committee colleagues) suggested several books, one of which had a chapter characterizing "homosexuality as aberrant human behavior."[156] Several faculty members complained, and though the university cleared Savage of wrongdoing, he later claimed he had been constructively discharged in violation of his rights to free speech and academic freedom.[157] The district court denied his claims because Savage was acting pursuant to his official duties; the court also rejected his argument that he fell under the academic freedom exception in *Garcetti* because he was not engaged in scholarship or teaching.[158]

The Third Circuit rejected a professor's allegations that his First Amendment rights had been violated in *Gorum v. Sessoms*.[159] Gorum came under fire after an audit revealed that he had changed students' grades in violation of university policy; though the university committee investigating him did not recommend termination, the university eventually dismissed him after the University President, Sessoms, intervened.[160] Gorum claimed he had been dismissed in retaliation for exercising free speech critical of Sessoms, for acting as an advisor to a student athlete, and for rescinding an invitation for Sessoms to speak at a prayer breakfast.[161] In affirming the district court's grant of summary judgment for the defendants, the court concluded that Gorum had acted pursuant to his official duties and so his speech was not protected under *Garcetti*.[162] Though the court acknowledged the issues left unresolved by *Garcetti*, it held that any academic freedom exception would not apply because Gorum's speech was unrelated to scholarship or teaching.[163]

Finally, in *Abcarian v. McDonald*,[164] the Seventh Circuit rejected the claims of a medical school professor because his speech was also not teaching or scholarship.[165] Abcarian, the head of the department of surgery in a medical school, had a contentious relationship with his supervisors where they "clashed over a number of issues including risk management, faculty recruitment, compensation and fringe benefits."[166] Abcarian believed the university was engaged in an elaborate scheme to ruin his professional

155. 716 F. Supp. 2d 709, 710, 718 (S.D. Ohio 2010).
156. *Id.* at 710–11.
157. *Id.* at 713–18.
158. *Id.* at 717–18.
159. 561 F.3d 179, 182 (3d Cir. 2009).
160. *Id.* at 182–83.
161. *Id.* at 183–84.
162. *Id.* at 182, 186.
163. *Id.* at 186.
164. 617 F.3d 931 (7th Cir. 2010).
165. *Id.* at 933, 938 & n.5.
166. *Id.* at 933.

reputation, including a conspiracy to force him to agree to a medical malpractice settlement.[167] Abcarian sued, alleging that the defendants had violated his First Amendment right to free speech, an allegation the district court rejected after applying *Garcetti*.[168] The Seventh Circuit affirmed, concluding that Abcarian spoke pursuant to his official duties.[169] Abcarian, as head of the surgery department, "had significant authority and responsibility over a wide range of issues."[170] The court addressed the academic freedom exception in a footnote, but concluded that the topics of his speech (risk management, fees, etc.) "involved administrative policies that were much more prosaic than would be covered by principles of academic freedom."[171]

2. Problems with the Limited-Exception Category.—The two cases where the courts limited the potential exception in *Garcetti* based on the class of persons demonstrate a tendency to limit the potential academic freedom exception to the specific words in the *Garcetti* opinion. I believe the trend of restricting any potential academic freedom exception to professors is likely to continue. The language of *Garcetti* makes it difficult to argue that secondary school teachers should receive additional protection. Justice Souter's dissent specifically refers to "the teaching of a public university professor" and "academic freedom in public colleges and universities."[172] Given the language of *Garcetti*, this case law, and the historical and theoretical foundations of academic freedom, it seems most probable that any right to academic freedom would be specific to professors.

The three cases where the courts limited the exception by the scope of official duties present a more interesting and difficult case. As these three cases demonstrate, courts have taken a very strict view of what official duties might exempt a person from *Garcetti*'s rule. These cases all rely heavily on the language in *Garcetti*, and if the speech in question is not directly related to "scholarship or teaching," courts tend to reject those claims. Though the courts acknowledge the potential that speech in the academic workforce might fall outside the scope of *Garcetti*, these cases suggest that courts currently conceive of only a very narrow exception to *Garcetti*.

Granted, no court has applied *Garcetti* to protect speech related to scholarship or teaching because no court has been asked to decide such a case.[173] A recent district court decision, *Van Heerden v. Board of*

---

167. *Id.* at 933–34.

168. Abcarian v. McDonald, No. 08 C 3843, 2009 WL 596575, at *6 (N.D. Ill. Mar. 9, 2009).

169. *Abcarian*, 617 F.3d at 937–38.

170. *Id.* at 937.

171. *Id.* at 937, 938 n.5.

172. Garcetti v. Ceballos, 547 U.S. 410, 438 (2006) (Souter, J., dissenting).

173. *Kerr v. Hurd*, 694 F. Supp. 2d 817 (S.D. Ohio 2010), raised the issue of whether a medical school professor could be disciplined for teaching delivery methods with which the school administration disagreed. *See supra* text accompanying notes 127–29. But the court in that case

Supervisors of Louisiana State University & Agricultural & Mechanical College,[174] is perhaps the case that has come the closest to evaluating whether a professor's academic speech, made as part of his job, is protected under the First Amendment. Van Heerden, a university engineering professor, claimed that he had been terminated in violation of his First Amendment rights.[175] In the wake of Hurricane Katrina, van Heerden publically criticized the Army Corps of Engineers for the levee failure, both in an individual capacity and in his role as member of a team of scientists tasked with determining the cause of the extensive flooding.[176] University administrators warned van Heerden to stop his public criticism, and when he refused, his contract was not renewed.[177] Applying *Garcetti* to assess van Heerden's claim, the court held that he was not acting pursuant to his official duties when he criticized the levee failure, and so it denied the university's motion for summary judgment.[178]

This case is particularly interesting because it is far from clear that van Heerden was not acting pursuant to his official duties when he spoke. Van Heerden worked at the university as "a hurricane expert," he served on the team of scientists investigating the levee failure as part of a partnership between the university and the state, and his initial job description required him to serve on that research team.[179] The court's decision that van Heerden was not acting in his official capacity relies heavily on the fact that the university tried to distance itself from van Heerden when his public comments became critical.[180] It seems likely, if not probable, that another court could view a professor's research as part of a team, on which the professor is required to serve, as pursuant to his official duties.

This would have been an ideal case for the court to recognize an exception to *Garcetti*, and the court seemed primed to do so. In dicta, the court extolled the importance of academic freedom and noted that *Garcetti* "could lead to a whittling-away of academics' ability to delve into issues or express opinions that are unpopular, uncomfortable or unorthodox."[181] In the end, though, the court did not recognize an academic freedom exception to *Garcetti* and instead made a somewhat tortured reading of pursuant to official duties in order to protect the professor's speech. This unwillingness to recognize an exception to *Garcetti* only serves to highlight further the

---

chose not to apply *Garcetti* and instead used *Pickering–Connick* to evaluate the professor's claim. *See supra* text accompanying notes 131–34. So in that sense, no case decided under *Garcetti* has protected speech related to teaching or scholarship.

174. No. 3:10–CV–155–JJB–CN, 2011 WL 5008410 (M.D. La. Oct. 20, 2011).

175. *Id.* at *1–2.

176. *Id.* at *1.

177. *Id.* at *1–2.

178. *Id.* at *7.

179. *Id.* at *5–6.

180. *Id.* at *5.

181. *Id.* at *6.

need for the Supreme Court to step forward and articulate a specific theory of academic freedom that will protect the speech of professors.

IV. Developing a Constitutional Right to Academic Freedom

The Court has consistently claimed that the idea of academic freedom implicates additional constitutional considerations not found in other First Amendment cases, but it has never explained what those additional considerations are.[182] Thus, as shown above, courts trying to assess First Amendment claims of professors, teachers, and universities have been left to rely on the Court's rhetoric or their own circuit precedent to decide these cases. Most often, these courts do not posit what a constitutional right to academic freedom would encompass; instead, the decisions typically define academic freedom in the negative, holding that, whatever the right to academic freedom might be, the issues raised in the case at hand do not implicate it.

The methods that courts have used in academic freedom cases after *Garcetti*—applying no exception, using a *Pickering–Connick* exception, or creating a very limited exception to *Garcetti* based on scope of official duties—do not work. In the no-exception category, courts have completely ignored the language in *Garcetti* that suggested academic freedom claims may be different than other public employee speech cases.[183] Though the *Pickering–Connick* exception improves on that problem by at least acknowledging that cases arising in academic settings present additional concerns, the *Pickering–Connick* exception is still not an answer to the issue left unresolved in *Garcetti*. *Pickering* and *Connick* were both general First Amendment claims, and applying general First Amendment principles to academic freedom claims ignores the fact that there is something distinctive about the speech of public university professors.[184] Finally, the cases in the limited-exception category seem to create an exception to *Garcetti* that is too limited; no court has recognized or defined an academic freedom exception to *Garcetti*, and absent guidance from the Supreme Court, they may never do so.[185]

The concerns about academic freedom raised in the wake of *Garcetti* are not new concerns, but because *Garcetti* threatens to restrict the rights of professors in a new way, it has become even more important for the Supreme Court to flesh out what the right to academic freedom actually entails. Using the post-*Garcetti* cases as a guide, I believe that developing a coherent theory

---

182. *See supra* text accompanying notes 3–7; *see also* J. Peter Byrne, *Academic Freedom: A "Special Concern of the First Amendment,"* 99 YALE L.J. 251, 257 (1989) ("The Court has been far more generous in its praise of academic freedom than in providing a precise analysis of its meaning.").

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

184. *See supra* subpart III(B).

185. *See supra* subpart III(C).

of academic freedom requires addressing two primary concerns. First, is it correct or feasible to talk about a constitutional right to academic freedom or is it more appropriate to conceive only of professional academic freedom? Second, what speech should be covered under a right to academic freedom? Does academic freedom simply protect in-class speech or are there other types of speech that academic freedom should protect? In developing my proposed theory of academic freedom, I will address these questions in turn.

A. Constitutional Academic Freedom or Professional Academic Freedom

The AAUP's 1915 *Declaration of Principles* undoubtedly set the standard for professional academic freedom in the United States.[186] Since that time, however, the idea of academic freedom has become more pervasive, dating back to a Supreme Court opinion in the 1950s that rooted the idea of academic freedom in the First Amendment.[187] Though Van Alstyne grounds his specific theory of academic freedom in a constitutional right, some recent scholarship has questioned whether it is appropriate to even consider a constitutional standard of academic freedom. In their recent work *For the Common Good: Principles of American Academic Freedom*, Matthew Finkin and Robert Post argue that the concept of academic freedom conceptualized in "the 1915 *Declaration* differs fundamentally from the individual First Amendment rights that present themselves so vividly to the contemporary mind."[188] They focus their entire book on the concept of professional academic freedom, essentially arguing for a return to the fundamental values of academic freedom as espoused in the 1915 *Declaration*.[189] They are not alone in moving away from constitutional academic freedom. In his review of Finkin and Post's book, Stanley Fish praises their decision to focus on professional academic freedom.[190] Fish takes an even stronger stance, characterizing constitutional academic freedom as "a non-topic, a quixotic, doomed effort to transform the professional concerns of scholars and teachers into constitutional rights."[191] Joan DelFattore, after examining academic freedom post-*Garcetti*, also advocates "a renewed reliance . . . on professional standards that were once the only protection for academic freedom and that remain its best hope."[192]

I think these scholars underestimate the relevance and importance of constitutional academic freedom. As William Van Alstyne explains in *The*

186. *See* Areen, *supra* note 135, at 954 (observing that "modern scholars consider [the 1915 *Declaration*] the seminal statement of American academic freedom").

187. *See supra* note 4.

188. MATTHEW W. FINKIN & ROBERT C. POST, FOR THE COMMON GOOD: PRINCIPLES OF AMERICAN ACADEMIC FREEDOM 7 (2009).

189. *Id.* at 8–9.

190. Stanley Fish, *Academic Freedom: How Odd Is That?*, 88 TEXAS L. REV. 171, 183 (2009).

191. *Id.*

192. JOAN DELFATTORE, KNOWLEDGE IN THE MAKING: ACADEMIC FREEDOM AND FREE SPEECH IN AMERICA'S SCHOOLS AND UNIVERSITIES 244–67 (2010).

Specific Theory of Academic Freedom and the General Issue of Civil Liberty, the AAUP's original notion of academic freedom was quickly "pressed into the larger field of civil liberties."[193] The Supreme Court seized on the notion of academic freedom, incorporating the phrase into its sweeping statements about the importance of universities in ensuring the free exchange of ideas and advancement of knowledge.[194] The professional notion of academic freedom has not remained a separate, static idea. Instead, "the two standards have converged as to the scope of academic freedom."[195] Though Fish paints constitutional and professional academic freedom as a dichotomy, the intertwining of those two ideas over time does not mean that choosing to recognize a constitutional right to academic freedom requires abandoning all aspects of professional academic freedom.

Most importantly, the Supreme Court directly raised the question as to whether or not constitutional academic freedom exists in *Garcetti*. The dissent supports the idea that the Constitution protects academic freedom. Justice Souter wrote that he "ha[d] to hope that [the] majority does not mean to imperil First Amendment protection of academic freedom in public colleges and universities."[196] His language implies that First Amendment protection for academic freedom already exists. The majority, however, seems uncertain: "There is some argument that expression related to academic scholarship or classroom instruction implicates additional constitutional interests."[197] Quite simply, *Garcetti* forces the issue: does the First Amendment protect academic freedom or not? Decisions in other courts after *Garcetti* have shown that the constitutional protection of academic freedom is far from certain; focusing on professional academic freedom risks leaving the work of colleges and universities unprotected under the First Amendment. When the Supreme Court eventually decides how *Garcetti* applies in the academic workforce, constitutional academic freedom will offer better protection than none at all.

B. Speech Protected Under an Individual Right of Academic Freedom

The question then becomes what, exactly, this constitutional right to academic freedom should protect. Both the majority and the dissent in *Garcetti* seem to acknowledge that academic freedom would encompass teaching.[198] The majority also specifically mentions scholarship when it mentions activities that may implicate additional constitutional concerns.[199] While I agree any right to academic freedom should protect professors'

---

193. Van Alstyne, *supra* note 10, at 62–63.
194. *See supra* text accompanying notes 3–7.
195. Areen, *supra* note 135, at 985.
196. Garcetti v. Ceballos, 547 U.S. 410, 438 (2006) (Souter, J., dissenting).
197. *Id.* at 425 (majority opinion).
198. *Id.*; *id.* at 438 (Souter, J., dissenting).
199. *Id.* at 425 (majority opinion).

teaching and scholarship, I believe a right to academic freedom should go further and also protect the speech of professors outside the classroom when the professor uses his or her professional expertise in speech that supports the academic function of the university.

The specific theory of academic freedom rests on the idea that there is something unique about the role professors play in society: they are expected to challenge existing ideas carefully and responsibly and to share their work.[200] Empowering professors to engage in critical inquiry means more than simply giving professors freedom in the classroom.[201] Many facets of university life contribute to a professor's ability to think and act critically. It depends on the professor's freedom to conduct research freely, to engage with well-qualified and diverse colleagues, and to contribute to discussions about the university's curriculum.

I proposed above that a constitutional right to academic freedom should protect professors when they speak based on their professional expertise and that speech supports the academic function of the university. The difficulty, though, comes in defining what speech supports the academic function of a university. Though bright-line rules are often tempting because they can offer clear-cut solutions to difficult questions, I believe a categorical rule would be inappropriate here.[202] As the discussion of post-*Garcetti* academic freedom cases shows, courts rarely face academic freedom cases that are simple; none of those decisions presented a situation where a professor claimed his or her right to academic freedom had been violated because of classroom speech or published scholarship.[203] When considering these more complicated cases that are actually reaching the courts, a nuanced approach would allow courts to consider speech in its entirety rather than allowing one single factor to determine a case.

Still, giving some substance to the concept of academic function is necessary. Given the language in the majority and dissenting opinions of *Garcetti*, academic function should include teaching and scholarship.[204] But

---

200. Van Alstyne, *supra* note 10, at 77.

201. David M. Rabban, *A Functional Analysis of "Individual" and "Institutional" Academic Freedom Under the First Amendment*, 53 LAW & CONTEMP. PROBS. 227, 295 (1990).

202. *Garcetti* itself shows the danger of strict rules. The Court imposed a threshold question, whether the speech is pursuant to official duties, and provided no leeway for courts to protect speech if the answer to that question is yes. *Garcetti*, 547 U.S. at 421. The dissent in *Garcetti* recognized this problem and, while it acknowledged that "necessary judicial line-drawing sometimes looks arbitrary, any distinction obliges a court to justify its choice." *Id.* at 430 (Souter, J., dissenting).

203. *See supra* Part III.

204. For a recent suggestion that an academic freedom exception to *Garcetti* should encompass teaching and scholarship, see Carol N. Tran, *Recognizing an Academic Freedom Exception to the* Garcetti *Limitation on the First Amendment Right to Free Speech*, 45 AKRON L. REV. 949, 986–87 (2012). Tran argues that the exception to *Garcetti* "should be limited to faculty academic speech, which only includes scholarship and teaching." *Id.* at 987. In my opinion, an exception as limited as the one Tran proposes would leave professors' speech unprotected on other issues that are vital to the university, such as discussions of curriculum and faculty hiring.

a right to academic freedom that is limited to teaching, or even one limited to teaching and scholarship, would leave professors' speech on many important matters unprotected. The specific theory of academic freedom posits that a professor's right to academic freedom stems from his or her professional expertise in a specific area of study.[205] Professors clearly employ this expertise when they teach in a classroom or publish research on their specific subject matter, but they also bring their professional expertise to bear on other factors of university life. Universities are supposed to foster discussion and promote critical inquiry, and those functions of a university involve more than just teaching and scholarship and include things like having a strong curriculum and talented faculty. My proposed academic freedom exception to *Garcetti* would go beyond simply protecting professors when they talk in a classroom or publish research and would ensure that they are protected when they further the academic function of the university. This exception should include protecting professors' speech when they talk about the university's curriculum; when they challenge hiring decisions based on professional standards, that speech should be protected as well. Examples of borderline cases may help to clarify further what speech I believe academic freedom should protect.

For the first example, consider a situation where the professor's speech concerns the hiring decisions of the university.[206] A philosophy professor, who serves on the hiring committee, interviews candidates for an assistant professorship. The administration wants to offer the position to a candidate who the professor believes shows serious deficiencies in scholarship and is not qualified for the job. The question becomes, if the professor speaks out about the administration's chosen candidate, is the professor's speech protected? Under my theory of academic freedom, the answer would be yes. Admittedly, the professor's speech is not concerned with teaching or with his own research. But, as noted above, academic freedom depends on more than that. It relies, in part, on a professor's ability to engage with well-qualified professors; this strengthens the professor's own scholarship and the scholarship of his colleagues. It is also vital to the university's mission of teaching students to think critically that students learn from serious scholars. The professor's speech supports the academic function of the university in both of these ways, and so it should be constitutionally protected.

Next, consider an English professor at a public university. This professor specializes in nineteenth-century British literature, and she takes

205. *See supra* notes 10–11, 16–21 and accompanying text.

206. This situation is not a far-fetched hypothetical. In *Hong v. Grant*, the professor claimed that his right to free speech had been violated, in part, because of negative comments he made about professors as part of the tenure process. 516 F. Supp. 2d 1158, 1160, 1162 & n.4 (C.D. Cal. 2007). He also alleged that he had been retaliated against because he complained to the administration that the university relied too heavily on lecturers, instead of full professors, to teach undergraduate classes. *Id.* at 1160, 1163.

issue with the curriculum choices of the university.[207] The English department has moved away from offering courses in the classics, choosing instead to offer classes on topics such as the poetry of modern rap music and writing for the blogosphere. The professor complains to the dean about the curriculum, and when raises are announced, the professor is the only member of the English department not to receive a raise. Did the university violate the professor's rights when it punished her for her speech? I would answer yes. As a scholar, the professor has a duty to ensure that she and her colleagues maintain standards of professional integrity; as a teacher, she is responsible for ensuring the university produces students who have been taught to think, challenge, and question. Shaping the university's curriculum, though not specific to classroom speech or research, is an important part of a professor's professional responsibility; professors should be free to share their views on their areas of expertise without fear of retaliation.

Finally, consider a chemistry professor at a major public university. The school has a large, successful, and well-funded football program. One day, the professor walks into her large chemistry class and delivers a fifty-minute speech criticizing the administration for dedicating so much of the school's budget to the football team. Students complain, and the professor is disciplined. Is her speech protected? Under my proposed standard of protecting speech that supports the academic function of the university, her speech would not be protected. The professor is not relying on her professional expertise to criticize the university. Though she could argue that her speech supports the academic function of the university because she believes the university should redirect athletic funding to support student research, her speech does not relate to the university's mission the way the professors' speech did in the previous two examples. She is not teaching her students to be critical thinkers or preparing them to be responsible, engaged citizens. She is airing a personal complaint from her classroom platform, and academic freedom should not protect that kind of speech.

As with most judicial rules, deciding what speech supports the academic function of the university will not always provide clear-cut answers. But a strict categorical rule would be inappropriate. The post-*Garcetti* cases and the examples above show that the more difficult questions revolve around subjects somewhere between speech in the classroom and speech completely

---

207. Like the previous hypothetical, this example is intended to be a realistic one. In *Evans-Marshall*, the Sixth Circuit considered a teacher's claims that her First Amendment rights had been violated for speech related to her curriculum; the court held her speech was not protected. Evans-Marshall v. Bd. of Educ., 624 F.3d 332, 340 (6th Cir. 2010). Though the teacher in *Evans-Marshall* was a high school teacher, *id.* at 334, and secondary school teachers' academic freedom claims are often evaluated differently from those of professors, it is easy to imagine that a professor could face similar challenges in a university setting.

unrelated to scholarship.[208] As courts develop case law to flesh out the theory of academic freedom, they will need some leeway to explore this admittedly difficult-to-define middle area. But, as summarized by Professor David Rabban, "[t]he key issue . . . is not how a particular close case should be resolved, but the recognition that some intramural speech on matters beyond an individual's teaching and scholarship should be protected by a first amendment right of academic freedom."[209] *Garcetti*'s restriction of the general free speech rights of public employees makes it more important than ever for the Court to acknowledge a right to academic freedom that protects professors' speech made in support of the university's academic function.

V. Conclusion

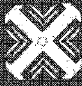
With *Garcetti v. Ceballos*, the Supreme Court put the status of academic freedom in the United States in a state of uncertainty. Subsequent court decisions have shown courts struggling to define the right to academic freedom in its aftermath and have cast a spotlight on the issues inherent in our understanding of academic freedom. The Court now faces an opportunity to articulate, for the first time, what the right to academic freedom truly entails. In order to preserve the unique function of American universities, the Court should recognize a constitutional right to academic freedom that protects professors' speech, as long as it is related to the academic purpose of the university. Without a constitutional protection rooted in the First Amendment, professors may lose their freedom to promote critical discussion and fail to fulfill their important role in our democracy.

—Lauren K. Ross

208. *See* Rabban, *supra* note 201, at 294 (observing that "disputes over university policies and personalities have far outnumbered classic academic freedom cases involving the content of teaching or scholarship" in an article written more than fifteen years before *Garcetti*).

209. *Id.* at 295.

* * *



JAMAIL CENTER FOR LEGAL RESEARCH
TARLTON LAW LIBRARY
THE UNIVERSITY OF TEXAS SCHOOL OF LAW

The Tarlton Law Library Oral History Series features interviews with outstanding alumni and faculty of The University of Texas School of Law.

Oral History Series

No. 1 - *Joseph D. Jamail, Jr.* 2005. $20
No. 2 - *Harry M. Reasoner* 2005. $20
No. 3 - *Robert O. Dawson* 2006. $20
No. 4 - *J. Leon Lebowitz* 2006. $20
No. 5 - *Hans W. Baade* 2006. $20
No. 6 - *James DeAnda* 2006. $20
No. 7 - *Russell J. Weintraub* 2007. $20
No. 8 - *Oscar H. Mauzy* 2007. $20
No. 9 - *Roy M. Mersky* 2008. $25

Forthcoming:

Gloria Bradford, Patrick Hazel, James W. McCartney,
Michael Sharlot, Ernest E. Smith, John F. Sutton, Jr.

Other Oral Histories Published by the Jamail Center for Legal Research

Robert W. Calvert (Texas Supreme Court Trilogy, Vol. 1). 1998. $20

Joe R. Greenhill, Sr. (Texas Supreme Court Trilogy, Vol. 2). 1998. $20

Gus M. Hodges (Tarlton Law Library Legal History Series, No. 3). 2002. $20

Corwin Johnson (Tarlton Law Library Legal History Series, No. 4). 2003. $20

W. Page Keeton (Tarlton Legal Bibliography Series, No. 36). 1992. $25

Jack Pope (Texas Supreme Court Trilogy, Vol. 3). 1998. $20

Order online at http://tarlton.law.utexas.edu/ click on Publications
or contact Publications Coordinator,
Tarlton Law Library, UT School of Law,
727 E. Dean Keeton St., Austin, TX 78705

phone (512) 471-6228; *fax* (512) 471-0243;
email tarltonbooks@law.utexas.edu

THE UNIVERSITY OF TEXAS SCHOOL OF LAW PUBLICATIONS
What the students print here changes the world

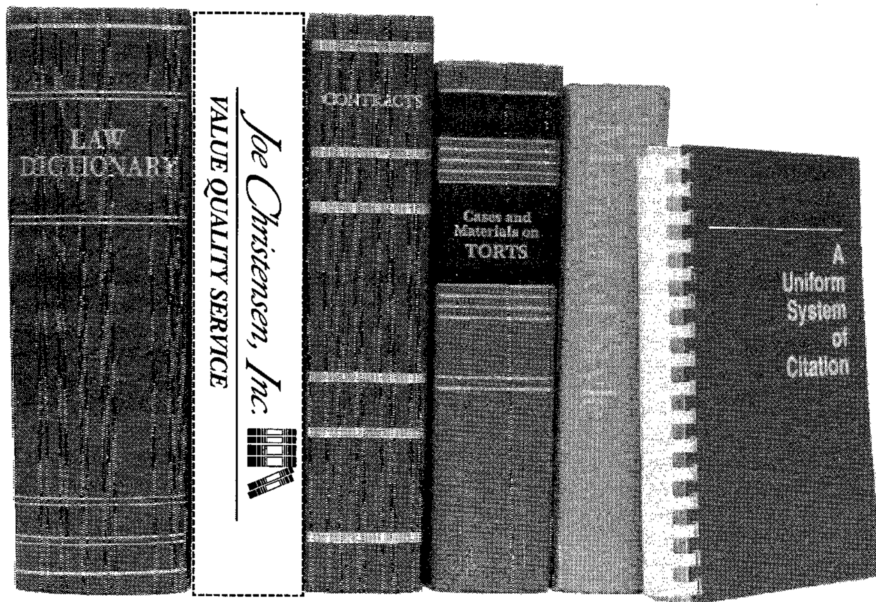
Journal	domestic/foreign
Texas Law Review http://www.TexasLRev.com	$47.00 / $55.00
Texas International Law Journal http://www.tilj.org	$45.00 / $50.00
Texas Environmental Law Journal http://www.texenrls.org/publications_journal.cfm	$40.00 / $50.00
American Journal of Criminal Law http://www.ajcl.org	$30.00 / $35.00
The Review of Litigation http://www.thereviewoflitigation.org	$30.00 / $35.00
Texas Journal of Women and the Law http://www.tjwl.org	$40.00 / $45.00
Texas Intellectual Property Law Journal http://www.tiplj.org	$25.00 / $30.00
Texas Hispanic Journal of Law & Policy http://www.thjlp.org	$30.00 / $40.00
Texas Journal On Civil Liberties & Civil Rights http://www.txjclcr.org	$40.00 / $50.00
Texas Review of Law & Politics http://www.trolp.org	$30.00 / $35.00
Texas Review of Entertainment & Sports Law http://www.tresl.net	$40.00 / $45.00
Texas Journal of Oil, Gas & Energy Law http://www.tjogel.org	$30.00 / $40.00

Manuals:

The Greenbook: Texas Rules of Form 12th ed. ISBN 1-878674-08-0
Manual on Usage & Style 11th ed. ISBN 1-878674-55-2

To order, please contact:
The University of Texas School of Law Publications
727 E. Dean Keeton St.
Austin, TX 78705 U.S.A.
Publications@law.utexas.edu

ORDER ONLINE AT:
http://www.texaslawpublications.com



We Complete the Picture.

In 1932, Joe Christensen founded a company based on Value, Quality and Service. Joe Christensen, Inc. remains the most experienced Law Review printer in the country.

Our printing services bridge the gap between your editorial skills and the production of a high-quality publication. We ease the demands of your assignment by offering you the basis of our business—customer service.

1540 Adams Street
Lincoln, Nebraska 68521-1819
Phone: 1-800-228-5030
FAX: 402-476-3094
email: sales@christensen.com

Value
Quality
Service

Your Service Specialists

* * *

* * *

* * *

* * *

* * *

* * *

* * *

Texas Law Review

The Greenbook: Texas Rules of Form

Twelfth Edition

A comprehensive guide for Texas citation, newly revised in 2010.

Texas Law Review Manual on Usage & Style

Twelfth Edition

A pocket reference guide on style for all legal writing.

Newly revised and released in Fall 2011

School of Law Publications
University of Texas at Austin
727 East Dean Keeton Street
Austin, Texas USA 78705
Fax: (512) 471-6988 Tel: (512) 232-1149
Order online: http://www.utexas.edu/law/publications

TEXAS
TLR
LAW REVIEW